Johann Radon Institute for Computational and Applied Mathematics Austrian Academy of Sciences (ÖAW)

ÖAW AUSTRIAN ACADEMY OF SCIENCES

RICAM
JOHANN · RADON · INSTITUTE
FOR COMPUTATIONAL AND APPLIED MATHEMATICS

A family of four-variable expanders with quadratic growth

M. Makhul

RICAM-Report 2018-13

A FAMILY OF FOUR-VARIABLE EXPANDERS WITH QUADRATIC GROWTH

MEHDI MAKHUL*

ABSTRACT. We prove that if $g(x, y)$ is a polynomial of constant degree d that $y_2 - y_1$ does not divide $g(x_1, y_1) - g(x_2, y_2)$, then for any finite set $A \subset \mathbb{R}$

$$|X| \gg_d |A|^2, \quad \text{where } X := \left\{ \frac{g(a_1, b_1) - g(a_2, b_2)}{b_2 - b_1} : a_1, a_2, b_1, b_2 \in A \right\}.$$

We will see this bound is also tight for some polynomial $g(x, y)$.

1. INTRODUCTION

Throughout this paper, when we write $X \gg Y$, this means that $X \geq cY$, for some absolute constant $c > 0$.

The sum set of a subset $A \subset \mathbb{R}$ is defined as $A + A := \{a + b : a, b \in A\}$. The product set is defined in a similar way, $AA := \{ab : a, b \in A\}$.

The Erdős-Szemerédi [ES83] conjecture states that, for all $\epsilon > 0$ and for any finite set $A \subset \mathbb{N}$,

$$\max\{|A + A|, |AA|\} \geq c(\epsilon)|A|^{2-\epsilon}.$$

It is natural to extend this conjecture for other settings (such as $\mathbb{R}$), and also to change the polynomials $F(x, y) = x + y$ and $F(x, y) = xy$ defining the sum and product sets to other polynomials or rational functions. In recent years much research has been done in this direction.

For many such functions, the images of sets are known to always grow. For example, the authors of [MRNS15] have studied several multivariable polynomials, including the function

$$G(x_1, x_2, x_3, x_4, x_5) = x_1(x_2 + x_3 + x_4 + x_5).$$

More precisely they showed that, for any finite set $A \subset \mathbb{R}$,

$$|A(A + A + A + A)| \gg \frac{|A|^2}{\log |A|},$$

where $A(A + A + A + A) := \{x_1(x_2 + x_3 + x_4 + x_5) : x_i \in A\}$.

In [MRNS17], the authors studied a more complicated function, namely

$$H(x_1, x_2, x_3, x_4, x_5) = (x_1 + x_2 + x_3 + x_4)^2 + \log x_5.$$

They showed that, for any finite $A \subset \mathbb{R}$,

$$|\{(a_1 + a_2 + a_3 + a_4)^2 + \log a_5 : a_i \in A\}| \gg \frac{|A|^2}{\log |A|}.$$

In the same circle of ideas, [BRN15] investigated the rational function $F(x_1, x_2, x_3, x_4) = \frac{x_1 + x_2}{x_3 + x_4}$, showing that for any finite set $A \subset \mathbb{R}$, we have

$$|F(A, A, A, A)| \geq 2|A|^2 - 1.$$

Our result is a generalization of the method of [MRNS15, Corollary 3.1], where they used the Szemerédi-Trotter Theorem to prove that for any finite set $A \subset \mathbb{R}$:

$$\left| \frac{A - A}{A - A} \right| \gg |A|^2.$$

A stronger version of this result, with a multiplicative constant 1, follows from an earlier geometric result of Ungar [Ung82].

In this article we consider a certain class of rational functions of four variables. Suppose that $g(x, y)$ is a polynomial of two variables of degree d. Let

$$F(x_1, x_2, y_1, y_2) = \frac{g(x_1, y_1) - g(x_2, y_2)}{y_2 - y_1}$$

be a four-variable rational function in terms of x_1, x_2, y_1, y_2. The main theorem of this paper is the following result concerning the growth of F.

Theorem 1.1. *Suppose that $g(x, y)$ is a polynomial of degree d, that $y_2 - y_1$ does not divide $g(x_1, y_1) - g(x_2, y_2)$, and that $A \subset \mathbb{R}$ is a finite set. Then*

$$|X| \gg_d |A|^2, \quad \text{where } X := \left\{ \frac{g(a_1, b_1) - g(a_2, b_2)}{b_2 - b_1} : a_1, a_2, b_1, b_2 \in A \right\}.$$

Notice that the following example shows that the condition that the denominator cannot be a divisor of the numerator is necessary.

Example 1.2. *Suppose that $g(x, y) = y^2$ and $A = \{1, 2, \dots n\}$. Then*

$$X = \left\{ \frac{b_1^2 - b_2^2}{b_2 - b_1} : b_1, b_2 \in A \right\}$$

equals $-\{b_2 + b_1 : b_i \in A\}$ *and has cardinality* $O(n)$.

Furthermore, notice that the condition rules out a degenerate case where the polynomial $g(x, y)$ does not depend on x.

On the other hand, it is known that for some polynomials g, the result of Theorem 1.1 is tight. For example, if we define $g(x_1, y_1) = x_1$ then Theorem 1.1 recovers the result of [MRNS15] and [Ung82]. This is known to be tight, since for the set $A = \{1, \dots, N\}$,

$$\left| \frac{A - A}{A - A} \right| = O(N^2).$$

However, we are not aware of any other polynomials g for which the bound in Theorem 1.1 is tight, and whether or not the bound can be improved for some particular g is an interesting question.

Our main result has some similarities with a result of Raz, Sharir and Solymosi [RSS15] concerning the growth of two variable polynomials. Their result states that, if F is a two variable polynomial with bounded degree, then for any $A, B \subset \mathbb{R}$ with $|A| = |B| = n$,

$$|F(A, B)| \gg_d n^{4/3},$$

provided that F satisfies a non-degeneracy condition. This condition states that F cannot be of one of the following forms

(1) $F(u, v) = f(g(u) + h(v))$,
(2) $F(u, v) = f(g(u) \cdot h(v))$.

This result gave an improvement upon an earlier result of Elekes and Ronyai [ER00].

1.1. **The Szemerédi-Trotter Theorem.** The essential ingredient used to prove our result is a corollary of the *Szemerédi-Trotter* Theorem [ST83], which gives a bound for the number of lines in the plane containing at least a fixed number of points k from a given finite set, that is, the number of *k-rich* lines.

Theorem 1.3 (Szemerédi-Trotter)**.** *Suppose that P is a set of n points and $\mathcal{L}$ is a set of m lines in $\mathbb{R}^2$. Then*

$$\mathcal{I}(P, \mathcal{L}) \ll n^{\frac{2}{3}} m^{\frac{2}{3}} + n + m. \tag{1}$$

Corollary 1.4. *Let $k, n \geq 2$ be natural numbers and fix $d \in \mathbb{N}$ such that $8d \leq k \leq d\sqrt{n}$. Let $\mathcal{L}$ be a set of n lines in the plane, and let $t_{\geq k}$ denote the number of points in the plane contained in at least k elements of $\mathcal{L}$, where each line appears with multiplicity at most d. Then*

$$t_{\geq k} = O_d\left(\frac{n^2}{k^3}\right).$$

2. Main Results

Suppose that $A, B \subset \mathbb{R}$ are finite, and $g(x_1, y_1)$ is a polynomial of degree d. We associate an element of $A \times B$ with a line via

$$A \times B \ni (a, b) \longleftrightarrow l_{a,b} : y = bx - g(a, b).$$

Consider $\mathcal{L} = \{\ell_{a,b} : a, b \in A \times B\}$ as a multi-set. Then $\mathcal{L}$ is a set of $|A||B|$ lines, such that each line appears at most d times. We also define the quantity

$$n(x, y) = \left|\{(a, b) \in A \times B : (x, y) \in l_{a,b}\}\right|,$$

which is interpreted geometrically as the number of lines of $\mathcal{L}$ that pass through (x, y).

Lemma 2.1. *Suppose that $d \in \mathbb{N}$ is fixed. Suppose that $A, B, X \subset \mathbb{R}$ are finite and satisfy $|X| \leq \frac{|A||B|}{4d^2}$, with $0 \notin X$. Then*

$$\sum_{x \in X} \sum_y n^2(x, y) \ll |A|^{\frac{3}{2}} |B|^{\frac{3}{2}} |X|^{\frac{1}{2}}. \tag{2}$$

Proof. The amount of t-rich points is given by

$$R_t := \left\{(x, y) \in \mathbb{R}^2 : n(x, y) \geq t\right\}.$$

We first show that

$$|R_t| \ll \frac{|A|^2 |B|^2}{t^3}.$$

First, we bound $n(x, y)$ for a given point (x, y). For fixed $b_0 \in B$ we obtain a line with slope b_0 passing through (x, y) and a one variable polynomial equation $g(a, b_0)$. Since each line is determined uniquely, by its slope and one point on it

(for fixed b_o and (x, y) the equation $g(a, b_0) = 0$ has at most d distinct solutions), we have

$$n(x, y) \leq d|B|.$$

With a similar argument for fixed $a \in A$ we obtain a univariate polynomial equation. Since each line is determined uniquely by its y-intercept and one point on it we have:

$$n(x, y) \leq d|A|.$$

These together imply:

$$n(x, y) \leq d(\min\{|A|, |B|\}) \leq (d|A|d|B|)^{\frac{1}{2}} = d|\mathcal{L}|^{\frac{1}{2}}.$$

This implies there are no points incident to more than $d\sqrt{|\mathcal{L}|}$ lines in $\mathcal{L}$, and by applying Corollary 1.4 we get:

$$|R_t| \ll \frac{|\mathcal{L}|^2}{t^3} \leq \frac{|A|^2|B|^2}{t^3}.$$

Let $\Delta > 2d$ be an integer to be specified later. We have

$$(3) \qquad \sum_{x \in X} \sum_{y} n^2(x, y) \leq \sum_{x \in X} \sum_{n(x,y) \leq \Delta} n^2(x, y) + \sum_{\substack{(x,y) \\ n(x,y) > \Delta}} n^2(x, y).$$

The first term is bounded by $\Delta |A||B||X|$, in fact

$$(4) \quad \sum_{x \in X} \sum_{n(x,y) \leq \Delta} n^2(x, y) \leq \Delta \sum_{x \in X} \sum_{y} n(x, y) = \Delta |A||B| \sum_{x \in X} 1 = \Delta |A||B||X|.$$

For second term we have:

$$
\begin{aligned}
(5) \qquad \sum_{\substack{(x,y) \\ n(x,y) > \Delta}} n^2(x, y) &= \sum_{j \geq 1} \sum_{2^{j-1}\Delta \leq n(x,y) \leq 2^j \Delta} n^2(x, y) \ll \\
&\ll \sum_{j \geq 1} \frac{|A|^2|B|^2}{(2^j \Delta)^3} \cdot (2^j \Delta)^2 = \frac{|A|^2|B|^2}{\Delta} \sum_{j \geq 1} \frac{1}{2^j} = \frac{|A|^2|B|^2}{\Delta}.
\end{aligned}
$$

For an optimal choice, set the parameter $\Delta = \left\lceil \frac{(|A||B|)^{1/2}}{|X|^{1/2}} \right\rceil > 2d$. Combining the bounds from (3) and (4) and (5), it follows that

$$\sum_x \sum_y n^2(x, y) \ll |A|^{\frac{3}{2}}|B|^{\frac{3}{2}}|X|^{\frac{1}{2}}. \qquad\qquad \square$$

Proof of Theorem 1.1. Consider:

$$|X| = \left| \left\{ (x, a_1, a_2, b_1, b_2) : x = \frac{g(a_1, b_1) - g(a_2, b_2)}{b_1 - b_2}, a_i, b_i \in A \right\} \right|$$

$$= \left| \left\{ (x, a_1, a_2, b_1, b_2) : b_1 x - g(a_1, b_1) = b_2 x - g(a_2, b_2) \right\} \right| =$$

$$\sum_x \sum_y n^2(x, y) \ll |A|^3 |X|^{\frac{1}{2}}.$$

On the other hand, $|X| \geq |A|^4$. Hence we obtain:

$$|A|^4 \ll |A|^3 |X|^{\frac{1}{2}}, \quad \text{hence} \quad |X| \gg |A|^2. \qquad\qquad \square$$

Corollary 2.2. *Suppose that $P = A \times A$ is a set of $|A|^2$ points. Let l be the y-axis. Suppose that $B(P)$ is the set of all bisectors determined by P. Then $|B \cap l| \gg |A|^2$.*

Proof. By a simple calculation we can see that the equation of the bisector determined by two points (x_1, y_1) and (x_2, y_2) in the s, t plane is:

$$s = \frac{2(x_1 - x_2)t + (x_2^2 - x_1^2) + (y_2^2 - y_1^2)}{2(y_2 - y_1)}.$$

Inserting $t = 0$, the hitting point has coordinate

$$\left(0, \frac{(x_2^2 - x_1^2) + (y_2^2 - y_1^2)}{2(y_2 - y_1)} \right).$$

Setting $g(x, y) = -2(x^2 - y^2)$, we obtain the result by Theorem 1.1. $\square$

As we mentioned, this bound is tight for some polynomials, for instance $g(x, y) = x$. However, we expect that if $F(x_1, x_2, y_1, y_2)$ is a generic rational function satisfying the condition of the Theorem 1.1 we have $|X| \gg |A|^3$.

References

[BRN15]   Antal Balog and Oliver Roche-Newton, *New sum-product estimates for real and complex numbers*, Discrete Comput. Geom. **53** (2015), no. 4, 825–846. MR 3341581

[ER00]   György Elekes and Lajos Rónyai, *A combinatorial problem on polynomials and rational functions*, J. Combin. Theory Ser. A **89** (2000), no. 1, 1–20. MR 1736139

[ES83]   Paul Erdös and Endre Szemerédi, *On sums and products of integers*, Studies in pure mathematics, Birkhäuser, Basel, 1983, pp. 213–218. MR 820223

[MRNS15]   Brendan Murphy, Oliver Roche-Newton, and Ilya Shkredov, *Variations on the sum-product problem*, SIAM J. Discrete Math. **29** (2015), no. 1, 514–540. MR 3323540

[MRNS17]   Brendan Murphy, Oliver Roche-Newton, and Ilya D. Shkredov, *Variations on the Sum-Product Problem II*, SIAM J. Discrete Math. **31** (2017), no. 3, 1878–1894. MR 3691216

[RSS15]   Orit E. Raz, Micha Sharir, and József Solymosi, *On triple intersections of three families of unit circles*, Discrete Comput. Geom. **54** (2015), no. 4, 930–953. MR 3416906

[ST83]   Endre Szemerédi and William T. Trotter, Jr., *Extremal problems in discrete geometry*, Combinatorica **3** (1983), no. 3-4, 381–392. MR 729791

[Ung82]   Peter Ungar, *2N noncollinear points determine at least 2N directions*, J. Combin. Theory Ser. A **33** (1982), no. 3, 343–347. MR 676751

(Mehdi Makhul) Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Linz and Research Institute for Symbolic Computation (RISC), Johannes Kepler University, Linz

E-mail address: mmakhul@risc.jku.at