LECTURE NOTES ON ADDITIVE COMBINATORICS

TUOMAS ORPONEN

ABSTRACT. These are the lecture notes for the course *Topics in additive combinatorics and incidence geometry*, given at the University of Helsinki in the Autumn semester of 2015. The material is harvested and brought together from various origins, see the list of references.

The students of the course were asked to choose a topic, give a presentation on it, and supply notes in LaTeX. These notes are, or will be, gathered in the appendices at the end of the document. As of 5/11/2015, Laura Venieri's notes have been added.

CONTENTS

1. Freiman's theorem

1.1. **The Cauchy-Davenport inequality.** In the first part of the lectures we will study the question: if $|A+A| \sim |A|$, then what can be said of the structure of A? The following easy proposition gives an indication of what is to be expected:

Proposition 1.1 (Cauchy-Davenport in $\mathbb{R}$)**.** *Let $A, B \subset \mathbb{R}$ be finite sets with $|A| = m$ and $|B| = n$. Then $|A+B| \geq m+n-1$, and equality holds, if and only if A and B are arithmetic progressions with the same gap.*

Proof. Write

$$A = \{a_1, \ldots, a_m\} \quad \text{and} \quad B = \{b_1, \ldots, b_n\}$$

with $a_1 < a_2 < \ldots < a_m$ and $b_1 < \ldots < b_n$. Then, consider the element $a_i + b_j \in A + B$ for some fixed $1 \leq i \leq m$ and $1 \leq j \leq n$. Observe that the set $A + B$ contains at least $i + j - 1$ elements less or equal to $a_i + b_j$, namely the i elements

$$a_k + b_1, \qquad 1 \leq k \leq i,$$

plus the $j - 1$ elements

$$a_i + b_k, \qquad 1 < k \leq j.$$

Similarly, there are $m + n - i - j$ strictly larger elements, namely the $m - i$ elements

$$a_k + b_j, \qquad i < k \leq m$$

plus the $n - j$ elements

$$a_m + b_k, \qquad j < k \leq n.$$

So, altogether, $A + B$ contains at least $(i + j - 1) + (m + n - i - j) = m + n - 1$ elements.

Next, assume that indeed $|A + B| = m + n - 1$, and write $A + B = \{c_1, \ldots, c_{m+n-1}\}$ with $c_1 < c_2 < \ldots < c_{m+n-1}$. Then, with $1 \leq i \leq m$ and $1 < j \leq n$ as before, we have

$$a_i + b_j = c_{i+j-1},$$

because, as we observed, $A+B$ contains at least $i+j-1$ elements less or equal to a_i+b_j, and also $m + n - i - j$ strictly larger ones. In particular,

$$a_1 + b_{i+1} = c_{i+1} = a_2 + b_i, \qquad 1 \leq i < n,$$

which implies $b_{i+1} - b_i = a_2 - a_1$ for $1 \leq i < n$. By definition, B is an arithmetic progression with gap $a_2 - a_1$. Similarly, $a_{i+1} + b_1 = c_{i+1} = a_i + b_2$ for $1 \leq i < m$, and this gives the same conclusion about A. □

1.2. **Freiman's theorem in $\mathbb{Z}$.** From the previous section, we know that if $|A+A| = 2|A|-1$, then A is an arithmetic progression. What if we relaxed the condition $|A + A| = 2n - 1$ to $|A + A| \leq C|A|$ for some constant $C > 0$ – which should be thought to be small in comparison with $|A|$. Then A need no longer resemble an arithmetic progression in the sense above. Indeed, fix integers $x_0, x_1, \ldots, x_d \in \mathbb{Z}$ and natural numbers $m_1, \ldots, m_d \in \mathbb{N}$, and consider

$$P := \left\{ x_0 + \sum_{j=1}^{d} \lambda_j x_j : 0 \leq \lambda_j \leq m_j - 1 \right\}.$$

A set of this form is called a *d-dimensional arithmetic progression*. Further, P is called *proper*, if $|P| = m_1 \cdots m_d$.

Exercise 1.2. If P is a proper d-dimensional arithmetic progression, then $|P + P| \leq 2^d |P|$.

Of course, if $P' \subset P$ is a large subset, $|P'| \sim |P|$, the previous exercise implies that $|P'+P'| \lesssim 2^d|P'|$. A deep result of Freiman states that the converse is also true – with worse constants, perhaps:

Theorem 1.3 (Freiman). *Suppose that $A \subset \mathbb{Z}$ satisfies $|A+A| \leq C|A|$. Then A is contained in a d-dimensional arithmetic progression P of size $|P| \leq K|A|$, where d and K only depend on C.*

Remark 1.4. Freiman's theorem also holds in $\mathbb{R}$ – or $\mathbb{C}$, or any *torsion-free group* – and the statement is easily reduced to $\mathbb{Z}$, at least as long as one does not care about best constants. This will be an exercise.

Remark 1.5. In Freiman's theorem, it is further possible to require that the arithmetic progression be proper: by a result of Gowers-Walters, any d-dimensional arithmetic progression P can be contained in a proper one P' with $|P'| \leq d^{d^3}|P|$. We will not prove the Gowers-Walters theorem in these lectures, however.

Unless otherwise noted, our presentation of Theorem 1.3 mixes ingredients from the lecture notes of B. Green [14], I. Ruzsa [29] and T. Tao [32].

2. Proof of Freiman's theorem

2.1. **Structure of the proof and Ruzsa calculus.** In order to explain the outline of the proof, we first need the following lemma of Ruzsa:

Lemma 2.1 (Ruzsa's covering lemma). *Let $A, B \subset \mathbb{R}$ be finite sets. Then, there exists a set $X \subset \mathbb{R}$ of cardinality $|X| \leq |A+B|/|A|$ such that*

$$B \subset X + A - A.$$

In other words, we can cover B by at most $|A+B|/|A|$ translates of $A-A$.

Proof. Let $X \subset B$ be maximal with the property that the sets $x+A$, $x \in X$, are disjoint. Since each of the disjoint sets $x+A$ is contained in $A+B$ and has size $|A|$, we have

$$|X| \leq \frac{|A+B|}{|A|}.$$

Furthermore, by maximality, if $b \in B$, then $(b+A) \cap (x+A) \neq \emptyset$ for some $x \in X$. This means that $b \in x+A-A$, and so $B \subset X+A-A$. □

Now, let us turn back to the proof of Freiman's theorem. Assume, for a moment, that we have found a proper d-dimensional arithmetic progression P such that d only depends on C and $|P+A| \sim_C |A| \sim_C |P|$. Then, by Ruzsa's covering lemma, $A \subset X+P-P$ for some set X with

$$|X| \leq |P+A|/|P| \lesssim_C 1.$$

But since P is a d-dimensional arithmetic progression, also $P-P$ is a d-dimensional arithmetic progression, and hence $X+P-P$ is an arithmetic progression of dimension $|X|+d \sim_C 1$ and size at most $2^d|X||P| \sim_C |A|$.

So, the main task is to find a d-dimensional arithmetic progression P with $|P+A| \sim_C |A| \sim_C |P|$. To this end, we will demonstrate that there exists a d-dimensional arithmetic progression $P \subset 2A-2A$ with $|P| \sim_C |2A-2A|$. This will suffice because of the following inequalities of Plünnecke-Ruzsa:

Theorem 2.2 (Plünnecke-Ruzsa inequalities). *Assume that $A \subset \mathbb{R}$ with $|A + A| \leq C|A|$. Then*

$$|kA \pm lA| \leq C^{k+l}|A|, \qquad k, l \in \mathbb{N}.$$

We will prove these inequalities in a moment – using a recent simple approach of Petridis [27] – but let us first see how they help with Freiman's theorem. With $k = 3$ and $l = 2$ the inequality gives

$$|A| \leq |2A - 2A| \sim_C |P| \leq |P + A| \leq |2A - 2A + A| = |3A - 2A| \lesssim_C |A|$$

Hence, $|A| \sim_C |P| \sim_C |P + A|$, and the proof of Freiman's theorem will be complete.

Proof of Theorem 2.2. We will prove the following statement for two finite sets $A, B \subset \mathbb{R}$: if $|A + B| \leq C|A|$, then $|kB \pm lB| \leq C^{k+l}|A|$ for all $k, l \in \mathbb{N}$. The claim then follows by choosing $B = A$.

We start by choosing a subset $A' \subset A$ such that the ratio $|A' + B|/|A|$ is minimised. Let C' be this ratio, so that $C' \leq C$,

$$|A' + B| = C'|A'|,$$

and

$$|A'' + B| \geq C'|A''|, \qquad A'' \subset A'. \tag{2.3}$$

With this notation, we next prove the following claim:

Claim 2.4. *If $R \subset \mathbb{R}$, then $|A' + B + R| \leq C'|A' + R|$.*

Proof of Claim. We proceed by induction on the cardinality $|R|$. For $|R| = 1$,

$$|A' + B + R| = |A' + B| \leq C'|A'| = C'|A' + R|.$$

Then, suppose that the claim holds for all sets R with $|R| \leq r$, and we are given a set R' with $|R'| = r + 1$. Write $R' = R \cup \{x\}$ for some $x \in \mathbb{R}$, so that $|R| \leq r$. We wish to bound the cardinality of $A' + B + R'$, and the first instinct is probably to write

$$A' + B + R' = (A' + B + R) \cup (A' + B + x), \tag{2.5}$$

and then estimate $|A' + B + R'| \leq |A' + B + R| + |A' + B| \leq C'|A' + R| + C'|A'|$ by induction. Clearly, this is not quite good enough, and the inefficiency stems from the fact that the sets $A' + B + R$ and $A' + B + x$ can overlap quite a bit. To mend this, we write $A' + B + R'$ as a union of two sets, which do not overlap quite so much. Let

$$A'' := \{a \in A' : a + x \in A' + R\} \subset A',$$

and observe that $A'' + B + x \subset A' + B + R$. Thus,

$$A' + B + R' \subset (A' + B + R) \cup [(A' + B + x) \setminus (A'' + B + x)],$$

which turns out to be critically better than (2.5). Namely, now we can use induction, and the minimality hypothesis for A', to estimate

$$|A' + B + R'| \leq |A' + B + R| + |A' + B| - |A'' + B| \leq C'|A' + R| + C'|A'| - C'|A''|,$$

Finally, observe that

$$|A' + R| + |A'| - |A''| = |A' + R| + |(A' + x) \setminus (A'' + x)| = |A' + R'|,$$

because the union $(A' + R) \cup [(A' + x) \setminus (A'' + x)]$ is disjoint by definition of A''. This completes the proof of the claim. □

Apply the claim with $R = B$ to obtain $|A' + 2B| \leq C'|A' + B| = (C')^2|A'|$, and more generally

$$|A' + kB| = |A' + B + (k-1)B| \leq C'|A' + (k-1)B| \leq \ldots \leq (C')^k|A'|.$$

We are almost done, but we need one additional easy lemma:

Lemma 2.6 (Ruzsa's triangle inequality)**.** *For any three finite sets $U, V, W \subset \mathbb{R}$,*

$$|U||V - W| \leq |U + V||U + W|.$$

Proof. Consider the mapping $\varphi\colon U \times (V - W) \to (U + V) \times (U + W)$ defined by

$$\varphi(u, x) = (u + v(x), u + w(x)),$$

where $v(x) \in V$ and $w(x) \in W$ are some points such that $v(x) - w(x) = x$. The claim follows, if we manage to check that φ is injective, so assume that $\varphi(u, x) = \varphi(u', x')$. Then $v(x) - v(x') = u' - u = w(x) - w(x')$, which gives

$$x = v(x) - w(x) = v(x') - w(x') = x'.$$

Then of course $v(x) = v(x')$, and so $u + v(x) = u' + v(x')$ implies $u = u'$. □

Exercise 2.7. Prove the following following inequalities of Freiman-Pigarev:

$$|A + A|^{3/4} \leq |A - A| \leq |A + A|^{4/3}.$$

Now we can finish the proof of the "minus sign" version of the Plünnecke-Ruzsa inequalities. Recall that we already established $|A' + kB| \leq (C')^k|A'|$ for $k \in \mathbb{N}$. Applying the triangle inequality to $U = A'$, $V = kB$ and $W = lB$, we have

$$|A'||kB - lB| \leq |A' + kB||A' + lB| \leq (C')^{k+l}|A'|^2.$$

In particular, $|kB - lB| \leq (C')^{k+l}|A'| \leq C^{k+l}|A|$ as claimed. To establish the "plus sign version", we obviously need a corresponding "plus sign" version of the triangle inequality above:

Proposition 2.8. *For any three finite sets $U, V, W \subset \mathbb{R}$,*

$$|U||V + W| \leq |U + V||U + W|.$$

Proof. The argument is due to Todd Cochrane, and I picked it up from the comments in Tim Gowers' blog. It is surprisingly different from the proof with the minus sign! Let $U' \subset U$ be a subset minimising the ratio $|U' + V|/|U'|$, and call this ratio C, whence in particular

$$C \leq \frac{|U + V|}{|U|}. \tag{2.9}$$

Then we are in a position to apply Claim 2.4 to $R = W$, so that

$$|U' + V + W| \leq C|U' + W|,$$

and finally, using also (2.9),

$$\begin{aligned} |U||V + W| &\leq |U||U' + V + W| \\ &\leq C|U||U' + W| \\ &\leq |U + V||U' + W| \\ &\leq |U + V||U + W|, \end{aligned}$$

as claimed. □

This completes the proof of the Plünnecke-Ruzsa inequalities. □

For future reference, we stitch together the "+" and "−" versions of the triangle inequality:

Corollary 2.10. *For any three sets U, V, W in any abelian group,*

$$|U||V \pm W| \leq |U+V||U+W|.$$

The following exercise is a (simple) lemma of Bourgain, which we will need later:

Exercise 2.11. Assume that A_1, A_2, A_3 are subsets of an Abelian group, satisfying

$$|A_1 \cap A_3| \geq \frac{|A_1|}{K} \quad \text{and} \quad |A_2 \cap A_3| \geq \frac{|A_2|}{K}$$

and

$$|A_i + A_i| \leq K|A_i|, \qquad i \in \{1,2,3\}.$$

Then

$$|A_1 + A_2| \leq K^5|A_3|.$$

2.2. **Freiman homomorphisms and reduction to $\mathbb{Z}_q$.** During the proof of Freiman's theorem, the constant C will be the one in $|A+A| \leq C|A|$, and from this point on, we will abbreviate $\sim_C$ and $\lesssim_C$ to $\sim$ and $\lesssim$. Recall that the aim is to find a d-dimensional arithmetic progression $P \subset 2A - 2A$ with $|P| \sim |2A-2A|$. It turns out – in the next section – that we can do this under the assumption that $A \subset \mathbb{Z}_q$, where q is a prime with $|A| = \alpha q$ and $\alpha > 0$ is a constant depending only on C. So, the task of the present section is to reduce matters to that case. To this end, we will use *Freiman homomorphisms*.

Definition 2.12. Let G, H be Abelian groups, and let $A \subset G$. A mapping $\phi \colon A \to H$ is called a *Freiman homomorphism of order k*, if the following holds. Whenever $x_1, \ldots, x_k$, $x_{k+1}, \ldots, x_{2k} \in A$, and

$$x_1 + \ldots + x_k = x_{k+1} + \ldots + x_{2k},$$

then

$$\phi(x_1) + \ldots + \phi(x_k) = \phi(x_{k+1}) + \ldots + \phi(x_{2k}).$$

If ϕ has an inverse, which is also a Freiman homomorphism of order k, then ϕ is called a *Freiman isomorphism of order k*. Two sets $A \subset G$ and $B \subset H$ are *Freiman k-isomorphic*, denoted $A \cong_k B$, if there exists a Freiman isomorphism of order k between A and B.

The next proposition, the proof of which I found in lecture notes of Ruzsa, implies that any set $A \subset \mathbb{Z}$ with small doubling (i.e. $|A+A| \leq C|A|$) has a large subset, which is Freiman k-isomorphic to a subset of $\mathbb{Z}_q$, where $|A| \sim q$.

Proposition 2.13 (Ruzsa)**.** *Assume that $A \subset \mathbb{Z}$ and $k \geq 2$ is a natural number. Then, for any $q \geq 4|kA - kA|$ there exists a set $A' \subset A$ with $|A'| \geq |A|/k$, which is Freiman k-isomorphic to a set in $\mathbb{Z}_q$.*

Proof. Define

$$\varphi(x) := [\xi x] \pmod q, \qquad x \in \mathbb{Z},$$

where $[\xi x]$ denotes the integer part of ξx, and $\xi \in [0, q]$ is a real number to be chosen later. Write $\{\xi x\}$ for the fractional part of ξx, so that

$$\xi x = [\xi x] + \{\xi x\}, \tag{2.14}$$

and define

$$A_j := A_j(\xi) := \left\{ x \in A : \frac{j-1}{k} \leq \{\xi x\} < \frac{j}{k} \right\}, \qquad 1 \leq j \leq k.$$

For a fixed ξ, one of the sets A_j must satisfy $|A_j| \geq m/k$, and we will choose $A' = A_j$ for this j.

In order to verify that φ restricted to A_j is a Freiman isomorphism of order k, for a suitable ξ, we must demonstrate that

$$\varphi(x_1) + \ldots + \varphi(x_k) = \varphi(x_{k+1}) + \ldots + \varphi(x_{2k}) \tag{2.15}$$

is equivalent to

$$x_1 + \ldots + x_k = x_{k+1} + \ldots + x_{2k} \tag{2.16}$$

for all $x_1, \ldots, x_{2k} \in A_j$. By definition of φ, (2.15) means that

$$[\xi x_1] + \ldots + [\xi x_k] \equiv [\xi x_{k+1}] + \ldots + [\xi x_{2k}] \pmod{q}. \tag{2.17}$$

Before getting properly started, let us quickly observe (the rather trivial fact) that the equivalence of (2.15) and (2.16) implies the injectivity of φ, so that the inverse φ^{-1} exists. Indeed, if $\varphi(x_1) = \varphi(x_2)$ for some $x_1, x_2 \in A_j$, then $\varphi(x_1) + \ldots + \varphi(x_1) = \varphi(x_2) + \ldots + \varphi(x_2)$, where the sum is k-fold on both sides. Now the implication from (2.15) to (2.16) gives that $kx_1 = kx_2$, and so $x_1 = x_2$.

Now, we proceed to verify the equivalence of (2.15) and (2.16). In both implications, it is useful to write, using (2.14),

$$\begin{aligned} \sum_{i=1}^{k} ([\xi x_i] - [\xi x_{i+k}]) &= \sum_{i=1}^{k} (\xi x_i - \{\xi x_i\} - (\xi x_{i+k} - \{\xi x_{i+k}\})) \\ &= \xi \sum_{i=1}^{k} (x_i - x_{i+k}) - \sum_{i=1}^{k} (\{\xi x_i\} - \{\xi x_{i+k}\}). \end{aligned} \tag{2.18}$$

Now, assume (2.16). Then the first sum on line (2.18) vanishes. The second sum is also small: recalling that $x_i, x_{i+k} \in A_j$, the absolute value of each difference $\{\xi x_i\} - \{\xi x_{i+k}\}$ is strictly less than $1/k$. Hence, the absolute value of the whole second sum on line (2.18) is strictly less than one. But the left hand side is an integer, so it must equal zero. This is even stronger than (2.17) (because we did not need to take $\pmod q$ to make the left and right hand sides agree).

Next, assume (2.17). This means that the left hand side of equation (2.18) is lq for some integer l, whereas the right hand side has the form $\xi y + \delta$ for some $y \in kA - kA$ and $|\delta| < 1$ (using again that $x_i, x_{i+k} \in A_j$). We want to infer that $y = 0$ (which is precisely (2.16)), so we need to choose $\xi \in [0, q]$ so that the equation $lq = \xi y + \delta$, $y \in (kA - kA) \setminus \{0\}$, can never occur for any $l \in \mathbb{Z}$ or $y \in (kA - kA) \setminus \{0\}$. First of all, if the equation holds, and recalling that $\xi \in [0, q]$, we have

$$|l|q = |\xi y + \delta| \leq |\xi y| + 1 \leq |y|q + 1,$$

so we only need to consider $|l| \leq |y| + 1$. Now, in terms of ξ, the equation is equivalent to

$$\xi = \frac{lq - \delta}{y}, \qquad |l| \leq |y| + 1, \ y \in (kA - kA) \setminus \{0\}.$$

Recalling that $|\delta| < 1$, the possible solutions ξ can be found in the union of the intervals $((lq/y - 1)/|y|, (lq/y + 1)/|y|)$, where $|l| \leq |y| + 1$ and $y \in (kA - kA) \setminus \{0\}$. These

intervals have total length $< |kA - kA| \cdot (2/|y|) \cdot (1 + |y|) \leq 4|kA - kA|$. Hence, as soon as $4|kA - kA| \leq q$, there exists a point $\xi \in [0, q]$, which is not contained in any of these intervals, and then, with this choice of ξ and the corresponding choice of $A' = A_j(\xi)$, the restriction of φ to A' is a Freiman isomorphism of order k. □

Recall that the aim of this section is to reduce the proof of Freiman's theorem to the case where $A \subset \mathbb{Z}_q$ and q is a prime with $|A| \geq \alpha N$. Assume for a moment that we have already established the theorem in that case. More precisely, assume that we have proved the following result (which is the content of the next sections):

Theorem 2.19. *Assume that that q is prime, and $A \subset \mathbb{Z}_q$ is a set with cardinality $|A| \geq \alpha q$ such that $|A + A| \leq C|A|$. Then $2A - 2A$ contains a proper d-dimensional arithmetic progression P of size $|P| \geq \beta q$, where $d = d(\alpha, C) \in \mathbb{N}$ and $\beta = \beta(\alpha, C) > 0$ only depend on α and C.*

Now, let us see how to combine Theorem 2.19 with Proposition 2.13 and the Plünnecke-Ruzsa inequalities to complete the proof of Freiman's theorem in $\mathbb{Z}$.

Proof of Theorem 1.3. The assumption of Freiman's theorem is that $A \subset \mathbb{Z}$ with $|A + A| \leq C|A|$. It follows from the Plünnecke-Ruzsa inequalities that $|8A - 8A| \leq C^{16}|A|$, and by the previous proposition, for any prime $4C^{16}|A| < q < 8C^{16}|A|$,[1] a subset $A' \subset A$ of cardinality $|A'| \geq |A|/8$ is Freiman 8-isomorphic to a set $X \subset \mathbb{Z}_q$ of cardinality

$$|X| = |A'| \geq |A|/8 \geq \frac{q}{64C^{16}}.$$

Now, since also

$$|X + X| = |A' + A'| \leq |A + A| \leq C|A| \leq 8C|X|,$$

where the first equation is a consequence of A' and X being Freiman isomorphic, Theorem 2.19 can be applied to X: there exists a proper d-dimensional arithmetic progession $P' \subset 2X - 2X$ of size $|P'| \geq \beta q$, where $d = d((64C^{16})^{-1}, 8C)$ and $\beta = \beta((64C^{16})^{-1}, 8C)$ only depend on C.

Next, this arithmetic progression is pulled back inside $2A' - 2A'$. To this end, we verify that $2A' - 2A' \cong_2 2X - 2X$. This is quite mechanical, so the reader may wish to skip the next few paragraphs. Let $\varphi \colon A' \to X$ be the Freiman isomorphism of order 8. We claim that the mapping $\psi \colon 2A' - 2A' \to 2X - 2X$ given by

$$\psi(x_1 + x_2 - x_3 - x_4) = \varphi(x_1) + \varphi(x_2) - \varphi(x_3) - \varphi(x_4)$$

is a well-defined Freiman 2-isomorphism between $2A' - 2A'$ and $2X - 2X$. To check that ψ is well-defined, assume that $x_1 + x_2 - x_3 - x_4 = y_1 + y_2 - y_3 - y_4$. Then $x_1 + x_2 + y_3 + y_4 = y_1 + y_2 + x_3 + x_4$, so that $\varphi(x_1) + \varphi(x_2) + \varphi(y_3) + \varphi(y_4) = \varphi(y_1) + \varphi(y_2) + \varphi(x_3) + \varphi(x_4)$. This means that

$$\psi(x_1 + x_2 - x_3 + x_4) = \psi(y_1 + y_2 - y_3 - y_4),$$

so the value of ψ is independent of the representation of a point in $2A' - 2A'$.

To check that ψ is a Freiman 2-isomorphism, we fix points $z_1, \ldots, z_4 \in 2A' - 2A'$ and represent them as $z_i = x_1^i + x_2^i - x_3^i - x_4^i$. Then, by the definition of ψ,

$$\psi(z_1) + \psi(z_2) = \psi(z_3) + \psi(z_4),$$

[1] Such a prime exists by Bertrand's postulate, which states that there always exists a prime strictly between n and $2n$.

if and only if

$$[\varphi(x_1^1)+\varphi(x_2^1)-\varphi(x_3^1)-\varphi(x_4^1)]+[\varphi(x_1^2)+\varphi(x_2^2)-\varphi(x_3^2)-\varphi(x_4^2)]$$
$$=[\varphi(x_1^3)+\varphi(x_2^3)-\varphi(x_3^3)-\varphi(x_4^3)]+[\varphi(x_1^4)+\varphi(x_2^4)-\varphi(x_3^4)-\varphi(x_4^4)].$$

By rearranging the terms and using the Freiman 8-isomorphism property of φ, this is eventually equivalent with $z_1+z_2=z_3+z_4$ as desired.

We have established that $\psi\colon 2A'-2A' \cong_2 2X-2X$, and now we claim that $P:=\psi^{-1}(P')\subset 2A'-2A'\subset 2A-2A$ is a proper d-dimensional arithmetic progression of size

$$|P|=|P'|\geq \beta q > 4\beta C^{16}|A|\geq 4\beta C^{12}|2A-2A|.$$

This is the content of the following lemma:

Lemma 2.20. *Let G,H be Abelian groups, and let $\psi\colon P\cong_2 \psi(P)\subset H$ be a Freiman 2-isomorphism, where $P\subset G$ is a proper d-dimensional arithmetic progression of size K. Then $\psi(P)$ is a proper d-dimensional arithmetic progression of size K.*

Proof. Write

$$P=\left\{x_0+\sum_{j=1}^d \lambda_j x_j : 0\leq \lambda_j\leq m_j-1\right\},$$

and define

$$x_0':=\psi(x_0)\quad\text{and}\quad x_j':=\psi(x_0+x_j)-\psi(x_0).$$

We claim that

$$\psi(P)=\left\{x_0'+\sum_{j=1}^d \lambda_j x_j' : 0\leq \lambda_j\leq m_j-1\right\},$$

and this is achieved by demonstrating that

$$\psi\left(x_0+\sum_{j=1}^d \lambda_j x_j\right)=x_0'+\sum_{j=1}^d \lambda_j x_j'. \tag{2.21}$$

The properness of $\psi(P)$ then simply follows from the fact that ψ is injective. We prove (2.21) by induction on $r=\sum_j \lambda_j$. For $r\in\{0,1\}$, this is clear by definition: if $r=0$, then $x_0+\sum\lambda_j x_j=x_0$ and $\psi(x_0)=:x_0'$. And if $r=1$, then $x_0+\sum\lambda_j x_j=x_0+x_j$ for some j, whence $\psi(x_0+x_j)=\psi(x_0)+x_j'=x_0'+x_j'$.

Assume next that the claim holds for some $r\geq 1$ (and smaller values), and fix

$$p:=x_0+\sum_{j=1}^d \lambda_j x_j\in P,\qquad \sum_{j=1}^d \lambda_j=r+1\geq 2.$$

Now, one can either find two elements $\lambda_{j_1},\lambda_{j_2}\geq 1$, $j_1\neq j_2$, or then just one element $\lambda_{j_1}\geq 2$. In the former case, define $\lambda':=\lambda_{j_1}$ and $\lambda'':=\lambda_{j_2}$. In the latter case, write $j_2:=j_1$ and choose $\lambda',\lambda''\geq 1$ arbitrarily so that $\lambda'+\lambda''=\lambda_{j_1}$. With this notation, define

$$u:=p-\lambda' x_{j_1},\quad v:=p-\lambda'' x_{j_2}\quad\text{and}\quad w:=p-\lambda' x_{j_1}-\lambda'' x_{j_2},$$

so that u, v, w are all points in P with the corresponding $\sum \lambda_j$ -sums being $\leq r$. We denote the λ_j-coefficients of u, v, w by $\lambda_j^u, \lambda_j^v, \lambda_j^w$. Since $p + w = u + v$ and ψ is a Freiman homomorphism of order 2, we find

$$\begin{aligned}\psi(p) &= \psi(u) + \psi(v) - \psi(w)\\ &= \left(x_0' + \sum_{j=1}^d \lambda_j^u x_j'\right) + \left(x_0' + \sum_{j=1}^d \lambda_j^v x_j'\right) - \left(x_0' + \sum_{j=1}^d \lambda_j^w x_j'\right) = x_0' + \sum_{j=1}^d \lambda_j x_j',\end{aligned}$$

and the induction is complete. □

Continuing where we left off before the proof of the lemma, we have now established that $P := \psi^{-1}(P') \subset 2A - 2A$ is a proper d-dimensional arithmetic progression of size $|P| \geq 4\beta C^{12}|2A - 2A|$. This yields (as we already saw in Section 2.1)

$$\begin{aligned}|A| \leq |2A - 2A| &\leq (4\beta C^{12})^{-1}|P| \leq (4\beta C^{12})^{-1}|P + A|\\ &\leq (4\beta C^{12})^{-1}|2A - 2A + A|\\ &= (4\beta C^{12})^{-1}|3A - 2A| \leq (4\beta C^7)^{-1}|A|\end{aligned}$$

by another application of the Plünnecke-Ruzsa inequalities. Thus, we have shown that $|A| \sim |P| \sim |P + A|$, and so $A \subset X + P - P$ with $|X| \lesssim 1$ by Ruzsa's covering lemma 2.1. Here $P - P$ is a d-dimensional arithmetic progression of size $|P - P| \leq 2^d|P| \lesssim |A|$, and X is an $|X|$-dimensional arithmetic progression of size $|X| \lesssim 1$. All in all, A can be covered by a $(d+|X|)$-dimensional arithmetic progression of dimension d and size $\lesssim |A|$. The proof of Freiman's theorem is complete. □

2.3. **Fourier analysis on $\mathbb{Z}_q$.** In the previous sections, we reduced the proof of Freiman's theorem in $\mathbb{Z}$ to the following statement in $\mathbb{Z}_q$:

Theorem 2.22. *Assume that that q is prime, and $A \subset \mathbb{Z}_q$ is a set with cardinality $|A| \geq \alpha q$ such that $|A + A| \leq C|A|$. Then $2A - 2A$ contains a proper d-dimensional arithmetic progression P of size $|P| \geq \beta q$, where $d = d(\alpha, C) \in \mathbb{N}$ and $\beta = \beta(\alpha, C) > 0$ only depend on α and C.*

The purpose of this section is to prove Theorem 2.22 using a Fourier-analytic technique. The proof has two parts: the first if to locate something called a *Bohr neighbourhood* $B(K, \delta) \subset 2A-2A$, and the second is to find a d-dimensional arithmetic progression inside $B(K, \delta)$. The Fourier analysis appears in the first part. Here is the definition of a Bohr neighbourhood:

Definition 2.23. Let $K \subset \mathbb{Z}_q$ and $\delta > 0$. The *Bohr δ-neighbourhood of K* is

$$B(K, \delta) := \left\{\xi \in \mathbb{Z}_q : \left\|\frac{\xi x}{q}\right\| \leq \delta \text{ for all } \xi \in K\right\}.$$

Here $\|\cdot\|$ stands for the distance to the nearest integer.

Remark 2.24. It makes no difference whether or not we interpret the multiplication rx above as $\xi x \in \mathbb{Z}$ or $\xi x (\mathrm{mod}\, q)$. Since

$$\frac{\xi x}{q} = \frac{lq + \xi x (\mathrm{mod}\, q)}{q} = l + \frac{\xi x (\mathrm{mod}\, q)}{q}$$

for some $l \in \mathbb{Z}$, we clearly have $\|rx/q\| = \|rx (\mathrm{mod}\, q)/q\|$.

Observe further that $\|\xi x/q\| = 0$, if and only if $e^{-2\pi i\xi x/q} = 1$. Hence, given $\kappa > 0$, there exists $\delta > 0$ depending only on κ such that $\|\xi x/q\| < \delta$ implies $|e^{-2\pi i\xi x/q} - 1| < \kappa$. In particular,

$$B(K,\delta) \subset \tilde{B}(K,\kappa) := \{x \in \mathbb{Z}_q : |e^{-2\pi i\xi x/q} - 1| < \kappa \text{ for all } \xi \in K\}.$$

Now, the two propositions, which combine to give Theorem 2.22 are the following:

Proposition 2.25 (Bohr neighbourhood inside $2A-2A$)**.** *Assume that q is prime, and $A \subset \mathbb{Z}_q$ is a set with cardinality $|A| \geq \alpha q$ such that $|A+A| \leq C|A|$. Then, $2A-2A$ contains a Bohr neighbourhood $B(K,\delta)$, where $|K| \leq 8C/\alpha$ and $\delta \in (0,1/2)$ is independent of C.*

Proposition 2.26 (Arithmetic progression inside $B(K,\delta)$)**.** *Let $K \subset \mathbb{Z}_q$ be a set with $|K| = k$, and let $\delta \in (0,1/2)$. Then, $B(K,\delta)$ contains a proper d-dimensional arithmetic progression P with $d = k$ and $|P| \geq (\delta/k)^k q$.*

As we stated earlier, the first proposition will be proved using Fourier analysis on $\mathbb{Z}_q$ (and the second one will be established in the next and final section). Here is the basic definition:

Definition 2.27 (The discrete Fourier transform)**.** Given any function $f\colon \mathbb{Z}_q \to \mathbb{R}$, we define the Fourier transform of f by

$$\hat{f}(\xi) := \sum_{x\in\mathbb{Z}_q} f(x)e^{2\pi ix\xi/q}, \qquad \xi \in \mathbb{Z}_q.$$

To save some typing, define $\omega := e^{2\pi i/q}$, so that $e^{2\pi ix\xi/q} = \omega^{x\xi}$. The following standard properties of $f \mapsto \hat{f}$ are an excercise:

Proposition 2.28. *Let $f,g\colon \mathbb{Z}_q \to \mathbb{R}$. Then*

(i) $f(x) = q^{-1}\sum_\xi \hat{f}(\xi)\omega^{-x\xi}$.

(ii) $\sum_x f(x)g(x) = q^{-1}\sum_\xi \hat{f}(\xi)\overline{\hat{g}(\xi)}$.

(iii) $\widehat{(f*g)}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$, *where $f*g$ is the convolution*

$$f*g(x) := \sum_{y\in\mathbb{Z}_q} f(y)g(x-y).$$

(iv) *If $g(x) = \overline{f(-x)}$, then $\hat{g}(\xi) = \overline{\hat{f}(\xi)}$.*

Most of the time, the relevant functions will be characteristic functions of sets $B \subset \mathbb{Z}_q$, or their convolutions. Abusing notation, we simply write B instead of χ_B or $\mathbf{1}_B$ for the characteristic function of B. Thus

$$B(x) = \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the Fourier transform of the characteristic function is abbreviated to $\hat{B}$.

Before starting the proof of Proposition 2.25 in earnest, we establish a few useful identities and equivalences. First, it follows immediately from Proposition 2.28(ii) applied to $f = B = g$ that

$$\sum_{\xi\in\mathbb{Z}_p} |\hat{B}(\xi)|^2 = q\sum_{x\in\mathbb{Z}_p} B(x) = q|B|. \tag{2.29}$$

Next, we will establish the following equivalence:

$$x \in 2B - 2B \quad \Longleftrightarrow \quad \sum_{\xi \in \mathbb{Z}_q} |\hat{B}(\xi)|^4 \omega^{-x\xi} > 0. \tag{2.30}$$

To begin with, observe that the expression $[B * B * (-B) * (-B)](x)$ counts the number of quadruples $(x_1, x_2, x_3, x_3) \in B^4$ such that $x_1 + x_2 - x_3 - x_4 = x$, since

$$\begin{aligned}
[B * B * (-B) * (-B)](x) &= \sum_{x_1 \in \mathbb{Z}_q} B(x_1)[B * (-B) * (-B)](x - x_1) \\
&= \sum_{x_1 \in B} \sum_{x_2 \in \mathbb{Z}_q} B(x_2)[(-B) * (-B)](x - x_1 - x_2) \\
&= \sum_{x_1 \in B} \sum_{x_2 \in B} \sum_{x_3 \in \mathbb{Z}_q} B(-x_3) B(x_1 + x_2 + x_3 - x) \\
&= \sum_{x_1 \in B} \sum_{x_2 \in B} \sum_{x_3 \in B} B(x_1 + x_2 - x_3 - x) \\
&= |\{(x_1, x_2, x_3) \in B^3 : x_1 + x_2 - x_3 - x \in B\}| \\
&= |\{(x_1, x_2, x_3, x_4) \in B^4 : x_1 + x_2 - x_3 - x_4 = x\}|.
\end{aligned}$$

Obviously, this number is non-zero, if and only if $x \in 2B - 2B$. Further, by Proposition 2.28(iii)-(iv), the Fourier transform of $B * B * (-B) * (-B)$ is

$$[B * B * (-B) * (-B)]\hat{} = \hat{B}\hat{B}\overline{\hat{B}}\overline{\hat{B}} = |\hat{B}|^4,$$

so by Proposition 2.28(i),

$$[B * B * (-B) * (-B)](x) = q^{-1} \sum_{\xi \in \mathbb{Z}_q} |\hat{B}(\xi)|^4 \omega^{-x\xi}.$$

This gives (2.30). An important special case is $x = 0$, namely

$$\sum_{\xi \in \mathbb{Z}_q} |\widehat{B}(\xi)|^4 = q|\{(x_1, \ldots, x_4) \in B^4 : x_1 + x_2 = x_3 + x_4\}|. \tag{2.31}$$

With these preliminaries, we are prepared to prove Proposition 2.25.

Proof of Proposition 2.25. Recall that $A \subset \mathbb{Z}_q$ satisfies $|A| \geq \alpha q$ and $|A + A| \leq C|A|$. Re-defining α if necessary, we assume that in fact $|A| = \alpha q$. Put $\epsilon = \alpha/\sqrt{8C}$ and

$$K := \{\xi \in \mathbb{Z}_q : |\hat{A}(\xi)| \geq \epsilon q\}.$$

Then $K \neq \emptyset$, because at least $0 \in K$. The plan is to show that

$$\tilde{B}(K, \kappa) := \{x \in \mathbb{Z}_q : |\omega^{-x\xi} - 1| < \kappa \text{ for all } \xi \in K\}$$

is contained in $2A - 2A$ for some small enough absolute constant $\kappa > 0$; recall from Remark 2.24 that $2A - 2A \supset \tilde{B}(K, \kappa)$ then contains a Bohr neighbourhood $B(K, \delta)$ for some $\delta > 0$ only depending on κ. This will complete the proof.

By (2.30), it suffices to demonstrate that

$$\sum_{\xi \in \mathbb{Z}_q} |\hat{A}(\xi)|^4 \omega^{-x\xi} > 0, \qquad x \in \tilde{B}(K, \kappa).$$

To this end, the definition of K and (2.29) give

$$\sum_{\xi\in\mathbb{Z}_q\setminus K}|\hat{A}(\xi)|^4<\epsilon^2q^2\sum_{\xi\in\mathbb{Z}_q\setminus K}|\hat{A}(\xi)|^2\le\epsilon^2q^3|A|=\frac{\alpha^3q^4}{8C}. \tag{2.32}$$

Recalling (2.31),

$$\sum_{\xi\in\mathbb{Z}_q}|\hat{A}(\xi)|^4=q\,\mathrm{card}\{(x_1,\dots,x_4)\in A^4:x_1+x_2=x_3+x_4\}.$$

It is an exercise to check that $|A+A|\le C|A|$ implies that the right hand side is at least $q|A|^3/C=\alpha^3q^4/C$.

Exercise 2.33. Assume that $|A+A|\le C|A|$, and prove that

$$\mathrm{card}\{(x_1,\dots,x_4)\in A^4:x_1+x_2=x_3+x_4\}\ge\frac{|A|^3}{C}.$$

In particular, using (2.32), we have

$$\sum_{\xi\in K}|\hat{A}(\xi)|^4\ge\sum_{\xi\in\mathbb{Z}_q}|\hat{A}(\xi)|^4-\sum_{\xi\in\mathbb{Z}_q\setminus K}|\hat{A}(\xi)|^4\ge\frac{\alpha^3q^4}{2C}. \tag{2.34}$$

Finally, if $x\in\tilde{B}(K,\kappa)$ and $\kappa<1/100$, say, then certainly $\mathrm{Re}\,\omega^{-x\xi}\ge1/2$ for all $\xi\in K$, which gives, using (2.32)-(2.34),

$$\begin{aligned}\sum_{\xi\in\mathbb{Z}_q}|\hat{A}(\xi)|^4\omega^{-\xi x}&=\mathrm{Re}\sum_{\xi\in\mathbb{Z}_q}|\hat{A}(\xi)|^4\omega^{-\xi x}\\&\ge\frac{1}{2}\sum_{\xi\in K}|\hat{A}(\xi)|^4-\left|\sum_{\xi\in\mathbb{Z}_q\setminus K}|\hat{A}(\xi)|^4\omega^{-x\xi}\right|\\&\ge\frac{\alpha^3q^4}{4C}-\frac{\alpha^3q^4}{8C}>0.\end{aligned}$$

As explained above, this proves that $\tilde{B}(K,\kappa)\subset2A-2A$. The last thing to verify is that $|K|\le8C/\alpha$. Fortunately, this follows immediately from (2.29):

$$|K|\le\frac{1}{\epsilon^2q^2}\sum_{\xi\in K}|\hat{A}(\xi)|^2\le\frac{q|A|}{\epsilon^2q^2}=\frac{\alpha}{\epsilon^2}=\frac{8C}{\alpha}.$$

This completes the proof of Proposition 2.25. □

2.4. **Lattice theory and the conclusion of the proof of Freiman's theorem.** It remains to prove Proposition 2.26, that is, to verify that the Bohr neighbourhood $B(K,\delta)$ found in Proposition 2.25 contains a large d-dimensional arithmetic progression. To this end, we need some tools from lattice theory. A set $\Lambda\subset\mathbb{R}^n$ is called a *lattice*, if it is a discrete subgroup and is not contained in any $n-1$-dimensional subspace of $\mathbb{R}^n$. A *fundamental domain* of Λ is a measurable subset $F\subset\mathbb{R}^n$ with the property that the sets $F+x$, $x\in\Lambda$, cover $\mathbb{R}^n$ without overlap. The determinant of the lattice, denoted by $|\Lambda|$, is the Lebesgue measure of a fundamental domain F – and then one can/should check that at least one F always exists, and the definition of $|\Lambda|$ does not depend on the particular choice of F. We will not do this here, but the idea is that every lattice Λ can be written as $\Lambda=\{\sum\lambda_iv_i:$

$\lambda_i \in \mathbb{Z}\}$, where $\{v_1, \ldots, v_n\} \subset \Lambda$ is a linearly independent collection of vectors (a *basis* for the lattice), and then $|\Lambda| = |\det(v_1, \ldots, v_n)|$. For those interested in the details, see e.g. the lecture notes on the webpage [8].

If $U \subset \mathbb{R}^n$ is an open, convex set – which will simply be $U = \{x \in \mathbb{R}^n : \|x\|_\infty < 1\}$ in our application – denote by $\lambda_k = \lambda_k(U, \Lambda)$ the positive real number

$$\lambda_k = \inf\{\lambda > 0 : \lambda U \text{ contains } k \text{ linearly independent vectors in } \Lambda\}.$$

A theorem of Minkowski gives an upper bound for the numbers λ_k in terms of $|\Lambda|$ and the Lebesgue measure of U:

Theorem 2.35 (Minkowski)**.** *Suppose that $U \subset \mathbb{R}^n$ is open, convex and centrally symmetric. Then*

$$\lambda_1 \lambda_2 \cdots \lambda_n |U| \leq 2^n |\Lambda|.$$

Proof. See Green's lecture notes, [14, Theorem 22]. □

If two lattices Λ_1, Λ_2 satisfy $\Lambda_1 \subset \Lambda_2$, it is natural to view Λ_1 as a subgroup of Λ_2. In particular, one can speak of the cosets of Λ_1 inside Λ_2, namely

$$\lambda + \Lambda_1, \qquad \lambda \in \Lambda_2,$$

and their number is denoted by $[\Lambda_1 : \Lambda_2]$. The determinants of Λ_1 and Λ_2 are have a simple relation in terms of $[\Lambda_1 : \Lambda_2]$:

Lemma 2.36 (Volume packing lemma)**.** *If $\Lambda_1 \subset \Lambda_2$, then $|\Lambda_1| = |\Lambda_2|[\Lambda_1 : \Lambda_2]$.*

Proof of Proposition 2.26. Recall that q is prime, and let $B(K, \delta) \subset \mathbb{Z}_q$ be the Bohr neighbourhood inside which we are supposed to locate an arithmetic progression. Enumerate the elements in K as $K = \{r_1, \ldots, r_k\}$ and let $\mathbf{r} := (r_1, \ldots, r_k) \in \mathbb{R}^k$; here we commit the usual abuse of notation, so $r_j \in \{0, \ldots, q-1\}$ is the integer corresponding to $r_j \in \mathbb{Z}_q$. As a technical point, we may assume that $r_j \neq 0$ for all $1 \leq j \leq k$, since

$$B(K, \delta) = \{x \in \mathbb{Z}_q : \|r_j x/q\| \leq \delta \text{ for all } 1 \leq j \leq k\},$$

is obviously the same set as $B(K', \delta)$, where K' is K minus the zero elements.

Let Λ_1 be the lattice $q\mathbb{Z}^k$, and let Λ_2 be the lattice $\Lambda_2 = \mathbf{r}\mathbb{Z} + q\mathbb{Z}^k$. The generic element of Λ_2 has the form $j\mathbf{r} + q\mathbf{m}$, where $j \in \mathbb{Z}$ and $\mathbf{m} \in \mathbb{Z}^k$, so the cosets of Λ_1 inside Λ_2 have the form

$$[j\mathbf{r} + q\mathbf{m}] + q\mathbb{Z}^k = j\mathbf{r} + q\mathbb{Z}^k.$$

The cosets with $0 \leq j \leq q-1$ are all disjoint, in fact, since the values of $jr_1 (\mathrm{mod}\, q)$, for instance, are distinct for $0 \leq j \leq q-1$ by the primality of q and the fact that $r_1 \neq 0$. Also, these are all the cosets: if $j \notin \{0, \ldots, q-1\}$, then $j = j_0 + sq$ for some $j_0 \in \{0, \ldots, q-1\}$ and $s \in \mathbb{Z}$, so that

$$j\mathbf{r} + q\mathbb{Z}^k = j_0\mathbf{r} + sq\mathbf{r} + q\mathbb{Z}^k = j_0\mathbf{r} + q\mathbb{Z}^k.$$

We may now infer from the Volume packing lemma 2.36 that

$$|\Lambda_2| = \frac{|\Lambda_1|}{[\Lambda_1 : \Lambda_2]} = \frac{q^k}{q} = q^{k-1}.$$

Next, let $U = \{x \in \mathbb{R}^k : \|x\|_\infty < 1\}$, and $\lambda_j := \lambda_j(U, \Lambda_2)$. By definition of λ_j, the set $\overline{\lambda_j U \cap \Lambda_2}$ contains j linearly independent vectors in Λ_2. We can choose such vectors inductively, starting with $\mathbf{b}^1 \in \overline{\lambda_1 U \cap \Lambda_2}$. Next, choose $\mathbf{b}^2 \in \overline{\lambda_2 U \cap \Lambda_2}$, which is linearly

independent of $\mathbf{b}^1$, then $\mathbf{b}^3 \in \overline{\lambda_3 U \cap \Lambda_2}$ independent of $\mathbf{b}^1, \mathbf{b}^2$ and so on. This eventually produces k linearly independent vectors $\mathbf{b}^1, \dots, \mathbf{b}^k$ with $\mathbf{b}^j \in \overline{\lambda_j U \cap \Lambda_2}$.

Since $\mathbf{b}^j \in \Lambda_2$, it can be written as

$$\mathbf{b}^j = b^j \mathbf{r} + q\mathbf{m}^j = (b^j r_1 + q m_1^j, \dots, b^j r_k + q m_k^j), \quad 0 \le b^j \le q-1, \ \mathbf{m}^j \in \mathbb{Z}^k.$$

Here

$$|b^j r_i + q m_i^j| \le \lambda_j \tag{2.37}$$

by virtue of $\mathbf{b}^j \in \overline{\lambda_j U}$. Now consider the k-dimensional arithmetic progression

$$P = \left\{ \sum_{j=1}^k n_j b^j : |n_j| \le \frac{\delta q}{k\lambda_j} \right\}.$$

Then $P \subset B(K, \delta)$, because if $r_i \in K$, $1 \le i \le k$, we have

$$\left\| \frac{r_i \left(\sum_{j=1}^k n_j b^j \right)}{q} \right\| = \left\| \frac{\sum_{j=1}^k n_j (b^j r_i + q m_i^j)}{q} \right\| \le \frac{1}{q} \sum_{j=1}^k |n_j| |b^j r_i + q m_i^j| \le \frac{1}{q} \sum_{j=1}^k n_j \lambda_j \le \delta$$

by (2.37) and the restriction $|n_j| \le \delta q/(k\lambda_j)$. So, it suffices to show that P is proper and estimate its size from below. Note that

$$\sum_{j=1}^k n_j \mathbf{b}^j = \left(\sum_{j=1}^k n_j (b^j r_1 + q m_1^j), \dots, \sum_{j=1}^k n_j (b^j r_k + q m_k^j) \right).$$

Again by (2.37) and the restriction on $|n_j|$, the absolute values of each coordinate of this vector is bounded by δq. Now, in order to prove that P is proper (that is to say: different sums in the definition of P result in distinct points) assume that

$$\sum_{j=1}^k n_j b^j = \sum_{j=1}^k n_j' b^j$$

for some $|n_j|, |n_j'| \le \delta q/(k\lambda_j)$. Then

$$\sum_{j=1}^k (n_j - n_j') \mathbf{b}^j = \left(\sum_{j=1}^k (n_j - n_j') q m_1^j, \dots, (n_j - n_j') q m_k^j \right) \in q\mathbb{Z}^k,$$

and

$$\left\| \sum_{j=1}^k (n_j - n_j') \mathbf{b}^j \right\|_\infty \le \left\| \sum_{j=1}^k n_j \mathbf{b}^j \right\|_\infty + \left\| \sum_{j=1}^k n_j \mathbf{b}^j \right\|_\infty \le 2\delta q < q,$$

since $\delta < 1/2$. The previous two equations force $\sum n_j \mathbf{b}^j = \sum n_j' \mathbf{b}^j$, and since the vectors $\mathbf{b}^j$ are linearly independent, $n_j = n_j'$ for $1 \le j \le k$. We have established that P is proper, and now it is a simple task to estimate its size from below, using Minkowski's theorem 2.35 and the previously shown fact that $|\Lambda_2| = q^{k-1}$:

$$|P| \ge \prod_{j=1}^k \frac{\delta q}{k\lambda_j} = \left(\frac{\delta q}{k} \right)^k \prod_{j=1}^k \frac{1}{\lambda_j} \ge \left(\frac{\delta q}{k} \right)^k \frac{|U|}{2^k |\Lambda_2|} = \left(\frac{\delta q}{k} \right)^k \frac{1}{q^{k-1}} = \left(\frac{\delta}{k} \right)^k q.$$

This completes the proof of Proposition 2.26, and that of Freiman's theorem. □

Exercise 2.38. Prove Freiman's theorem in $\mathbb{R}$ by exhibiting a Freiman isomorphism of arbitrary order between $A \subset \mathbb{R}$ and a certain subset $B \subset \mathbb{Z}$. You may first wish to find a linear isomorphism ψ between the integer span

$$\operatorname{span}(A) := \left\{ \sum_{x \in A} r_x x : r_x \in \mathbb{Z} \right\} \subset \mathbb{R}$$

and $\mathbb{Z}^n$ for a suitable $n = n(A) \in \mathbb{N}$. After that, you just need to check that φ restricted to $\psi(A)$,

$$\varphi(r_1, \ldots, r_n) := \sum_{j=1}^{n} r_j M^{j-1},$$

is a Freiman m-isomorphism for $M = M(A, m) \in \mathbb{N}$ large enough. As a further remark, using a similar proof, you could also prove Freiman's theorem in $\mathbb{R}^d$ – or more generally in any *torsion-free* group G (but the exercise only concerns $\mathbb{R}$).

3. The Balog-Szemerédi-Gowers lemma

In this section, we still remain close to the theme of the previous ones. Now we understand quite well the structure of sets $A \subset \mathbb{R}$ with $|A + A| \sim |A|$. In applications, however, one often has somewhat weaker information available than this. Here is an example. Equip $\mathbb{Z}$ with the counting measure μ, and assume that $A \subset \mathbb{Z}$ is a set such that the $L^2(\mu)$-norm of $A * A$ is large. Can we say something about the structure of A?

This question is related to the size of $A + A$, because the support of $A * A$ is contained in $A + A$. In fact, the assumption $|A + A| \sim |A|$ forces $\|A * A\|_{L^2(\mu)}$ to be large:

$$|A|^2 = \|A * A\|_{L^1(\mu)}^2 \leq |A + A|^{1/2} \|A * A\|_{L^2(\mu)} \sim |A|^{1/2} \|A * A\|_{L^2(\mu)},$$

so $\|A * A\|_{L^2(\mu)}^2 \gtrsim |A|^3$. This is about as large as the $L^2(\mu)$-norm of $A * A$ can get, because

$$A * A(x) = \sum_{y \in \mathbb{Z}} A(y) A(x - y) = \sum_{a \in A} A(x - a) \leq |A|,$$

so that always $\|A * A\|_{L^2(\mu)}^2 \leq |A|^3$.

This hopefully demonstrates that there is a strong analogue between the assumptions

$$|A + A| \sim |A| \quad \text{and} \quad \|A * A\|_{L^2(\mu)}^2 \sim |A|^3.$$

However, the first one is more restrictive. For instance, consider the case $A = B \cup C \subset \mathbb{Z}$, where $|B| = |C| = n$, B is contained in a low-dimensional arithmetic progression of size $\sim n$, and C contains n arbitrary points. Then the second assumption is always satisfied – simply because $\|A * A\|_{L^2(\mu)}^2 \geq \|B * B\|_{L^2(\mu)}^2 \gtrsim n^3$ – whereas the validity of the first one depends on B. In a sense, the second assumption is more robust.

The example above shows that the strongest structural conclusion for A, which can be deduced from the assumption $\|A * A\|_{L^2(\mu)}^2 \geq C|A|^3$, is the following: A contains a large subset A', which is covered by a generalised arithmetic progression, whose dimension and size are bounded by constants depending only on C. This begins to sound a lot like the conclusion of Freiman's theorem, except for the "large subset" part. In order to apply Freiman's theorem directly, we would need another result, which states that there

exists a large subset $A_0 \subset A$ with the property that $|A_0 + A_0| \leq C_0|A_0|$. Such a result is available, in fact, and is known as the Balog-Szemerédi-Gowers lemma/theorem:

Theorem 3.1 (BSGT). *Let A be a finite non-empty set in an abelian group G. Write*

$$T(A) := |\{(a_1, a_2, a_3, a_4) \in A^4 \colon a_1 - a_2 = a_3 - a_4\}|,$$

and suppose that $T(A) \geq \gamma|A|^3$ for some constant $\gamma \in (0, 1)$. Then there exists a subset $A_0 \subset A$ with $|A_0| \geq (\gamma^2/40)|A|$ such that

$$|\{(x_1, \dots, x_4, y_1, \dots, y_4) \in A^8 : a_1 - a_2 = x_1 + \cdots + x_4 - y_1 - \cdots - y_4\}| \geq 2^{-28}\gamma^{10}|A|^7$$

for any $a_1, a_2 \in A_0$. In particular,

$$|A_0 - A_0| < 2^{28}\gamma^{-10}|A|.$$

Remark 3.2. If μ is the counting measure on G, then $T(A) = \|A * A\|^2_{L^2(\mu)}$ according to the computation leading to (2.31). Note that, according to Exercise 2.33, the conclusion $|A_0 - A_0| \lesssim |A_0|$ also gives

$$T(A) \geq T(A_0) \gtrsim |A_0|^3 \gtrsim |A|^3.$$

So, the Balog-Szemerédi-Gowers lemma is, in a sense, an "if and only if" result.

3.1. **Proof of the Balog-Szemerédi-Gowers lemma.** Before starting the proof, we should mention that this version of the theorem is due to W. T. Gowers [13] from 1998. The original version due to A. Balog and E. Szemerédi [3], from 1994, has worse constants. The proof of Theorem 3.1 proceeds through a few lemmas. The intuition behind the first lemma is the following. Suppose that A is a finite set, and $A_1, \dots, A_n \subset A$ are subsets with cardinality at least $\delta|A|$, where n large compared to δ. Then, the average intersection of $A_i \cap A_j$ contains at least $\delta^2|A|$ elements. The lemma below says that one can actually find a large selection $\mathcal{A}$ among the subsets A_i, $1 \leq i \leq n$, such that "most" intersections of pairs of sets $A_i, A_j \in \mathcal{A}$ have "almost" the expected cardinality.

Lemma 3.3. *Let A be a finite set with $m := |A|$, and suppose that $A_1, \dots, A_n \subset A$ are subsets of size at least $|A_i| \geq \delta m$, where $\delta > 0$ is a constant. Then, there exists a set of indices $I \subset \{1, \dots, n\}$ such that $|I| \geq \delta n/2$ and*

$$|\{(i, j) \in I \times I : |A_i \cap A_j| \leq 0.03\delta^2 m\}| < \frac{|I|^2}{25}.$$

Proof. For $a \in A$, let $I_a = \{i \in \{1, \dots, n\} : a \in A_i\}$. Then,

$$\begin{aligned}
\sum_{i,j=1}^{n} |A_i \cap A_j| &= \sum_{i,j=1}^{n} \sum_{a \in A} (A_i \cap A_j)(a) = \sum_{a \in A} \sum_{i=1}^{n} A_i(a) \left(\sum_{j=1}^{n} A_j(a) \right) \\
&= \sum_{a \in A} \sum_{i=1}^{n} A_i(a) \cdot |I_a| = \sum_{a \in A} |I_a|^2 \overset{\text{C-S}}{\geq} |A|^{-1} \left(\sum_{a \in A} |I_a| \right)^2 \\
&= \frac{1}{m} \left(\sum_{a \in A} \sum_{i=1}^{n} A_i(a) \right)^2 = \frac{1}{m} \left(\sum_{i=1}^{n} |A_i| \right)^2 \geq \delta^2 m n^2.
\end{aligned}$$

This means exactly that the average intersection $A_i \cap A_j$ has at least $\delta^2 m$ elements. We claim that we can take $I = I_a$ for some $a \in A$. Let us choose a at random. Taking expectations with respect to this random choice, we have

$$\begin{aligned}\mathbb{E}_{a \in A}|I_a|^2 &= \frac{1}{m}\sum_{a \in A}|\{(i,j) : i,j \in I_a\}| = \frac{1}{m}\sum_{a \in A}|\{(i,j) : a \in A_i \cap A_j\}| \\ &= \frac{1}{m}\sum_{i,j=1}^{n}\sum_{a \in A}(A_i \cap A_j)(a) = \frac{1}{m}\sum_{i,j=1}^{n}|A_i \cap A_j| \geq \delta^2 n^2,\end{aligned}$$

using the previous estimate in the final inequality. On the other hand, we have the upper bound

$$\begin{aligned}\mathbb{E}_{a \in A}|\{(i,j) \in I_a \times I_a : |A_i \cap A_j| \leq 0.03\delta^2 m\}| &= \frac{1}{m}\sum_{a \in A}\sum_{\substack{i,j=1 \\ |A_i \cap A_j| \leq 0.03\delta^2 m}}^{n}(I_a \times I_a)(i,j) \\ &= \frac{1}{m}\sum_{\substack{i,j=1 \\ |A_i \cap A_j| \leq 0.03\delta^2 m}}^{n}\sum_{a \in A}(A_i \cap A_j)(a) \\ &= \frac{1}{m}\sum_{\substack{i,j=1 \\ |A_i \cap A_j| \leq 0.03\delta^2 m}}^{n}|A_i \cap A_j| \leq 0.03\delta^2 n^2.\end{aligned}$$

Since $25 \cdot 0.03 = 0.75$, using linearity of expectation now leads to

$$\mathbb{E}_{a \in A}\left\{|I_a|^2 - 25 \cdot |\{(i,j) \in I_a \times I_a : |A_i \cap A_j| \leq 0.03\delta^2 m\}|\right\} \geq \frac{\delta^2 n^2}{4}.$$

In particular, there has to exist $a \in A$ such that

$$|I_a|^2 - 25 \cdot |\{(i,j) \in I_a \times I_a : |A_i \cap A_j| \leq 0.03\delta^2 m\} \geq \frac{\delta^2 n^2}{4}.$$

For this a, we have $|I_a| \geq \delta n/2$ and

$$|\{(i,j) \in I_a \times I_a : |A_i \cap A_j| \leq 0.03\delta^2 m\}| \leq \frac{|I_a|^2}{25} - \frac{\delta^2 n^2}{100} < \frac{|I_a|^2}{25},$$

as required. □

In fact, we will eventually use the lemma in the following form:

Lemma 3.4. *Let A, B be two finite non-empty sets with $m := |A|$. Suppose that to every $b \in B$ there corresponds a set $N(b) \subset A$ of cardinality $|N(b)| \geq \delta m$, where $\delta > 0$ is a constant. Then, there exists a set $B' \subset B$ with $|B'| \geq \delta|B|/2$ such that*

$$|\{(b,b') \in B' \times B' : |N(b) \cap N(b')| \leq 0.03\delta^2 m\}| < \frac{|B'|^2}{25}.$$

Of course, this lemma is precisely the same as the one above; now the sets A_i are simply denoted by $N(b)$. Next, we introduce a simple lemma in graph theory.

Lemma 3.5. *Let G be an undirected graph on a vertex set V (it is allowed that G has loops, i.e. that $v \to v$ is an edge). Suppose that the average degree[2] of G satisfies $\bar{v} \geq (1-\lambda)|V|$. Then V contains at least $(1-\sqrt{\lambda})|V|$ vertices of degree higher than $(1-\sqrt{\lambda})|V|$.*

Proof. Let $k := |\{v \in V : \deg(v) > (1-\sqrt{\lambda})|V|\}|$. Then

$$\bar{d} = \frac{1}{|V|} \sum_{\deg(v) \leq (1-\sqrt{\lambda})|V|} \deg(v) + \frac{1}{|V|} \sum_{\deg(v) > (1-\sqrt{\lambda})|V|} \deg(v) \leq (1-\sqrt{\lambda})(|V|-k) + k,$$

which gives

$$(1-\lambda)|V| \leq \bar{d} \leq (1-\sqrt{\lambda})(|V|-k) + k = (1-\sqrt{\lambda})|V| + \sqrt{\lambda}k,$$

and so

$$k \geq \left(\frac{\sqrt{\lambda}-\lambda}{\sqrt{\lambda}}\right)|V| = (1-\sqrt{\lambda})|V|.$$

□

Combining the results so far leads to our final lemma:

Lemma 3.6. *Let A, B be finite sets with $m := |A|$; as in Lemma 3.4, suppose that to every element $b \in B$ there exists a set $N(b) \subset A$ of cardinality $|N(b)| \geq \delta m$. Then there exist subsets $B_0 \subset B' \subset B$ with $|B_0| \geq 4|B'|/5 > 2\delta|B|/5$, such that for any $b_0 \in B_0$ we have*

$$|\{b' \in B' : |N(b_0) \cap N(b')| > 0.03\delta^2 m\}| > \frac{4|B'|}{5}.$$

Proof. Start by finding $B' \subset B$ as in Lemma 3.4. Thus, we have $|B'| \geq \delta|B|/2$, and

$$|\{(b,b') \in B' \times B' : |N(b) \cap N(b')| \leq 0.03\delta^2 m\}| < \frac{|B'|^2}{25}. \tag{3.7}$$

Next, construct a graph G on the vertex set $V = \{N(b') : b' \in B'\}$ by joining $N(b')$ to $N(b'')$, if and only if $|N(b') \cap N(b'')| > 0.03\delta^2 m$. In particular, each set $N(b')$ is joined to itself (this is called a *loop edge*). How many non-loop edges are there in G? Well, there are $\binom{|B'|}{2}$ two-element subsets $\{b', b''\} \subset B'$, and to every one of these sets there corresponds a non-loop edge **except** if $|N(b') \cap N(b'')| \leq 0.03\delta^2 m$. According to (3.7), there are at most $|B'|^2/50$ subsets $\{b', b''\}$ with the latter property (observing that no pair (b', b') can have the property), so the number of non-loop edges is at least $\binom{|B'|}{2} - |B'|^2/50$. Thus, adding the $|B'|$ loop edges, the average degree of G is

$$\frac{|B'| + 2 \cdot |\{\text{non-loop edges in } G\}|}{|V|} \geq \frac{|B'| + 2 \cdot \binom{|B|'}{2} - |B'|^2/25}{|B'|} = \frac{24 \cdot |B'|}{25}.$$

Applying Lemma 3.5 with $\lambda = 1/25$, we find a set $B_0 \subset B'$ with at least $(1-\sqrt{\lambda})|B'| = 4|B'|/5$ elements such that if $b_0 \in B_0$, then the degree of b_0 is higher than $4|B'|/5$. In other words,

$$|\{b' \in B' : |N(b_0) \cap N(b')| > 0.03\delta^2 m\}| > \frac{4|B'|}{5}, \quad b_0 \in B_0,$$

just as we wanted. □

Now we are prepared to prove the Balog-Szemerédi-Gowers theorem.

[2]The 'degree of v' means the same as the 'number of edges starting/ending at v'.

Proof of Theorem 3.1. Recall that A is a finite subset of an abelian group, and write $m := |A|$. Also, recall that

$$v_A(d) = |\{(a_1, a_2) \in A \times A : d = a_1 - a_2\}|, \qquad d \in G,$$

and

$$T(A) := |\{(a_1, a_2, a_3, a_4) \in A^4 : a_1 - a_2 = a_3 - a_4\}|.$$

We are assuming that $T(A) \geq \gamma |A|^3$ for some constant $\gamma \in (0, 1)$. We say that an element $d \in G$ is a *popular difference*, if $v_A(d) \geq \gamma m/2$; write D for the set of all popular differences. We construct a graph Γ on the vertex set A by joining a_1 to a_2, if and only if $a_1 - a_2 \in D$ (since $0 \in D$, we also loop a to a for every $a \in A$). Observe that the graph so constructed is undirected: if a_1 is joined to a_2, then also a_2 is joined to a_1 by virtue of the equation $-D = D$. To estimate the average degree $\bar{d}$ of Γ, we first write

$$\begin{aligned} \bar{d} &:= \frac{1}{m} \sum_{a \in A} \deg(a) := \frac{1}{m} \sum_{a \in A} |\{a' \in A : a' - a \in D\}| \\ &= \frac{1}{m} |\{(a, a') \in A \times A : a' - a \in D\}| = \frac{1}{m} \sum_{d \in D} v_A(d). \end{aligned}$$

Next, using the equation

$$T(A) = \sum_{d \in A - A} (v_A(d))^2,$$

justified during the proof of Proposition **??**, we estimate

$$\begin{aligned} \gamma m^3 \leq T(A) = \sum_{d \in A - A} (v_A(d))^2 &= \sum_{d \in (A-A) \setminus D} (v_A(d))^2 + \sum_{d \in D} (v_A(d))^2 \\ &\leq \frac{\gamma m}{2} \sum_{d \in A - A} v_A(d) + m \sum_{d \in D} v_A(d) = \frac{\gamma m^3}{2} + m^2 \bar{d}. \end{aligned}$$

This gives

$$\bar{d} \geq \frac{\gamma m}{2}.$$

We then apply the graph-theoretic Lemma 3.5 to Γ with $\lambda = 1 - \gamma/2$; observing that $\bar{d} \geq (1 - \lambda)m$, we obtain a subset $B \subset A$ with $|B| \geq (1 - \sqrt{\lambda})m > \gamma m/4$ such that $\deg(b) > \gamma m/4$ for any $b \in B$. To rephrase this, if $N(b)$ stands for the neighbourhood of b in Γ – including b – then $|N(b)| > \gamma m/4$ for any $b \in B$.

Then, we apply Lemma 3.6 with $\delta = \gamma/4$ to the collection of sets $N(b) \subset A$, $b \in B$. Thus, we find subsets $A_0 \subset A' \subset B$ such that

$$|A_0| \geq \frac{4|A'|}{5} \geq \frac{2\delta |B|}{5} > \frac{\gamma^2 m}{40}, \tag{3.8}$$

and

$$|\{a' \in A' : |N(a_0) \cap N(a')| > 0.03(\gamma/4)^2 m\}| > \frac{4|A'|}{5} \tag{3.9}$$

for any $a_0 \in A_0$. We now claim that

$$|\{(x_1, \ldots, x_4, y_1, \ldots, y_4) \in A^8 : a_1 - a_2 = x_1 + \cdots + x_4 - y_1 - \cdots - y_4\}| \geq 2^{-28} \gamma^{10} |A|^7$$

for any pair of elements $a_1, a_2 \in A_0$. According to (3.8), this will finish the proof of Theorem 3.1, except for the easy 'in particular' part. Fix $a_1, a_2 \in A_0$. According to (3.9), the set

$$S = \{a' \in A' : |N(a_i) \cap N(a')| > 0.03(\gamma/4)^2 m \text{ for } i = 1, 2\}$$

is the intersection of two subsets of A' with cardinality $> 4|A'|/5$, so $|S| > 3|A'|/5$, using the equation $|U \cap V| = |U| + |V| - |U \cup V|$. Now, at the end, things become a bit complicated. Fix $a' \in S$. Note that, if $a \in N(a_1) \cap N(a')$, then $a_1 - a$ and $a' - a$ are both popular differences, so

$$|\{(x_1, y_1) \in A \times A : a_1 - a = x_1 - y_1\}| \geq \frac{\gamma m}{2}$$

and

$$|\{(x_1', y_1') \in A \times A : a' - a = x_1' - y_1'\}| \geq \frac{\gamma m}{2}.$$

This means that

$$|\{(x_1, y_1, x_1', y_1') \in A^4 : a_1 - a' = (x_1 - y_1) - (x_1' - y_1')\}| \geq \left(\frac{\gamma m}{2}\right)^2.$$

Notice how the point a has disappeared completely from above; since $a \in N(a_1) \cap N(a')$ can be chosen in $> 0.03(\gamma/4)^2 m$ ways, we may actually multiply $(\gamma m/2)^2$ above with this constant to end up with

$$|\{(x_1, y_1, x_1', y_1') \in A^4 : a_1 - a' = (x_1 - y_1) - (x_1' - y_1')\}| > 0.03 \cdot 2^{-6}\gamma^4 m^2.$$

An important point used above is this: the same quadruple (x_1, y_1, x_1', y_1') cannot appear multiple times for different choices of a, since a is *determined* by each quadruple it produces; e.g. $a = a_1 - x_1 + y_1$. Using the same reasoning with a_2 instead of a_1 yields

$$|\{(x_2, y_2, x_2', y_2') \in A^4 : a_2 - a' = (x_2 - y_2) - (x_2' - y_2')\}| > 0.03 \cdot 2^{-6}\gamma^4 m^2.$$

Thus, corresponding to each $a' \in S$, we find

$$(0.03 \cdot 2^{-6}\gamma^4 m^2)^2$$

octuples $(x_1, y_1, x_1', y_1', x_2, y_2, x_2', y_2') \in A^8$ such that

$$a_1 - a_2 = (x_1 - y_1) - (x_1' - y_1') - (x_2 - y_2) + (x_2' - y_2'). \tag{3.10}$$

Again, we may freely choose $a' \in S$; there are at least $3|A'|/5$ such choices, so we have, all in all,

$$> (0.03 \cdot 2^{-6}\gamma^4 m^2)^2 \cdot \frac{3|A'|}{5} > 2^{-28}\gamma^{10} m^7 \tag{3.11}$$

octuples $(x_1, y_1, x_1', y_1', x_2, y_2, x_2', y_2') \in A^8$ such that (3.10) holds. As before, it is important that a' be determined by each octuple it produces; consequently, the same octuple is not produced by multiple choices of a'. The number on the r.h.s of (3.11) is as claimed in Theorem 3.1, so all that remans of the proof is the 'in particular' part.

To prove that $|A_0 - A_0| < 2^{28}\gamma^{-10}|A|$, note that, altogether, there are exactly $|A|^8$ octuples of the form $(x_1, \ldots, x_4, y_1, \ldots, y_4) \in A^8$. If $a_1 - a_2 \neq a_1' - a_2'$, then the sets of octuples corresponding to the differences $a_1 - a_2$ and $a_1' - a_2'$ (via (3.10) or the corresponding equation in the statement of Theorem 3.1) are clearly disjoint. Given that there are $|A_0 - A_0|$ distinct differences, and each difference generates at least $2^{-28}\gamma^{10}|A|^7$ octuples, we end up with

$$|A_0 - A_0| \cdot 2^{-28}\gamma^{10}|A|^7 \leq |A|^8.$$

This completes the proof. □

4. Existence of arithmetic progressions in subsets of $\mathbb{Z}$

In this section, we discuss Roth's theorem on the existence of 3-term arithmetic progressions in dense subsets of $\mathbb{Z}$. This result is the simplest non-trivial case of the following fundamental theorem of Szemerédi:

Theorem 4.1 (Szemerédi's theorem)**.** *Assume that $\delta > 0$ and $k \in \mathbb{N}$, $k \geq 3$. Then, there is a number $n(\delta, k) \in \mathbb{N}$ with the following property: if $n \geq n(\delta, k)$, and $A \subset \{1, \ldots, n\}$ has $|A| \geq \delta n$, then A contains a k-term arithmetic progression.*

In this section, "arithmetic progression" should be interpreted in the classical sense, so the theorem asserts the existence of a set of the form $\{x, x + y, x + 2y, \ldots, x + (k - 1)y\} \subset A$, $y \geq 1$. There are essentially three different proofs of this theorem, which was originally conjectured by Erdős and Turán in 1936. The first was found in 1974 by E. Szemerédi, who used a complicated combinatorial argument. The second proof is due to H. Furstenberg from 1977, using ergodic theory. The third proof from 2001 is due to W. T. Gowers, and employs Fourier analysis. Gowers' proof generalises – in a very nontrivial way! – a previous argument by Roth from 1953 (also based on Fourier analysis), which already gave Szemerédi's theorem in the case $k = 3$. So, in these lecture notes we will, in fact, prove Roth's theorem, and in doing so we will follow Gowers' argument from 2001.

4.1. Heuristics of the proof.

A large random subset of $\{1, \ldots, N\}$ contains plenty of 3-term arithmetic progressions in expectation:

Exercise 4.2. Prove that a random subset $A \subset \{1, \ldots, N\}$ with $|A| = \delta N$ contains roughly $\delta^3 N^2$ 3-term arithmetic progressions in expectation, if N is large compared to δ.

The idea in the proof of Roth's theorem is to define a concept of "pseudorandomness", and to establish the following alternative:

(i) If the set A is pseudorandom to begin with, then it behaves like a random set in the sense that there are roughly $\delta^3 N^2$ 3-term arithmetic progressions inside A.[3]
(ii) If the set A is not pseudorandom and has $|A| = \delta' N$, $\delta' \geq \delta$, then there is a long arithmetic progression $P \subset \{1, \ldots, N\}$ such that $|A \cap P| \geq (\delta' + \delta^2/1000)|P|$.

After these claims have been established, a simple iteration will give Roth's theorem: if A is pseudorandom, (i) gives the claim. If not, then forget about A and consider $A_1 := A \cap P$, defined in (ii), and observe that P can be identified with $\{1, \ldots, |P|\}$. If A_1 is pseudorandom in $\{1, \ldots, |P|\}$, we are again done by (i), and if not, iterate with (ii). Eventually, we either end up in alternative (i), or then find a set $A_m \subset A$ with $|A_m| > (2/3)|P'|$ for some (long, depending on δ, ϵ) arithmetic progression $P' \subset \{1, \ldots, N\}$. Then, it is a simple matter to check that A_m – hence A – contains a 3-term arithmetic progression.[4]

4.2. Recapping Fourier analysis and basic definitions.

As in the proof of Freiman's theorem, most of the work will take place in $\mathbb{Z}_q$ instead of $\{1, \ldots, N\}$. Here q is a large prime.

[3] Of course this is roughly what we are claiming even without the assumption of pseudorandomness, but the point is that in the pseudorandom case the existence of arithmetic progressions can be easily proved with a direct argument.

[4] If A_m is a subset of $\{1, \ldots, k\}$ with $|A| > (2/3)k$, then A contains $\{j, j + 1, j + 2\}$ for some j.

Gowers likes to define the Fourier transform with a different sign than we did before, namely

$$\hat{f}(\xi) = \sum_{x \in \mathbb{Z}_q} f(r)\omega^{-x\xi}.$$

Here $f\colon \mathbb{Z}_q \to \mathbb{C}$ and $\omega^{-x\xi} = e^{-2\pi i x\xi/q}$ as before. We will adopt this convention in order to minimise sign errors caused by T.O. With this convention, the Fourier inversion and Plancherel formulae look like this:

Proposition 4.3. *Let $f, g\colon \mathbb{Z}_q \to \mathbb{C}$, and $B \subset \mathbb{Z}_q$.*

(i) $f(x) = q^{-1}\sum \hat{f}(\xi)\omega^{x\xi}$.
(ii) $\sum f(x)\overline{g(x)} = q^{-1}\sum \hat{f}(\xi)\overline{\hat{g}(\xi)}$, *so that* $\sum |f(x)|^2 = q^{-1}\sum |\hat{f}(\xi)|^2$.

As before, the indicator function of a set $B \subset Z_q$ is denoted by B. The notion of "pseudorandomness" of A from the previous subsection will not be officially defined anywhere, but if it were, it would require that

$$|\hat{A}(\xi)| \leq \alpha q \quad \text{for all } \xi \neq 0,$$

where α is some small constant (in particular much smaller than $\delta = |A|/q$). Under this assumption, the existence of roughly $\delta^3 q^2$ 3-term arithmetic progressions is not hard to show directly; in fact we will do so at the very end of the proof.

Definition 4.4 (Balanced indicator)**.** If $B \subset \mathbb{Z}_q$ is a set with $|B| = \delta q$, we define the *balanced indicator function* f_B by $f_B(x) = B - \delta$, that is,

$$f_B(x) := \begin{cases} 1-\delta, & \text{if } x \in B, \\ -\delta, & \text{if } x \notin B. \end{cases}$$

Observe that $\widehat{f_B}(\xi) = \hat{B}(\xi)$ for all $\xi \neq 0$. Of course, $\hat{B}(0) = \sum B(x) = |B|$ and $\widehat{f_B}(0) = \sum f(x) = 0$.

4.3. **Proof of Roth's theorem on 3-term arithmetic progressions.** Below, the *diameter* of a set $B \subset \mathbb{Z}_q$, denoted $\operatorname{diam}(B)$, is the smallest natural number s such that

$$B \subset \{n, n+1, \dots, n+s\} \pmod q$$

for some $n \in \mathbb{Z}_q$. For instance

$$\operatorname{diam}(\{0, q-1\}) = 1.$$

Lemma 4.5. *Let $r, s \in \{1, \dots, q\}$ with $rs \geq q$, and let $\phi\colon \{0, 1, \dots, r-1\} \to \mathbb{Z}_q$ be the linear mapping $\phi(x) = \xi x \pmod q$ for some $\xi \in \mathbb{Z}_q$. Then, the set $\{0, 1, \dots, r-1\}$ can be partitioned into arithmetic progressions $P_1, \dots, P_M$ such that*

$$\operatorname{diam}(\phi(P_j)) \leq s \quad \text{and} \quad \left(\frac{rs}{4q}\right)^{1/2} \leq |P_j| \leq \left(\frac{rs}{q}\right)^{1/2}$$

for $1 \leq j \leq M$.

Proof. We may assume that $rs > 4q$, since otherwise we may prove the lemma by using arithmetic progressions of length one. To avoid using floor and ceiling functions excessively, we will – unrealistically – assume that various quantities below are integer-valued. One such example is

$$t := \left(\frac{rq}{4s}\right)^{1/2}.$$

Two of the numbers $\phi(0), \phi(1), \ldots, \phi(t) \in \{0, \ldots, q-1\}$ must be within q/t of each other, say $|\phi(x_1) - \phi(x_2)| \leq q/t$, where $x_1 < x_2$. Setting $u := x_2 - x_1$, we have

$$0 < u \leq t \qquad \text{and} \qquad |\phi(u)| \leq q/t. \tag{4.6}$$

Now, split $\{0, 1, \ldots, r-1\}$ into congruence classes modulo u; these congruence classes are arithmetic progressions in $\{0, 1, \ldots, r-1\}$, with length in $[r/u - 1, r/u + 1]$.[5] The sets P_j are then defined by partitioning these congruence classes further into blocks of consecutive elements, where each block has length roughly (but at most) st/q. Here

$$\frac{st}{q} = \frac{s(rq/4s)^{1/2}}{q} = \frac{1}{2} \cdot \left(\frac{rs}{q}\right)^{1/2}, \tag{4.7}$$

which means that this procedure yields arithmetic progressions of the desired length. Of course, for this to make sense, we need to know that $st/q \geq 1$ (which follows from (4.7) and $rs > 4q$), and that $st/q \leq r/u - 1$, which is the following computation:

$$\frac{st}{q} = \frac{1}{2} \cdot \left(\frac{rs}{q}\right)^{1/2} = \frac{1}{4} \cdot \frac{r}{(rq/4s)^{1/2}} = \frac{1}{4} \cdot \frac{r}{t} \overset{(4.6)}{\leq} \frac{1}{4} \cdot \frac{r}{u} \leq \frac{r}{u} - 1.$$

Now, each P_j has the form $P_j = \{m, m+u, m+2u, \ldots, m+ku\}$ with $k \leq (st/q - 1)$, so $\phi(P_j)$ is an arithmetic progression with gap $\leq q/t$ by (4.6). The claim about diameters then follows from $|P_j| \cdot (t/q) \leq (st/q) \cdot (q/t) = s$. □

Lemma 4.8. *Let $f\colon \{0, 1, \ldots, r-1\} \to \mathbb{D}$, let $\phi\colon \{0, 1, \ldots, r-1\} \to \mathbb{Z}_q$ be a linear function as in the previous lemma, and let $\alpha \in (0, 1)$. If $q, r \geq 8\pi/\alpha$, and*

$$\left|\sum_{x=0}^{r-1} f(x)\omega^{-\phi(x)}\right| \geq \alpha r,$$

then there is a partition of $\{0, 1, \ldots, r-1\}$ into $M \leq (32\pi r/\alpha)^{1/2}$ arithmetic progressions $P_1, \ldots, P_M$ such that

$$\sum_{j=1}^{M} \left|\sum_{x \in P_j} f(x)\right| \geq \frac{\alpha r}{2},$$

and $(\alpha r/32\pi)^{1/2} \leq |P_j| \leq (\alpha r/6\pi)^{1/2}$ for all $1 \leq j \leq M$.

Proof. Choose an integer s between $\alpha q/(8\pi)$ and $\alpha q/(6\pi)$. Then $rs \geq q$, so the previous lemma partitions $\{0, 1, \ldots, r-1\}$ into arithmetic progressions $P_1, \ldots, P_M$ such that

$$\left(\frac{\alpha r}{32\pi}\right)^{1/2} \leq \left(\frac{rs}{4q}\right)^{1/2} \leq |P_j| \leq \left(\frac{rs}{q}\right)^{1/2} \leq \left(\frac{\alpha r}{6\pi}\right)^{1/2}$$

[5] Here $r/u \geq r/t = 2(rs/q)^{1/2} > 4$.

and $\mathrm{diam}(\phi(P_j)) \le s$ for all $1 \le j \le M$. The upper bound for M follows from the lower bound for $|P_j|$. By the triangle inequality,

$$\sum_{j=1}^{M} \left| \sum_{x \in P_j} f(x)\omega^{-\phi(x)} \right| \ge \left| \sum_{x=0}^{r-1} f(x)\omega^{-\phi(x)} \right| \ge \alpha r. \tag{4.9}$$

Fix an arbitrary element $x_j \in P_j$. If $x \in P_j$ is another element, then $|\phi(x) - \phi(x_j) + kq| \le s$ for some $k \in \mathbb{Z}$ by $\mathrm{diam}(\phi(P_j)) \le s$. Now

$$|\omega^{-\phi(x)} - \omega^{-\phi(x_j)}| = |1 - \omega^{\phi(x)-\phi(x_j)+kq}| = |1 - e^{2\pi i(\phi(x)-\phi(x_j)+kq)/q}|.$$

Here $|2\pi(\phi(x) - \phi(x_j) + kq)/q| \le |2\pi s/q| \le \alpha/3$, so writing $\rho := 2\pi(\phi(x) - \phi(x_j) + kq)/q$, we have

$$|1 - e^{i\rho}| = \left| 1 - \sum_{k=0}^{\infty} \frac{(i\rho)^k}{k!} \right| \le \sum_{k=1}^{\infty} (\alpha/3)^k \le \frac{1}{1-(\alpha/3)} - 1 = \frac{\alpha}{3-\alpha} \le \frac{\alpha}{2}.$$

This proves that $|\omega^{-\phi(x)} - \omega^{-\phi(x_j)}| \le \alpha/2$, and consequently

$$\begin{aligned}
\sum_{j=1}^{M} \left| \sum_{x \in P_j} f(x) \right| &= \sum_{j=1}^{M} \left| \sum_{x \in P_j} f(x)\omega^{-\phi(x_j)} \right| \\
&\ge \sum_{j=1}^{M} \left| \sum_{x \in P_j} f(x)\omega^{-\phi(x)} \right| - \sum_{j=1}^{M} \left| \sum_{x \in P_j} f(x)[\omega^{-\phi(x_j)} - \omega^{-\phi(x)}] \right| \\
&\ge \sum_{j=1}^{M} \left| \sum_{x \in P_j} f(x)\omega^{-\phi(x)} \right| - \sum_{j=1}^{M} (\alpha/2)|P_j| \ge \frac{\alpha r}{2}
\end{aligned}$$

by (4.9), and $|f(x)| \le 1$, and the fact that the progressions P_j partition $\{0, \dots, r-1\}$. This completes the proof of the lemma. □

The next corollary says that if A is **not** "pseudorandom", then there exists a long arithmetic progression P, where A has increased density. This is essentially the second alternative from Subsection 4.1.

Corollary 4.10. *Let $A \subset \mathbb{Z}_q$, and assume that $|\hat{A}(\xi)| \ge \alpha q$ for some $\xi \ne 0$. Then, there exists an arithmetic progression $P \subset \{0, 1, \dots, q-1\}$ of length at least $(\alpha q/32\pi)^{1/2}$ such that $|A \cap P| \ge (\delta + \alpha/10)|P|$.*

Proof. Let ϕ be the linear function $\phi(x) := x\xi$, so that

$$\left| \sum_{x=0}^{q-1} f_A(x)\omega^{-\phi(x)} \right| = |\widehat{f_A}(\xi)| = |\hat{A}(\xi)| \ge \alpha q.$$

Now, the previous lemma states that $\{0, 1, \dots, q-1\}$ can be partitioned into $M \le (32\pi q/\alpha)^{1/2}$ arithmetic progressions $P_1, \dots, P_M$ such that

$$\sum_{j=1}^{M} \left| \sum_{x \in P_j} f_A(x) \right| \ge \frac{\alpha q}{2},$$

and $(\alpha q/32\pi)^{1/2} \leq |P_j| \leq (\alpha q/6\pi)^{1/2}$ for all $1 \leq j \leq M$. Let J be the set of indices $1 \leq j \leq M$ such that $\sum_{x \in P_j} f(x) \geq 0$. Then

$$\sum_{j \in J} \sum_{x \in P_j} f_A(x) \geq \frac{\alpha q}{4}. \tag{4.11}$$

Indeed, we have

$$\sum_{j \in J} \sum_{x \in P_j} f_A(x) = \sum_{x \in \mathbb{Z}_q} f_A(x) - \sum_{j \notin J} \sum_{x \in P_j} f_A(x) = -\sum_{j \notin J} \sum_{x \in P_j} f_A(x),$$

so if (4.11) failed, we would get the contradiction

$$\sum_{j=1}^{M} \left| \sum_{x \in P_j} f_A(x) \right| = \sum_{j \in J} \sum_{x \in P_j} f_A(x) - \sum_{j \notin J} \sum_{x \in P_j} f_A(x) < \frac{\alpha q}{2}.$$

Now, (4.11), and the bounds for M and $|P_j|$, imply that

$$\sum_{x \in P_j} f_A(x) \geq \frac{\alpha q}{4M} \geq \frac{\alpha q}{4} \cdot \left(\frac{\alpha}{32\pi q} \right)^{1/2} = \frac{\alpha}{4\sqrt{16/3}} \cdot \left(\frac{\alpha q}{6\pi} \right)^{1/2} \geq \frac{\alpha |P_j|}{10}$$

for some $j \in J$. Writing out the definition of f_A, this means that

$$|A \cap P_j| - \delta |P_j| = \sum_{x \in A \cap P_j} (1 - \delta) + \sum_{x \in P_j \setminus A} (-\delta) = \sum_{x \in P_j} f_A(x) \geq \frac{\alpha |P_j|}{10},$$

or $|A \cap P_j| \geq (\delta + \alpha/10)|P_j|$, as desired. □

We are ready to prove the main result of the section:

Proof of Roth's theorem. Assume that $A_0 \subset \{0, \ldots, q_0 - 1\}$ with $|A| = \delta_0 q_0$. There will be several cases treated below, but in each one the strategy will be the following: either we find a long arithmetic progression, wherein A_0 has increased density, and iterate the proof immediately from the beginning – or then (in the last case) there is a direct argument showing that A_0 contains plenty of 3-term progressions.

Case 1. (reduction to prime field) Assume that $q = |P|$ is the length of some subprogression P found in the course of the proof – so $q = q_0$ on the first round – and assume that $A := A_0 \cap P$ has $|A| \geq \delta q$ with $\delta_0 \leq \delta < 1/2$. Also, assume that $P = \{0, 1, \ldots, q - 1\}$. If q is not prime – or even if it is – we may always find a prime p between $q/3$ and $2q/3$ by Bertrand's postulate. Then, either

$$|A \cap \{0, \ldots, p - 1\}| \geq \delta(1 - \delta/200)p \quad \text{or} \quad |A \cap \{0, \ldots, p - 1\}| < \delta(1 - \delta/200)p.$$

The latter case is actually easier, because then

$$|A \cap \{p, \ldots, q - 1\}| \geq \delta(q - (1 - \delta/200)p) = \delta((q - p) + \delta p/200) \geq \delta(1 + \delta/400)(q - p),$$

which means that we can iterate inside the progression $P' = \{p, \ldots, q - 1\}$, where A has density at least $\delta(1 + \delta/400) > \delta$.

Case 2. (reduction to the middle third) So, we only need to worry about the case $|A \cap \{0,\dots,p-1\}| \geq \delta' p$, where p is prime and $\delta' = \delta(1-\delta/200)$.[6] The density is temporarily a little smaller than δ, but in this case we will find a subprogression inside which the density increases again markedly above δ. Let $B := A \cap [p/3, 2p/3)$. If $|B| < \delta' p/5$, then either $A \cap [0, p/3)$ or $A \cap [2p/3, p)$ has cardinality at least $2\delta' p/5 = (6\delta'/5)(p/3)$. Since $6\delta'/5 = 6\delta(1-\delta/200)/5 > \delta(1+\delta/1000)$, in this case we can continue the iteration inside either $A \cap [0, p/3)$ or $A \cap [2p/3, p)$. So, we may assume that

$$|B| \geq \delta' p/5 \tag{4.12}$$

in the sequel. This will be used to guarantee that an arithmetic progression inside $A \cap \{0,\dots,p-1\}$ actually lies inside the "middle third" B, so that it must be a genuine progression (and not something like $\{p-2, p-1, 0\}$).

Case 3. (non-pseudorandom case) Write $A_p := A \cap \{0,\dots,p-1\}$, and let $\alpha := (\delta')^2/10$. If $|\widehat{A_p}(\xi)| > \alpha p$ for some $\xi \in \{1,\dots,p-1\}$ (where the Fourier transform is taken with respect to $\mathbb{Z}_p$), then the previous corollary states that there exists an arithmetic progression $P' \subset \{0,\dots,p-1\}$ of cardinality $|P'| \geq (\alpha p/32\pi)^{1/2}$ such that

$$\begin{aligned} |A_p \cap P'| &\geq (\delta' + (\delta')^2/100)|P'| \\ &= (\delta(1-\delta/200) + \delta^2(1-\delta/200)^2/100)|P'| \\ &\geq (\delta + \delta^2/500)|P'|. \end{aligned}$$

So, in this case we may continue the iteration inside P'.

Case 4. (pseudorandom case) The remaining case is where $|\widehat{A_p}(\xi)| \leq \alpha p$ for all $\xi \in \{1,\dots,p-1\}$. Now we will directly find plenty of 3-term arithmetic progressions inside A_p. Observe that these correspond to the triples $(x,y,z) \in A_p^3$ such that $x+z-2y=0$, so the strategy will be to find plenty of such triples. For technical reasons alluded to above, we will in fact search for such triples inside $A_p \times B^2$. Here we use the very useful formula

$$\sum_{\xi=0}^{p-1} \omega^{x\xi} = \begin{cases} p, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

With this in mind, the number of triples $(x,y,z) \in A \times B^2$ such that $x+z-2y=0$ is

$$\begin{aligned} \frac{1}{p} \sum_{x \in A_p} \sum_{y \in B} \sum_{z \in B} \sum_{\xi=0}^{p-1} \omega^{\xi(2y-x-z)} &= \frac{1}{p} \sum_{\xi=0}^{p-1} \widehat{A_p}(\xi) \hat{B}(-2\xi) \hat{B}(\xi) \\ &\geq \frac{1}{p}|A_p||B|^2 - \frac{1}{p} \max_{\xi \neq 0} |\widehat{A_p}(\xi)| \left(\sum_{\xi \neq 0} |\hat{B}(-2\xi)|^2 \right)^{1/2} \left(\sum_{\xi \neq 0} |\hat{B}(\xi)|^2 \right)^{1/2} \\ &\geq \delta'|B|^2 - \alpha|B|p = |B|(\delta'|B| - \alpha p) \geq \left(\frac{\delta' p}{5}\right)\left(\frac{(\delta')^2 p}{5} - \frac{(\delta')^2 p}{10}\right) = \frac{(\delta')^3 p^2}{50}. \end{aligned}$$

[6] I'm not sure whether the primality assumption gets used anywhere in the proof. Gowers seems to think so, and I chose to play it safe and write the argument in this way, so the assumption is readily available. I'm grateful to anyone, who could point out to me, where primality is needed.

using Cauchy-Schwarz, Proposition 4.3(ii), (4.12) and the definition of α. Any triple $(x, y, z) \in A_p \times B^2$ with $x + z = 2y$ clearly corresponds to a genuine arithmetic progression in A_p (i.e. the quad-sum above does not accidentally count triples like $(x, y, z) = (p-2, p-1, 0)$). The only issue remaining is that we have not ruled out the possibility $x = y = z$, but because there are at most p such triples, and $(\delta')^3 p^2/50 > p$ for large p, this is no real issue either. So, in this last remaining case, we have found a genuine three-term arithmetic progression inside A_p, and the proof of Roth's theorem is complete. □

Remark 4.13. The proof above gives the following more quantitative result: if $\delta > 0$, and if $n \geq \exp\exp(C\delta^{-1})$ for some absolute constant $C \geq 1$, then any set $A \subset \{1, \ldots, n\}$ of size at least δn contains a 3-term arithmetic progression.

5. Sum-product theory

The typical example of a set $A \subset \mathbb{R}$ with $|A + A| \sim |A|$ is a generalised arithmetic progression – and Freiman's theorem says that the typical example is roughly the only example. Then, what is the typical example of a set $A \subset \mathbb{R}$ with $|A \cdot A| \sim |A|$? By analogy, this would be the geometric progression

$$G = G(x, y) = \{x \cdot y^k : k \in \mathbb{Z}\},$$

or some d-dimensional variant thereof. It is clear intuitively that arithmetic and geometric progressions "look quite different", so any set cannot "simultaneously look like both". Since $|A + A| \sim |A|$ implies that "A looks like an arithmetic progression" by Freiman, this heuristic suggests that $|A \cdot A| \gg |A|$. Erdős and Szemerédi took this heuristic very seriously and conjectured the following in 1983:

Conjecture 5.1 (Sum-product conjecture)**.** *For any $\epsilon > 0$, there is a constant $c(\epsilon) > 0$ such that*

$$\max\{|A + A|, |A \cdot A|\} \geq c_\epsilon |A|^{2-\epsilon}.$$

Observe that $|A \cdot A| \leq |A|^2$ for any set $A \subset \mathbb{R}$, so the conjecture is the strongest possible. This conjecture has been the subject of vigorous research recently; you will get a good idea of this by googling "sum-product theorem", and hence a comprehensive bibliography is omitted. As far as I know, the best bound as of 2015 is the following by S. Konyagin and I. Skhredov [24]:

Theorem 5.2 (Konyagin and Skhredov (2015))**.** *For any set $A \subset \mathbb{R}$,*

$$\max\{|A + A|, |A \cdot A|\} \gtrsim |A|^{4/3+c},$$

where $c > 0$ is an absolute constant.

The proof of Konyagin and Skhredov builds heavily on a breakthrough result of Solymosi [31] from 2008:

Theorem 5.3. *For any set $A \subset \mathbb{R}$,*

$$\max\{|A + A|, |A \cdot A|\} \gtrsim \frac{|A|^{4/3}}{\log |A|}.$$

We will prove Solymosi's result below. According to some, "the proof is short enough to fit on a napkin", but one still needs quite small handwriting.

5.1. **Solymosi's sum-product theorem.** We start with a heuristic proof, then give the full details. They heuristic proof actually suggests that

$$\max\{|A+A|, |A/A|\} \gtrsim |A|^{4/3},$$

where $A/A = \{b/a : a, b \in A\}$, but making the argument rigorous will eventually cost us the $1/\log|A|$ -factor. It does not matter much, whether we study $A \cdot A$ or A/A, but the latter has a clearer geometric interpretation. Assume that A consists entirely of positive reals, so A/A is certainly well-defined. Observe that any pair $(a, b) \in A \times A \subset \mathbb{R}^2$ is contained on the half-line

$$L_{a,b} = \{(t, bt/a) : t > 0\} \subset \mathbb{R}^2.$$

So, $A \times A$ is contained on precisely $N := |A/A|$ such half-lines, see Figure 1. Enumerate

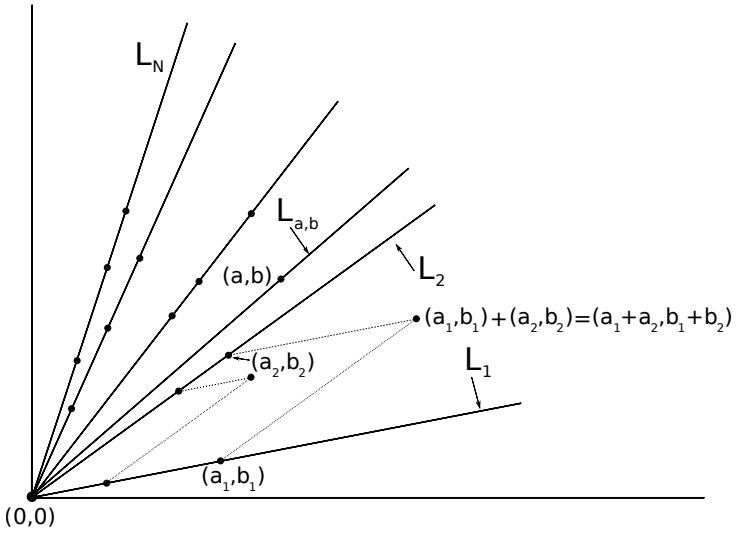

FIGURE 1. The proof of Solymosi's theorem.

this lines as $\{L_1, L_2, \dots, L_N\}$ in increasing order according to slope (so L_1 is the "most horizontal" one and L_N is the "most vertical one"). Assume that each L_j contains equally many points of $A \times A$: this number is of course $|A|^2/|A/A|$. Then, fix a pair L_j, L_{j+1} of consecutive half-lines, and consider vector sums of the form

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) \in (A + A) \times (A + A), \tag{5.4}$$

where $(a_1, b_1) \in L_j$ and $(a_2, b_2) \in L_{j+1}$. As the points on L_j and L_{j+1} vary, such vector sums are never repeated, because the lines (rather: the spanning vectors of) L_j and L_{j+1} are linearly independent. The conclusion is that, for each pair L_j, L_{j+1}, we obtain exactly

$$|L_j \cap (A \times A)| \cdot |L_{j+1} \cap (A \times A)| = \frac{|A|^2}{|A/A|} \cdot \frac{|A|^2}{|A/A|} = \frac{|A|^4}{|A/A|^2}$$

vector sums of the form (5.4) in $(A+A)\times(A+A)$. There are $|A/A|-1$ pairs of consecutive half-lines, and the vector sums corresponding to distinct pairs are clearly distinct (as they

lie "between the lines"), so we obtain the lower bound

$$|A+A|^2 = |(A+A) \times (A+A)| \geq (|A/A|-1) \cdot \frac{|A|^4}{|A/A|^2} \sim \frac{|A|^4}{|A/A|}.$$

In other words, $|A/A||A+A|^2 \gtrsim |A|^4$, and this yields $\max\{|A+A|, |A/A|\} \gtrsim |A|^{4/3}$. The heuristic proof of Solymosi's theorem is complete; next, we show how to get rid of the assumption that all the lines L_j contain an equal number of points in $A \times A$, and how to deduce a statement about $A \cdot A$ rather than A/A.

Detailed proof of Theorem 5.3. The connection between the sets $A \cdot A$ and A/A can be seen by studying the concept of *multiplicative energy*

$$E_\times(A) = |\{(a,b,c,d) \in A^4 : ad = bc\}|.$$

(This is the multiplicative analogue of the "additive energy" $T(A)$ we saw in connection with the Balog-Szemerédi-Gowers lemma). Similarly to the heuristic proof above, we are aiming for the estimate

$$|A+A|^2 \gtrsim \frac{|A|^4}{|A \cdot A| \log |A|},$$

and this will follow from the following two inequalities:

$$|A+A|^2 \gtrsim \frac{E_\times(A)}{\log |A|} \geq \frac{|A|^4}{|A \cdot A| \log |A|}. \tag{5.5}$$

The latter one is a simple application of Cauchy-Schwarz:

$$\begin{aligned} |A|^2 &= \sum_{r \in A \cdot A} |\{(a,b) \in A^2 : ab = r\}| \\ &\leq |A \cdot A|^{1/2} \left(\sum_{r \in A \cdot A} |\{(a,b) \in A^2 : ab = r\}|^2 \right)^{1/2} \\ &= |A \cdot A|^{1/2} |\{(a,b,c,d) \in A^4 : ad = bc\}|^{1/2} = |A \cdot A|^{1/2} E_\times(A)^{1/2}. \end{aligned}$$

So, it remains to prove the first inequality in (5.5). Assume that $A \subset (0,\infty)$, and $\log |A| \in \mathbb{N}$. Now, the benefit of considering the multiplicative energy becomes clear: $E_\times(A)$ is clearly the same number as the "divisive energy"

$$|\{(a,b,c,d) \in A^4 : a/c = b/d\}| = \sum_{r \in A/A} |\{(a,b) \in A^2 : b/a = r\}|^2,$$

and now we are back in the picture of the "heuristic proof", counting pairs $(a,b) \in \mathbb{R}^2$, which lie on a common half-line through the origin. Write further

$$E_\times(A) = \sum_{j=0}^{\log |A|-1} \sum_{\substack{r \in A/A \\ 2^j \leq |\{(a,b) \in A^2 : b/a = r\}| \leq 2^{j+1}}} |\{(a,b) \in A^2 : a/b = r\}|^2.$$

Now, we may fix $j \in \{0, 1, \ldots, \log |A| - 1\}$ such that

$$\frac{E_\times(A)}{\log |A|} \lesssim 2^{2j} |\{r \in A/A : 2^j \leq |\{(a,b) \in A^2 : b/a = r\}| \leq 2^{j+1}\}|. \tag{5.6}$$

It remains to estimate the size of the set $R := \{r \in A/A : 2^j \leq |\{(a,b) \in A^2 : b/a = r\}| \leq 2^{j+1}\}|$. The geometric interpretation is something we have seen before: the quantity $N := |R|$ counts the number of half-lines containing roughly 2^j points of $A \times A$. Let $L_1, \ldots, L_N$ be an enumeration of the half-lines $L_r = \{(t, rt) \in \mathbb{R}^2 : t > 0\}$, with $r \in R$, and in increasing order of slope (as in the "heuristic proof"). Then,

$$2^j \leq |L_i \cap (A \times A)| \leq 2^{j+1}, \qquad 1 \leq i \leq N.$$

For each pair L_i, L_{i+1} of consecutive half-lines, consider the collection of vector sums of the form

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) \in (A + A) \times (A + A),$$

where $(a_1, b_1) \in L_i$ and $(a_2, b_2) \in L_{i+1}$. As we discussed in the "heuristic proof", these vectors sums are distinct, and their total number (for all pairs of consecutive half-lines) is

$$\sum_{i=0}^{N-1} |L_i \cap (A \times A)||L_{i+1} \cap (A \times A)| \sim 2^{2j} N.$$

As the vector sums are contained in $(A \times A) + (A \times A) = (A + A) \times (A + A)$, we can combine the estimate above with (5.6) to infer that

$$|A + A|^2 \gtrsim 2^{2j} N \gtrsim \frac{E_\times(A)}{\log |A|}.$$

The proof of (5.5) – and Theorem 5.3 – is complete. □

5.2. **The Bourgain-Katz-Tao Sum-product theorem in prime fields.** In this subsection, we consider the sum-product problem in a the finite field $\mathbb{F} = \mathbb{Z}_p$, p prime, where a geometric argument such as the one above no longer works. In this context, we prove the following theorem of J. Bourgain, N. Katz and T. Tao:

Theorem 5.7 (Sum-product theorem in finite fields)**.** *Let $\mathbb{F} = \mathbb{Z}_p$, and assume that $A \subset \mathbb{F}$ satisfies $|\mathbb{F}|^\delta < |A| < |\mathbb{F}|^{1-\delta}$ for some $\delta > 0$. Then,*

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

where $\epsilon > 0$ depends only on δ.

The proof below follows mostly the original article [6], but also incorporates some simplifications from [4]. In particular, the proof of the following proposition is from [4]:

Proposition 5.8. *Assume that $A \subset \mathbb{F}$, where $\mathbb{F}$ is an arbitrary field, and*

$$|A + A| \lesssim |A|^{1+\epsilon} \quad \text{and} \quad |A \cdot A| \lesssim |A|^{1+\epsilon}.$$

Then, there exists a subset $A_1 \subset A$ with $|A_1| \gtrsim |A|^{1-\epsilon}$ such that $|kA_1^k| \lesssim_k |A|^{1+C_k\epsilon}$ for all $k \in \mathbb{N}$.

Before giving the proof, we recall Exercise 2.11:

Exercise 5.9. Assume that B_1, B_2, B_3 are subsets of an Abelian group, satisfying

$$|B_1 \cap B_3| \geq \frac{|B_1|}{K} \quad \text{and} \quad |B_2 \cap B_3| \geq \frac{|B_2|}{K}$$

and

$$|B_i + B_i| \leq K|B_i|, \qquad i \in \{1, 2, 3\}.$$

Then

$$|B_1 + B_2| \leq K^5 |B_3|.$$

Proof of Proposition 5.8. We start by repeating a Cauchy-Schwarz estimate from the proof of Theorem 5.3, and using the assumption $|A \cdot A| \sim |A|$:

$$\begin{aligned} |A|^2 &= \sum_{r \in A \cdot A} |\{(a,b) \in A^2 : ab = r\}| \\ &\leq |A \cdot A|^{1/2} |\{(a,b,c,d) \in A^4 : ab = cd\}|^{1/2} \\ &\lesssim |A|^{(1+\epsilon)/2} \left(\sum_{a,c \in A} |\{(b,d) \in A^2 : ab = cd\}| \right)^{1/2}. \end{aligned}$$

Here $|\{(b,d) \in A^2 : ab = cd\}| = |\{b \in A : ab \in cA\}| = |aA \cap cA|$. Consequently,

$$\frac{1}{|A|} \sum_{a \in A} \left[\sum_{c \in A} |aA \cap cA| \right] \gtrsim |A|^{2-\epsilon},$$

whence there exists $a_0 \in A$ such that

$$\sum_{c \in A} |a_0 A \cap cA| \geq \kappa |A|^{2-\epsilon}$$

for some constant $\kappa > 0$. Furthermore, the set

$$A_1 := \{c \in A : |a_0 A \cap cA| \geq \kappa |A|^{1-\epsilon}/2\}$$

satisfies

$$|A_1| \geq \frac{\kappa |A|^{1-\epsilon}}{2},$$

because otherwise

$$\kappa |A|^{2-\epsilon} < \sum_{c \in A_1} |A| + \sum_{A \setminus A_1} \kappa |A|^{1-\epsilon}/2 \leq \frac{\kappa |A|^{2-\epsilon}}{2} + \frac{\kappa |A|^{2-\epsilon}}{2} = \kappa |A|^{2-\epsilon}.$$

Now, we claim that A_1 is the subset we are after; the size is correct, at least.

Claim 5.10. *Let $y_1, y_2 \in A_1^k$. Then*

$$|y_1 A + y_2 A| \lesssim_k |A|^{1+5^k \epsilon}.$$

Proof of claim. To establish a basis for induction, fix $x_1, x_2 \in A_1$. Then, let

$$B_1 := x_1 A, \quad B_2 := x_2 A \quad \text{and} \quad B_3 := a_0 A.$$

Since $|A + A| \lesssim |A|^{1+\epsilon}$, we have $|B_i + B_i| \lesssim |A|^{\epsilon} |B_i|$ for $i \in \{1,2,3\}$. Also, by definition of A_1,

$$|B_i \cap B_3| = |a_0 A \cap x_i A| \geq \frac{|B_i|}{2\kappa^{-1} |A|^{\epsilon}}, \quad i \in \{1,2\},$$

so the hypotheses of Exercise 5.9 are valid with $K = 2\kappa^{-1} |A|^{\epsilon}$. It follows that

$$|x_1 A + x_2 A| = |B_1 + B_2| \leq K^5 |B_3| = 32 \kappa^{-5} |A|^{5\epsilon} |A|,$$

which proves the case $k = 1$ of the claim.

For $k \geq 2$, fix $y_1 = x_1^1 \cdots x_k^1 \in A_1^k$ and $y_2 = x_1^2 \cdots x_k^2 \in A_1^k$. To apply Exercise 5.9 again, we set

$$B_1 := y_1 A, \quad B_2 := y_2 A \quad \text{and} \quad B_3 := x_1^1 \cdots x_{k-1}^1 a_0 A \cup x_1^2 \cdots x_{k-1}^2 a_0 A.$$

It is clear that $|B_1 + B_1| \leq |A|^\epsilon |B_1|$ and $|B_2 + B_2| \leq |A|^\epsilon |B_2|$. By induction, we may also assume that

$$|(x_1^1 \cdots x_{k-1}^1 A) + (x_1^2 \cdots x_{k-1}^2 A)| \lesssim_k |A|^{5^{k-1}\epsilon} |A|,$$

which gives $|B_3 + B_3| \lesssim_k |A|^{5^{k-1}\epsilon} |A|$. Again, it follows from the definition of A_1 that

$$|B_i \cap B_3| \geq |x_k^i A \cap a_0 A| \geq \frac{|B_i|}{2\kappa^{-1}|A|^\epsilon}, \quad i \in \{1, 2\}.$$

So, applying Exercise 5.9 with $K \sim_k |A|^{5^{k-1}\epsilon}$ gives

$$|y_1 A + y_2 A| = |B_1 + B_2| \leq K^5 |B_3| \sim_k |A|^{5^k \epsilon} |A|.$$

This proves the claim. □

Observe that the previous claim also holds, if $y_1, y_2 \in A_1^{-1} A_1^k$. This follows by writing $y_1 = x_1^{-1} \tilde{y}_1$ and $y_2 = x_2^{-1} \tilde{y}_2$ and doing the following manipulation:

$$|y_1 A + y_2 A| = \left| \frac{\tilde{y}_1 A}{x_1} + \frac{\tilde{y}_2 A}{x_2} \right| = \left| \frac{x_2 \tilde{y}_1 A + x_1 \tilde{y}_2 A}{x_1 x_2} \right| \lesssim_k |A|^{1+5^{k+1}\epsilon}. \tag{5.11}$$

We can now complete the proof of Proposition 5.8. It suffices to show that

$$|A_1^k + A_1^k| \lesssim_k |A|^{1+C_k\epsilon} \lesssim |A|^{(C_k+1)\epsilon} |A_1^k|, \tag{5.12}$$

because then

$$|kA_1^k| \lesssim_k |A|^{k(C_k+1)\epsilon} |A_1^k| \lesssim_k |A|^{k(C_k+k+1)\epsilon} |A|,$$

by two applications of the Plünnecke-Ruzsa inequalities (both in the additive and multiplicative groups of $\mathbb{F}$).

To prove (5.12), write

$$\chi_{A_1^k(x)} \leq \frac{1}{|A_1|} \sum_{y \in A_1^{-1} A_1^k} \chi_{yA_1}(x),$$

Indeed, if $x \in A_1^k$, then $x \in yA_1$, whenever $y \in xA_1^{-1}$, and there are $|xA_1^{-1}| = |A_1|$ such choices of $y \in xA_1^{-1} \subset A_1^{-1} A_1^k$.

Now, assume that $x \in A_1^k + A_1^k$, so that $x = y + (x - y)$, where $y, x - y \in A_1^k$. Then

$$\begin{aligned} \chi_{A_1^k + A_1^k}(x) &\leq \chi_{A_1^k}(y) \chi_{A_1^k}(x - y) \\ &\leq \frac{1}{|A_1|^2} \sum_{y_1, y_2 \in A_1^{-1} A_1^k} \chi_{y_1 A_1}(y) \chi_{y_2 A_1}(x - y) \\ &\leq \frac{1}{|A_1|^2} \sum_{y_1, y_2 \in A_1^{-1} A_1^k} \chi_{y_1 A_1 + y_2 A_1}(x). \end{aligned}$$

This, along with (5.11), implies that

$$|A_1^k + A_1^k| \leq \frac{1}{|A_1|^2} \sum_{y_1, y_2 \in A_1^{-1} A_1^k} |y_1 A + y_2 A| \lesssim_k \frac{|A^{-1} A^k|^2}{|A_1|^2} |A|^{1+5^{k+1}\epsilon}. \tag{5.13}$$

Here, using the multiplicative version of the Plünnecke-Ruzsa inequalities,

$$|A^{-1}A^k| \lesssim_k |A|^{(k+1)\epsilon}|A| \leq \frac{2|A|^{(k+2)\epsilon}}{\kappa}|A_1|.$$

Plugging this into (5.13) proves (5.12) and completes the argument for Proposition 5.8. □

Recall that we are in the process of proving the the sum-product theorem in finite fields, Theorem 5.7. By assumption, $\mathbb{F}$ is a finite field, and $A \subset \mathbb{F}$ is a set with $|\mathbb{F}|^{\delta} < |A| < |\mathbb{F}|^{1-\delta}$ for some positive parameter $\delta > 0$.

Lemma 5.14. *Assume that $|A| \geq |\mathbb{F}|^{\delta} > 2$. There exist a positive integer $k \sim 1/\delta$, and invertible elements $\xi_1, \ldots, \xi_k \in \mathbb{F}$ such that the linear mapping $\Lambda \colon A^k \to \mathbb{F}$ defined by*

$$\Lambda(x_1, \ldots, x_k) := x_1\xi_1 + \ldots + x_k\xi_k$$

is a surjection, that is, $\mathbb{F} = \xi_1 A + \ldots + \xi_k A$.

Proof. We find the elements $\xi_j \in \mathbb{F}^* := \mathbb{F} \setminus \{0\}$ one by one, using the following claim iteratively:

Claim 5.15. *Let $A, B \subset \mathbb{F}$ be non-empty. Then, there exists an element $\xi \in \mathbb{F}^*$ such that*

$$|A + \xi B| \geq \min\{|A||B|/2, |\mathbb{F}|/10\}. \tag{5.16}$$

Proof of Claim. We prove that if $|A||B| \leq (|\mathbb{F}| - 1)/2$, then we can find $\xi \in \mathbb{F}^*$ such that $|A+\xi B| \geq |A||B|/2$. Why is this sufficient also if $|A||B| > (|\mathbb{F}|-1)/2$? In this case, choose any subsets $A' \subset A$ and $B' \subset B$ with $|\mathbb{F}|/4 \leq |A'||B'| \leq (|\mathbb{F}| - 1)/2$. Then, we can find $\xi \in \mathbb{F}^*$ such that

$$|A + \xi B| \geq |A' + \xi B'| \geq |A'||B'|/2 \geq |\mathbb{F}|/10.$$

So, for the remainder of the proof, assume that $|A||B| \leq (|\mathbb{F}| - 1)/2$. We use the following standard inequality:

$$\left|\bigcup_{j \in J} B_j\right| \geq \sum_{j \in J} |B_j| - \frac{1}{2}\sum_{j \neq j'} |B_j \cap B_{j'}|.$$

Given an arbitrary element $\xi \in \mathbb{F}^*$, this gives

$$\begin{aligned} |A + \xi B| = \left|\bigcup_{a \in A} a + \xi B\right| &\geq \sum_{a \in A} |a + \xi B| - \frac{1}{2}\sum_{a \neq a'} |(a + \xi B) \cap (a' + \xi B)| \\ &\overset{(*)}{=} |A||B| - \frac{1}{4}\sum_{a \neq a'}\sum_{b \neq b'} \chi_{a+b\xi = a'+b'\xi} \\ &= |A||B| - \frac{1}{4}\sum_{a \neq a'}\sum_{b \neq b'} \chi_{(a-a')/(b-b') = \xi}. \end{aligned}$$

Here $(*)$ follows by observing that to each point in $(a+\xi B) \cap (a' + \xi B)$, there corresponds a unique two-element set $\{b, b'\} \subset B$ with $b \neq b'$ such that $a + b\xi = a' + b'\xi$, and the sum over the pairs (b, b'), $b \neq b'$, counts each $\{b, b'\}$ twice.

Next, we average the inequality over the elements $\xi \in \mathbb{F}^*$ to obtain

$$\frac{1}{|\mathbb{F}^*|}\sum_{\xi\in\mathbb{F}^*}|A+\xi B| \geq |A||B| - \frac{1}{4}\sum_{a\neq a'}\sum_{b\neq b'}\frac{1}{|\mathbb{F}|-1} \geq |A||B| - \frac{1}{4}\frac{|A|^2|B|^2}{|\mathbb{F}|-1}.$$

Using $|A||B| \leq (|\mathbb{F}|-1)/2$, Claim 5.15 follows. □

To prove Lemma 5.14, we require the finite field version of the Cauchy-Davenport inequality – our very first proposition – which says that

$$|A+B| \geq \min\{|A|+|B|-1, |\mathbb{F}|\} \tag{5.17}$$

for all non-empty sets $A, B \subset \mathbb{F}$. Despite the resemblance to Proposition 1.1, the inequality (5.17) is a bit harder; we will omit the proof here, but you can find various arguments in the lecture notes [32] of Tao.

Using Claim 5.15, choose $\xi_1 \in \mathbb{F}^*$ such that

$$|A+\xi_1 A| \geq \min\{|A|^2/2, |\mathbb{F}|/10\}.$$

If the $\min$ is $|\mathbb{F}|/10$, stop. Otherwise choose another element $\xi_2 \in \mathbb{F}^*$ such that

$$|(A+\xi_1 A)+\xi_2 A| \geq \min\{|A|^3/4, |\mathbb{F}|/10\}.$$

Again, if the $\min$ is $|\mathbb{F}|/10$, stop. Otherwise, iteration eventually gives

$$|A+\xi_1 A+\ldots+\xi_k A| \geq \min\{|A|^{k+1}/2^k, |\mathbb{F}|/10\}.$$

Since $|A| > 2$, we have $|A|^\rho = 2$ for some $\rho \in (0, 0.99)$, so $|A|^{k+1}/2^k \geq |A|^{(1-\rho)(k+1)} \geq |\mathbb{F}|^{0.1\delta(k+1)}$. This quantity becomes larger than $|\mathbb{F}|/10$ in $k \sim 1/\delta$ steps, and then we have

$$|A+\xi_1 A+\ldots+\xi_k A| \geq |\mathbb{F}|/10.$$

Finally, we sum $A+\xi_1 A+\ldots+\xi_k A$ with itself a few times, until (5.17) shows that the resulting sumset has cardinality precisely $|\mathbb{F}|$. This completes the proof of the lemma. □

Having now found an initial surjection $\Lambda\colon A^k \to \mathbb{F}$, the next step is to modify Λ to a surjection from a certain lower-order product set to $\mathbb{F}$. Here we need the assumption $\mathbb{F} = \mathbb{Z}_p$.

Lemma 5.18. *Let $B \subset \mathbb{F} = \mathbb{Z}_p$ be non-empty, assume that $k > 1$, and suppose that $\Lambda\colon B^k \to \mathbb{F}$ is a linear surjection,*

$$\Lambda(x_1,\ldots,x_k) = x_1\xi_1 + \ldots + x_k\xi_k.$$

Then, there exists a linear surjection $\tilde{\Lambda}\colon \tilde{B}^{k-1} \to \mathbb{F}$ of the same form, where

$$\tilde{B} = B\cdot(B-B) - B\cdot(B-B).$$

Proof. The mapping Λ is a surjection, but it cannot be a bijection, because otherwise $p = |B|^k$, contradicting the primality of p. So, we find two distinct elements $(b_1,\ldots,b_k)$ and $(b_1',\ldots,b_k')$ such that

$$(b_1-b_1')\xi_1 + \ldots + (b_k-b_k')\xi_k = 0. \tag{5.19}$$

For instance, assume that $b_k \neq b_k'$. Then, using the assumption on Λ, and (5.19), we get

$$\begin{aligned}\mathbb{F} &= (b_k - b_k')[\xi_1 B + \ldots + \xi_k B] = \xi_1(b_k - b_k')B + \ldots + \xi_k(b_k - b_k')B \\ &= [(b_k - b_k')\xi_1 B + \ldots + (b_k - b_k')\xi_{k-1}B] - ((b_1 - b_1')\xi_1 + \ldots + (b_{k-1} - b_{k-1}')\xi_{k-1})B \\ &\subset [(b_k - b_k')B - (b_1 - b_1')B]\xi_1 + \ldots + [(b_k - b_k')B - (b_{k-1} - b_{k-1}')B]\xi_{k-1} \\ &\subset \tilde{B}\xi_1 + \ldots + \tilde{B}\xi_{k-1}.\end{aligned}$$

This means that $\tilde{\Lambda}(x_1, \ldots, x_{k-1}) = x_1\xi_1 + \ldots + x_{k-1}\xi_{k-1}$ is a surjection on $\tilde{B}^{k-1}$. □

We are ready to finish the proof of Theorem 5.7:

Proof of Theorem 5.7. Assume that $\max\{|A \cdot A|, |A + A|\} \lesssim |A|^{1+\epsilon}$ for some small $\epsilon \in (0, 1/2)$. By Proposition 5.8, we may find a subset $A_1 \subset A$ such that $|A_1| \gtrsim |A|^{1-\epsilon} \geq |\mathbb{F}|^{(1-\epsilon)\delta} \geq |\mathbb{F}|^{\delta/2}$, and

$$|kA_1^k| \lesssim_k |A|^{1+C_k\epsilon} \lesssim |A_1|^{(1+C_k\epsilon)/(1-\epsilon)} \leq |A_1|^{1+2(C_k+1)\epsilon}, \qquad k \in \mathbb{N}. \tag{5.20}$$

We aim for a lower bound on ϵ, depending only on δ. Let $\Lambda \colon A_1^k \to \mathbb{F}$ be the initial surjection given by Lemma 5.14, where $k \lesssim 1/\delta$. Since $|A| \leq |\mathbb{F}|^{1-\delta}$, we have $k > 1$. Let F be the "set-valued function"

$$F(B) := B \cdot (B - B) - B \cdot (B - B), \qquad B \subset \mathbb{F}.$$

By Lemma 5.18, we can find a linear surjection $[F(A_1)]^{k-1} \to \mathbb{F}$. Another iteration gives a linear surjection $[F^2(A_1)]^{k-2} \to \mathbb{F}$, and exactly $k-1$ iterations gives a linear surjection

$$F^{k-1}(A_1) \to \mathbb{F}.$$

In particular,

$$|F^{k-1}(A_1)| \geq |\mathbb{F}|.$$

On the other hand,

$$F(B) = B \cdot (B - B) - B \cdot (B - B) \subset 2(B \cdot B) - 2(B \cdot B), \tag{5.21}$$

so

$$|F(A_1)| \leq |2(A_1 \cdot A_1) - 2(A_1 \cdot A_1)| \lesssim |A_1|^{8(C_2+1)\epsilon}|A_1 \cdot A_1| \lesssim |A|^{[8(C_2+1)+1]\epsilon}|A|.$$

by (5.20) with $k = 2$, the Plünnecke-Ruzsa inequalities and the assumption $|A \cdot A| \lesssim |A|^{1+\epsilon}$. Similarly, applying (5.21) with $B = F^{j-1}(A_1)$, $1 \leq j \leq k-1$, one can show that

$$|F^j(A_1)| \lesssim |A|^{D_j\epsilon}|A|,$$

where D_j is some large constant depending only on the (absolute) constants C_j in (5.20). The precise argument is an easy induction. With $j = k-1$, we obtain

$$|\mathbb{F}| \leq |F^{k-1}(A_1)| \leq |A|^{D_{k-1}\epsilon}|A| \lesssim |\mathbb{F}|^{(1-\delta)(1+D_{k-1}\epsilon)}.$$

Since $k \lesssim 1/\delta$, here D_k depends only on δ, and we see that $\epsilon > 0$ must have a lower bound depending only on δ. The proof is complete. □

Remark 5.22. Soon after Theorem 5.7 appeared in 2003, S. Konyagin [23] observed that the assumption $|A| \geq |\mathbb{F}|^{\delta}$ is unnecessary. Later, in 2007, M. Garaev [11] obtained a quantitative version of Theorem 5.7, which was subsequently refined by Katz and Shen, Bourgain and Garaev [5] and Shen [30]. Shen's result,

$$\max\{|A+A|, |A \cdot A|\} \gtrsim \frac{A|^{13/12}}{\log^{1/3}|A|},$$

is the world record in the prime field setting, as far as I know, although M. Rudnev [28] proved an even slightly better bound for sets $A \subset \mathbb{F}$ with $|A| < |\mathbb{F}|^{1/2}$ in 2011. Katz and Shen [22] have also established a quantitative sum-product theorem in finite fields **not** of prime order, which has probably been improved by many other authors by today.

6. Incidence geometry

Here is a stereotypical problem in incidence geometry: you are given a set of points P, and a family of geometric objects $\mathcal{L}$ such as lines, circles, hyperplanes – you name it – and the problem is to bound the number of *incidences*

$$I(P, \mathcal{L}) := \{(p, l) \in P \times \mathcal{L} : p \in L\}.$$

between P and $\mathcal{L}$ from above. Often there are further restrictions on P and $\mathcal{L}$, requiring that not all point in P lie on a single line, or not all lines in $\mathcal{L}$ lie on a single plane etc.

It may appear somewhat inaccurate to use the title "Incidence geometry" only for this section, since the sum-product problem is very much a question in incidence geometry – at least in Solymosi's perspective! On the other hand, there are hardly any geometric objects "$\mathcal{L}$" visible in the proof of the Bourgain-Katz-Tao theorem, and that argument seems to belong purely to the realm of additive combinatorics. For exactly this reason, the sum-product problem makes for such a nice transition between the topics of the early course to the ones below: the two proofs in the previous section demonstrate how closely these areas are interconnected.

6.1. **Tools from topology.** When counting incidences, it is often very efficient to partition space – and the points in P therein – into "cells", then count incidences in the cells separately, and finally sum up the results. We will see a basic example of this shortly, in the proof of the Szemerédi-Trotter theorem. This theorem was found in 1983, and the idea of "cell partitioning" was present already then; back in those days, the geometric objects in $\mathcal{L}$ were used for the purpose of partitioning, but we will not speak more of that. A breakthrough idea from around 2011 – due to L. Guth and N. Katz [18], based on an earlier related idea of Z. Dvir [9] from 2005 – has been the use of zero-sets of polynomials (also known as algebraic varieties) for the purpose of partitioning. The technique not only tidies up the details of the partitioning procedure immensely, but also adds tools from algebraic geometry to the equation.

Those students, who have attended Topology I at the University of Helsinki should know the following very basic case of polynomial cell partitioning. The *ham sandwich theorem*, due to Banach from the 30's, states that if U_1, U_2 are finite volume open sets in $\mathbb{R}^2$, then some line bisects both U_1 and U_2 simultaneously, that is, cuts the volumes of U_1 and U_2 into half. Apparently, the right way to think about this result is the following: the line is the zero-set of a degree one polynomial in $\mathbb{R}^2$. This motivates the following theorem:

Theorem 6.1 (Polynomial ham sandwich theorem for open sets)**.** *Let $V = V(n,d)$ be the vector space of real polynomials in $\mathbb{R}^n$ of degree at most n. If $U_1, \ldots, U_N$ are finite sets of finite volume, and $N < \dim V = \binom{d+n}{n}$, then there exists a polynomial $f \in V$ such that the zero set of f, denoted by $Z(f) := \{x \in \mathbb{R}^n : f(x) = 0\}$, bisects every set U_i.*

To be precise, the claim says that

$$\mathcal{H}^n(U_i \cap \{f > 0\}) = \frac{\mathcal{H}^n(U_i)}{2} = \mathcal{H}^n(U_i \cap \{f < 0\}), \quad 1 \leq i \leq N.$$

Theorem 6.1 is easy to prove, IF you know the Borsuk-Ulam theorem:

Theorem 6.2 (Borsuk-Ulam)**.** *Let $n \geq 1$, and write $S^n := \{x \in \mathbb{R}^{n+1} : |x| = 1\}$. Suppose that $\phi \colon S^n \to \mathbb{R}^n$ is a continuous map such that $\phi(-x) = -\phi(x)$ for $x \in S^n$. Then $0 \in \phi(S^n)$.*

Unfortunately, we have no chance to prove the Borsuk-Ulam theorem, which is a rather deep result in algebraic topology. We will use it nevertheless:

Proof of Theorem 6.1. Assume without loss of generality that $N = \dim V - 1$ (otherwise add dummy sets). Then, identify $V(n,d)$ with $\mathbb{R}^{N+1}$, so that and let $S^N \subset V$ be the unit sphere of V obtained from this identification. Define $\phi \colon S^N \to \mathbb{R}^N$ by $\phi(f) = (\phi_1(f), \ldots, \phi_N(f))$, where

$$\phi_j(f) = \mathcal{H}^n(U_i \cap \{f > 0\}) - \mathcal{H}^n(U_i \cap \{f < 0\}).$$

One can easily check (using the boundedness of the volumes of the sets U_i, for details see [16, p. 4]) that ϕ is continuous, and clearly $\phi(-f) = -\phi(f)$ for all $f \in S^N$. It follows from the Borsuk-Ulam theorem that $\phi(f) = 0$ for some polynomial $f \in S^N$, and then $Z(f)$ bisects all the sets U_i. □

In the leader to this subsection, we stated that it is efficient to partition the points $P \subset \mathbb{R}^n$ into cells. Now, what is the relation between the open sets in Theorem 6.1 and the set P? In analogy with the definition above, we say that $Z(f)$ bisects P, if

$$|P \cap \{f < 0\}| \leq \frac{|P|}{2} \quad \text{and} \quad |P \cap \{f > 0\}| \leq \frac{|P|}{2}.$$

There is a innocent-looking but rather major difference to the definition for open sets: now, a large portion of the set P can lie on $Z(f)$ (which was completely negligible earlier). With this deifnition, we have the following analogue of Theorem 6.1

Lemma 6.3 (Polynomial ham sandwich theorem for finite sets)**.** *Let $P_1, \ldots, P_N$ be finite sets in $\mathbb{R}^n$. If $N < \dim V(n,d) = \binom{d+n}{n}$, there exists a non-zero polynomial $f \in V(n,d)$ such that $Z(f)$ bisects all the sets P_j, $1 \leq j \leq N$.*

Proof. Let $P_i(\delta)$ be the open δ-neighbourhood of P_i. By Theorem 6.1, a non-zero polynomial $f_\delta \in V(n,d)$ bisects every $P_i(\delta)$. Moreover, recall from the proof of Theorem 6.1 that $|f_\delta| = 1$ for every $\delta > 0$, where $|\cdot|$ is the norm we obtained by identifying $V(n,d)$ with $\mathbb{R}^{\dim V(n,d)}$. Since the unit sphere of $V(n,d)$ is compact, we can find a subsequence $(f_{\delta_j})_{j \in \mathbb{N}}$ converging to a polynomial $f \in V(n,d)$ with $|f| = 1$. In particular, f is non-zero. We claim that f bisects every set P_i.

It is easy to check that $f_{\delta_j} \to f$ uniformly on compact sets, using the fact that the the coefficients of the polynomials f_{δ_j} converge to those of f.[7] Now, if f did not bisect P_i, then $\{f > 0\}$, say, would contain more than half of the points in P_i. Call these points P_i^+. For $\epsilon > 0$ sufficiently small, it follows that $f > 0$ on the closure of the set $P_i^+(\epsilon)$ (which is a union of disjoint ϵ-balls for $\epsilon > 0$ small enough). By uniform convergence on compact sets, $f_{\delta_j} > 0$ on the set $P_i^+(\epsilon)$ for large enough $j \in \mathbb{N}$. Once $\delta_j < \epsilon$, this implies that $\{f_{\delta_j} > 0\}$ contains all the balls constituting $P_i^+(\delta_j)$, and the total mass of these balls is strictly larger than $\mathcal{H}^n(P_i(\delta))/2$. Thus, f_{δ_j} does not bisect $P_i(\delta)$, a contradiction. □

Remark 6.4. Observe that $\binom{d+n}{n} \geq d^n/n!$, so

$$N < \left(\frac{C^n}{n!}\right) N \leq \binom{CN^{1/n} + n}{n},$$

as soon as $C > (n!)^{1/n}$. In particular, this holds if $C = 2(n!)^{1/n}$. Consequently, given N sets $P_1, \ldots, P_N$ as in Lemma 6.3, we can find a non-zero polynomial of degree $d \leq CN^{1/n}$, which bisects each set P_i.

Now we are fully prepared to partition the set P into cells:

Theorem 6.5 (Polynomial cell partitions). *Let $P \subset \mathbb{R}^n$ be a finite set, and let $d \in \mathbb{N}$ be a fixed degree. Then, there exists a non-zero polynomial f of degree at most d such that $\mathbb{R}^n \setminus Z(f)$ can be partitioned into open sets $U_1, \ldots, U_m$, $m \leq d^n$, such that $\partial U_k \subset Z(f)$ and $|P \cap U_k| \lesssim |P|d^{-n}$ for each $1 \leq k \leq m$.*

Remark 6.6. It is worth noting that Theorem 6.5 gives no lower bound for the number of points of P in the open sets U_j, and indeed there is none to be had: it will be possible in practical situations that P is entirely contained in $Z(f)$.

Proof. The proof is an iteration of Lemma 6.3. First, choose a first degree polynomial f_1, which bisects the points in P, and let

$$P_1^1 := P \cap \{f_1 > 0\} \quad \text{and} \quad P_2^1 := P \cap \{f_1 < 0\}.$$

Then $|P_1| \leq |P|/2$ and $|P_2| \leq |P|/2$. Next, assume that we have already defined a sequence of polynomials $f_1, \ldots, f_j$ and associated sets $P_1^j, \ldots, P_{2^j}^j \subset P$ such that $|P_i^j| \leq |P|/2^j$ for $1 \leq i \leq 2^j$, and each P_i^j has the form $P_i^j = P \cap U_i^j$, where U_i^j is one of the 2^j sets specified by requiring that each f_i is either strictly negative or positive in U_i^j (for example, U_i^3 might look like $\{f_1 < 0, f_2 > 0, f_3 > 0\}$). Moreover, we assume that the degree of each f_i is bounded by $C2^{(i-1)/n}$, where $C \geq 1$ only depends on n. It follows that the degree of $F_j := f_1 \cdots f_j$ is bounded by $\sum_{i \leq j} C2^{i/n} \leq C'2^{j/n}$, where $C' \geq 1$ also only depends on n.

If j is the largest integer such that $C'2^{j/n} \leq d$, we stop the process. Then $2^j = cd^n \leq d^n$, where $c \in (0, 1)$ again only depends on n. In this case, we let $f = F_j$, and relabel the sets U_i^j, $1 \leq i \leq 2^j$, as $U_1, \ldots, U_m$. Then $m = 2^j \leq d^n$ as claimed, and it is clear that

[7] This follows from the fact that all norms in the finite dimensional vector space $V(n, d)$ are equivalent, and one possible choice for $|f|$ is the maximum absolute value of the coefficients of f. Then, convergence in this norm means precisely that the coefficients converge.

$\partial U_k \subset Z(f)$ for $1 \leq k \leq m$. Moreover, with $U_k = U_i^j$, we have $|P \cap U_k| = |P_i^j| \leq |P|/2^j = c^{-1}|P|d^{-n}$, so the proof is complete in this case.

In the opposite case, where $C'2^{(j+1)/n} \leq d$, we continue with the construction of f_{j+1}. By Remark 6.4, we can find a polynomial f_{j+1} with degree $\leq C2^{j/n}$, such that $Z(f_{j+1})$ bisects each of the finite sets $P_1^j, \ldots, P_{2^j}^j$. Thus, for $1 \leq i \leq 2^j$, we have

$$|P \cap [U_i^j \cap \{f_{j+1} > 0\}]| = |P_i^j \cap \{f_{j+1} > 0\}| \leq |P_i^j|/2 \leq |P|/2^{j+1}$$

and similarly

$$|P \cap [U_i^j \cap \{f_{j+1} < 0\}]| \leq |P|/2^{j+1}$$

The sets U_i^{j+1}, $1 \leq i \leq 2^{j+1}$, are defined to be all sets of the form $U_i^j \cap \{f_{j+1} > 0\}$ and $U_i^j \cap \{f_{j+1} < 0\}$, where $1 \leq i \leq 2^j$. Then, as required by the induction, we set $P_i^{j+1} := P \cap U_i^{j+1}$. By the inequalities above, the sets P_i^{j+1} satisfy the correct cardinality estimates, and the proof of the theorem is complete. □

6.2. **The Szemerédi-Trotter theorem via polynomial cell decompositions.** The material in this subsection is largely taken from T. Tao's blog entry dated 18/02/2011.

Given a set of points $P \subset \mathbb{R}^2$ and family of lines $\mathcal{L}$ also in $\mathbb{R}^2$ in arbitrary position, the Szemerédi-Trotter theorem gives a sharp bound on the number of incidences $I(P, \mathcal{L}) = \{(p, l) \in P \times \mathcal{L} : p \in L\}$:

Theorem 6.7 (Szemerédi-Trotter)**.** *For any set of points $P \subset \mathbb{R}^2$, and any set of lines $\mathcal{L}$ in $\mathbb{R}^2$, we have*

$$|I(P, L)| \lesssim |P|^{2/3}|L|^{2/3} + |P| + |L|.$$

As we advertised in the previous section to motivate the cell decompositions, the strategy of proof will be to split P into cells using Theorem 6.5, perform a "trivial estimate" for incidences inside each cell individually, and finally add up the outcomes. This way, however, we are only counting the incidences between $\mathcal{L}$ and the points of P contained in the union of the (open) cells: a further – fortunately simple – argument will be needed to handle the points of P contained in the boundary of the cells – which is, as we know, the zero-set of a relatively low-degree polynomial.

We start with the "trivial estimate" used to bound incidences in individual cells:

Lemma 6.8. *For any set of points $P \subset \mathbb{R}^2$ and any set of lines $\mathcal{L}$ in $\mathbb{R}^2$, we have*

$$|I(P, \mathcal{L})| \lesssim |P||\mathcal{L}|^{1/2} + |\mathcal{L}|.$$

and

$$|I(P, \mathcal{L})| \lesssim |P|^{1/2}|\mathcal{L}| + |P|.$$

Proof. We only prove the first inequality; the proof of the second is essentially the same, with the roles of P and $\mathcal{L}$ interchanged. By the definition of incidences, and Cauchy-Schwarz,

$$\begin{aligned}|I(P,\mathcal{L})| &= \sum_{L\in\mathcal{L}} |P\cap L| \\ &\leq |\mathcal{L}|^{1/2}\left(\sum_{L\in\mathcal{L}} |\{(p,q)\in P^2 : p,q\in L\}|\right)^{1/2} \\ &= |\mathcal{L}|^{1/2}\left(\sum_{p,q\in P} |\{L\in\mathcal{L} : p,q\in L\}|\right)^{1/2} \\ &\leq |\mathcal{L}|^{1/2}\left(\sum_{p\in P} |\{L : p\in L\}| + \sum_{p\neq q} |\{L : p,q\in L\}|\right)^{1/2}.\end{aligned}$$

The first sum inside the brackets is simply $I(P,\mathcal{L})$. To estimate the second sum, note that at most one line can pass through two distinct point in $\mathbb{R}^2$. So, the second sum is at most $|P|(|P|-1)\leq |P|^2$. All in all,

$$|I(P,\mathcal{L})| \lesssim |\mathcal{L}|^{1/2}|I(P,\mathcal{L})|^{1/2} + |P||\mathcal{L}|^{1/2},$$

which gives the claim after rearranging terms. □

We are ready to prove Theorem 6.7:

Proot of Theorem 6.7. Before starting the proof in earnest, we dismiss a few simple special cases, namely that either $|P| > |\mathcal{L}|^2/10$ or $|P| < 10|\mathcal{L}|^{1/2}$. In the first case, for instance, the previous lemma gives

$$|I(P,\mathcal{L})| \lesssim |P|^{1/2}|\mathcal{L}| + |P| \lesssim |P|,$$

and we are done. The second case is handled similarly, using the other inequality in the previous lemma. So, we assume that

$$10|\mathcal{L}|^{1/2} \leq |P| \leq |\mathcal{L}|^2/10 \tag{6.9}$$

in the future.

We apply the cell partition from Theorem 6.5 with a degree $d\in\mathbb{N}$ to be optimised later: we obtain a real-valued polynomial f on $\mathbb{R}^2$ with $\deg(f)\leq d$, and the sets $U_1,\dots,U_m$, $m\leq d^2$, such that $\partial U_k\subset Z(f)$ and $|P\cap U_k|\lesssim |P|d^{-2}$ for all $1\leq k\leq m$. Writing $Z := Z(f)$, we can count the incidences $I(P,\mathcal{L})$ as follows:

$$|I(P,\mathcal{L})| = \sum_{k=1}^{m} |I(P\cap U_k, \mathcal{L}\cap U_k)| + |I(P\cap Z,\mathcal{L})|,$$

where $\mathcal{L} \cap U_k$ stands for the lines in $\mathcal{L}$, which intersect U_k. By the Lemma 6.8 and Cauchy-Schwarz,

$$\begin{aligned}\sum_{k=1}^{m} |I(P \cap U_k, \mathcal{L} \cap U_k)| &\lesssim \sum_{k=1}^{m} (|P \cap U_k||\mathcal{L} \cap U_k|^{1/2} + |\mathcal{L} \cap U_k|) \\ &\lesssim \frac{|P|}{d^2} \sum_{k=1}^{m} |\mathcal{L} \cap U_k|^{1/2} + \sum_{k=1}^{m} |\mathcal{L} \cap U_k| \\ &\leq \frac{|P|}{d} \left(\sum_{k=1}^{m} |\mathcal{L} \cap U_k| \right)^{1/2} + \sum_{k=1}^{m} |\mathcal{L} \cap U_k|.\end{aligned}$$

Next comes one of main highlights of the proof: in order to estimate the sum $\sum |\mathcal{L} \cap U_k|$, we wish to find an upper bound for the number of cells U_k that any individual line in $\mathcal{L}$ can intersect. If this bound is $C \geq 1$, say, then

$$\sum_{k=1}^{m} |\mathcal{L} \cap U_k| = \sum_{L \in \mathcal{L}} |\{k : L \cap U_k \neq \emptyset\}| \leq C|\mathcal{L}|,$$

and the proof can proceed. Now, we claim that $C \leq d+1$. Indeed, if L intersected at least $d+2$ of the sets U_k, then it would intersect $Z(f)$ in at least $d+1$ different places. Restricting f to the line L would, hence, produce a one-variable polynomial of degree $\leq d$, which has $d+1$ zeroes. This would force the said restriction to be the zero polynomial, and consequently $L \subset Z(f)$ – a clear contradiction!

By the estimates above,

$$|I(P, \mathcal{L})| \lesssim \frac{|P|}{d}(d|\mathcal{L}|)^{1/2} + d|\mathcal{L}| + |I(P \cap Z, \mathcal{L})| = \frac{|P||\mathcal{L}|^{1/2}}{d^{1/2}} + d|\mathcal{L}| + |I(P \cap Z, \mathcal{L})|.$$

It remains to estimate $|I(P \cap Z, \mathcal{L})|$ and optimise the degree d. Towards the former task,

$$|I(P \cap Z, \mathcal{L})| = |I(P \cap Z, \mathcal{L}_1)| + |I(P \cap Z, \mathcal{L}_2)|,$$

where $\mathcal{L}_1 = \{L \in \mathcal{L} : L \subset Z\}$ and $\mathcal{L}_2 = \mathcal{L} \setminus \mathcal{L}_1$. By the argument above, every line in $\mathcal{L}_2$ can only meet Z – hence $P \cap Z$ – in at most $d+1$ different places. Consequently, $|I(P \cap Z, \mathcal{L}_2)| \leq (d+1)|\mathcal{L}_2| \leq (d+1)|\mathcal{L}|$.

To estimate $|I(P \cap Z, \mathcal{L}_1)|$, we use induction: we may assume that Theorem 6.7 holds for all sets of lines with cardinality strictly smaller than $|\mathcal{L}|$. This is useful, because $|\mathcal{L}_1| \leq d$. Indeed, a generic line in $\mathbb{R}^2$ hits all the lines in $\mathcal{L}_1$ (which are contained in Z) so $|\mathcal{L}_1| > d$ would imply that a generic line in $\mathbb{R}^2$ hits Z in more than d places. This would force the generic line in $\mathbb{R}^2$ to be contained in Z, and consequently $f \equiv 0$.

So, if we are able to choose $d < |\mathcal{L}|$, we have $|\mathcal{L}_1| < |\mathcal{L}|$, and so

$$|I(P \cap Z, \mathcal{L}_1)| \lesssim |P|^{2/3}|\mathcal{L}_1|^{2/3} + |P| + |\mathcal{L}_1|.$$

Finally, let $d = |P|^{2/3}/|\mathcal{L}|^{1/3}$ (or the closest integer). Then $1 < d < |\mathcal{L}|/2$ by our starting assumption (6.9), and hence

$$|I(P, \mathcal{L})| \lesssim \frac{|P||\mathcal{L}|^{1/2}}{|P|^{1/3}/|\mathcal{L}|^{1/6}} + \frac{|P|^{2/3}}{|\mathcal{L}|^{1/3}}|\mathcal{L}| + |I(P \cap Z, \mathcal{L}_1)| \lesssim |P|^{2/3}|\mathcal{L}|^{2/3} + |P| + |\mathcal{L}|.$$

The proof is complete. □

Exercise 6.10. Demonstrate that the Szemerédi-Trotter theorem is sharp. Hint: take P to be a grid.

We close this section with a useful – but almost immediate – corollary of the Szemerédi-Trotter bound:

Corollary 6.11 (Corollary to Theorem 6.7)**.** *Let $k > 1$, let $P \subset \mathbb{R}^2$ be a finite set, and let $\mathcal{L}$ be a finite set of lines in $\mathbb{R}^2$ such that $|L \cap P| \geq k$ for all $L \in \mathcal{L}$. Then,*

$$|\mathcal{L}| \lesssim \frac{|P|^2}{k^3} + \frac{|P|}{k}.$$

Proof. Since $k > 1$, we have the trivial bound $|\mathcal{L}_k| \leq |P|^2$, and this already implies the corollary for $k \leq k_0$, say (if the implicit constant is chosen large enough, depending on k_0). So, we may assume that $k \geq k_0$ for some absolute $k_0 \in \mathbb{N}$ to be picked momentarily.

Clearly $|I(P, \mathcal{L})| \geq k|\mathcal{L}|$, and so

$$|\mathcal{L}| \leq \frac{|I(P,\mathcal{L})|}{k} \leq C\frac{|P|^{2/3}|\mathcal{L}|^{2/3} + |P| + |\mathcal{L}|}{k},$$

where $C \geq 1$ is the absolute constant from the Szemerédi-Trotter bound. Assuming that $k \geq k_0 \geq 2C$, we see that the third term on the right hand side of the inequality above is $C|\mathcal{L}|/k \leq |\mathcal{L}|/2$. So, we may flip it to the left hand side, arriving at $|\mathcal{L}| \leq 2C(|P|^{2/3}|\mathcal{L}|^{2/3} + |P|)/k$. Next, if $|P|^{2/3}|\mathcal{L}|^{2/3} \leq |P|$, we have $|\mathcal{L}| \leq 4C|P|/k$, which is good enough. In the opposite case, we have $|\mathcal{L}| \leq 4C|P|^{2/3}|\mathcal{L}|^{2/3}/k$, which also yields the required bound by rearranging terms. □

6.2.1. *An application of Szemerédi-Trotter: Beck's theorem.* The Szemerédi-Trotter bound is a fundamental tool in (planar) incidence geometry and as such has many applications; for instance, one can prove the sum-product bound $\max\{|A + A|, |A \cdot A|\} \gtrsim |A|^{5/4}$ by applying Szemerédi-Trotter to $P = (A \cdot A) \times (A + A)$ and the $|A|^2$ lines $L_{a,b} := \{(x, y) \in \mathbb{R}^2 : y = a^{-1}x + b\}$, where $a \in A$ and $b \in A$. Then each such line has $\geq |A|$ indicences with P, since

$$\{(aa', a' + b) : a' \in A\} \subset P \cap L_{a,b},$$

and then one just plugs all the information into Corollary 6.11. Of course, this bound is weaker than Solymosi's elementary bound in Theorem 5.3.

We next demonstrate another application of Szemerédi-Trotter, a result of Beck [1] from 1983:

Theorem 6.12. *Let $P \subset \mathbb{R}^2$ be a finite set, and let $\mathcal{L}$ be the set of all lines containing at least 2 points of P (the lines "spanned" by P). Then, at least one of the following two alternatives hold:*

(i) $|\mathcal{L}| \gtrsim |P|^2$.
(ii) *There exists $L \in \mathcal{L}$ with $|L \cap P| \gtrsim |P|$.*

Proof. For $k \geq 1$, let $\mathcal{L}_k := \{L \in \mathcal{L} : 2^k \leq |L \cap P| < 2^{k+1}\}$, so by Corollary 6.11,

$$|\mathcal{L}_k| \leq C\frac{|P|^2}{2^{3k}} + \frac{|P|}{2^k} \tag{6.13}$$

for some absolute constant $C \geq 1$. Moreover, $\mathcal{L}$ is contained in the union of the set $\mathcal{L}_k$, $k > 1$.

Assume that (ii) fails, so $\mathcal{L}_k = \emptyset$ for $2^k > c|P|$, where c is some small constant. Then, each pair of distinct points $(p,q) \in P \times P$ lies on some line $L \in \mathcal{L}_k$ with $2^k \leq c|P|$. To write this down in symbols, let

$$(P \times P)_k := \{(p,q) \in P \times P : p \neq q \text{ and } p, q \in L \text{ for some line } L \in \mathcal{L}_k\}.$$

Thus each pair $(p,q) \in P \times P$ with $p \neq q$ belongs to $(P \times P)_k$ for some $2^k \leq c|P|$, and in particular

$$\left| \bigcup_{2 \leq 2^k \leq c|P|} (P \times P)_k \right| \geq \frac{|P|^2}{2}. \tag{6.14}$$

On the other hand, by (6.13), we have

$$|(P \times P)_k| \leq 2^{2k} |\mathcal{L}_k| \leq C \left[\frac{|P|^2}{2^k} + 2^k |P| \right],$$

whence

$$\left| \bigcup_{1/c \leq 2^k \leq c|P|} (P \times P)_k \right| \leq \frac{|P|^2}{10},$$

if c is small enough. Combining this with (6.14) gives

$$\frac{|P|^2}{5} \leq \left| \bigcup_{2 \leq 2^k \leq 1/c} (P \times P)_k \right| \leq \left(\frac{1}{c} \right)^2 \left| \bigcup_{2 \leq 2^k \leq 1/c} \mathcal{L}_k \right| \leq \left(\frac{1}{c} \right)^2 |\mathcal{L}|.$$

This gives (i) and completes the proof. □

6.3. **Finite field Kakeya and the Joints problem.** In this subsection, we give two further applications of the polynomial method. The main tool will be the "limiting case" of the polynomial ham sandwich theorem, Lemma 6.3, where the sets $P_1, \ldots, P_N$ are singletons. However, we now need this result in the generality of arbitrary fields (instead of just $\mathbb{R}$), and that stops us from deriving it as an immediate corollary of previous results:

Lemma 6.15. *Let $P \subset \mathbb{F}^n$ be a finite set, where $\mathbb{F}$ is a field, and let $V(n,d)$ be the vector space of $\mathbb{F}$-valued polynomials in $\mathbb{F}^n$ of degree at most d. If $|P| < \dim V(n,d) = \binom{d+n}{n}$, there exists non-zero polynomial $f \in V(n,d)$ such that $P \subset Z(f)$.*

Proof. Let V_P be the vector space of all functions $V \to \mathbb{F}$. Then $\dim V_P = |P| < \dim V(n,d)$, so the restriction mapping $f \mapsto f|_P$ from $V(n,d)$ to V_P is not injective. Hence, $f_1|_P = f_2|_P$ for two distinct $f_1, f_2 \in V(n,d)$, and then $f = f_1 - f_2$ is the polynomial we were after. □

6.3.1. *The Kakeya problem.* Suppose that a Borel set $E \subset \mathbb{R}^n$ contains a unit line segment with all possible orientations. What is the best lower bound for $\dim E$? The claim that the answer should be "n" is known as the *Kakeya conjecture* – one of the most famous open problems in both geometric measure theory and Euclidean harmonic analysis. Despite considerable effort, this is only verified in $\mathbb{R}^2$ (by A. Córdoba in 1977), whereas in higher dimensions only partial results are available. I will not give a full bibliography here, but in $\mathbb{R}^3$, for instance, the world record is $\dim E \geq 2.5$, due to T. Wolff [33] from 1995. Later, in 2000, N. Katz, I. Laba and T. Tao [20] obtained $\overline{\dim}_B E \geq 2.5 + \epsilon$, where $\overline{\dim}_B$ is the *upper box-dimension*, a quantity larger than Hausdorff dimension.

It was T. Wolff's idea to study the "finite field Kakeya-problem" as a toy model for the actual problem:

Conjecture 6.16 (Finite field Kakeya conjecture)**.** *Let $\mathbb{F} = \mathbb{Z}_q$, where q is prime. A set $K \subset \mathbb{F}^n$ is called a* Kakeya set*, if K contains a line of the form $\{x + tv : t \in \mathbb{F}\}$ for all $v \in \mathbb{F}^n \setminus \{0\}$. The finite field Kakeya conjecture states that $\operatorname{card} K \gtrsim_n q^n$ for all Kakeya sets $K \subset \mathbb{Z}_q$.*

This problem seemed just as intractable as the real Kakeya problem, until Z. Dvir [9] solved it using polynomials in 2005 (this event probably marks the birth of the "polynomial method"):

Theorem 6.17 (Dvir)**.** *Conjecture 6.16 is true. In fact, $|K| \geq (q-1)^n/n!$.*

Proof. Assume that $|K| < (q-1)^n/n! \leq \binom{(q-1)+n}{n}$. Hence, Lemma 6.15 states that we may find a non-zero polynomial $f \in V(n, q-1)$ such that $K \subset Z(f)$. Let $d \leq q-1$ be the degree of f,[8] and write

$$f = f_d + g = \sum_{\alpha_1+\ldots+\alpha_n=d} c_{(\alpha_1,\ldots,\alpha_n)} x_1^{\alpha_1} \cdots x_n^{\alpha_n} + g,$$

so that f_d consists of the monomials of degree d and g contains the lower-order terms. Note that f_d is not identically zero.

Now, consider the restriction of f to a line of the form $\{x + tv : t \in \mathbb{F}\}$:

$$\begin{aligned} f(x+tv) &= \sum_{\alpha_1+\ldots+\alpha_n=d} c_{(\alpha_1,\ldots,\alpha_n)} (x_1+tv_1)^{\alpha_1} \cdots (x_n+tv_n)^{\alpha_n} + g(x+tv) \\ &=: \left(\sum_{\alpha_1+\ldots+\alpha_n=d} c_{(\alpha_1,\ldots,a_n)} v_1^{\alpha_1} \cdots v_n^{\alpha_n} \right) t^d + h_{x,v}(t) \\ &= f_d(v) t^d + h_{x,v}(t). \end{aligned}$$

Here $h_{x,v}$ is a polynomial of degree at most $d-1$ in the variable t. Since $t \mapsto f(x+tv)$ is a one-variable polynomial of degree $d < q$, and it vanishes identically (since $\{x + tv : t \in \mathbb{F}\} \subset K$), we infer that $t \mapsto f(x+tv)$ is the zero polynomial. In particular, the coefficient of the order d term is zero, and this coefficient happens to be $f_d(v)$. So, we have proven that $f_d(v) = 0$ for all $v \in \mathbb{F} \setminus \{0\}$. It is also clear that $f_d(0) = 0$, so $f_d \equiv 0$. So, the degree of f is, in fact, strictly lower than d, and this is a contradiction. The proof is complete. □

6.3.2. *The Joints problem.* As a further demonstration of the usefulness Lemma 6.15, we discuss the *Joints problem*: assume that $\mathcal{L}$ is a collection of lines in $\mathbb{R}^3$, and let P be the set of "joints", that is, points, where three non-coplanar lines in $\mathcal{L}$ meet (thus the said three lines are not permitted to share a common plane). How large can P be in terms of $|\mathcal{L}|$? I am not sure of the motivation of this problem, but it was raised in 1990 and remained open for 20 year, attracting several people to prove partial results.

To formulate a reasonable conjecture, consider a collection of N planes in $\mathbb{R}^3$ in generic position. Then, every intersection of two planes determines a line, and every intersection

[8] By this, we mean the lowest possible degree of any representation of f. In the finite field setting, one has to be a bit careful here, because for instance $x \mapsto x^q$ coincides with $x \mapsto x$ in $\mathbb{Z}_q$ by Fermat's little theorem. The following familiar fact is still true in this setting: if $f \in V(1,d)$ has strictly more than d zeroes, then all of the coefficients of f vanish.

of three planes determines a joint. Thus, a collection of $\sim N^2$ lines can determine $\sim N^3$ joints, and hence a reasonable conjecture could be $|P| \lesssim |\mathcal{L}|^{3/2}$. This was proven by L. Guth and N. Katz [19] in 2010:

Theorem 6.18. *A set of N lines in $\mathbb{R}^3$ can determine at most $32N^{3/2}$ joints.*

Proof. It suffices to prove the following claim: If P is the collection of joints, then there exists a line $L \in \mathcal{L}$ with $|L \cap P| \leq 10|P|^{1/3}$. Indeed, if this is the case, then we may bound the number of joints iteratively: first, choose a line $L_1 \in \mathcal{L}$ with $|L_1 \cap P| \leq 10|P|^{1/3}$. The family of lines $\mathcal{L} \setminus \{L_1\}$ again generates some joints, say $P_1 \subset P$, and we note that $P = P_1 \cup [L_1 \cap P]$. Again, we may find a line $L_2 \in \mathcal{L} \setminus \{L_1\}$ with $|L_2 \cap P_1| \leq 10|P_1|^{1/3}$. Then, $\mathcal{L} \setminus \{L_1, L_2\}$ generates some joints $P_2 \subset P$, and $P = P_2 \cup [L_2 \cap P_1] \cup [L_1 \cap P]$. Iterating, we eventually end up with $\mathcal{L} \setminus \{L_1, \ldots, L_{N-2}\}$ containing just two lines, and hence generating no joints. At this stage, we estimate

$$|P| \leq |L_{N-2} \cap P_{N-3}| + |L_{N-3} \cap P_{N-4}| + \ldots + |L_1 \cap P| \leq 10N|P|^{1/3},$$

and this implies the theorem, since $10^{3/2} \leq 32$.

To prove the claim, we make a counter-assumption: every line in $\mathcal{L}$ contains strictly more than $10|P|^{1/3}$ joints in P. Choose a non-zero polynomial $f \in V(3, d)$ such that $d \leq 10|P|^{1/3}$ and $P \subset Z(f)$ (recall that there always exists such a polynomial with $\deg(f) \leq C|P|^{1/3}$, and now simply the best constant C is smaller than $2(3!)^{1/3} < 10$ by Remark 6.4). Assume that f has the least degree among all such non-zero polynomials. Now, the assumption that $|L \cap P| > 10|P|^{1/3}$ for each $L \in \mathcal{L}$ forces f to vanish identically on each line $L \in \mathcal{L}$. In particular, for each $p \in P$, f vanishes on three non-coplanar lines meeting at p. This forces $\nabla f(p) = 0$, and hence $\partial_i f$ is a polynomial vanishing on P for every $i \in \{1, 2, 3\}$. Since $\partial_i f$ has lower degree than f, we conclude that $\partial_i f \equiv 0$ for all $i \in \{1, 2, 3\}$, and hence $f \equiv$ constant. This is absurd, and the proof is complete. □

6.4. **A generalised Loomis-Whitney inequality.** A classical inequality of Loomis and Whitney [25] from 1949 states the Lebesgue measure of a set in $\mathbb{R}^n$ can be estimated by the Lebesgue measures of its coordinate projections. More precisely,

$$\mathcal{H}^n(E) \lesssim \prod_{j=1}^{n} \mathcal{H}^{n-1}(\pi_j(E))^{1/(n-1)},$$

where π_j is the projection onto the plane $\{x_j = 0\}$. This is not a hard theorem: the length of Loomis and Whitney's paper is two pages, introduction included.

Now, assume for a moment that each projection $\pi_j(E)$ is the union of a collection $\mathcal{B}_j$ of balls of diameter one. Then, any pre-image $\pi_j^{-1}(B)$, $B \in \mathcal{B}_j$, is a tube of width one, and

$$E \subset \bigcap_{j=1}^{n} \bigcup_{B \in \mathcal{B}_j} \pi_j^{-1}(B).$$

So, an essentially equivalent reformulation of the Loomis-Whitney inequality is the following: given n collections of tubes $\mathcal{T}_j$ of width one such that the tubes in $\mathcal{T}_j$ are parallel to the x_j-axis, then the Lebesgue measure of $\bigcap_{j=1}^{d} \bigcup_{T \in \mathcal{T}_j} T$ is bounded by $\prod_{j=1}^{n} |\mathcal{T}_j|^{1/(n-1)}$.

Things get significantly more difficult, if the tubes in $\mathcal{T}_j$ are allowed to be slightly tilted. The following theorem is due to J. Bennett, A. Cabery and T. Tao [2] and L. Guth [16]:

Theorem 6.19. *Suppose that $\mathcal{T}_j$, $1 \leq j \leq n$, are collections of tubes of width one, such that each tube in $\mathcal{T}_j$ makes an angle of $< (100n)^{-1}$ with the x_j-axis. Assume that $|\mathcal{T}_j| \leq A$ for all $1 \leq j \leq n$. Then the set*

$$E := \bigcap_{j=1}^{n} \bigcup_{T \in \mathcal{T}_j} T$$

satisfies $\mathcal{H}^n(E) \lesssim_n A^{n/(n-1)}$.

The original proof of Bennett-Carbery-Tao was hard, using something called "heat flow monotonicity", and gave a slightly weaker result. Subsequently, a much simpler proof – and exactly the statement above – using the polynomial method was discovered by L. Guth, and this is the argument we present below. Unlike in some previous applications, the polynomial method is not indispensable here, and may not even give the shortest (self-contained) argument: recently, L. Guth [15] has found yet another proof based on an elementary "induction of scales" technique, which is hardly more difficult than the next argument.

Before starting the proof, we introduce the concept of *directed volume*. Let $S \subset \mathbb{R}^n$ be an $(n-1)$-dimensional smooth hypersurface (such as a zero-set of a polynomial). Given a unit vector $v \in S^{n-1}$, we write

$$V_S(v) := \int_S |v \cdot N(x)| \, d\mathcal{H}^{n-1}(x)$$

for the directed volume of S in direction v. Here $N(x)$ is the unit normal of S at x. Larry Guth claims – and I could not find a justification for this anywhere, unfortunately – that

$$V_S(v) = \int_{v^\perp} |S \cap \pi_{v^\perp}^{-1}(y)| \, d\mathcal{H}^{n-1}(y),$$

where $|\cdot|$ stands for cardinality as usual, and $\pi_{v^\perp}$ is the projection onto $v^\perp$. Using these two formulae for directed volume, we prove two simple lemmas:

Lemma 6.20. *Let $T \subset \mathbb{R}^n$ be a tube or width r, and let v be a unit vector parallel to a line contained in T. Then, if f is any polynomial of degree d, the v-directed volume of $Z(f)$ inside T is bounded as follows:*

$$V_{Z(f) \cap T}(v) \lesssim r^{n-1} d.$$

Proof. Let $B(x, r/2) := \pi_{v^\perp}(T) \subset v^\perp$. Then, using the second formula for directed volume,

$$V_{Z(f) \cap T}(v) = \int_{B(x,r/2)} |Z(f) \cap \pi_{v^\perp}^{-1}(y)| \, d\mathcal{H}^{n-1}(y)$$

If $|Z(f) \cap \pi_{v^\perp}^{-1}(y)| > d$, then $\pi_{v^\perp}^{-1}(y) \subset Z(f)$ by the already familiar "restrict f to a line" argument. Since $\mathcal{H}^n(Z(f)) = 0$, this can happen only for a $\mathcal{H}^{n-1}$-null set of lines $\pi_{v^\perp}^{-1}(y)$, and the for the rest of the lines the integrand above is bounded by d. Since $\mathcal{H}^{n-1}(B(x, r/2)) \sim r^{n-1}$, the desired estimate follows. □

Lemma 6.21. *Let $v_1, \ldots, v_n \in S^1$, and assume that the angle between v_j and the j-axis is at most $(100n)^{-1}$. Then*

$$\mathcal{H}^{n-1}(S) \lesssim \sum_{j=1}^{n} V_S(v_j).$$

Proof. Let $N(x) \in S^1$ be a unit normal of S at x. Using the defining formula for directed volume, we obtain

$$\mathcal{H}^{n-1}(S) = \int_S |N(x)| \, d\mathcal{H}^{n-1}(x) \lesssim \int_S \sum_{j=1}^n |N(x) \cdot v_j| \, d\mathcal{H}^{n-1}(x) = \sum_{j=1}^n V_S(v_j)$$

as desired. □

Proof of Theorem 6.19. Let $Q_1, \ldots, Q_V$ be the dyadic cubes of side-length one, which intersect E. It suffices to prove that $V \lesssim_n A^{n/(n-1)}$.

To this end, we use the polynomial ham sandwich theorem, Theorem 6.1, to find a polynomial f of degree $d \lesssim V^{1/n}$, which bisects all the interiors of the cubes Q_j. In order for the polynomial f to bisect Q_j, the zero set $Z(f)$ must have significant surface are inside Q_j, namely

$$\mathcal{H}^{n-1}(Z(f) \cap Q_j) \gtrsim 1.$$

Fixing one of the cubes Q_j, let $T_i(Q_j) \in \mathcal{T}_i$ be a tube intersecting Q_i (by definition of E, there is at least one such a tube for every $1 \leq i \leq n$). If $v(T_i(Q_j))$ is a unit vector parallel to $T_i(Q_j)$, Lemma 6.21 gives

$$\sum_{i=1}^n \sum_{j=1}^V V_{Z(f) \cap Q_j}(v(T_i(Q_j))) = \sum_{j=1}^V \sum_{i=1}^n V_{Z(f) \cap Q_j}(v(T_i(Q_j))) \gtrsim \sum_{j=1}^V \mathcal{H}^{n-1}(Z(f) \cap Q_j) \gtrsim V.$$

Thus, for some $i \in \{1, \ldots, n\}$, we have

$$\sum_{T \in \mathcal{T}_i} \sum_{\substack{1 \leq j \leq V \\ T_i(Q_j) = T}} V_{Z(f) \cap Q_j}(v(T)) = \sum_{j=1}^V V_{Z(f) \cap Q_j}(v(T_i(Q_j))) \gtrsim_n V,$$

and since $|T_i| \leq A$, there exists a tube $T \in \mathcal{T}_i$ such that

$$\sum_{\substack{1 \leq j \leq V \\ T_i(Q_j) = T}} V_{Z(f) \cap Q_j}(v(T)) \gtrsim_n \frac{V}{A}. \tag{6.22}$$

Unwrapping the notation, this means that there are many cubes Q_j intersecting T such that $V_{Z(f) \cap Q_j}(v(T))$ is large. All of the cubes Q_j intersecting T are disjoint and contained in $2T$, so the right hand side of (6.22) is a lower bound for $V_{Z(f) \cap 2T}(v(T))$. This can be compared with the upper bound given by Lemma 6.20:

$$\frac{V}{A} \lesssim_n V_{Z(f) \cap 2T}(v(T)) \lesssim d \lesssim V^{1/n}.$$

Hence $V \lesssim_n A^{n/(n-1)}$, as claimed. □

7. Possible topics for presentations

Here are some topics for your consideration. The presentations should be preferably written down in LaTeX, and the length could be about ten pages. So, if some of the topics below require way more than ten pages, you can skip some of the intermediate steps in the proof.

- Applications of sum-product theory to exponential sums. For instance, you could prove the following theorem of Bourgain: *If $H \subset \mathbb{Z}_p \setminus \{0\}$ is a multiplicative subgroup, and $|H| > p^\epsilon$ for some $\epsilon > 0$, then*

$$\max_{\gcd(\xi,p)=1} |\widehat{H}(\xi)| = \max_{\gcd(\xi,p)=1} \left| \sum_{x \in H} e^{2\pi i \xi x/p} \right| \lesssim |H|^{1-\delta},$$

  where δ only depends on ϵ. The proof can be found in Bourgain's lecture notes *Sum-product theorems and applications*, which you can access for free online. Bourgain's notes may be hard to read, so you may wish to look elsewhere, too. You can skip some of the preliminary steps in the proof, like the proof of the particular version of the Balog-Szemerédi-Gowers lemma etc.
- The following theorem of Guth and Katz (answering a question of Bourgain): *Let $\mathcal{L}$ be a set of N^2 lines in $\mathbb{R}^3$, and let $P \subset \mathbb{R}^3$ be a finite set. Suppose that no more than N lines on $\mathcal{L}$ lie on a common plane, and each line in $\mathcal{L}$ contains at least N points of P. Then $|P| \gtrsim N^3$.* The proof is the main result of [19].
- Another incidence theorem of Guth and Katz in $\mathbb{R}^3$: *Let $\mathcal{L}$ be a set of N^2 lines in $\mathbb{R}^3$ such that no more than N lie in a common plane, and no more than $\lesssim N$ lie in a common regulus. Then, the number of intersection points of at least k lines in $\mathcal{L}$ is $\lesssim N^3/k^2$.* This is one of the main results in [18], see [18, Theorem 2.10] in particular, and is a key ingredient in solving the *Erdős distinct distance problem* in the plane (the problem and the connection is explained very clearly in [18]). In fact, if you don't care to find out what a regulus is, you could also prove slightly weaker theorem, which has an easier – and probably clearer – proof, check out Guth's recent paper [17].
- Explicit bounds in finite-field sum-product theorems. What is the state of the art? Some results were mentioned in Remark 5.22 above.
- Edgar and Miller's paper [10] on Borel subrings of the reals.
- Find your own topic! If you're not tempted by the proposals above, cook up something else and discuss with me.

APPENDIX A. GUTH'S "INDUCTION ON SCALES" PROOF OF THE NON-ENDPOINT MULTILINEAR KAKEYA INEQUALITY, BY LAURA VENIERI

A.1. **Introduction.** The multilinear Kakeya inequality yields an estimate of the measure of the intersection of cylindrical tubes in $\mathbb{R}^n$ pointing in different directions.

This inequality is a multilinear version of the so called linear Kakeya conjecture, which is a quantitative version of the Kakeya conjecture. The latter states that every Borel set in $\mathbb{R}^n$ containing a segment of unit length in every direction must have Hausdorff dimension n. There exist such sets of Lebesgue measure zero. They are known as Kakeya (or Besicovitch) sets because of a question that Kakeya asked in 1917: what is the smallest area in which a unit line segment can be rotated 180 degrees in the plane? Besicovitch proved that this can be done in arbitrarily small area.

The Kakeya conjecture was proved only in the plane by Davies but it is still an open problem for $n \geq 3$.

The linear Kakeya conjecture (which is stronger than the Kakeya conjecture) states the following. Given $0 < \delta < 1$, we define a δ-tube as a rectangular box in $\mathbb{R}^n$ with one side of length 1 and the others of length δ. Let $T_1, \ldots, T_N$ be a collection of δ-tubes whose

sides of length 1 have directions that form a δ-separated set of points in the unit sphere. Then the conjecture states that

$$\left\| \sum_{j=1}^{N} \chi_{T_j} \right\|_{L^{n/(n-1)}(\mathbb{R}^n)} \leq C_{n,\epsilon} \delta^{-\epsilon} (\delta^{n-1} N)^{(n-1)/n},$$

for every $\epsilon > 0$, where χ_{T_j} denotes the characteristic function of T_j and $C_{n,\epsilon}$ is a constant depending only on n and ϵ.

Also this conjecture is still open for $n \geq 3$. Chapters 11, 22 and 23 in [26] contain a discussion of the known partial results and connections to other problems.

In the multilinear Kakeya inequality we consider n families of tubes of width 1 such that tubes in the jth family have directions close to the x_j-axis. More precisely, let $l_{j,a}$, $j = 1, \ldots, n$, $a = 1, \ldots, N_j$, be lines in $\mathbb{R}^n$ and let $T_{j,a}$ be their 1 neighbourhoods.

Theorem A.1. *(Multilinear Kakeya) Suppose that each line $l_{j,a}$, $a = 1, \ldots, N_j$, makes an angle $\leq (10n)^{-1}$ with the x_j-axis. Then for any $\epsilon > 0$ and $S \geq 1$, we have*

$$\int_{Q_S} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}} \right)^{1/(n-1)} \leq C_\epsilon S^\epsilon \prod_{j=1}^{n} N_j^{1/(n-1)}, \tag{A.2}$$

where Q_S denotes a cube of side length S (and edges parallel to the coordinate axes).

In this case lines within the same family can be parallel, but lines in different families have almost orthogonal directions.

The multilinear Kakeya inequality was conjectured by Bennett, Carbery and Tao (2006) in [2], where they proved it except for the endpoint case. In particular, they proved the following, where the tubes $T_{j,a}$ are δ-tubes whose directions belong to some sufficiently small fixed neighbourhood of e_j, the jth vector of the standard basis of $\mathbb{R}^n$.

Theorem A.3. *(Near-optimal multilinear Kakeya) If $\frac{n}{n-1} < q \leq \infty$, then there exists a constant $C = C_{q,n}$ such that*

$$\left\| \prod_{j=1}^{n} \sum_{a=1}^{N_j} \chi_{T_{j,a}} \right\|_{L^{q/n}(\mathbb{R}^n)} \leq C \prod_{j=1}^{n} (\delta^{n/q} N_j). \tag{A.4}$$

They formulated the inequality using δ-tubes but it is equivalent, just by scaling, to formulate it for tubes of width 1 and arbitrary (possibly infinite) length. The proof was hard and used something called heat flow and its monotonicity properties.

The endpoint case of the conjecture was proved by Guth (2010) in [16]. His proof (seen partially in the lectures) relies on the polynomial method.

Theorem A.1 is a slightly weaker version: in the strongest form the factor S^ϵ does not appear. It was proved by Guth in [15] (2015). Here we will show that proof, which is quite short and relies on multiscale analysis.

The idea of the proof is first to reduce to the case when the angle $(10n)^{-1}$ is replaced by a very small angle δ, then use multiscale analysis, working at a sequence of scales δ^{-k}, $k = 1, 2, \ldots$ up to an arbitrary scale S. To get from one scale to the next one we use the Loomis-Whitney inequality ([25]), which gives (A.2) in the case when the lines $l_{j,a}$ are parallel to the x_j-axis.

In Section 2 we recall the Loomis-Whitney inequality (1949) and its proof then in Section 3 we prove Theorem A.1. In Section 4 we mention some applications of the multilinear Kakeya inequality.

A.2. **The Loomis-Whitney inequality.** The Loomis-Whitney inequality bounds the volume of a set in terms of the measures of its projections onto the coordinate hyperplanes.

Let π_j be the orthogonal projection onto the hyperplane $\{x_j = 0\}$.

Theorem A.5. *Let $U \subset \mathbb{R}^n$ be open. Then*

$$\mathcal{L}^n(U) \leq \prod_{j=1}^{n} \mathcal{L}^{n-1}(\pi_j(U))^{1/(n-1)}. \tag{A.6}$$

We use $\mathcal{L}^n$ to denote the Lebesgue measure in $\mathbb{R}^n$.

We will show the proof of Theorem A.5 given in [25]. Before that, we will state a more general form of the Loomis-Whitney inequality, which gives the multilinear Kakeya inequality in the case when the lines $l_{j,a}$ are parallel to the x_j- axis.

In the following $P_j : \mathbb{R}^n \to \mathbb{R}^{n-1}$ denotes the map that forgets the j-th coordinate,

$$P_j(x_1, \ldots, x_n) = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n).$$

Theorem A.7. *Let $f_j : \mathbb{R}^{n-1} \to \mathbb{R}$ be measurable functions. Then*

$$\int_{\mathbb{R}^n} \prod_{j=1}^{n} f_j(P_j(x))^{1/(n-1)} \leq \prod_{j=1}^{n} \|f_j\|_{L^1(\mathbb{R}^{n-1})}^{1/(n-1)}. \tag{A.8}$$

The proof can be found for example in [12] (Theorem 5.7.1).

If the line $l_{j,a}$ is parallel to the x_j-axis, then it can be written as $P_j(x) = y_a$ for some $y_a \in \mathbb{R}^{n-1}$. Thus we have $\sum_{a=1}^{N_j} \chi_{T_{j,a}}(x) = \sum_{a=1}^{N_j} \chi_{B(y_a,1)}(P_j(x))$, where $B(y_a, 1)$ denotes the ball with center y_a and radius 1. Applying (A.8) with $f_j = \sum_{a=1}^{N_j} \chi_{B(y_a,1)}$, we have

$$\int_{\mathbb{R}^n} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}}(x) \right)^{1/(n-1)} \leq \prod_{j=1}^{n} \left\| \sum_{a=1}^{N_j} \chi_{B(y_a,1)} \right\|_{L^1(\mathbb{R}^{n-1})}^{1/(n-1)} = C_n \prod_{j=1}^{n} N_j^{1/(n-1)},$$

which is the multilinear Kakeya inequality.

Let us now prove Theorem A.5. We will reduce the proof of to the following combinatorial lemma.

Lemma A.9. *Let S be a collection of N pairwise non overlapping cubes in $\mathbb{R}^n$ with fixed side length and edges parallel to the coordinate axes. For $j = 1, \ldots, n$ let S_j be the collection of $(n-1)$-dimensional cubes that are projections onto $\{x_j = 0\}$ of cubes in S. Let N_j be the number of cubes in S_j. Then*

$$N \leq \prod_{j=1}^{n} N_j^{1/(n-1)}. \tag{A.10}$$

Proof. The proof proceeds by induction on n. For $n = 2$ the claim holds trivially. Assume that (A.10) holds for $n - 1$. Projecting the cubes in S onto the x_1 axis we obtain intervals $I_1, \ldots, I_m$. Suppose that for a fixed i there are a_i cubes projecting onto I_i. Now project

these cubes onto $\{x_j = 0\}$ for $j = 2, \ldots, n$ and let a_{ij} be the number of such projections for each j. Then we have

$$\sum_{i=1}^{m} a_i = N \tag{A.11}$$

and

$$\sum_{i=1}^{m} a_{ij} = N_j, \qquad j = 2, \ldots, n. \tag{A.12}$$

By the inductive hypothesis we have for every $i = 1, \ldots, m$,

$$a_i^{n-2} \leq \prod_{j=2}^{n} a_{ij}.$$

Since $\{x_1 = 0\}$ is orthogonal to the x_1 axis, we have $a_i \leq N_1$ for every $i = 1, \ldots, m$. Thus

$$a_i^{n-1} \leq N_1 \prod_{j=2}^{n} a_{ij}. \tag{A.13}$$

Using (A.11) and (A.13) we get

$$N = \sum_{i=1}^{m} a_i \leq \sum_{i=1}^{m} N_1^{1/(n-1)} \prod_{j=2}^{n} a_{ij}^{1/(n-1)}.$$

Applying repeatedly Hölder's inequality, we then get

$$N \leq N_1^{1/(n-1)} \prod_{j=2}^{n} (\sum_{i=1}^{m} a_{ij})^{1/(n-1)},$$

thus by (A.12)

$$N \leq \prod_{j=1}^{n} N_j^{1/(n-1)}.$$

□

To prove Theorem A.5 observe that given any $\epsilon > 0$ there exists $\delta > 0$ so small that if we take a partition of $\mathbb{R}^n$ into cubes of side length δ (and sides parallel to the coordinate axes), then $\mathcal{L}^n(U \setminus Q) < \epsilon$, where Q is the union of cubes contained in the interior of U. Let N be the number of cubes contained in Q and N_j be the number of cubes in the projection of Q onto the coordinate hyperplanes. Then by (A.10) we have

$$\mathcal{L}^n(Q)^{n-1} = (N\delta^n)^{n-1} \leq \delta^{n(n-1)} \prod_{j=1}^{n} N_j = \prod_{j=1}^{n} (N_j \delta^{n-1}) \leq \prod_{j=1}^{n} \mathcal{L}^{n-1}(\pi_j(U)).$$

Since ϵ is arbitrary, (A.6) follows.

A.3. **Proof of the multilinear Kakeya inequality.**

A.3.1. *Reduction to small angle.* The first step of the proof of Theorem A.1 consists in reducing to the case when the lines $l_{j,a}$ make a very small angle with the x_j-axis, that is reducing to prove the following.

Theorem A.14. *Let $\epsilon > 0$. Then there exists $\delta > 0$ such that the following holds. Suppose that each line $l_{j,a}$, $a = 1, \dots, N_j$ makes an angle $\leq \delta$ with the x_j-axis. Then for any $S \geq 1$, we have*

$$\int_{Q_S} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}} \right)^{1/(n-1)} \leq C_\epsilon S^\epsilon \prod_{j=1}^{n} N_j^{1/(n-1)}, \tag{A.15}$$

where Q_S denotes a cube of side length S.

Let us see why this implies Theorem A.1. Suppose that each line l_j, a makes an angle $\leq (10n)^{-1}$ with the x_j-axis, that is the direction of $l_{j,a}$ is contained in the spherical cap $B(e_j, (10n)^{-1}) \cap S^{n-1}$, where e_j is the j-th coordinate vector.

Given ϵ, Theorem A.14 gives δ. Divide the spherical cap into M smaller caps $B(u_k^j, \delta/10) \cap S^{n-1}$, $u_k^j \in S^{n-1}$, where $M \lesssim \delta^{1-n} \lesssim_\epsilon 1$. Denote by $v_{j,a}$ the unit vector giving the direction of $l_{j,a}$. Then we have

$$\begin{aligned}\int_{Q_S} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}} \right)^{1/(n-1)} &\lesssim \sum_{k=1}^{M} \int_{Q_S} \prod_{j=1}^{n} \Big(\sum_{v_{j,a} \in B(u_k^j, \delta/10)} \chi_{T_{j,a}} \Big)^{1/(n-1)} \\ &\lesssim_\epsilon \max_k \prod_{j=1}^{n} \Big(\sum_{v_{j,a} \in B(u_k^j, \delta/10)} \chi_{T_{j,a}} \Big)^{1/(n-1)}.\end{aligned}$$

If $u_k^j = e_j$ then we can apply Theorem A.14 because the lines make an angle at most δ with the x_j-axis. Otherwise we need to perform a linear change of coordinates that maps u_k^j to e_j. Since the angle between $l_{j,a}$ and e_j is at most $(10n)^{-1}$, this linear change of coordinates distorts length by a factor of at most 2 and volume by at most 2^n. Then the integral in the new coordinates can be bounded using (A.15).

A.3.2. *Multiscale analysis.* Now we will prove Theorem A.14. Instead of proving it directly at scale S, the idea is to work at scales $\delta^{-1}, \delta^{-2}, \dots$ up to any scale S. To get from one scale to the other we use the Loomis-Whitney inequality. This is done in the following lemma, in which we denote by $T_{j,a}^W$ the W neighbourhood of the line $l_{j,a}$.

Lemma A.16. *Suppose that $l_{j,a}$ are lines that make an angle at most δ with the x_j-axis. If $S \geq \delta^{-1} W$, and Q_S is any cube of side length S, then*

$$\int_{Q_S} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^W} \right)^{1/(n-1)} \leq C_n \delta^n \int_{Q_S} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^{\delta^{-1}W}} \right)^{1/(n-1)}. \tag{A.17}$$

Proof. Divide the cube Q_S into subcubes of side length between $(20n\delta)^{-1}W$ and $(10n\delta)^{-1}W$. Then it suffices to show that (A.17) holds when we integrate over any such cube Q.

Observe that the intersection of any tube $T_{j,a}^W$ and Q is contained in a $2W$ tube with direction parallel to the x_j-axis (since the side length of Q is at most $(10n\delta)^{-1}W$). More

precisely, there exists a tube $\tilde{T}_{j,a}^{2W}$ with direction parallel to the x_j-axis such that for every $x \in Q$, $\chi_{T_{j,a}^W}(x) \leq \chi_{\tilde{T}_{j,a}^{2W}}(x)$. Thus

$$\int_Q \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^W} \right)^{1/(n-1)} \leq \int_Q \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{\tilde{T}_{j,a}^{2W}} \right)^{1/(n-1)}.$$

To estimate this we can use the Loomis-Whitney inequality. We get by (A.8)

$$\int_Q \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{\tilde{T}_{j,a}^{2W}} \right)^{1/(n-1)} \leq C_n W^n \prod_{j=1}^n N_j(Q)^{1/(n-1)},$$

where $N_j(Q)$ denotes the number of tubes $T_{j,a}^W$ that intersect Q.

But if $T_{j,a}^W \cap Q \neq \emptyset$ then $T_{j,a}^{\delta^{-1}W} \cap Q = Q$ because the diameter of Q is at most $(10\delta)^{-1}W$. Hence

$$\begin{aligned} C_n W^n \prod_{j=1}^n N_j(Q)^{1/(n-1)} &\leq C_n \frac{W^n}{\mathcal{L}^n(Q)} \int_Q \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^{\delta^{-1}W}} \right)^{1/(n-1)} \\ &= C_n \delta^n \int_Q \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^{\delta^{-1}W}} \right)^{1/(n-1)}, \end{aligned}$$

which completes the proof. □

To prove Theorem A.14 we apply this lemma repeatedly.

Proof of Theorem A.14. We are given $\epsilon > 0$ and we need to choose δ such that the theorem holds. We will first work with a given δ and then specify how to choose it.

We first prove the theorem when $S = \delta^{-M}$ for an integer $M > 0$. Applying Lemma A.16 M times we get

$$\begin{aligned} \int_{Q_S} \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}} \right)^{1/(n-1)} &\leq C_n \delta^n \int_{Q_S} \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^{\delta^{-1}}} \right)^{1/(n-1)} \\ &\leq \cdots \leq C_n^M \delta^{Mn} \int_{Q_S} \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^{\delta^{-M}}} \right)^{1/(n-1)}. \end{aligned}$$

Since $\sum_{a=1}^{N_j} \chi_{T_{j,a}^{\delta^{-M}}} \leq N_j$, we get

$$\begin{aligned} C_n^M \delta^{Mn} \int_{Q_S} \prod_{j=1}^n \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}^{\delta^{-M}}} \right)^{1/(n-1)} &\leq C_n^M \delta^{Mn} \delta^{-Mn} \prod_{j=1}^n N_j^{1/(n-1)} \\ &= C_n^M \prod_{j=1}^n N_j^{1/(n-1)}. \end{aligned}$$

Since $S = \delta^{-M}$, we have $M = \frac{\log S}{\log \delta^{-1}}$, thus

$$C_n^M = S^{\frac{\log C_n}{\log \delta^{-1}}}.$$

Choosing δ so that $\frac{\log C_n}{\log \delta^{-1}} \leq \epsilon$, we obtain

$$\int_{Q_S} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}} \right)^{1/(n-1)} \leq S^{\epsilon} \prod_{j=1}^{n} N_j^{1/(n-1)},$$

which is the desired bound.

For an arbitrary $S \geq 1$, we can cover Q_S with $c_\delta = C_\epsilon$ cubes of side length δ^{-M} for a certain integer $M > 0$. Then we get

$$\int_{Q_S} \prod_{j=1}^{n} \left(\sum_{a=1}^{N_j} \chi_{T_{j,a}} \right)^{1/(n-1)} \leq C_\epsilon S^{\epsilon} \prod_{j=1}^{n} N_j^{1/(n-1)},$$

which proves Theorem A.14. □

A.4. **Some applications.** We give here some examples of applications of the multilinear Kakeya inequality.

(1) In [2] the multilinear Kakeya inequality was applied to prove a multilinear restriction estimate. The restriction problem is an important problem in harmonic analysis, asking when the restriction of the Fourier transform of a function f to the unit sphere S^{n-1} makes sense. We will not go into the details here but Bennett, Carbery and Tao considered a multilinear version of the problem, proving the equivalence of multilinear Kakeya and multiliar restriction.

(2) Another application given in [2] regards the joint problems. As we have seen in the lectures, it asks what is the maximum number of joints determined by a collection L of N lines in $\mathbb{R}^3$ in terms of N. We recall that a joint is the intersection of three lines in L which are not coplanar. The conjectured bound $N^{3/2}$ was proved by Guth and Katz in [19] in 2010. Bennett, Carbery and Tao had used the multilinear Kakeya inequality to progress towards the conjecture provided that the joints are sufficiently transverse.

For $0 < \theta \leq 1$, they say that three lines are θ-transverse if the parallelepiped generated by the unit vectors parallel to the lines has volume at least θ. Then they define a θ-transverse joint as the intersection of three θ-transverse lines and prove the following.

Theorem A.18. *For any $0 < \theta \leq 1$, the number of θ-transverse joints is*

$$\leq C_\epsilon N^{3/2+\epsilon} \theta^{-1/2-\epsilon}$$

for any $\epsilon > 0$.

Proof. We only give the proof in the case $\theta \approx 1$ to show how the multilinear Kakeya inequality is used.

Cover the unit sphere S^2 with $O(1)$ finitely overlapping spherical caps of radius $\theta/1000$. If three lines l, l', l'' are θ-transverse then their directions lie in three

distinct caps, call them $C_i, C_{i'}, C_{i''}$. Then it is enough to show that

$$\#J_{i,i',i''} = \#\{p \in \mathbb{R}^3 : p \in l, l', l'' \text{ for some } l \in L_i, l' \in L_{i'}, l'' \in L_{i''}\} \lesssim N^{3/2}$$

for each such transverse triple $C_i, C_{i'}, C_{i''}$, where L_j denotes the collection of lines in L whose directions are in C_j.

By rescaling, we can assume that the joints are all contained in the ball of radius $1/1000$ centred at the origin. For a small $\delta > 0$, let T_l denote the δ-tube with axis $l \in L$ and center the closest point of l to the origin. If $p \in J_{i,i',i''}$ then for $j = i, i', i''$

$$\sum_{l \in L_j} \chi_{T_l}(x) \geq 1$$

when $|x - p| < c\delta$, where $c > 0$ is a constant depending on the transversality constant of $(C_i, C_{i'}, C_{i''})$. Since the number of joints is finite, if δ is small enough the balls $\{x \in \mathbb{R}^3 : |x - p| < c\delta\}$ are disjoint for $p \in J_{i,i',i''}$. Thus

$$\left\| \left(\sum_{l \in L_i} \chi_{T_l} \right) \left(\sum_{l \in L_{i'}} \chi_{T_l} \right) \left(\sum_{l \in L_{i''}} \chi_{T_l} \right) \right\|_{L^{q/3}(\mathbb{R}^3)} \geq C_q (\#J_{i,i',i''})^{3/q} \delta^{9/q}$$

for any $\frac{3}{2} < q \leq \infty$. But applying Theorem A.3 the left-hand side can be estimated by

$$\leq C_q (\delta^{3/q} \#L_i)(\delta^{3/q} \#L_{i'})(\delta^{3/q} \#L_{i''}),$$

hence we get

$$\#J_{i,i',i''} \leq C_q (\#L_i \#L_{i'} \#L_{i''})^{q/3}.$$

Since $\#L_i, \#L_{i'}, \#L_{i''} \leq N$ and q can get arbitrarily close to $3/2$ we get the claim. □

(3) In [16] Guth obtains, as a corollary of his method, a 'planiness' estimate for unions of tubes in $\mathbb{R}^n$, proving the following.

Theorem A.19. *(Box estimate) There exists a constant $C(n) > 0$ such that the following holds. If $X \subset \mathbb{R}^n$ is a collection of cylinders of radius 1 and length $L >> 1$ then for every $x \in X$ there exists a rectangular box $B(x)$ with the following properties:*

(a) *$B(x)$ is centred at x, it can be oriented in any direction and $\mathcal{L}^n(B(x)) \leq C(n)\mathcal{L}^n(X)$;*
(b) *for every cylinder $T \subset X$, if $x \in T$ is a random point then $T \subset B(x)$ with probability at least $9/10$.*

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), p. 281–297

[2] J. BENNETT, A. CARBERY AND T. TAO: *On the multilinear restriction and Kakeya conjectures*, Acta Math. **196**, Issue 2 (2006), p. 261–302

[3] A. BALOG AND E. SZEMERÉDI: *A statistical theorem of set addition*, Combinatorica **14** (1994), p. 263–268

[4] J. BOURGAIN: *Mordell's exponential sum estimate revisited*, J. Amer. Math. Soc. **18** (2005), p. 477–499

[5] J. BOURGAIN AND M. GARAEV: *On a variant of sum-product estimates and explicit exponential sum bounds in prime fields*, Math. Proc. Cambridge Philos. Soc. **146** (2009), p. 1–21

[6] J. BOURGAIN, N. KATZ, T. TAO: *A sum-product estimate in finite fields, and applications*, Geom. Funct. Anal. **14**, Issue 1 (2004), p. 27–57

[7] A. CÓRDOBA: *The Kakeya maximal function and the spherical summation of multipliers*, Amer. J. Math. **99** (1977), p. 1–22

[8] D. DADUSH: *Lattices, Convexity & Algorithms*, available at http://cs.nyu.edu/courses/spring13/CSCI-GA.3033-013/
[9] Z. DVIR: *On the size of Kakeya sets in finite fields*, J. Amer. Math. Soc. **22** (2009), p. 1093–1097
[10] G. A. EDGAR AND C. MILLER: *Borel Subrings of the Reals*, Proc. Amer. Math. Soc. **131**, No. 4 (2003), p. 1121–1129
[11] M. GARAEV: *An explicit sum-product estimate in* $\mathbb{F}_p$, Int. Math. Res. Not. (2007), No. 11
[12] D. J. H. GARLING: *Inequalities: an journey into linear analysis*, Cambridge University Press, Cambridge, 2007
[13] W. T. GOWERS: *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, GAFA **8** (1998), p. 529–551
[14] B. GREEN: *Structure Theory of Set Addition*, available at http://people.maths.ox.ac.uk/greenbj/papers/icmsnotes.pdf
[15] L. GUTH: *A short proof of the multilinear Kakeya inequality*, to appear in Math. Proc. Cambridge Phil. Soc. (2015), available at arXiv:1409.4683
[16] L. GUTH: *The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture*, Acta Math. **205** (2010), p. 263–286
[17] L. GUTH: *Distinct distance estimates and low degree polynomial partitioning*, available at arXiv:1404.2321
[18] L. GUTH AND N. KATZ: *On the Erdős distinct distance problem in the plane*, Ann. Math. **181** (2015), p. 155–190
[19] L. GUTH AND N. KATZ: *Algebraic methods in discrete analogs of the Kakeya problem*, Adv. Math. **225** (2010), p. 2828-2839
[20] N. KATZ, I. LABA AND T. TAO: *An Improved bound on the Minkowski dimension of Besicovitch sets in* $\mathbb{R}^3$, Ann. of Math. **152**, No. 2 (2000), p. 383–446
[21] N. KATZ AND C.-Y. SHEN: *A slight improvement to Garaev's sum-product estimate*. Proc. Amer. Math. Soc. **136**, No. 7 (2008), p. 2499–2504
[22] N. KATZ AND C.-Y. SHEN: *Garaev's inequality in Finite Fields not of prime order*, Online J. Anal. Comb., Issue 3 (2008)
[23] S. KONYAGIN: *A sum-product estimate in fields of prime order*, available at arXiv:0304217
[24] S. KONYAGIN AND I. SKHREDOV: *On sum sets of sets, having small product set*, available at arXiv:1503.05771
[25] L. LOOMIS AND H. WHITNEY: *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc. **55** (1949), p. 961–962
[26] P. MATTILA: *Fourier analysis and Hausdorff dimension*, Cambridge University Press, Cambridge, 2015
[27] G. PETRIDIS: *Upper bounds on the Cardinality of Higher Sumsets*, Acta Arith. **158**, No. 4 (2013), p. 299-319. See also Tim Gowers' blog post at https://gowers.wordpress.com/2011/02/10/a-new-way-of-proving-sumset-estimates/
[28] M. RUDNEV: *An improved Sum-Product Inequality in Fields of Prime Order*, Int. Math. Res. Not. (2012), Issue 16, p. 3693-3705
[29] I. RUZSA: *Sumsets and structure*, available at http://www.math.cmu.edu/ af1p/Teaching/AdditiveCombinatorics/Additive-Combinatorics.pdf
[30] C.-Y. SHEN: *An extension of Bourgain and Garaev's sum-product estimates*, Acta Arith. **135** (2008), p. 351–536
[31] J. SOLYMOSI: *Bounding multiplicative energy by the sumset*, Adv. Math. **222**, Issue 2 (2009), p. 402-408
[32] T. TAO: *Some highlights of arithmetic combinatorics*, available at http://www.math.ucla.edu/ tao/254a.1.03w/
[33] T. WOLFF: *An improved bound for Kakeya type maximal functions*, Rev. Mat. Iberoam. **11** (1995), p. 651–674

UNIVERSITY OF HELSINKI, DEPT. OF MATHEMATICS AND STATISTICS
E-mail address: tuomas.orponen@helsinki.fi