ADDITIVE STRUCTURE, RICH LINES, AND EXPONENTIAL SET-EXPANSION

A Thesis
Presented to
The Academic Faculty

by

Evan Borenstein

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosopy in the
Mathematics

Georgia Institute of Technology
August 2009

ADDITIVE STRUCTURE, RICH LINES, AND EXPONENTIAL SET-EXPANSION

Approved by:

Ernest Croot, Advisor
Mathematics
Georgia Institute of Technology

Neil Lyall
Mathematics
University of Georgia

Prasad Tetali
Mathematics
Georgia Institute of Technology

Kevin Costello
Mathematics
Georgia Institute of Technology

XingXing Yu
Mathematics
Georgia Institute of Technology

Date Approved: 13 August 2009

ACKNOWLEDGEMENTS

I would like to thank my advisor, Ernie Croot, for years of inspiration and guidance. He has been a rich source of fascinating questions and ideas. I always left our conversations, mathematical and otherwise, with renewed motivation and intellectual curiosity.

I also owe a special thanks to Andrew Granville. He agreed to mentor me when I was an over-eager college freshman. And he allowed me to sit in on the dissertation defense of his doctoral student, who happened to become my advisor four years later!

I would like to acknowledge a number of people whose conversations, emails, papers, and expository notes have aided in the writing of this thesis. This list includes but is not limited to Prasad Tetali, Neil Lyall, Kevin Costello, XingXing Yu, Joszef Solymosi, Terry Tao, Alex Iosevich, Boris Bukh, P.M. Wood, Harald Helfgott, and Noah Forman.

I would also like to thank my parents, Hedy and Aaron, who have been extremely supportive of this endeavor.

And lastly, I have to thank Jeff Floyd, my high school math team coach, for getting me hooked.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

SUMMARY v

I NOTATION 1

II INTRODUCTION 5

2.1 Structural Results 5

2.2 Combinatorial Estimates 10

III REVIEW OF LITERATURE 18

IV ON A CERTAIN GENERALIZATION OF THE BALOG-SZEMERÉDI-GOWERS THEOREM 22

4.1 Proof of Theorem 5.1 22

4.1.1 Notation and basic assumptions 22

4.1.2 Lengths of iterations and the choice of δ and k 24

4.1.3 The iteration part of the argument 24

4.1.4 The sets H' and H'' 27

4.1.5 The final leg of the proof 29

V ON RICH LINES IN GRIDS 32

5.1 Proof of the main theorem 32

5.1.1 Producing new rich lines from old ones 32

5.1.2 Passing to a set of rich lines with usable properties 34

5.1.3 The sequence Θ_i 44

5.1.4 A growth lemma 48

5.1.5 Continuation of the proof 50

VI ON BIVARIATE SET FUNCTIONS AND EXPONENTIAL EXPANSION 52

6.1 Tools 53

6.2 Proof of Theorem 6.1 58

6.3 Proof of Theorem 6.2 62

SUMMARY

We will survey some of the major directions of research in arithmetic combinatorics and their connections to other fields. We will then discuss three new results. The first result will generalize a structural theorem from Balog and Szemerédi. The second result will establish a new tool in incidence geometry, which should prove useful in attacking combinatorial estimates. The third result evolved from the famous sum-product problem, by providing a partial categorization of bivariate polynomial set functions which induce exponential expansion on all finite sets $A \subset \mathbb{R}$.

CHAPTER I

NOTATION

Set Arithmetic

For sets A and B define their sumset as

$$A + B = \{a + b | a \in A, b \in B\},$$

and their difference set as

$$A - B = \{a - b | a \in A, b \in B\}.$$

Likewise, define their product set and divisor set as

$$A \cdot B = \{ab | a \in A, b \in B\},$$

and

$$A/B = \{\frac{a}{b} | a \in A, b \in B\},$$

respectively.

Define dilation

$$k \cdot A = \{ka | a \in A\},$$

which is not to be confused with iterated set addition:

$$kA = \underbrace{A + A + \ldots A}_{k \text{ times}}.$$

The following result relates estimates on sumsets and estimates on iterated-sumsets.

Lemma 1 *(Plünnecke's Inequality)[31]*

If A and B are two additive sets in an ambient abelian group and $|A+B| \leq K|B|$, then for any positive integer k, we have

$$|kA - mA| \leq K^{k+m}|B|.$$

One way to measure the additive structure *between* two sets A and B is to count the the number of **additive quadruples**

$$(a, a', b, b') \in A^2 \times B^2 : a + b = a' + b'.$$

We call the total number of additive quadruples the *additive energy* between A and B and label it $E(A, B)$. We call $E(A, A)$ the internal additive energy of A.

Discrete Fourier Transforms

Associate with a set, A, the function

$$A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & x \notin A \end{cases}$$

and define the Fourier coefficients of A with respect to a prime p as

$$\widehat{A}(\zeta) = \sum_{a \in A} e^{\frac{2\pi\imath\zeta a}{p}} \text{ for } \zeta \in \{0, 1, \ldots, p-1\}.$$

We define convolution of two functions A and B as

$$A * B(n) = \sum_{a+b=n} A(a)B(b),$$

and observe that

$$A \hat{*} B(n) = \hat{A}(n)\hat{B}(n)$$

Also useful will be Parseval's Identity, which in this context gives us

$$p|A| = \sum_{\zeta=0}^{p-1} |\hat{A}(\zeta)|^2.$$

Graph Theory

Let $G = G(V, E)$ denote an undirected graph consisting of a set of vertices V and a set of undirected edges connecting two vertices.

For two sets of vertices $X, Y \subseteq V$, define

$$E(X, Y)$$

to be the number of edges of G with one vertex in X and one vertex in Y. Next define the density between those vertex sets as

$$\Delta(X, Y) = \frac{E(X, Y)}{|X||Y|}.$$

We will call a pair of disjoint vertex sets V_1 and V_2 'ϵ-regular' if, given any vertex sets

$$X \subseteq V_1, Y \subseteq V_2,$$

satisfying

$$|X| \geq \epsilon|V_1|, |Y| \geq \epsilon|V_2|,$$

we have

$$|\Delta(X, Y) - \Delta(V_1, V_2)| \leq \epsilon.$$

This is to say that edges between V_1 and V_2 are relatively well distributed.

Geometry

Within an abelian group, we can define an arithmetic progression P of length $k \geq 3$, with common difference d as

$$P = \{a + bn | 0 \leq b \leq k - 1\}.$$

Its natural generalization will prove even more useful. So we call the set

$$Q = \{a_0 + a_1 n_1 + a_2 n_2 + \dots a_k n_k | 1 \leq n_i \leq N_i\}$$

a Generalized Arithmetic Progression of dimension k and volume $N_1 N_2 \dots N_k$. Keep in mind that the volume of a GAP can be greater than its number of elements, because some elements could have multiple representations.

Incidence geometry will have many applications. Given a set of points P and curves L in the Euclidean plane, we call the occurrence of a point on a curve an *incidence*, and label the total incidences between P and L as $I(P, L)$. For our purposes we will most often be interested in incidences between curves and grids. For sets A and B define a *grid* as the set of points

$$A \times B = \{(a, b) | a \in A, b \in B\}.$$

In a grid $A \times A$, with $|A| = n$, we will call a curve '$n^{1-\delta}$-rich' if the curve is incident to at least $n^{1-\delta}$ points in the grid.

CHAPTER II

INTRODUCTION

Arithmetic Combinatorics addresses the combinatorial properties of sets inside algebraic structures, most typically abelian groups, rings, or finite fields. We will motivate the discussion of our new results by briefly surveying two interrelated lines of research. In the process, we will illuminate how arithmetic combinatorics has both led to, and flourished under progress from seemingly disparate fields, such as number theory, harmonic analysis, discrete geometry, and graph theory.

2.1 Structural Results

Much of arithmetic combinatorics concerns the extent to which structure must occur in all sufficiently large, or sufficiently dense sets. Put another way:

Large sets of disorder must contain regions of order.

The occurrence of "structure" can take on many different forms:

- Sárközy [26] showed that if $A \subseteq [1, N]$ and

$$A|| \gg N\left(\frac{(\log\log N)^2}{\log N}\right)^{1/3},$$

  then $A - A$ **must** contain a square.

- Vu and Nguyen showed [23] that after proper dilation, any zero-sum-free subset A of $\mathbb{Z}_p$ has the form

$$A = A' \cup A'',$$

  where the elements of A' are small ($\sum_{x\in A'} |x| < p$) and $|A''| \leq p^{6/13+o(1)}$.

- Any sufficiently large subset of the vertices on an n dimensional hypercube contain two connected vertices. This is equivalent to saying that for sufficiently large subsets of ${\mathbb{F}_2}^n$, there must be two elements separated by a hamming distance of 1.

- Freiman showed [16] that if A is an additive set satisfying $|A+A| < k|A|$, then there exists r and n dependent only on k, such that A is contained in a proper GAP of rank at most r, and size at most $n|A|$. Clearly sets with small doubling are highly structured.

- In the early 1900's Schur showed that for every integer $r > 0$ and every r-coloring of the set $\mathbb{N}$ of natural numbers, there is a monochromatic triple $(x, y, z) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, where x, y, and z are distinct and $x + y = z$.

- In 1927, Van der Waerden [33] showed that every r-coloring of $\mathbb{N}$ contains a monochromatic arithmetic progression of arbitrary length.

Van der Waerden's work inspired vast efforts to quantify the critical densities at which sets of integers must contain arithmetic progressions. In fact, arithmetic progressions have remained the most commonly researched form of structure. Let us define

$$r_k(N) = \text{ the size of the largest subset of } \{1, \ldots, N\} \text{ which avoids k-AP's.}$$

It is immediately clear that $r_k(N) \geq r_{k+1}(N)$ for all k, and that $r_3(N)$ in fact counts the largest subset of $[1, N]$ with *no* arithmetic progressions. Establishing tight bounds has proven untenable and heuristic arguments provide very little guidance. Lower bounds have followed from complex constructions. Recalling the enigmatic properties of the Cantor set, it is natural to consider the set

$$S = \{n \in [1, N] \mid \text{where } n \text{ has no 1's in its ternary expansion}\}.$$

Clearly the average of any two elements in S must have a 1 in its ternary expansion, so S has no 3-AP's.This gives the the lower bound $N^{\log_3 2}$.In 1946, Behrend gave a much stronger lower bound.

Behrend's Construction [3]. Recognizing that the surface of a sphere does not contain AP's, we seek an appropriate map from $\mathbb{Z}^m$ into $\mathbb{Z}$ which preserves that property. Let n, M to be large integers to be determined later. Consider sets of the form

$$S(r) = \{x \in \{1, \dots M\}^n : x_1^2 + \dots {x_n}^2 = r^2\}.$$

As r^2 ranges over the integers from n to nM^2, these sets cover the cube $[1, M]^n$. By the pigeonhole principle, there must be one such set S with

$$|S| \geq \frac{M^n}{n(M^2 - 1)} > M^{n-2}/n.$$

Now we want to map the vectors of S into $[1, N]$. Define

$$P(x) = P(x_1, \dots x_n) = \frac{1}{2M} \sum_{k=1}^{n} x_k (2M)^k.$$

Clearly P is one-to-one. We also observe that because S has no AP's, and because there is no 'carrying' when adding two elements, $P(S)$ has no AP's. Also, $P(S) \subset [1, (2M)^n]$. So we need $(2M)^n$ to be no bigger than N. Setting $M = \lfloor N^{1/n}/2 \rfloor$ and $n = \sqrt{\log N}$ gives us the lower bound $Ne^{-C\sqrt{\log N}}$.

Erdős and Turan conjectured that $r_k(n) = o(n)$. Roth confirmed the conjecture for $k = 3$. His proof went roughly as follows.

1. Take a set A of density δ inside $[1, N]$ and assume it has no non-trivial 3-AP's. Embed it inside $\mathbb{Z}_p$, for some appropriate prime p.

2. Observe that the quantity

$$\frac{1}{p}\sum_{k\in\mathbb{Z}_p}\hat{A}^2(k)\hat{A}(-2k)$$

counts the number of 3-AP's in A. Since we took A to be AP-free, there must be at least one large Fourier coefficient aside from $\hat{A}(0)$.

3. This large Fourier coefficient permits the existence of a long AP in Z_p on which A has increased density $\delta + \delta^2$.

4. Loop back to step 2.

If $\delta = O(1)$, then the above algorithm must loop enough times to generate a progression on which A has density greater than 1, a clear contradiction. Hence $\delta = o(1)$.

The best known bounds to date,

$$\frac{N\log^{1/4}N}{2^{2\sqrt{2}}\sqrt{\log_2 N}} \leq r_3(N) \leq N(\log N)^{-2/3+o(1)},$$

are due to Elkin and Bourgain, respectively (both unpublished).

Before continuing, it is interesting to note the difference between the conditions forcing arithmetic progressions in $\mathbb{Z}$ and the conditions forcing geometric progression in $\mathbb{Z}$, best illustrated by the following point. The square free integers clearly cannot contain a 3-GP, $\{n, nr, nr^2\}$, yet they have a density of $\frac{6}{\pi^2}$.

In 1975, Szemerédi generalized Roth's result to progressions of arbitrary length.

Theorem 2.1 (Szemerédi's Theorem) *Every subset of the integers with positive upper density contains arbitrarily long arithmetic progressions.*

Perhaps even more monumental than his theorem was the lemma he proved in the process.

Lemma 2 *For every $\epsilon > 0$, and $m \geq 1$, there exist constants $K > 0$ and $M > 0$ such that the following holds: If G is a graph having vertex set V satisfying $|V| = k \geq K$, then there exists an integer m', with*

$$m \leq m' \leq M,$$

and a partition of G into vertex sets $V_0, V_1, \ldots V_{m'}$, with the following properties:

1. *$|V_0| \leq \epsilon |V|$. We call this the exceptional set.*
2. *$|V_1| = |V_2| = \ldots |V_{m'}|$.*
3. *All but at most $\epsilon (m')^2$ of the distinct pairs $(V_i, V_j), 1 \leq i, j \leq m'$ are ϵ-regular.*

Szemerédi's Theorem has since been proven by various different methods without the use of the Regularity Lemma- first by Furstenberg [17] using ergodic theory and then by Gowers [19] by using Fourier analysis and combinatorics. And yet Szemerédi's Regularity Lemma remains a monumental achievement in its own right, spawning many applications in both Ramsey theory and combinatorics. One such application of particular interest is the following widely used structural result by Balog and Szemerédi.

Theorem 2.2 (Balog-Szemerédi Theorem) *[1] Let $A \subseteq \mathbb{Z}^D$ satisfy*

$$E(A, A) \geq \frac{|A|^3}{K}.$$

Then there exist constants c and C dependent only on K such that there exists a large subset $A' \subseteq A$ such that

$$|A'| \geq c|A| \text{ and } |A' + A'| \leq C|A|.$$

As just one example of how this theorem can be used, we will sketch a proof of the following theorem which originally required its development.

Theorem 2.3 (Balog-Szemerédi) *[1]*

For any $k \in \mathbb{N}, c > 0$, there exists an $N(c,k)$ such that, if $A \subset \mathbb{Z}$ has size

$$|A| \geq N(c,k),$$

and contains $c|A|^2$ 3-term arithmetic progressions, then A must contain a k-term arithmetic progression.

Sketch of Proof Basic Fourier analysis, similar to that used by Roth, shows that A's high number of 3-AP's forces A to have high additive energy. The Balog-Szemerédi theorem thus shows that A has a positive density subset A' with small doubling constant. From there, Freiman's theorem tells us that A' is a positive density subset of a GAP with low rank. It follows that we must have a long progression P with positive density intersection with A. So by Szemerédi's theorem, A must have a long arithmetic progression with length dependant on c and k.

In chapter 3, we will discuss generalizations of the Balog-Szemerédi theorem, one of which we will make use of in chapter 6. In chapter 4, we will establish a new generalization.

2.2 Combinatorial Estimates

Another major line of research in arithmetic combinatorics is the study of the effect of arithmetic operations on the size of sets.

Given two sets of real numbers, A and B, we can easily see that

$$|A| + |B| - 1 \leq |A + B| \leq |A||B| + \min(|A|, |B|).$$

The upper bound comes from the number of unique ways to choose an element from A and an element from B. The lower bound comes from the following observation. Order the elements of A as $a_1 < a_2 < \ldots a_r$ and the elements of B as $b_1 < b_2 < ...b_s$ It follows that $A + B$ at least contains the $r + s - 1$ distinct elements

$$a_1 + b_1 < a_1 + b_2 < a_1 + b_3 < \cdots < a_1 + b_s < a_2 + b_s < \ldots a_r + b_s.$$

Under what conditions is the lower bound attained? Observe that by symmetry, $A + B$ also contains the $r + s - 1$ distinct elements

$$a_1 + b_1 < a_2 + b_1 < a_3 + b_1 < \cdots < a_s + b_1 < a_s + b_2 < \ldots a_s + b_r.$$

But if $|A + B| = |A| + |B| - 1$ then the above two listings must be the same.

$$a_1 + b_2 = a_2 + b_1, a_1 + b_3 = a_2 + b_2 \cdots \Rightarrow a_2 - a_1 = b_2 - b_1 = b_3 - b_2 = \ldots,$$

showing us that A and B both must be arithmetic progressions with the same common difference.

Cauchy first addressed this question about sumsets in $\mathbb{F}_p$, establishing the lower bound $\min(p, |A|+|B|-1)$ in 1813. Davenport later rediscovered this bound in 1935. This result is considered the first in arithmetic combinatorics, and today all known proofs still require some non-trivial idea, such as Fourier methods.

Let us now restrict ourselves to individual sets. Note the large gap between the lower and upper bounds:

$$2|A| - 1 \leq |A + A| \leq \frac{|A|(|A| + 1)}{2}.$$

This gap implores further investigation. For a random set, A, sums $a_i + b_j$ are unlikely to overlap, so we expect $|A + A|$ will be on the order of $|A|^2$. For a specific example, set

$$A = \{2^0, 2^1, 2^2, \ldots 2^{n-1}\}$$

Since every sum $a_i + a_j, i < j$ has a unique binary representation, $|A+A| = |A|(|A|+1)/2$.

It is thus natural to say that a set A with *small doubling* has high additive structure (with $2|A| - 1$ being smallest possible). What other sets have small doubling?

If we next take A to be any set of n elements from a progression of length kn, then we immediately get $|A + A| \leq 2k|A|$. If we take A to be a large subset of a proper GAP, then it is again clear that $A + A$ will not be significantly larger than A. Freiman's theorem, introduced earlier, says that such sets are the only kinds which have small doubling.

It is important to observe that all of these properties of set addition have immediate analogues in set multiplication. For instance, if A is a geometric progression, then $|A \cdot A| = 2|A| - 1$. And we can use a geometric version of Freiman's theorem to see that the only sets A for which $|A \cdot A| \leq k|A|$ are large subsets of proper generalized *geometric* progressions.

It is natural to ask whether $A + A$ and $A \cdot A$ can simultaneously be 'small'. In other words, can A be relatively closed with respect to addition and multiplication? In [13], Erdős and Szemerédi showed that it in fact cannot.

Theorem 2.4 *There is some absolute constant $\epsilon > 0$ such that if A is a set of real numbers, $|A| \geq 2$, then*

$$\max(|A + A|, |A \cdot A|) > |A|^{1+\epsilon}.$$

Efforts to quantify this *sum-product phenomenon*, as well as generalize it to other settings, are a cornerstone of arithmetic combinatorics. In chapter 3, we will survey that progress, but first, we will focus on Elekes's effort to determine the correct value of ϵ and its elegant connection to incidence geometry. We begin our foray into incidence geometry with the following result:

Theorem 2.5 *(Crossing Number Inequality)[31] Let $G = G(V, E)$ be a graph with $|E| \geq 4|V|$. Let the crossing number, $Cr(G)$, denote the minimum possible number of intersections between edges of a graph drawn in the plane. Then the crossing number $Cr(G) > \frac{|E|^3}{64|V|^2}$.*

Proof. First observe that any graph G can be made planar (crossing number zero) by removing at most $Cr(G)$ edges. Combining this fact with Euler's characteristic formula tells us that

$$Cr(G) \geq |E| - 3|V|$$

for arbitrary graphs $G(E, V)$.

Now we fix $G = G(V, E)$ with $|E| \geq 4|V|$. Let V' be a random subset of V consisting of vertices chosen independently with probability p to be determined later. Let $G' = G'(V', E')$ be the subgraph of G induced by the vertices of V'. Now we apply the above inequality to G', and utilize the linearity of expectation to see

$$\mathbb{E}(Cr(G')) \geq \mathbb{E}(|E'|) - 3\mathbb{E}(|V'|).$$

Since each vertex of V has probability p of being in included in V', it follows again by linearity of expectation that

$$\mathbb{E}(|V'|) = p|V|,$$

and

$$\mathbb{E}(|E'|) = p^2|E|.$$

Now we consider a drawing of G with the minimal number of crossings, $Cr(G)$. Since each crossing spawns from four vertices, there is only a p^4 probability that a crossing survives when we pass to G'. So by one more applications of linearity of expectation we have

$$\mathbb{E}(Cr(G')) \leq p^4 Cr(G).$$

We now conclude that

$$Cr(G) \geq p^{-2}|E| - 3p^{-3}|V|.$$

By setting $p := \frac{4|V|}{|E|}$ we achieve our desired inequality.

Theorem 2.6 *(Szemerédi-Trotter theorem) Let P be a finite set of points and L be a finite set of lines, both in $\mathbb{R}^2$. Then we have*

$$I(P, L) \leq 4|P|^{2/3}|L|^{2/3} + 4|P| + |L|$$

Proof. Without loss of generality we can disregard lines $l \in L$ which do not contain any points in P, since they do not contribute anything to the left-hand side. Thus we assume that every line in L contains at least one point in P.

We proceed by embedding this information into a graph. Let $G = G(P, E)$ be the graph whose vertices are the points in P. Next, we connect two points (vertices) a and b if and only if the open line segment from a to b lies in a line in L and contains no points in P.

Now we count $|E|$, the number of edges, in two different ways. First, note that if a line l in L contains $k \geq 1$ points in p, then l contributes $k - 1$ edges to E Summing over $l \in L$, we see

$$|E| + |L| = I(P, L).$$

Secondly, we observe that G has a drawing in the plane, with the vertices in P represented by distinct points, and with each edges $(a, b) \in E$ represented by a line seqment from a to b. Since no two lines can intersect in more than one point, we conclude that

$$Cr(G) \leq |L|^2$$

Now, either $|E| \leq 4|P|$, or we can apply the crossing number inequality, by which $Cr(G) \geq \frac{|E|^3}{64|P|^2}$. It follows that

$$|E| \leq \max(4|P|, 4|P|^{2/3}|L|^{2/3}),$$

and our desired inequality follows. ■

We can now use this tool to prove Elekes's result.

Theorem 2.7 *Let A be a finite non-empty set of reals. Then*

$$|A+A| + |A \cdot A| \gg (|A|^{5/4})$$

Proof. Define a grid

$$P = \{(a,b) | a \in A+A, b \in A \cdot A\},$$

with cardinality $|A+A||A \cdot A|$. Consider the set L of lines of the form

$$\{(x,y) | y = a(x-b); a, b \in A\}$$

It is clear that L has $|A|^2$ elements. Additionally, each such line contains at least $|A|$ points in P, namely the points of the form $(b+c, ac)$ with $c \in A$. Thus $I(P,L) \geq |A|^3$. We apply the Szemerédi-Trotter theorem to conclude

$$|A|^3 \leq O((|A+A||A \cdot A|)^{2/3}(|A|^2)^{2/3} + |A+A||A \cdot A| + |A|^2).$$

From here, elementary algebra leads us to our claim. ■

The key fact that Elekes needed for his proof, and which is a weak corollary of the Szemerédi-Trotter incidence theorem, at least as far as just getting a non-trivial bound of the sort

$$|A+A| \cdot |A \cdot A| \ \gg \ |A|^{2+\varepsilon},$$

is the following basic claim.

Claim 1. There are absolute constants $\varepsilon > 0$ and $\delta > 0$ such that if A and B are sets of n real numbers, and n is sufficiently large (in terms of ε and δ), then any set of at least $n^{2-\varepsilon}$ distinct lines contains a member that hits the grid in fewer than $n^{1-\delta}$ points. In other words, one cannot have a collection of $n^{2-\varepsilon}$ lines whereby all are "$n^{1-\delta}$-rich" in the grid $A \times B$.

Actually, Elekes's proof only needs the following even weaker claim.

Claim 2. There exist absolute constants $\varepsilon > 0$ and $\delta > 0$ so that the following holds for all integers n sufficiently large: Suppose that A and B are sets of real numbers of size n, and that one has a family of lines such that

- There are at least $n^{1-\varepsilon}$ distinct slopes among them; and,
- every line is parallel to at least $n^{1-\varepsilon}$ others.

Then, at least one of the lines must hit the grid $A \times B$ in fewer than $n^{1-\delta}$ points. In other words, not all the lines can be $n^{1-\delta}$-rich in the grid.

In Chapter 5 we prove the following theorem, which shows that it is possible to considerably strengthen this second claim; futhermore, our theorem is not the sort that is quickly deducible from the Szemerédi-Trotter incidence theorem:

Theorem 5.1 *For every $\varepsilon > 0$, there exists $\delta > 0$ so that the following holds for all n sufficiently large: Suppose that A and B are sets of real numbers of size n, and that one has a family of lines such that*

- *There are at least n^{ε} distinct slopes among them; and,*
- *every line is parallel to at least n^{ε} others.*

Then, at least one of the lines must hit the grid $A \times B$ in fewer than $n^{1-\delta}$ points.

Our theorem is related to a conjecture of Solymosi (see [10, Conj. 3.10] for details), which we modify and extend to make it better fit the context of the above results.

Solymosi's Conjecture. For every $\varepsilon > 0$, there exists $\delta > 0$, such that the following holds for all integers n sufficiently large: Suppose A and B are sets of real numbers of size n, and suppose that one has a collection of n^{ε} lines in general position (that is, no pair is parallel, and no three meet at a point). Then, not all of the lines can be $n^{1-\delta}$-rich in the grid $A \times B$.

This conjecture of Solymosi easily implies our main theorem (Theorem 5.1) above, for if one has a family of lines as described by our theorem, then it is a simple matter

to select one line from each of $\gg n^{\varepsilon/3}$ groups of parallel lines in such a way that one produces a collection in general position (first, select a single line of slope λ_1; then, select a line of slope $\lambda_2 \neq \lambda_1$; then, select a line of slope $\lambda_3 \notin \{\lambda_1, \lambda_2\}$ such that the three lines do not have a common intersection point; then, select a line of slope $\lambda_4 \notin \{\lambda_1, \lambda_2, \lambda_3\}$...).

In chapter 6, we will broaden the concept of the sum-product phenomenon; but first, we need to introduce an alternative notation for describing operations on sets. Given a set A and a bivariate function $f(x, y)$, define

$$f(A, A) = \{f(x, y) | x, y \in A\}.$$

In this setting, clearly $A + A$ corresponds to the function $f(x, y) = x + y$ and $A \cdot A$ corresponds to $f(x, y) = xy$. Yet this latter notation is more versatile, permitting more complicated set functions such as $f(x, y) = x^2(y+1) + xy + 1$, which cannot be represented by the prior notation.

It is natural to deviate slightly from the extensive work on sum-product inequalities by studying the effect of other bivariate functions on the size of a set. We will call f a *set expander* if there exists an $\epsilon > 0$ such that for all sufficiently large A,

$$|f(A, A)| > |A|^{1+\epsilon}$$

Take, as an example, the function $f(x, y) = (x - 1)(y - 1)$. For $A = \{2^0 + 1, 2^1 + 1, \ldots 2^{n-1} + 1\}$, we have $f(A, A) = \{2^0, 2^1, \ldots 2^{2n-2}\}$. So clearly f is not a set expander. In chapter 6, we will categorize all degree 2 and degree 3 bivariate polynomial expanders.

CHAPTER III

REVIEW OF LITERATURE

The Balog-Szemerédi-Gowers theorem has a rich history, and is a very useful tool in additive combinatorics. It began with a paper by Balog and Szemerédi [1], and then was refined by Gowers [18] to the following basic result (actually, Gowers proved somewhat more than we bother to state here):

Theorem 3.1 *There exists an absolute constant $\kappa > 0$ such that the following holds for all finite subsets X and Y of size $n > n_0$ of an abelian group: Suppose that there are at least Cn^3 solutions to $x_1 + y_1 = x_2 + y_2$, $x_i \in X$ and $y_i \in Y$. Then, X contains a subset X', of size at least $C^{\kappa}n$, such that*

$$|X' + X'| \ \leq \ C^{-\kappa}n.$$

Sudakov, Szemerédi and Vu [28] proved a refinement of this theorem (Balog [2] independently obtained a similar result), given as follows:

Theorem 3.2 *Let n, C, K be positive numbers, and let A and B be two sets of n integers. Suppose that there is a bipartite graph $G(A, B, E)$ with at least n^2/K edges and $|A +_G B| \leq Cn$. Then one can find a subset $A' \subset A$ and a subset $B' \subset B$ such that $|A'| \geq n/16K^2$, $|B'| \geq n/4K$ and $|A' + B'| \leq 2^{12}C^3K^5n$.*

Remark. It is not difficult to show that this theorem, along with some lemmas and theorems of Ruzsa (the Ruzsa triangle inequality [31], and the Ruzsa-Plunnecke Theorem [25]), implies that we may take $\kappa < 20$ in Theorem 3.1.

In the same paper, Sudakov, Szemerédi and Vu [28, Theorem 4.3] proved the following powerful hypergraph version of the Balog-Szemerédi-Gowers Theorem:

Theorem 3.3 *For any positive integer k, there are polynomials $f_k(x,y)$ and $g_k(x,y)$ with degrees and coefficients depending only on k, such that the following holds. Let n, C, K be positive numbers. If $A_1, ..., A_k$ are sets of n positive integers, $H(A_1, ..., A_k, E)$ is the k-partite, k-uniform hypergraph with at least n^k/K edges, and $|\oplus_{Hi=1}^{k} A_i| \leq Cn$, then one can find subsets $A_i' \subset A_i$ such that*

- $|A_i'| \geq n/f_k(C,K)$ *for all* $1 \leq i \leq k$;
- $|A_1' + \cdots + A_k'| \leq g_k(C,K)n$.

The notation $\oplus_H$ means that the sum is restricted to the hypergraph H.

Beautiful and useful as it is, it would be nice if one had some control on the degrees of these polynomials f and g. And, for particular applications, it would be good to be able to control the rate of growth of sums $A_1' + \cdots + A_\ell'$, where ℓ is much smaller than k – it would be good to be able to bound the size of this sum from above by

$$C^{1+\varepsilon} K^{d_k} n, \tag{1}$$

where d_k depends only on k. Perhaps such a bound can be developed by modifying the proof of Sudakov, Szemerédi and Vu; however, in this chapter, we take a different tack, and produce an alternate proof of a related hypergraph Balog-Szemerédi-Gowers theorem, where such an upper bound as (1) will be implicit, though only for the case where $A_1 = \cdots = A_k$. In our proof, we will use some of the same standard tricks as Sudakov, Szemerédi and Vu do in their proof.

The notation we use to describe this theorem, and its proof, will be somewhat different from that used by Sudakov, Szemerédi and Vu. Furthermore, we will not attempt here to give the most general formulation of the theorem.

Our work in chapter 5 makes use of several standard methods in arithmetic combinatorics, though is quite intricate and technical. In particular, some of our approaches

are similar to those appearing in the well-known paper of Bourgain, Katz and Tao [5], as was pointed out to us by P. M. Wood. Even so, we do not assume any results more sophisticated than the Szemerédi-Trotter theorem. It was pointed out to us recently by T. Tao that perhaps we could make use of a particular sum-product ideas of Bourgain to give a simpler proof; however, we decided to present here our original approach.

It is possible that perhaps some of the ideas of Harald Helfgott [20] might allow us to give a shorter proof, as part of our argument can be phrased in terms of growth and generation in subgroups of $GL_2(\mathbb{R})$.

In 1997, Székely [29] generalized the Szemerédi and Trotter's incidence bound for general curves, which will be essential in chapter 6. In 2003, Tóth [32] extended the lines and incidences bounded to the complex plane. Many more results in incidence geometry can be found in [10],[4],[24],[12], and [6].

As discussed in the introduction, Erdős and Szemerédi first studied sum-product inequalities over $\mathbb{Z}$ in [13]. Nathanson [22] showed that one can set $\delta = 1/31$ to satisfy

$$|A + A| + |A \cdot A| \gg |A|^{1+\delta}.$$

Ford [15] further improved δ it to $1/15$ The proof of Elekes we discussed not only improved δ to $1/4$, it extended the result to $\mathbb{R}$. Solymosi [27] recently improved δ to $3/11$ and extended the study to complex numbers, by building on Elekes' connection to incidence geometry.

Bourgain, Katz and Tao [5] first studied the sum-product phenomenon over finite fields. In a recent preprint, Ernie Croot and Derrick Hart studied the problem over $\mathbb{C}[x]$. More along the lines of our work in chapter 6, Van Vu [34] characterized the bivariate polynomials P of degree k over $\mathbb{F}_q[x_1, x_2]$ such that for all $A \subseteq F_q$,

$$\max\{|A+A|, |P(A)|\} \geq |A| \min\{\delta(\frac{|A|^2}{k^4 q})^{1/4}, \delta(\frac{q}{k|A|})^{1/3}\}.$$

More directly, our work expands on the following work of Elekes, Nathanson, and Rusza [11], which will help us reduce our work in chapter 6.

Theorem 3.4 *Let $A \subset \mathbb{R}$ and let f be a strictly convex (or concave) function. Then*

$$|A + f(A)| \gg |A|^{5/4}.$$

CHAPTER IV

ON A CERTAIN GENERALIZATION OF THE BALOG-SZEMERÉDI-GOWERS THEOREM

In this chapter, we will prove the following generalization of the Balog-Szemerédi-Gowers theorem.

Theorem 4.1 *For every $0 < \varepsilon < 1/2$ and $c > 1$, there exists $\delta > 0$, such that the following holds for all k sufficiently large, and all sufficiently large finite subsets A of an additive abelian group: Suppose that*

$$S \subseteq A \times A \times \cdots \times A = A^k,$$

and let

$$\Sigma(S) := \{a_1 + \cdots + a_k : (a_1, ..., a_k) \in S\}.$$

If

$$|S| \geq |A|^{k-\delta}, \text{ and } |\Sigma(S)| < |A|^c,$$

then there exists

$$A' \subseteq A, \ |A'| \geq |A|^{1-\varepsilon},$$

such that

$$|\ell A'| = |A' + \cdots + A'| \leq |A'|^{c(1+\varepsilon\ell)}.$$

4.1 Proof of Theorem 5.1

4.1.1 Notation and basic assumptions

It will be advantageous to describe the proof in terms of strings. So, the set

$$S \subseteq A^k$$

will be thought of as a collection of strings of length k:

$$x_1 x_2 \cdots x_k,$$

where each $x_i \in A$.

Often, we split these strings up into substrings; for example, the string

$$x \;=\; x_1 \cdots x_k$$

can be written as a product of a "left substring ℓ of length $k/2$" (assume k is even) and a "right substring r of length $k/2$". So,

$$x \;=\; \ell r.$$

We may assume that

$$k \;=\; 2^n,$$

since if this is not the case, then we let k' be the largest power of 2 of size at most k, and proceed as follows: Given a string $x_1 \cdots x_k$ in S, we write it as a product $\ell_x r_x$, where

$$\ell_x \;:=\; x_1 \cdots x_{k'} \;\text{ and }\; r_x \;:=\; x_{k'+1} \cdots x_k.$$

Now, for some string y we will have that $r_x = y$ for at least $|S|/|A|^{k-k'}$ choices for $x \in S$. Letting S' denote the set of all strings ℓ_x with $r_x = y$, we will have

$$|S'| \;\geq\; |A|^{k'-\delta},$$

and clearly

$$|\Sigma(S')| \;\leq\; |\Sigma(\{\ell_x y \;:\; x \in S'\})| \;\leq\; |\Sigma(S)| \;\leq\; |A|^c.$$

So, we could just assume that our k had this value k' all along (remember, we get to choose k to be as large as needed to get the desired conclusion).

4.1.2 Lengths of iterations and the choice of δ and k

Our proof will be highly iterative, and will produce a sequence of sets

$$S_0 \ := \ S, \ S_1, \ S_2, \ ..., \text{ each } S_m \ \subseteq \ A^{k_m},$$

until one is found that has certain nice properties.

We will think of this process in terms of 'replacing' the set $S_m \subseteq A^{k_m}$ with a set $S_{m+1} \subseteq A^{k_{m+1}}$ that satisfies 'better' inequalities, specifically

$$|S_{m+1}| \ \geq \ A^{k_{m+1}-\delta_{m+1}}, \text{ and } |A|^{1-O(\delta)} \ \leq \ |\Sigma(S_{m+1})| \ \leq \ |\Sigma(S_m)|^{1-\varepsilon/400c},$$

where each $\delta_i \leq 5^i\delta$. Clearly, for $\delta > 0$ small enough, the number of such iterations we can take will be bounded from above in terms of ε and c. Furthermore, since at each step, k_{m+1} will be at least half the size of k_m, so long as the initial value of $k_0 = k$ is large enough in terms of c and ε, we will not run out of dimensions.

Since our theorem is a qualitative result, in that it does not even attempt to explain how δ or k depends on ε and c, there is no need to be more precise about just how small one needs take δ or how large to take k, in order for our iteration process to terminate and prove our theorem.

Let

$$S_0 \ := \ S, \ k_0 \ := \ k, \ \delta_0 \ := \ \delta, \text{ and set } m \ := \ 0.$$

4.1.3 The iteration part of the argument

Given a string x of length $k_m/2$, we let $R_m(x)$ denote the set of all strings y of length $k_m/2$ such that

$$xy \ \in \ S_m.$$

We analogously define $L_m(y)$ to be those strings x such that $xy \in S_m$.

We will now select an x, and therefore $R_m(x)$, very carefully, so that it satisfies certain useful properties: We begin with the inequality

$$\sum_x |R_m(x)| \;=\; |S_m| \;\geq\; |A|^{k_m - \delta_m}.$$

We now apply the following lemma.

Lemma 3 *Suppose that V is a set of n elements, and suppose that*

$$U_1, U_2, ..., U_r \;\subseteq\; V$$

satisfy

$$\sum_{i=1}^{r} |U_i| \;\geq\; rn^{1-\delta}.$$

Then, there exists $1 \leq j \leq r$ such that

$$\sum_{1 \leq i \leq r} |U_i \cap U_j| \;\geq\; rn^{1-2\delta}.$$

Proof of the lemma. Let $r(v)$ denote the number of sets U_i that contain the element $v \in V$. One easily sees that

$$\sum_{v \in V} r(v)^2 \;=\; \sum_{1 \leq i,j \leq r} |U_i \cap U_j|,$$

and

$$\sum_{v \in V} r(v) \;=\; \sum_{i=1}^{r} |U_i|.$$

So, the Cauchy-Schwarz inequality tells us that

$$\sum_{1 \leq i,j \leq r} |U_i \cap U_j| \;\geq\; \left(\sum_{i=1}^{r} |U_i| \right)^2 |V|^{-1} \;\geq\; r^2 n^{1-2\delta}.$$

Picking out any value i making the sum over j on the corresponding terms on the left-hand-side maximal, we see that

$$\sum_{j=1}^{r} |U_i \cap U_j| \;\geq\; rn^{1-2\delta},$$

as claimed. ■

From this lemma we easily deduce that there exists x such that

$$\sum_y |R_m(x) \cap R_m(y)| \ \geq \ |A|^{k_m - 2\delta_m}.$$

Next, we let

$$S_{m+1} \ := \ \{yz \in S_m \ : \ z \in R_m(x)\}, \tag{1}$$

and we observe that

$$|S_{m+1}| \ = \ \sum_y |R_m(x) \cap R_m(y)| \ \geq \ |A|^{k_m - 2\delta_m};$$

so, S_{m+1} is not too much smaller than S_m.

We now let

$$\delta_{m+1} \ := \ 2\delta_m, \text{ and } k_{m+1} \ := \ k_m,$$

and observe that S_{m+1} satisfies

$$|S_{m+1}| \ \geq \ |A|^{k_{m+1} - \delta_{m+1}},$$

and we in addition have that every element of S_{m+1} can be expressed as yz, where $z \in R_m(x) = R_{m+1}(x)$.

Now suppose that there is a string y of length $k_{m+1}/2$ such that if

$$|R_{m+1}(y)| \ \geq \ |A|^{k_{m+1}/2 - 2\delta_{m+1}},$$

then

$$|\Sigma(R_{m+1}(y))| \ \leq \ |\Sigma(S_{m+1})|^{1-\varepsilon/400c}.$$

If this occurs, then we let

$$S_{m+2} \ := \ R_{m+1}(y), \ k_{m+2} \ := \ k_{m+1}/2, \ \delta_{m+2} \ := \ 2\delta_{m+1},$$

and we reassign

$$m \ \leftarrow \ m + 2,$$

and then we start back at the very beginning of this subsection 4.1.3.

4.1.4 The sets H' and H''

When we come out of the iteration loops from the previous subsection, we finish with a set S_m having a number of highly useful properties, among them:

- $|S_m| \ \geq \ |A_m|^{k_m - \delta_m}$;
- For a particular string x of length $k_m/2$, each $R_m(y) \subseteq R_m(x)$; and,
- If we let H denote those strings h of length $k_m/2$ such that

$$|R_m(h)| \ \geq \ |A_m|^{k_m/2 - 2\delta_m},$$

then for every such h we will have that

$$|\Sigma(S_m)|^{1-\varepsilon/400c} \ < \ |\Sigma(R_m(h))| \ \leq \ |\Sigma(S_m)|.$$

One can easily show, using the lower bound for $|S_m|$, that for $|A|$ sufficiently large,

$$|H| \ > \ |A_m|^{k_m/2 - 2\delta_m}.$$

Since

$$\sum_{z \in R_m(x)} |\{h \in H \ : \ hz \in S_m\}| \ \geq \ |H| \cdot |A|^{k_m/2 - 2\delta_m},$$

we deduce that there exists $z \in R_m(x)$ such that there are at least

$$|H| \cdot |A|^{-2\delta_m} \ \geq \ |A|^{k_m/2 - 4\delta_m}$$

vectors $h \in H$ satisfying

$$hz \ \in \ S_m. \tag{2}$$

Fix one of these z, and let

$$H' \ \subseteq \ H$$

denote all those $h \in H$ such that (2) holds. Note that

$$|H'| \ \geq \ |A|^{k_m/2-4\delta_m}. \tag{3}$$

Next, let

$$H'' \ \subseteq \ H'$$

denote those $h \in H'$ such that there are at least

$$|H'| \cdot |\Sigma(H')|^{-1}/2 \tag{4}$$

other $h' \in H'$ satisfying

$$\Sigma(h') \ = \ \Sigma(h).$$

We have that

$$|H' \setminus H''| \ \leq \ |\Sigma(H')|(|H'| \cdot |\Sigma(H')|^{-1}/2) \ = \ |H'|/2$$

So,

$$|H''| \ \geq \ |H'|/2 \ \geq \ |A|^{k_m/2-5\delta_m}, \tag{5}$$

for $|A|$ sufficiently large.

We also note that

$$|\Sigma(H'')| \ \leq \ |\Sigma(H')| \ = \ |\Sigma(\{hz \ : \ h \in H'\})| \ \leq \ |\Sigma(S_m)|.$$

This is one of the places where it was essential to have that $z \in R_m(h)$ for all $h \in H'$.

Now suppose that, in fact,

$$|\Sigma(H'')| \ \leq \ |\Sigma(S_m)|^{1-\varepsilon/400c}. \tag{6}$$

If so, then we set

$$S_{m+1} \ := \ H'', \ k_{m+1} \ := \ k_m/2, \ \delta_{m+1} \ := \ 5\delta_m,$$

and then we update m to

$$m \ \leftarrow \ m+1,$$

and we repeat our iteration process again, starting in subsection 4.1.3.

On the other hand, if (6) does not hold, then we will have that

$$|\Sigma(S_m)|^{1-\varepsilon/400c} \ \leq \ |\Sigma(H'')| \ \leq \ |\Sigma(H')| \ \leq \ |\Sigma(S_m)| \tag{7}$$

4.1.5 The final leg of the proof

From the fact that

$$|\Sigma(\{hu \in S_m \ : \ h \in H'', u \in R_m(h)\})| \ \leq \ |\Sigma(S_m)|,$$

along with the fact that $R_m(h) \subseteq R_m(x)$ and

$$|\Sigma(S_m)|^{1-\varepsilon/400c} \ \leq \ |\Sigma(R_m(h))| \ \leq \ |\Sigma(R_m(x))| \ \leq \ |\Sigma(S_m)|,$$

as well as (7), we deduce that there are at least

$$|\Sigma(H'')|^2(\min_{h\in H''}|\Sigma(R_m(h))|^2)|\Sigma(S_m)|^{-1} \ \geq \ |\Sigma(S_m)|^{3-\varepsilon/100c}$$

quadruples

$$\sigma_1, \sigma_2 \ \in \ \Sigma(H''), \text{ and } \sigma_3, \sigma_4 \ \in \ \Sigma(R_m(x)),$$

such that

$$\sigma_1 + \sigma_3 \ = \ \sigma_2 + \sigma_4.$$

Now we apply Theorem 3.1, setting

$$X \ := \ \Sigma(H''), \text{ and } Y \ := \ \Sigma(R_m(x)).$$

Following the comment after Theorem 3.2, we have that there exists

$$\Sigma \ \subseteq \ \Sigma(H''), \ |\Sigma| \ \geq \ |\Sigma(H'')|^{1-\varepsilon/2c},$$

such that

$$|\Sigma + \Sigma| \ \leq \ |\Sigma|^{1+\varepsilon/2c}. \tag{8}$$

Let H''' denote the set of all

$$h \ \in \ H'',$$

such that

$$\Sigma(h) \ \in \ \Sigma.$$

By (3), (4), and (7), we have that

$$\begin{aligned}
|H'''| &\geq |\Sigma|(|H'| \cdot |\Sigma(H')|^{-1}/2) \\
&\geq |\Sigma(H'')|^{1-\varepsilon/2c}|H'| \cdot |\Sigma(S_m)|^{-1}/2 \\
&\geq |\Sigma(H'')|^{1-\varepsilon/2c}|\Sigma(H'')|^{-1/(1-\varepsilon/400c)}|H'|/2 \\
&\geq |\Sigma(H'')|^{-\varepsilon/c}|H'| \\
&\geq |A|^{k_m/2-4\delta_m-\varepsilon}.
\end{aligned}$$

By simple averaging, there is some vector

$$w \ \in \ A^{k_m/2-1},$$

such that there are at least

$$|A|^{1-4\delta_m-\varepsilon}$$

vectors $h \in H'''$ whose last $k_m/2-1$ coordinates are the vector w. The upshot of this is that if we let

$$A' \ := \ \{a \in A \ : \ aw \in H'''\},$$

then

$$|A'| \ \geq \ |A|^{1-4\delta_m-\varepsilon}, \tag{9}$$

and

$$A' + A' + 2\Sigma(w) \ \subseteq \ \Sigma(H''') + \Sigma(H''') \ = \ \Sigma + \Sigma.$$

Now we apply a weak form of the Ruzsa-Plunnecke Theorem [25], given as follows:

Theorem 4.2 *Suppose that X is some finite subset of an additive abelian group, such that*

$$|X + X| \ \leq \ C|X|.$$

Then, we have that

$$|kX| \ = \ |X + X + \cdots + X| \ \leq \ C^k|X|.$$

Using

$$X \ := \ \Sigma, \text{ and } C \ := \ |\Sigma|^{\varepsilon/2c},$$

we deduce that for ℓ even,

$$|\ell A'| \ \leq \ |\ell\Sigma| \ \leq \ |\Sigma|^{1+\varepsilon\ell/2c} \ \leq \ |A|^{c+\varepsilon\ell} \ \leq \ |A'|^{(c+\varepsilon\ell)/(1-4\delta_m-\varepsilon)}$$

By selecting $\delta > 0$ small enough (and therefore $\delta_m > 0$ small enough), relative to $\varepsilon > 0$, we can ensure that for $\varepsilon < 1/2$,

$$|\ell A'| \ \leq \ |A'|^{c(1+2\varepsilon\ell)}.$$

Of course, when $1/2 \leq \varepsilon < 1$ the inequality is trivial, as $c > 1$. Clearly, on rescaling ε appropriately, our theorem is proved.

CHAPTER V

ON RICH LINES IN GRIDS

In this chapter, we will prove the following theorem:

Theorem 5.1 *For every $\varepsilon > 0$, there exists $\delta > 0$ so that the following holds for all n sufficiently large: Suppose that A and B are sets of real numbers of size n, and that one has a family of lines such that*

- *There are at least n^{ε} distinct slopes among them; and,*
- *every line is parallel to at least n^{ε} others.*

Then, at least one of the lines must hit the grid $A \times B$ in fewer than $n^{1-\delta}$ points.

5.1 Proof of the main theorem

The first step in our proof is to reduce from the case of working with grids $A \times B$ to grids $A \times A$. This is easily handled by simply letting $C = A \cup B$, and then noting that the hypotheses of our theorem imply that we have a family of rich lines passing through the grid $C \times C$. Upon rescaling n to $|A \cup B| \leq 2n$, we see that we could have just assumed that our grid was $A \times A$ (or $C \times C$) all along.

5.1.1 Producing new rich lines from old ones

In our proof we will be combining together lots of pairs of rich lines, possibly of different slope: Given a line ℓ hitting $A \times A$ in some points, we let

$$\begin{aligned} X(\ell) &= \text{projection of } \ell \cap (A \times A) \text{ onto the x} - \text{axis;} \\ Y(\ell) &= \text{projection of } \ell \cap (A \times A) \text{ onto the y} - \text{axis.} \end{aligned}$$

If two lines

$$\ell \ : \ y \ = \ \lambda x + \mu \ \text{ and } \ \ell' \ : \ y \ = \ \lambda' x + \mu',$$

have the property that

$$|Y(\ell) \cap Y(\ell')| \ = \ \text{“large”},$$

then there will be lots of triples

$$(x, z, y) \ \in \ A \times A \times A$$

satisfying

$$\lambda x + \mu \ = \ y \ = \ \lambda' z + \mu'.$$

So, the new line

$$z \ = \ (\lambda/\lambda')x + (\mu - \mu')/\lambda'$$

also hits the grid $A \times A$ in many points.

A convenient way of keeping track of the new rich lines that we can produce from old ones is to use matrix notation: We form the association

$$y \ = \ \lambda x + \mu \quad \leftrightarrow \quad \begin{bmatrix} \lambda & \mu \\ 0 & 1 \end{bmatrix}.$$

Then, when we combine together lines as above, the new line we get will be the one associated to a certain product of matrices; specifically,

$$\begin{aligned} y \ = \ & (\lambda/\lambda')x + (\mu - \mu')/\lambda' \\ & \leftrightarrow \quad \begin{bmatrix} \lambda/\lambda' & (\mu - \mu')/\lambda' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda' & \mu' \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \lambda & \mu \\ 0 & 1 \end{bmatrix} \end{aligned}$$

A basic fact, which is an easy consequence of the Cauchy-Schwarz inequality, is the following lemma:

Lemma 4 *Given lines*

$$\ell_1,\ ...,\ \ell_K,$$

each hitting a grid

$$A\ \times\ A,$$

in at least

$$n^{1-\delta_0}\text{ points},$$

we have that at least

$$K^2 n^{-2\delta_0}/2$$

of the pairs (ℓ_i, ℓ_j) *have the property that*

$$|Y(\ell_i) \cap Y(\ell_j)|\ \geq\ n^{1-2\delta_0}/2. \tag{1}$$

If the lines $\ell_1, ..., \ell_K$ have slopes $\lambda_1, ..., \lambda_K$, respectively, then upon combining it with our preceeding observations, we deduce that there are lots of lines of slope λ_i/λ_j, for lots of pairs (i, j), such that each is at least $n^{1-2\delta_0}/2$ rich in the grid $A \times A$.

5.1.2 Passing to a set of rich lines with usable properties

Given $\varepsilon > 0$, we let $\delta' > 0$ denote some parameter that we will choose later. Then, given $\varepsilon, \delta' > 0$ we let $\delta > 0$ be some parameter chosen later. We will show that if $\delta > 0$ is small enough, and if (as stated in the hypotheses of our theorem) we had a set of lines involving n^ε slopes, each parallel to at least n^ε others, each $n^{1-\delta}$-rich in the grid, then in fact there would have to exist at least n^4 lines, each hitting $A \times A$ in at least two points. This clearly cannot happen, because there are fewer lines hitting the grid in two points than there are ordered pairs of points of the grid; there are n^2 points of the grid, and therefore n^4 ordered pairs. This will prove our theorem.

So, we assume that $\varepsilon > 0$ is given, and then we will select $\delta' > 0$ as small as needed, and then choose $\delta > 0$ even smaller later.

We begin by letting $L_1(\lambda)$ denote the set of our lines having slope λ. We note that

$$|L_1(\lambda)| \ \geq \ n^{\varepsilon},$$

where λ is one of the slopes of our set of lines. To make certain later estimates easier, we will trim our list of lines so that for each slope λ we have

$$|L_1(\lambda)| \ = \ \lceil n^{\varepsilon} \rceil.$$

Denote our initial set of slopes by Λ_1.

Using Lemma 4, we can easily deduce that there are at least

$$|\Lambda_1|^2 n^{-O(\delta)}$$

ordered pairs

$$(\lambda, \lambda') \ \in \ \Lambda_1 \times \Lambda_1,$$

for which there are at least

$$|L_1(\lambda)| \cdot |L_1(\lambda')| n^{-O(\delta)} \ \sim \ n^{2\varepsilon - O(\delta)}$$

pairs of lines

$$(\ell, \ell') \ \in \ L_1(\lambda) \times L_1(\lambda') \tag{2}$$

satisfying

$$|Y(\ell) \cap Y(\ell')| \ \geq \ n^{1-O(\delta)}. \tag{3}$$

Note that each of these intersections gives rise to a line having slope λ/λ' that hits $A \times A$ in $n^{1-O(\delta)}$ points.

When such a pair (λ, λ') has the above property we will say that it is "good for step 1". Note that our definition of "good" is dependent upon the implied constants in the big-ohs – for our purposes, the implied constants in the "good for step i" can all be taken to be 1000^i.

If a pair (λ, λ') is good for step 1, and if in addition we have that the number of distinct lines of slope λ/λ' produced by combining pairs (ℓ, ℓ') satisfying (2) and (3) is at least

$$n^{\varepsilon(1+\delta')}, \tag{4}$$

we will say that (λ, λ') is "very good for step 1".

Let us suppose that all but at least a fraction $n^{-O(\delta)}$ of the "good" pairs (λ, λ') are, in fact, "very good". Let Λ_2 denote these "very good" pairs, and note that we are saying

$$|\Lambda_2| \ \geq \ |\{\text{good pairs}\}| n^{-O(\delta)} \ \geq \ |\Lambda_1|^2 n^{-O(\delta)}.$$

For $\theta \in \Lambda_2$, say $\theta = (\lambda, \lambda')$, we let $L_2(\theta)$ denote those lines produced by combining together pairs of lines, one from $L_1(\lambda)$ and the other from $L_1(\lambda')$. Note that for all $\theta \in \Lambda_2$ we have, by (4), that

$$|L_2(\theta)| \ \geq \ n^{\varepsilon(1+\delta')}.$$

And, as with the set of lines $L_1(\lambda)$, we trim our set of lines (in an arbitrary manner) so that for every such θ we have that

$$|L_2(\theta)| \ = \ \lceil n^{\varepsilon(1+\delta')} \rceil.$$

It is easily deduced from Lemma 4 that there are at least

$$|\Lambda_2|^2 n^{-O(\delta)}$$

ordered pairs

$$(\theta, \theta') \ \in \ \Lambda_2 \times \Lambda_2,$$

for which there are at least

$$|L_2(\theta)| \cdot |L_2(\theta')| n^{-O(\delta)} \ \sim \ n^{2\varepsilon(1+\delta')-O(\delta)}$$

pairs of lines

$$(\ell, \ell') \ \in \ L_2(\lambda) \times L_2(\lambda')$$

satisfying

$$|Y(\ell) \cap Y(\ell')| \ \geq \ n^{1-O(\delta)}.$$

When such a pair (θ, θ') has the above property we will say that it is "good for step 2", and we say that it is "very good for step 2" if the set of rich lines that it produces has size at least

$$n^{\varepsilon(1+\delta')^2}.$$

We will repeat the above process we have started as above, by defining Λ_3 to be the set of all "very good for step 2" pairs $\beta = (\theta, \theta') \in \Lambda_2 \times \Lambda_2$, and we will let $L_3(\beta)$ be those lines produced by combining together ones from $L_2(\theta)$ with $L_(\theta')$, and then trimming the list so that

$$|L_3(\beta)| \ = \ \lceil n^{\varepsilon(1+\delta')^2} \rceil.$$

It is clear that we can continue the above process, producing sets

$$\Lambda_4, \ \Lambda_5, \ ..., \ \text{where } \Lambda_i \ \subseteq \ \Lambda_{i-1} \times \Lambda_{i-1},$$

and sets

$$L_3(\alpha_3), \ L_4(\alpha_4), \ ..., \ \text{where } \alpha_i \ \in \ \Lambda_i.$$

However, the process cannot go on for too long, since we always have the upper bound

$$|L_t(\alpha)| \ \leq \ n^4,$$

since the lines of $L_t(\alpha)$ will hit the grid in at least two points. In fact,

$$t \ \ll \ T \ := (1/\delta') \log(4/\varepsilon).$$

Well, the above sequence of Λ_j's and $L_j(\alpha_j)$'s is not quite what we want, because for later arguments we will need that the sequence terminates with $t > k$, for some

$k = k(\varepsilon)$ depending only on $\varepsilon > 0$. The way we get around this is as follows: Going back to how our sequences of Λ_j's and $L_j(\alpha_j)$'s are defined, if we are willing to allow the Λ_j, $j = 1, 2, ..., k$ to merely contain "good for step j" pairs, instead of "very good for step j" pairs, then the problem of stopping at time $t \leq k$ is avoided. There is the issue of how to trim the sets $L_2(\alpha_2), ..., L_k(\alpha_k)$ in the right way. To solve this problem, we merely trim them so that they each contain $n^{\varepsilon - O(\delta)}$ lines, which is easily guaranteed. Furthermore, by choosing $\delta' > 0$ small enough, we can still have that for $j > k$ and $\theta \in \Lambda_j$,

$$|L_j(\theta)| \ = \ \lceil n^{\varepsilon(1+\delta')^j} \rceil,$$

the reason being that for small $\delta' > 0$, the $(1 + \delta')^k$ can be made as close to 1 as needed.

Before unraveling what this all means, we make one more observation: An element $\theta \in \Lambda_i$ corresponds to a pair of elements of Λ_{i-1}, and each member of the pair itself corresponds to pairs of elements of Λ_{i-2}, and so on; so, in the end, an element of $\theta \in \Lambda_i$ in fact corresponds to a sequence of elements of Λ_1 of length 2^{i-1}. Say the sequence is

$$\lambda_1, \ ..., \ \lambda_{2^{i-1}}.$$

Then, the lines it corresponds to all have slope

$$\lambda_1 \cdots \lambda_{2^{i-2}} \ / \ \lambda_{2^{i-2}+1} \cdots \lambda_{2^{i-1}}.$$

When our above process terminates at time t satisfying

$$k \ < \ t \ \ll \ T,$$

we will have that the following all hold:

- First, for at least

$$|\Lambda_1|^{2^{t-2}} n^{-O_t(\delta)}$$

sequences

$$\lambda_1,\ ...,\ \lambda_{2^{t-2}}\ \in\ \Lambda_1$$

we will have a set of lines of slope

$$\lambda_1 \cdots \lambda_{2^{t-3}}\ /\ \lambda_{2^{t-3}+1} \cdots \lambda_{2^{t-2}}$$

that are $n^{1-O_t(\delta)}$-rich in our grid $A \times A$.

- Second, there are at least

$$|\Lambda_1|^{2^{t-1}} n^{-O_t(\delta)}$$

pairs of sequences

$$\lambda_1,\ ...,\ \lambda_{2^{t-1}} \in \Lambda_1,\ \text{and}\ \lambda'_1,\ ...,\ \lambda'_{2^{t-1}} \in \Lambda_1,$$

corresponding to a pair of elements

$$(\nu_1, \nu_2)\ \in\ \Lambda_{t-1} \times \Lambda_{t-1},$$

that are "good for step t" but not "very good for step t" (since otherwise we could continue the iteration for another step). For such a pair, suppose that our $n^{1-O_t(\delta)}$-rich lines corresponding to ν_1 are of the form

$$y\ =\ (\lambda_1 \cdots \lambda_{2^{t-3}} / \lambda_{2^{t-3}+1} \cdots \lambda_{2^{t-2}})x\ +\ B_{\nu_1}, \tag{5}$$

and those corresponding to ν_2 are of the form

$$y\ =\ (\lambda'_1 \cdots \lambda'_{2^{t-3}} / \lambda'_{2^{t-3}+1} \cdots \lambda'_{2^{t-2}})x + B_{\nu_2}. \tag{6}$$

Then, since the pair (ν_1, ν_2) is "good for step t", we have that there are

$$|B_{\nu_1}| \cdot |B_{\nu_2}| n^{-O_t(\delta)}$$

ordered pairs of lines, one corresponding to ν_1 and the other to ν_2, such that when combined, give us an $n^{1-O_t(\delta)}$-rich line of the form

$$y \;=\; \alpha x + (b_1 - b_2)/\beta,$$

where

$$\alpha \;=\; \lambda_1 \cdots \lambda_{2^{t-3}} \lambda'_{2^{t-3}+1} \cdots \lambda'_{2^{t-2}} / \lambda'_1 \cdots \lambda'_{2^{t-3}} \lambda_{2^{t-3}+1} \cdots \lambda_{2^{t-2}},$$

where

$$b_1 \;\in\; B_{\nu_1},\; b_2 \;\in\; B_{\nu_2},\; \text{and where } \beta \;=\; \lambda'_1 \cdots \lambda'_{2^{t-3}} / \lambda'_{2^{t-3}+1} \cdots \lambda'_{2^{t-2}}.$$

Furthermore, since the pair (ν_1, ν_2) is not "very good for step t", we have that the possibilities for the difference $b_1 - b_2$ is at most

$$n^{\varepsilon(1+\delta')^t} \;\leq\; |L_{t-1}(\nu_1)|^{1+\delta'} \;=\; |B_{\nu_1}|^{1+\delta'}.$$

What this means is that the "additive energy" between the sets B_{ν_1} and B_{ν_2} must be "large". In fact, because there are so many pairs (ν_1, ν_2), there must exist $\nu_1 \in \Lambda_{t-1}$ such that there are at least

$$|\Lambda_{t-1}| n^{-O_t(\delta)}$$

choices for $\nu_2 \in \Lambda_t$, such that we have the following lower bound for the additive energy:

$$\begin{aligned} E(B_{\nu_1}, B_{\nu_2}) \;&=\; |\{(b_1, b_2, b_3, b_4) \in B_{\nu_1} \times B_{\nu_1} \times B_{\nu_2} \times B_{\nu_2} \;:\; b_1 - b_3 \;=\; b_2 - b_4\}| \\ &\geq\; |B_{\nu_1}|^{3-O(\delta')}. \end{aligned}$$

We now require the following standard lemma.

Lemma 5 *Suppose that X and Y are sets of size M, such that*

$$E(X, Y) \;=\; |\{(x, x', y, y') \in X \times X \times Y \times Y \;:\; x - y = x' - y'\}| \;\geq\; cM^3.$$

Then, there is some translate u such that

$$|(X + u) \cap Y| \;\geq\; cM.$$

Proof of the Lemma. Another way of writing the additive energy is

$$E(X,Y) \ = \ \sum_{\substack{u\in X\\ v\in Y}} |(X-u)\cap (Y-v)|.$$

So, by simple averaging, among the M^2 pairs $(u,v) \in X\times Y$, there exists one for which

$$|(X-u+v)\cap Y| \ = \ |(X-u)\cap (Y-v)| \ \geq \ cM;$$

■

So, for some fixed $\nu_1 \in \Lambda_{t-1}$, and for $|\Lambda_{t-1}|n^{-O_t(\delta)}$ elements $\nu_2 \in \Lambda_{t-1}$, there exist translates $\tau(\nu_2)$ for which

$$|B_{\nu_1} \cap (B_{\nu_2} + \tau(\nu_2))| \ \geq \ |B_{\nu_1}|n^{-O_t(\delta')}.$$

We now arrive at the following basic claim.

Claim 3. Under the hypotheses of our theorem, there are distinct slopes

$$\theta_1, ..., \theta_N,$$

where

$$N \ > \ n^{\varepsilon - O(\delta)},$$

such that for

$$m \ = \ 2^{t-2},$$

at least $N^{m-O(\delta)}$ of the m-fold products $\theta_{i_1}\cdots\theta_{i_m}$, we have a set of $n^{1-O(\delta)}$-rich lines of the form

$$y \ = \ \theta_{i_1}\cdots\theta_{i_m}x \ + \ B(i_1,...,i_m),$$

where $B(i_1,...,i_m)$ is some set of slopes. We furthermore assume there is a set C of real numbers such that for each of these $> N^{m-O(\delta)}$ sets $B(i_1,...,i_m)$, there exists a

real number $\tau(i_1, ..., i_m)$, such that

$$|B(i_1, ..., i_m) \triangle (C + \tau(i_1, ..., i_m))| \ < \ |B(i_1, ..., i_m)| n^{-O(\delta)}. \tag{7}$$

Here, $S \triangle T$ denotes the symmetric difference between S and T.

Proof of the claim. Basically, we just need to show how these slopes θ_i link up with the lines in (5) and (6); further, we need to explain the presence of the δ here, rather than the δ' appearing earlier.

Let us first address the issue of the δ versus of the δ': Since we get to choose $\delta' > 0$ as small as desired relative to $\varepsilon > 0$, we can just as well rewrite it is $\delta > 0$.

As to the relationship between the θ_i's above and the λ_j's in (5), we will take

$$\{\theta_1, \ ..., \ \theta_N\} \ = \ \{\lambda_i\} \cup \{1/\lambda_i\}.$$

Then, for $m = 2^{t-2}$ we have that the lines of (5) have slope of the form $\theta_{i_1} \cdots \theta_{i_m}$. Furthermore, the fact that $t > k$ is what will allow us to take m as large as needed. ■

Now we combine together pairs of these rich lines – as discussed in subsection 5.1.1 – having the same slope, to produce many other rich lines having slope 1: Fix one of the slopes $\theta_{i_1} \cdots \theta_{i_m}$ leading to rich lines with the set of slopes $B(i_1, ..., i_m)$. Applying Lemma 4, we find that there are at least

$$|B(i_1, ..., i_m)|^2 n^{-O(\delta)}$$

ordered pairs

$$(b, b') \ \in \ B(i_1, ..., i_m) \times B(i_1, ..., i_m),$$

such that the line

$$y \ = \ x \ + \ (b - b')/\theta_{i_1} \cdots \theta_{i_m}$$

is $n^{1-O(\delta)}$-rich in the grid $A \times A$.

From (7), and a little bit of effort, we can easily deduce that at least $|B(i_1, ..., i_m)|^2 n^{-O(\delta)}$ of these pairs (b, b') have the property that there exists $(c, c') \in C \times C$ satisfying

$$(b, b') \;=\; (c + \tau(i_1, ..., i_m), c' + \tau(i_1, ..., i_m)).$$

For such pairs, we will have that

$$b - b' \;=\; c - c'.$$

By the pigeonhole principle, there exists at least one pair (in fact, lots of pairs) $(c, c') \in C \times C$, $c \neq c'$, such that at least $N^{m-O(\delta)}$ of the sequences $i_1, ..., i_m$ have the property that the line

$$y \;=\; x \;+\; (c - c')/\theta_{i_1} \cdots \theta_{i_m}$$

is $n^{1-O(\delta)}$-rich in the grid $A \times A$. Let us denote this constant $c - c'$ as ξ, so that our rich lines all look like

$$y \;=\; x \;+\; \xi\varphi_{i_1} \cdots \varphi_{i_m}, \text{ where } \varphi_i \;:=\; 1/\theta_i.$$

By combining together pairs of these lines, as discussed in subsection 5.1.1, we can form new ones of the form

$$y \;=\; x \;+\; \xi(\varphi_{i_1} \cdots \varphi_{i_m} - \varphi_{j_1} \cdots \varphi_{j_m}) \tag{8}$$

that are rich in the grid. If we then combine together pairs of *those* lines, we get ones of the form

$$\begin{aligned} y \;=\; x \;+\; \xi(&\varphi_{i_1} \cdots \varphi_{i_m} - \varphi_{j_1} \cdots \varphi_{j_m} \\ &+\varphi_{k_1} \cdots \varphi_{k_m} - \varphi_{\ell_1} \cdots \varphi_{\ell_m}). \end{aligned} \tag{9}$$

Continuing in this manner, we can generate lines of slope 1 with y-intercept equal to ξ times alternating sums of m-fold products of the φ_i's; and, at the tth iteration, these alternating sums have 2^t terms.

5.1.3 The sequence Θ_i

Now we take a digression for a few pages, and define and analyze a certain sequence of expressions: Starting with the set

$$\Theta \ := \ \{\varphi_i \ : \ i = 1, 2, ...\},$$

consider the sequence of sets (expressions)

$$\Theta_1 := \Theta.\Theta - \Theta.\Theta, \ \Theta_2 := \Theta_1.\Theta_1 - \Theta_1.\Theta_1, \tag{10}$$

and so on. If we formally expand out the expressions, we will get sums of the following type: Θ_1 consists of sums of the type

$$a_1a_2 - a_3a_4, \ a_i \in \Theta,$$

and Θ_2 consists of the sums

$$\begin{aligned} &a_1a_2a_5a_6 - a_3a_4a_5a_6 - a_1a_2a_7a_8 + a_3a_4a_7a_8 \\ &-a_9a_{10}a_{13}a_{14} + a_9a_{10}a_{15}a_{16} + a_{11}a_{12}a_{13}a_{14} - a_{11}a_{12}a_{15}a_{16}, \end{aligned} \tag{11}$$

where again each $a_i \in \Theta$. We will not bother to write down Θ_3! In general, at the jth iteration, the terms in the alternating sum will involve 4^j variables a_i, and the number of terms will be 2^{2^j-1}.

Later on, in another subsection, we will show that so long as $\delta > 0$ is small enough, upon expanding Θ_{t-2} into the alternating sum of products of variables $a_1, ..., a_{4^{t-2}}$, as in (10) and (11), at least

$$|\Theta|^{4^{t-2}} n^{-O_t(\delta)}$$

choices for these $a_i \in \Theta$ will produce a

$$\theta = \theta(a_1, ..., a_{4^{t-2}}) \in \Theta_{t-2}$$

so that the line

$$y \;=\; x \;+\; \xi\theta \tag{12}$$

is $n^{1-O_t(\delta)}$-rich in the grid $A \times A$. We will then use Lemma 6 to show that this is impossible for t large enough and $\delta > 0$ small enough. The fact that $t > k$, where k is chosen as large as desired (k is as appears in subsection 5.1.2), will allow us to reach our contradiction, thereby proving Theorem 5.1.

5.1.3.1 A certain inductive claim

The key fact that we will show and use to accomplish our goal is the following.

Claim 4. Suppose that $g(x_1, ..., x_u)$ is some polynomial in the variables $x_1, ..., x_u$, which are to be thought of as taking on values in the set Θ. Consider the expansion of

$$\Theta_j \Theta_j g(x_1, ..., x_u)$$

into the variables $a_1, ..., a_{2 \cdot 4^j}, x_1, ..., x_u \in \Theta$. [1] Suppose that there are at least

$$|\Theta|^{2 \cdot 4^j + u} n^{-O_{j,u}(\delta)}.$$

choices for these variables, producing a value

$$\gamma \;=\; \gamma(a_1, ..., x_u) \;=\; \Theta_j \Theta_j g(x_1, ..., x_u)$$

such that the line

$$y \;=\; x \;+\; \xi\gamma$$

is $n^{1-O_{j,u}(\delta)}$-rich in the grid $A \times A$. Then, there are at least

$$|\Theta|^{4^{j+1} + u} n^{-O_{j,u}(\delta)}$$

choices for the variables

$$b_1, ..., b_{4^{j+1}}, y_1, ..., y_u \;\in\; \Theta$$

[1] The first Θ_j is expanded into $a_1, ..., a_{4^j}$, and the second Θ_j is expanded into $a_{4^j+1}, ..., a_{2 \cdot 4^j}$.

such that the line

$$y \ = \ x \ + \ \xi\gamma', \ \gamma' = \gamma'(b_1, ..., y_u) \in \Theta_{j+1} g(y_1, ..., y_u)$$

is $n^{1-O_{j,u}(\delta)}$-rich in $A \times A$.

Proof of the claim. Under the hypotheses of the above claim, the pigeonhole principle implies that for at least

$$|\Theta|^{4^{j+1}+u} n^{-O_{j,u}(\delta)} \tag{13}$$

choices of variables

$$b_1, ..., b_{2 \cdot 4^j}, c_1, ..., c_{2 \cdot 4^j}, x_1, ..., x_u \ \in \ \Theta,$$

we will have that if we let

$$\gamma_1 \ := \ \gamma_1(b_1, ..., b_{2 \cdot 4^j}, x_1, ..., x_u) \ \in \ \Theta_j \Theta_j g(x_1, ..., x_u),$$

and

$$\gamma_2 \ := \ \gamma_2(c_1, ..., c_{2 \cdot 4^j}, x_1, ..., x_u) \ \in \ \Theta_j \Theta_j g(x_1, ..., x_u)$$

(note that the value of $x_1, ..., x_u$ here is the same as for γ_1), then both the lines

$$y \ = \ x \ + \ \xi\gamma_1 \ \text{ and } \ y \ = \ x \ + \ \xi\gamma_2$$

are $n^{1-O_{j,u}(\delta)}$-rich in $A \times A$. Furthermore, by dint of Lemma 4 and the comments following it, we will additionally have that for (13) many choices of the b_i's, c_i's, and x_i's, the pair of lines may be combined to produce the new line

$$y \ = \ x \ + \ \xi(\gamma_1 - \gamma_2),$$

which will also be $n^{1-O_{j,u}(\delta)}$-rich in $A \times A$.

This

$$\gamma_1 - \gamma_2 \ = \ (\Theta_j \Theta_j - \Theta_j \Theta_j) g(x_1, ..., x_u)$$

has the form $\Theta_{j+1} g(x_1, ..., x_u)$. Clearly this proves the claim. ■

A consequence of this claim, and an easy induction argument (to be described presently), is that if the number of choices for

$$x_1, ..., x_{2^Z} \ \in \ \Theta$$

for which

$$y \ = \ x \ + \ \xi x_1 \cdots x_{2^Z} \tag{14}$$

is $n^{1-O_Z(\delta)}$-rich in $A \times A$ is at least

$$|\Theta|^{2^Z} n^{-O_Z(\delta)}, \tag{15}$$

which it is by the properties of the set Θ described earlier, then there are at least

$$|\Theta|^{4^Z} n^{-O_Z(\delta)}$$

choices for $y_1, ..., y_{4^Z} \in \Theta$ such that the line

$$y \ = \ x + \xi\gamma, \ \gamma = \gamma(y_1, ..., y_{4^Z}) \in \Theta_Z$$

is $n^{1-O_Z(\delta)}$-rich in $A \times A$.

The way that this is proved is as follows: First, write the product

$$x_1 \cdots x_{2^Z} \ = \ (x_1 x_2)(x_3 x_4) \cdots (x_{2^Z - 1} x_{2^Z}).$$

Then, applying the claim to the pair $x_1 x_2$, and then $x_3 x_4$, and so on, we deduce that lots of variable choices make lines $y = x + \xi\alpha$, $\alpha \in \Theta_1 \cdots \Theta_1$ (2^{Z-1} copies here), rich in $A \times A$. Then, the claim is applied again to the products $\Theta_1\Theta_1$ (grouped in twos), leading to lines $y = x + \xi\beta$, $\beta \in \Theta_2 \cdots \Theta_2$ (2^{Z-2} copies here). Continuing, one reaches lines $y = x + \xi\gamma$, $\gamma \in \Theta_Z$, as claimed.

Combining this deduction with Claim 3, we deduce:

Claim 5. There are at least

$$N^{4^{t-2} - O_t(\delta)}$$

choices of variables $a_1, ..., a_{4^{t-2}} \in \Theta$ such that for $\theta = \theta(a_1, ..., a_{4^{t-2}}) \in \Theta_{t-2}$, the line

$$y \ = \ x \ + \ \xi\theta$$

is $n^{1-O_t(\delta)}$-rich in $A \times A$.

5.1.4 A growth lemma

Given a probability measure f supported on a finite set C, we let f^* denote a certain measure on $CC - CC$ given as follows:

$$f^*(x) \ := \sum_{c_1c_2 - c_3c_4 = x} f(c_1)f(c_2)f(c_3)f(c_4). \tag{16}$$

Lemma 6 *Suppose that C is a finite set of real numbers. Let f be a measure on C. Then,*

$$\max_x f^*(x) \ \ll \ (\max_x f(x))^{4/3}(\log |C|)^2.$$

5.1.4.1 Proof of Lemma 6

Let

$$M \ := \ \max_x f(x).$$

We begin by partitioning the set C into the disjoint sets, some of which may be empty:

$$C \ = \ C_1 \cup C_2 \cup \cdots \cup C_k \cup C_0,$$

where for $i \geq 1$,

$$C_i \ := \ \{c \in C \ : \ f(c) \in (2^{-i}M, 2^{-i+1}M]\},$$

where C_0 is the remaining elements of C, and where

$$k \ = \ \lfloor 5 \log |C| / \log 2 \rfloor + 1.$$

We define

$$f^*_{\alpha,\beta,\gamma,\delta}(x) \ := \sum_{\substack{c_1 \in C_\alpha, c_2 \in C_\beta, c_3 \in C_\gamma, c_4 \in C_\delta \\ c_1c_2 - c_3c_4 = x}} f(c_1)f(c_2)f(c_3)f(c_4).$$

We have that

$$f^*(x) \ = \ \sum_{0 \leq \alpha,\beta,\gamma,\delta \leq k} f^*_{\alpha,\beta,\gamma,\delta}(x).$$

To prove the theorem, then, all we need to do is get bounds on these individual terms, and then sum them up.

First, we can easily bound the total contribution of the terms where any of the α, β, γ, or δ is 0: The contribution of all such terms is clearly bounded from above by

$$\ll \sum_{x \in CC-CC} M 2^{-5 \log |C| / \log 2} \ \ll \ |C|^{-1}.$$

Now we handle the other terms. First, suppose that $1 \leq \alpha, \beta, \gamma, \delta \leq k$. Then, one easily sees from the fact f is a probability measure that

$$|C_i| \ \ll \ 2^i M^{-1}, \ i = \alpha, \beta, \gamma, \delta.$$

The size of $f^*_{\alpha,\beta,\gamma,\delta}(x)$ is

$$\ll \ M^4 2^{-\alpha-\beta-\gamma-\delta} |\{a \in C_\alpha, b \in C_\beta, c \in C_\gamma, d \in C_\delta \ : \ ab - cd = x\}|. \qquad (17)$$

To bound this last factor from above, we will apply Elekes's [9] idea of using the Szemerédi-Trotter incidence theorem [30] to prove sum-product inequalities. We begin with the Szemerédi-Trotter theorem:

Theorem 5.2 *Suppose that one has N points and L lines in the plane. Then, the number of incidences is bounded from above by*

$$O((NL)^{2/3} + N + L).$$

The way we apply this theorem is as follows: Consider the family of lines

$$ax + cy \ = \ z, \text{ where } a \in C_\alpha, \ c \in C_\gamma.$$

Note that there are $|C_\alpha| \cdot |C_\gamma|$ lines in total.

Each of these lines intersects the grid $C_\beta \times C_\delta$ in some number of points (or perhaps no points at all). The total number of incidences $(x, y) \in C_\beta \times C_\delta$ is the right-most factor of (17). From the Szemerédi-Trotter theorem, this number is

$$\begin{aligned} &\ll (|C_\alpha| \cdot |C_\beta| \cdot |C_\gamma| \cdot |C_\delta|)^{2/3} + |C_\beta| \cdot |C_\delta| + |C_\alpha| \cdot |C_\gamma| \\ &\ll 2^{2(\alpha+\beta+\gamma+\delta)/3} M^{-8/3} + 2^{\beta+\delta} M^{-2} + 2^{\alpha+\gamma} M^{-2}. \end{aligned}$$

The total weight $f(a)f(x)f(c)f(y)$ that each such representation $ax + cy = z$ gets is

$$\ll 2^{-\alpha-\beta-\gamma-\delta} M^4.$$

So,

$$f^*_{\alpha,\beta,\gamma,\delta}(z) \ll 2^{-(\alpha+\beta+\gamma+\delta)/3} M^{4/3} + 2^{-\alpha-\gamma} M^2 + 2^{-\beta-\delta} M^2.$$

It follows that for all $z \in CC - CC$,

$$f^*(z) \ll |C|^{-1} + M^{4/3} (\log |C|)^2 \ll M^{4/3} (\log |C|)^2.$$

The second inequality here comes from the fact that $M \geq |C|^{-1}$, which follows from the fact that f is a probability measure.

5.1.5 Continuation of the proof

We now define a sequence of functions by first letting

$$f_0(h) := \begin{cases} 1/N, & \text{if } h \in \Theta; \\ 0, & \text{if } h \notin \Theta. \end{cases}$$

(Note that f_0 is a probability measure.) Then, we inductively define

$$f_{i+1}(h) := f_i^*(h),$$

where f^* is as in (16). It is easy to see that these f_i are all also probability measures.

The connection between this function f and our sequence of Θ_i is as follows: For a given real number h we have that $f_j(h)$ is $|\Theta|^{-4^j}$ times the number of choices for

$$x_1, ..., x_{4^j} \ \in \ \Theta$$

such that

$$\theta \ = \ \theta(x_1, ..., x_{4^j}) \ \in \ \Theta_j$$

satisfies

$$\theta \ = \ h.$$

As will will see, the upper bound on $f_j(h)$ provided by Lemma 6 will produce for us a lower bound on the number of rich lines in our grid.

Now, Lemma 6 implies that for some constant $c > 0$, if

$$t \ \geq \ k \ := \ c \log(1/\varepsilon),$$

then for all h,

$$f^*_{t-2}(h) \ \leq \ 1/n^5$$

So, for each real number h, there are at most

$$n^{-5}|\Theta|^{4^{t-2}}$$

choices for $x_1, ..., x_{4^{t-2}} \in \Theta$ such that $\theta = \theta(x_1, ..., x_{4^{t-2}})$ equals h. Combining this with Claim 5, we quickly deduce that there are $n^{5-O_t(\delta)}$ distinct values of θ among these rich lines (of Claim 5). If $\delta > 0$ is small enough relative to ε, then we will see that this number exceeds n^4.

We have now reached a contradiction, since there can be at most n^4 lines that hit an $n \times n$ grid in at least two points each. Our theorem is now proved.

CHAPTER VI

ON BIVARIATE SET FUNCTIONS AND EXPONENTIAL EXPANSION

Recall that a bivariate polynomial f over $\mathbb{R}[x, y]$ is a *set-expander* if there exists an $\epsilon > 0$ such that

$$|f(A, A)| \gg |A|^{1+\epsilon}$$

for all finite sets $A \subset \mathbb{R}$.

In this chapter we will prove the following theorems.

Theorem 6.1 *A bivariate polynomial of degree* 2 *over* $\mathbb{R}[x, y]$ *is not a set expander if and only if it is expressible in the form*

$$(i) f(x, y) = g(x) + c \cdot g(y),$$

where g is a quadratic,

$$(ii) f(x, y) = a(x + r)(y + r) + c,$$

or

$$(iii) f(x, y) = g(x + ry),$$

where g is quadratic, for some $a, c, r \in \mathbb{R}$.

Theorem 6.2 *A bivariate polynomial of degree* 3 *over* $\mathbb{R}[x, y]$ *is not a set expander if and only if it is expressible in the form*

$$(i) f(x, y) = g(x) + c \cdot g(y),$$

where g is a degree 3 polynomial,

$$(ii) f(x,y) = a(x+r)^2(y+r) + c,$$

$$(iii) f(x,y) = a(x+r)(y+r)^2 + c,$$

or

$$(iv) f(x,y) = g(x+ry),$$

where g is a degree 3 polynomial, for some $a, c, r \in \mathbb{R}$.

Our general strategy will be to assume that $f|(A,A)| \leq |A|^{1+\epsilon}$. Then we will see that unless f is as prescribed in our hypotheses , we can produce large sets of 'rich' curves and apply an incidence theorem to induce a lower bound on ϵ.

6.1 Tools

Proposition 1 *A bivariate set function* $f(x,y)$ *is a set expander if and only if*

$$f(ax+b, ay+b) + c$$

is a set expander, for $a, b, c \in \mathbb{R}, a \neq 0$.

Proof: The proposition is made clear by associating the set A with the set $A' = \frac{A-b}{a}$. So $f(a*A'+b, a*A'+b) = f(A,A)$ Translation by c obviously has no effect on the size of a set. ■

Just as Solymosi used the Szemerédi-Trotter theorem to establish sum-product estimates, we too will take advantage of incidence geometry. But for our purposes, we need an estimate of incidences between grid-points and curves, which was established by Székeley.

Theorem 6.3 *(Generalized Szemerédi-Trotter theorem)[29] Let* P *be a finite collection of points in* $\mathbb{R}^2$*, and let* L *be a finite collection of curves in* $\mathbb{R}^2$*. Suppose that*

any two curves in L intersect in at most α points, and any two points in P are simultaneously incident to at most β curves. Then

$$|\{(p,l) \in P \times L : p \in l\}| = O(\alpha^{1/3}\beta^{1/3}|P|^{2/3}|L|^{2/3} + |L| + \beta|P|).$$

We will be using the **GST theorem** in conjunction with the following simple counting argument to produce large sets of rich curves.

Proposition 2 *Let f be a bivariate function such that each element of $f(A,A)$ has $O(|A|)$ representations as $f(a,b),(a,b) \in A^2$. If $\epsilon > 0$ satisfies*

$$|f(A,A)| < |A|^{1+\epsilon},$$

then there are $\Omega(|A|^{2-\epsilon})$ curves

$$g_{c,d} : f(x,c) - f(y,d) = 0,$$

which each contain $\Omega(|A|^{1-\epsilon})$ solutions $(x,y) \in A^2$.

Proof: First we count Q, the number of quadruples $(x,c,y,d) \in A^4$ satisfying

$$f(x,c) = f(y,d).$$

Let $h(k)$ denote the number of pairs $(a,b) \in A^2$ for which $f(a,b) = k$. Also, define

$$g(k) = \begin{cases} 1 & \text{if } f(a,b) = 1 \text{ for some } (a,b) \in A^2 \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $h(k)g(k) = g(k)$ and $g(k)^2 = g(k)$. Now apply the Cauchy-Schwartz Inequality to see

$$\left(\sum_{k\in f(A,A)}(h(k)g(k)\right)^2 \leq \left(\sum_{k\in f(A,A)}h(k)^2\right)\left(\sum_{k\in f(A,A)}g(k)^2\right)$$

$$\Rightarrow \quad \left(\sum_{k\in f(A,A)}g(k)\right)^2 \leq (Q)\left(\sum_{k\in f(A,A)}g(k)\right)$$

$$\Rightarrow \quad \left(|A|^2\right)^2 \leq Q|f(A,A)|$$

$$\Rightarrow \quad Q \geq \frac{|A|^4}{|f(A,A)|}$$

$$\Rightarrow \quad Q \geq |A|^{3-\epsilon}.$$

Let K denote the number of pairs (c,d) for which $g_{c,d}$ has at least $|A|^{1-\epsilon}/2$ solutions $(x,y) \in A^2$. Each (c,d) counted by K trivially contributes at most $M|A|$ solutions (x,y) (for some constant M dependent on f), while the remaining pairs (c,d) contribute at most $|A|^{1-\epsilon}/2$ solutions (x,y). Thus a trivial upper bound on the number of solutions to $f(x,c) = f(y,d)$ leads to the following inequalities:

$$K \cdot M|A| + (|A|^2 - K)|A|^{1-\epsilon}/2 \geq |A|^{3-\epsilon}$$

$$\Rightarrow \quad K(M|A| - |A|^{1-\epsilon}/2) \geq |A|^{3-\epsilon}/2$$

$$\Rightarrow \quad K \geq \frac{|A|^{3-\epsilon}/2}{M|A| - |A|^{1-\epsilon}}$$

$$\Rightarrow \quad K \geq (|A|^{2-\epsilon})/2M.$$

■

We can also use the **GST theorem** to give a slightly strengthened version of the result concerning concave (or convex) curves we introduced in the review of literature.

Corollary 1 *Let $A \in \mathbb{R}$ be a finite set and set I be an open interval containing A. If f is a continuous non-linear function which changes concavity only finitely many times on $\mathbb{R}$, then*

$$|A + f(A)| \gg |A|^{5/4}.$$

Proof: Consider the grid $G = (A + f(A)) \times (A + f(A))$, and the set of curves $C_{s,t} : \{(a + s, f(a) + t), a \in I\}$, for $s \in f(A)$ and $t \in A$. Now apply the Generalized Szemerédi-Trotter theorem. We have $|A + f(A)|^2$ grid points and $|A||f(A)|$ curves, each of which contains at least $|A|$ grid points. Since f only changes concavity a finite number of times, α and β are both absolute constants. Also because of our restrictions on f, we know $|f(A)| = \Theta(|A|)$. So we have

$$\begin{aligned} & |A||f(A)||A| \ll (|A + f(A)|^2)^{2/3}(|A||f(A))^{2/3} \\ \Rightarrow \quad & \|A|^3 \ll |A + f(A)|^{4/3}|A|^{4/3} \\ \Rightarrow \quad & |A + f(A)| \gg |A|^{5/4}. \end{aligned}$$

To make full use of our incidence bound estimates, we will need the following result from algebraic geometry.

Theorem 6.4 *(Bezout's Theorem) [21] Two algebraic curves of degree m and n intersect in at most mn points unless they have a common factor.*

■

Finally, we will need to use the following fact:

Lemma 7 *Bivariate linear functions over $\mathbb{R}[x, y]$ are not set-expanders. That is, for every $c \in \mathbb{R}$ and every $\epsilon > 0$, there exist arbitrarily large sets $A \subset \mathbb{R}$ such that*

$$|A + c \cdot A| \ll |A|^{1+\epsilon}.$$

Proof:Case 1: If c is rational, let $c = \frac{a}{b}$, where $gcd(a, b) = 1$. Then set

$$A = \{b, 2b, \ldots Nb\},$$

so that

$$A + c \cdot A = \{xb + ya : 1 \leq x, y \leq N\}.$$

It is clear that for all $\epsilon > 0$,

$$|A + c \cdot A| \leq (b + a)N \leq |A|^{1+\epsilon},$$

for A sufficiently large.

Case 2: Now assume c is algebraic and that its minimal polynomial over $\mathbb{Z}[x]$ with the smallest possible coefficients (in absolute value) is

$$m_0 x^k + m_1 x^{k-1} + \ldots m_{k-1} x + m_k$$

Now consider the set

$$A = \{x_0 + x_1 c + x_2 c^2 + \ldots x_{k-1} m_0 c^{k-1} : 1 \leq x_i, y_i \leq N\},$$

with size N^k. We can see that

$$\begin{aligned} A + c \cdot A &= \{x_0 + x_1 c + x_2 c^2 + \ldots x_{k-1} m_0 c^{k-1} \\ &\quad + y_0 c + y_1 c^2 + y_2 c^3 \ldots y_k m_0 k - 1 c^k : 1 \leq x_i, y_i \leq N\} \\ &= \{(x_0 - y_{k-1} m_k) + c(x_1 + y_0 - y_{k-1} m_{k-1}) + \ldots \\ &\quad c^{k-1}(x_{k-1} + y_{k-2} - y_{k-1} m_1) : 1 \leq x_i, y_i \leq N\}. \end{aligned}$$

It follows that for every $\epsilon > 0$

$$|A + cA| = O_c(N^k) = O_c|A| < |A|^{1+\epsilon},$$

for A sufficiently large.

Case 3: Lastly, assume c is transcendental. Then set

$$A = \{x_0 + x_1 c + \ldots x_{k-1} c^{k-1} : 1 \leq x_i \leq N\},$$

where k will be chosen later.

If any element of A had multiple representations of the above form, that would contradict the fact that c is transcendental. So $|A| = N^k$. And for any $\epsilon > 0$,

$$\begin{aligned} |A + cA| &= |\{x_0 + (x_1 + y_0)c + \ldots (x_{k-1} + y_{k-2})c^{k-1} + y_{k-1}c^k : 1 \leq x_i, y_i \leq N\} \\ &\leq (N)^{k+1}2^{k-1} \\ &= |A|N2^{k-1} \\ &< |A|N^2, \text{ for } k = \log_2 N \\ &< |A|^{1+\epsilon}, \end{aligned}$$

for N sufficiently large.

6.2 Proof of Theorem 6.1

Let us begin by assuming that a polynomial set function f has no crossterms. Then it can be written as

$$f(x, y) = g(x) + m(y).$$

Clearly, we can assume that neither g nor m contain a constant term, as translation by a constant will not change the size of any output set. Notice that if m is a scalar multiple of g, then we can set A to a pre-image (with respect to g) of an appropriate set prescribed in the proof of Lemma 7, and $|f(A, A)|$ will be small.

Now assume that m is not a multiple of g. So we have,

$$\begin{aligned} |f(A, A)| &= |g(A) + m(A)| \\ &= |A' + m(g^{-1}(A')|, \text{ where } A' = g(A) \\ &= |A' + h(A')|, \text{ where } h = mg^{-1}. \end{aligned}$$

Notice that if h was neither concave up nor concave down, then $h''(x) = 0$ and so $h'(x) = s$, where s is some constant. Yet we see

$$\begin{aligned} h'(x) &= \frac{f'(g^{-1}(x))}{g'(g^{-1}(x))} \\ \Rightarrow \quad f'(x) &= s \cdot g'(x) \\ \Rightarrow \quad f(x) &= s \cdot g(x) + t, \text{ where } s \text{ and } t \text{ are constants.} \end{aligned}$$

Since we assumed that neither h nor g contained a constant term, $t = 0$. And since we are considering f and g which are not scalar multiples of one another, it follows that h must have concavity. Because f and g are polynomials, h can only change concavity a finite number of times. So by Corollary 1 , we have

$$|f(A, A)| \gg |A|^{5/4}.$$

This takes care of form (i) for both Theorem 6.1 and 6.2.

Now we will handle the cases that have cross terms. The first, and most obvious application of Proposition 1 will be to assume, in each of our cases, that the leading term has coefficient 1 and that there is no constant term.

Case 1: $f(x, y) = xy + ax + by$

First notice that

$$\begin{aligned} f(x-b, y-b) + ab &= (x-b)(y-b) + a(x-b) + b(y-b) + ab \\ &= xy - bx - by + b^2 + ax + by - ab - b^2 + ab \\ &= x(y + a - b), \end{aligned}$$

so by appropriate use of Proposition 1, we can reduce case 1 to the case $f(x, y) = x(y + r)$. When $r = 0$, f corresponds to the form (ii), and by setting A equal to a geometric progression, f is easily seen not to be a set expander. Now let $r \neq 0$.

Without loss of generality, we can assume that A is a positive set and r is positive. So

$$\begin{aligned}
|A(A+r)| &= |\ln(A(A+r))| \\
&= |\ln(A) + \ln(A+r)| \\
&= |A' + \ln(e^{A'} + r)|, \text{ where } A' = \ln(A) \\
&= |A' + h(A')|, \text{ where } h(x) = \ln(e^x + r) \\
&\gg |A'|^{5/4} \\
&\gg |A|^{5/4},
\end{aligned}$$

by Corollary 1, since $h(x)$ is concave-up for $x > 0$.

Case 2: $f(x,y) = x^2 + a_1xy + a_2y^2 + a_3x + a_4y, a \neq 0$

Now assume that $|f(A,A)| \leq |A|^{1+\epsilon}$. Then by Proposition 2, there exists at least $\Omega(|A|^{2-\epsilon})$ pairs (c,d) for which

$$f(x,c) = f(y,d)$$

has at least $|A|^{1-\epsilon}/2$ solutions $(x,y) \in A^2$. Put another way, we have at least $\Omega(|A|^{2-\varepsilon})$ curves

$$\begin{aligned}
g_{c,d} \quad &: \quad f(x,c) - f(y,d) = 0 \\
&\Rightarrow \quad (x^2 - y^2) + a_1(cx - dy) + a_2(c^2 - d^2) + a_3(x - y) + a_4(c - d),
\end{aligned}$$

each of which intersect the grid $A \times A$ in at least $|A|^{1-\epsilon}/2$ points.

Before we can use the GST theorem, we need to be able to bound the number of intersections between any two curves. Fortunately, Bezout's theorem tells us that no two of these curves can intersect in more than $2^2 = 4$ points unless they share

a common factor. So those curves which share highe than 4 intersections must be reducible. If $g_{c,d}$ is reducible, it must be factorable into the form

$$(x + y + b_1)(x - y + b_2) = 0.$$

Expanding and comparing coefficients gives us

$$\begin{aligned} b_1 + b_2 &= a_1 c + a_3 && (1) \\ b_2 - b_1 &= -a_1 d - a_3 && (2) \\ b_1 b_2 &= a_2(c^2 - d^2) + a_4(c - d). && (3) \end{aligned}$$

Solving for b_1 and b_2 and plugging them into (3) gives us

$$\begin{aligned} & \frac{a_1^2}{4}(c^2 - d^2) + \frac{a_1 a_3 (c - d)}{2} = a_2(c^2 - d^2) + a_4(c - d) && (4) \\ \Rightarrow \quad & (\frac{a_1^2}{4} - a_2)(c^2 - d^2) + (\frac{a_1 a_3 (c - d)}{2} - a_4)(c - d) = 0. && (5) \end{aligned}$$

The only way that (5) can have $\omega(|A|)$ solutions (c, d) is if $a_2 = (a_1/2)^2$ and $\frac{a_1 a_3}{2} = a_4$. This reduces case 2 into form (iii), which does not produce an expander, by Lemma 7 . Otherwise, (5) will have only $\emptyset(|A|)$ solutions, giving us only $\emptyset(|A|)$ reducible curves. That still leaves us with $\Omega(|A|^{2-\epsilon}) - O(|A|) = \Omega(|A|^{2-\epsilon})$ irreducible curves to work with, and we call that set of curves L, while P is just the set of the $|A|^2$ points from the grid $A \times A$. We bound the number of incidences $I(P, L)$ above by the GST theorem and below by the simple count that each of $\Omega(|A|^{2-\epsilon})$ curves has at least $\Omega(A^{1-\epsilon})$ points on the grid. That is to say:

$$|A|^{2-\varepsilon}|A|^{1-\varepsilon} \ll I(P, L) \ll (|A|^2)^{2/3}(|A|^2)^{2/3}.$$

Clearly, then $\epsilon > 1/6$, and Theorem 6.1 is now proved

6.3 Proof of Theorem 6.2

While more complicated, this proof follows the same strategy as the proof of Theorem 6.1. First note that form (ii) and form (iii) are essentially the same (just replace $f(x, y)$ with $f(y, x)$). So in considering the case in which we have no x^3 and no y^3, we may assume that we *do* have an x^2y term, and by Proposition 1, we can assume its coefficient is 1.

Case 1: $f(x, y) = x^2y + a_1xy^2 + a_2xy + a_3x^2 + a_4y^2 + a_5x + a_6y$

Notice that the expansion of $f(x - a_3, y - a_3)$ contains no x^2 term. So by appropriate use of Proposition 1, we can cancel out the x^2 term, leaving us with

$$x^2y + a_1xy^2 + a_2xy + a_3y^2 + a_4x + a_5y.$$

So our curves

$$g_{c,d} : f(x, c) - f(y, d) = 0$$

take the form

$$(x^2c - y^2d) + a_1(xc^2 - yd^2) + a_2(xc - yd) + a_3(c^2 - d^2) + a_4(x - y) + a_5(c - d) = 0$$

If these equations factor, they must take the form

$$(x + b_1y + b_2)(cx + b_3y + b_4) = 0.$$

Expanding and matching coefficients leads to the following system of equations:

$$\begin{aligned}
xy &: cb_1 + b_3 = 0 && (6)\\
y^2 &: b_1b_3 = -d && (7)\\
x &: b_4 + cb_2 = (a_1c^2 + a_2c + a_4) && (8)\\
y &: b_2b_3 + b_1b_4 = -(a_1d^2 + a_2d + a_4) && (9)\\
constant &: b_2b_4 = a_3(c^2 - d^2) + a_5(c - d) && (10)
\end{aligned}$$

Equations (6) and (7) imply that in order for $g_{c,d}$ to be factorable, we must have $cd > 0$, in which case

$$\begin{aligned} b_1 &= \sqrt{d/c} \\ b_3 &= -\sqrt{cd}. \end{aligned}$$

Plugging those values into equation (9), we see

$$\begin{aligned} -(\sqrt{cd})b_2 + (\sqrt{d/c})b_4 &= -(a_1d^2 + a_2d + a_4) \\ \Rightarrow -cb_2 + b_4 &= -(a_1d^2 + a_2d + a_4)\sqrt{c/d}. \end{aligned}$$

Combining this with (8) gives us

$$\begin{aligned} b_2 &= \frac{1}{2c}((a_1c^2 + a_2c + a_4) + (a_1d^2 + a_2d + a_4)\sqrt{c/d}) \\ b_4 &= \frac{1}{2c}((a_1c^2 + a_2c + a_4) - (a_1d^2 + a_2d + a_4)\sqrt{c/d}). \end{aligned}$$

Multiplying the two equations above give us

$$b_2b_4 = \frac{1}{4c}((a_1c^2 + a_2c + a_4)^2 - (a_1d^2 + a_2d + a_4)^2(c/d)).$$

But we already have a formula for b_2b_4 from (10), so we have

$$4a_3(c^3d - cd^3) + 4a_5(c^2d - cd^2) = (a_1c^2 + a_2c + a_4)^2d - (a_1d^2 + a_2d + a_4)^2c.$$

The only way this equations could have $\omega(|A|)$ solutions $(c, d) \in A^2$ is if

$$a_1 = a_3 = a_4 = 0, (*)$$

and

$$a_2^2 = 4a_5(**).$$

In that case, we can greatly simplify our case into

$$\begin{aligned} f(x,y) &= x^2y + 2rxy + r^2y \\ &= (x+r)^2y. \end{aligned}$$

Notice the similarity to case 1 of Theorem 6.1. If $r = 0$, than f clearly does not exponentially expand $A = \{2^0, 2^1, \ldots 2^{n-1}\}$. This corresponds to the form (i). Now assume $r \neq 0$. Again, we can assume without loss of generality that A is positive and r is positive.

$$\begin{aligned} |A^2(A+r)| &= |\ln(A^2(A+r))| \\ &= |2\ln(A) + \ln(A+r)| \\ &= |A' + \ln(e^{A'/2} + r)|, \text{, where } A' = 2\ln(A) \\ &= |A' + h(A')|, \text{, where } h(x) = \ln(e^{x/2} + r) \\ &\geq c|A'|^{5/4} \\ &= c|A|^{5/4}, \end{aligned}$$

by Corollary 1, since h is concave up for $x > 0$. It follows that f is a set expander unless $r = 0$. On the other hand, if either $(*)$ or $(**)$ is not satisfied, than we proceed as in case 2 of Theorem 6.1. Because we assume $|f(A,A)| < |A|^{1+\epsilon}$, we have, $\Omega(|A|^{2-\epsilon}) - O(|A|) = \Omega(|A|^{2-\epsilon})$ irreducible curves, L, each of which intersect $\Omega(|A|^{1-\epsilon})$ grid-points $P \in A \times A$. By the **GST theorem**,

$$|A|^{2-\varepsilon}|A|^{1-\varepsilon} \ll I(P,L) \ll (|A|^2)^{2/3}(|A|^2)^{2/3}.$$

Cleary, $\epsilon > 1/6$, and this case is completed.

Next we will consider the cases that include an x^3 or y^3 term. Again, because of symmetry and Proposition 1, we can assume that f has an x^3 term with coefficient 1.

Case 2: $x^3 + a_1x^2y + a_2xy^2 + a_3y^3 + a_4x^2 + a_5xy + a_6y^2 + a_7x + a_8y$, where at least one of a_1, a_2, a_4 is non-zero.

Our curves $g_{c,d} : f(x,c) - f(y,d)$ can be written as

$$\begin{aligned} & (x^3 - y^3) + a_1(cx^2 - dy^2) + a_2(c^2x - d^2y) + a_3(c^3 - d^3) \\ + \quad & a_4(x^2 - y^2) + a_5(cx - dy) + a_6(c^2 - d^2) + a_7(x - y) + a_8(c - d) = 0. \end{aligned}$$

By exhausting a few possibilities, we see that if $g_{c,d}$ does factor, it must factor into the form:

$$(x - y + b_1)(x^2 + xy + y^2 + b_2x + b_3y + b_4) = 0.$$

Expanding and matching coefficients gives us

$$x^2 \quad : \quad b_1 + b_2 = a_1c + a_4 \tag{11}$$

$$y^2 \quad : \quad b_1 - b_3 = a_1d - a_4 \tag{12}$$

$$xy \quad : \quad b_1 + b_3 - b_2 = 0 \tag{13}$$

$$x \quad : \quad b_1b_2 + b_4 = (a_2c^2 + a_5c) \tag{14}$$

$$y \quad : \quad b_1b_3 - b_4 = -(a_2d^2 + a_5d) \tag{15}$$

$$constant \quad : \quad b_2b_4 = a_3(c^3 - d^3) + a_6(c^2 - d^2) + a_7(c - d). \tag{16}$$

Combining (11), (12), and (13) gives us

$$b_1 = \frac{a_1(c+d)}{3} \tag{17}$$

$$b_2 = \frac{a_1(2c+d)+3a_4}{3} \tag{18}$$

$$b_3 = \frac{a_1(c+2d)+3a_4}{3}. \tag{19}$$

Now add (14) to (15) and substitute (11), (12), and (13) to the left hand side to get

$$\frac{{a_1}^2}{3}(c^2-d^2)+\frac{2a_1a_4}{3}(c-d) = a_2(c^2-d^2)+a_3(c-d).$$

Next, subtract (15) from (14) to solve for b_4, and plug that formula (along with (18)) into (16) to get

$$\begin{aligned} & (\frac{a_1a_2}{6}-\frac{a_1^3}{54})(c^3-d^3)+\frac{a_1a_5}{6}(c^2-d^2) \\ + \quad & \frac{a_1a_7}{3}(c-d)+(\frac{{a_1}^3}{18}-\frac{a_1a_2}{6})c^2d+(\frac{a_1a_2}{6}-\frac{{a_1}^3}{18})cd^2 \\ = \quad & a_3(c^3-d^3)+a_6(c^2-d^2)+a_8(c-d). \end{aligned}$$

The equation $g_{c,d}$ is reducible if and only if $(c,d) \in A^2$ simultaneously satisfies both of the above two equations. Thus it is clear that $g_{c,d}$ is reducible for $\omega(|A|)$ pairs (c,d) only if the following conditions on f are met:

$$\begin{aligned} \frac{a_1^2}{3} &= a_2, \\ \frac{a_1a_5}{6} &= a_6, \\ \frac{a_1a_7}{3} &= a8, \\ \frac{a_1a_2}{9} &= a_3, \text{ and} \\ \frac{2a_1a_4}{3} &= a_5. \end{aligned}$$

In this case,

$$\begin{aligned} f(x,y) &= (x^3 + 3rx^2y + 3r^2xy^2 + r^3y^3) + m(x^2 + 2rxy + r^2y^2) + s(x + ry) \\ &= (x + ry)^3 + m(x + ry)^2 + s(x + ry), \end{aligned}$$

which, by one more application of Proposition 1, is equivalent to form (iv). If case 2 can not be reduced to that form, than $g_{c,d}$ is reducible for only $O(|A|)$ pairs (c, d), and by the same argument seen in previous cases, we have

$$|f(A, A)| > |A|^{1+1/6}.$$

And our theorem is proved.

We believe that the theorem could be generalized in the obvious way.

Conjecture 6.5 *A bivariate polynomial over $\mathbb{R}[x, y]$ is not a set expander if and only if it is of the form*

$$(i) f(x, y) = g(x) + c \cdot g(y),$$

where g is a polynomial,

$$(ii) f(x, y) = (x + r)^a (y + r)^b + c, \text{ where } a, b \in \mathbb{Z}^+,$$

or

$$(iii) f(x, y) = g(x + ry),$$

where g is a polynomial.

In order to prove this conjecture, we need to find a way to bypass the factorization methods we used in our proofs.

REFERENCES

[1] A. Balog and E. Szemerédi, *A statistical theorem of set addition*, Combinatorica **14** (1994), 263-268.

[2] A. Balog, *Many additive quadruples*, CRM Proceedings and Lecture Notes in Additive Combinatorics, **43** (2007).

[3] F. Behrend, *On the sets of integers which contain no three in arithmetic progression*, Proc. Nat. Acad. Sci.,**23** (1946),331-332.

[4] J. Beck, *On the lattice property of the plane and some problems of Dirac, Motzkin, and Erdős*, Combinatorica, **3-4** 1983 281-297.

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

[6] K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl. *Combinatorial complexity bounds for arrangements of curves and surfaces*, Discrete and Computational Geometry, **5** (1990) 99-106.

[7] E. Borenstein and E. Croot, *On a certain generalization of the Balog-Szemerédi-Gowers theorem*, submitted.

[8] E. Borenstein and E. Croot, *On rich lines in grids*, submitted

[9] G. Elekes, *On the number of sums and products*, Acta Arith. **81** (1997), 365-367.

[10] ——-, *Sums versus products in number theory, algebra and Erdős geometry*, Paul Erdős and his mathematics, II (Budapest, 1999), Bolyai Soc. Math. Stud., **11** János Bolyai Math. Soc., (2002), 241-290.

[11] G. Elekes, M. Nathanson, and I. Ruzsa. *Convexity and sumsets*, Journal of Number Theory, **83** (1999) 194-201.

[12] G. Elekes and E. Szabó. *Triple points of circle grids*, (to appear in Combinatorica)

[13] P. Erdős and E. Szemerédi, *On sums and products of integers*, Studies in pure mathematics, 213-218, Birkhäuser, Basel, 1983.

[14] P. Erdős, Turan, *On some sequences of integers*, J. London Math. Society, **11** (1936), 261-264.

[15] K. Ford, *Sums and products from a finite set of real numbers*, Ramanujan Journal **2** (1998), 1-2.

[16] G. A. Freiman, *Structure theory of set addition*, Astrisque, **258** (1999), 133.

[17] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerdi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204-256.

[18] T. Gowers, *A new proof of Szemerédi's Theorem for progressions of length four*, GAFA **8** (1998), 529-551.

[19] T. Gowers, *A new proof of Szemerdi's theorem*, GAFA **11** (2001), 465-588.

[20] H. Helfgott, *Growth and generation in* $SL_2(\mathbb{Z}/p\mathbb{Z})$, Ann. of Math. **167** (2008), 601-623.

[21] Kirwan, Frances. *Complex Algebraic Curves*. United Kingdom: Cambridge University Press. 1992

[22] M. Nathanson, *On sums and products of integers*, Proc. Am Math. Soc. **125** (1997), 9-16.

[23] N.H. Nguyen and V. Vu, *Classification theorems for sumsets modulo a prime*, *submitted*.

[24] J. Pach and M. Sharir, *On the number if incidences between points and curves*, Combinatorics, Probability and Computing, **7** (1998) 121-127.

[25] I. Ruzsa, *Arithmetic Progressions and the number of sums*, Periodica Math. Hung. **33** (1992), 105-111.

[26] A. Sárkőzy, *On difference sets of sequences of integers*, I., Acta Math. Acad. Sci. Hungar. **31** (1978), 125-149.

[27] J. Solymosi, *On sumsets and products sets of complex numbers*, *preprint*.

[28] B. Sudakov, E. Szeméredi, and V. Vu, *On a question of Erdős and Moser*, Duke Math Jour. **129** (2005), 129-155.

[29] L. Székeley, *Crossing numbers and hard Erd̋os problems in disrete geometry*, Combinatorics,Probability, and Computing, **6,3** (1997) 353-358.

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

[31] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Univ. Press, 2006.

[32] C. Tóth, *The Szemerédi-Trotter theorem in the complex plane*, *submitted*.

[33] B.L. van der Waerden, *Beweis Einer Baudetschen Vermutung*, Nieuw. Arch. Wisk. **15** (1927) 212-216.

[34] V. Vu, *Sum-product estimates via directed expanders*, Math. Res. Lett. **15** (2008), no. 2, 375-388.