A Szemerédi-type theorem for subsets of the unit cube

Vjekoslav Kovač (U. of Zagreb)

Joint work with *Polona Durcik* (Caltech)

Supported by *HRZZ* UIP-2017-05-4129 (MUNHANAP)

Georgia Tech, Atlanta, Jan 29, 2020

Ramsey theory (for analysts)

Searches for patterns in large arbitrary structures

discrete	vs.	**Euclidean**
Ramsey (1920s)		Erdős et al. (1973)
coloring theorems	vs.	**density theorems**
e.g., Van der Waerden's thm (1927)		e.g., Szemerédi's thm (1975)
non-arithmetic patterns	vs.	**arithmetic patterns**
e.g., spherical configurations		e.g., arithmetic progressions

This combination of choices: Cook, Magyar, and Pramanik (2017)

Szemerédi's theorem on $\mathbb{Z}$

A question posed by Erdős and Turán (1936)

For a positive integer $n \geq 3$ and a number $0 < \delta \leq 1/2$ is there a positive integer N such that each set $S \subseteq \{0, 1, 2, \ldots, N-1\}$ with at least δN elements must contain a nontrivial arithmetic progression of length n?

- $n = 3$ ✓ Roth (1953)
- $n = 4$ ✓ Szemerédi (1969)
- $n \geq 5$ ✓ Szemerédi (1975)

Szemerédi's theorem on $\mathbb{Z}$

Let $N(n, \delta)$ be the smallest such N

What are the best known bounds for $N(n, \delta)$?

- $N(3, \delta) \leq \exp(\delta^{-C})$ Heath-Brown (1987)
- $N(4, \delta) \leq \exp(\delta^{-C})$ Green and Tao (2017)
- $N(n, \delta) \leq \exp(\exp(\delta^{-C(n)}))$ Gowers (2001)

Szemerédi's theorem in $[0,1]^d$

Reformulation of Szemerédi's theorem with the above bounds

For $n \geq 3$ and $d \geq 1$ there exists a constant $C(n,d)$ such that for $0 < \delta \leq 1/2$ and a measurable set $A \subseteq [0,1]^d$ with $|A| \geq \delta$ one has

$$\int_{[0,1]^d} \int_{[0,1]^d} \prod_{i=0}^{n-1} \mathbb{1}_A(x+iy)\, \mathrm{d}y\, \mathrm{d}x \geq \begin{cases} \left(\exp(\delta^{-C(n,d)})\right)^{-1} & \text{when } 3 \leq n \leq 4 \\ \left(\exp(\exp(\delta^{-C(n,d)}))\right)^{-1} & \text{when } n \geq 5 \end{cases}$$

Gaps of progressions in $A \subseteq [0,1]^d$

What can be said about the following set?

$$\mathrm{gaps}_n(A) := \{y \in [-1,1]^d : (\exists x \in [0,1]^d)\big(x, x+y, \dots, x+(n-1)y \in A\big)\}$$

It contains a ball $\mathrm{B}(0,\varepsilon)$ for some $\varepsilon > 0$

A proof by Stromberg (1972):

- Assume that A is compact
- Find an open set U such that $U \supseteq A$ and $|U| \leq (1 + 1/2n)|A|$
- Take $\varepsilon := \mathrm{dist}(A, \mathbb{R}^d \setminus U)/n$
- For any $\|y\|_{\ell^2} < \varepsilon$: $x \in A \cap (A - y) \cap \cdots \cap (A - (n-1)y) \neq \emptyset$

Gaps of progressions in $A \subseteq [0,1]^d$

Number ε has to depend on “geometry” of A and not just on $|A|$

There is an obstruction already when we ask for much less

$\ell^2\text{-gaps}_n(A) :=$
$\big\{\lambda \in [0,\infty) : (\exists x,y)\big(x, x+y, \ldots, x+(n-1)y \in A,\ \|y\|_{\ell^2} = \lambda\big)\big\}$

Does $\ell^2\text{-gaps}_n(A)$ contain an interval of length depending only on n, d, and the measure $|A|$?

Gaps of progressions in $A \subseteq [0,1]^d$

No! Bourgain (1986):

$$A := \left\{ x \in [0,1]^d : (\exists m \in \mathbb{Z}) \big(m - 1/10 < \|\varepsilon^{-1} x\|_{\ell^2}^2 < m + 1/10 \big) \right\}.$$

- The parallelogram law:

$$\|x\|_{\ell^2}^2 - 2\|x+y\|_{\ell^2}^2 + \|x+2y\|_{\ell^2}^2 = 2\|y\|_{\ell^2}^2,$$

- $x, x+y, x+2y \in A \quad \Longrightarrow \quad m' - 2/5 < 2\|\varepsilon^{-1} y\|_{\ell^2}^2 < m' + 2/5$
- $|A| \gtrsim 1$ uniformly as $\varepsilon \to 0^+$

One can switch attention to other patterns (spherical configurations)
or ...

Gaps of 3-term progressions in other ℓ^p norms

For $p \in [1, \infty]$ define:

$$\ell^p\text{-gaps}_n(A) := \{\lambda \in [0, \infty) : (\exists x, y)\big(x, x+y, \ldots, x+(n-1)y \in A,\ \|y\|_{\ell^p} = \lambda\big)\}$$

Theorem (Cook, Magyar, and Pramanik (2017))

Take $n = 3$, $p \neq 1, 2, \infty$, $d \geq D(p)$, $\delta \in (0, 1/2]$, $A \subseteq [0, 1]^d$ measurable, $|A| \geq \delta$.
Then $\ell^p\text{-gaps}_3(A)$ contains an interval of length depending only on p, d, and δ

Gaps of longer progressions in other ℓ^p norms

Theorem (Durcik and K. (2020))

Take $n \geq 3$, $p \neq 1, 2, \ldots, n-1, \infty$, $d \geq D(n,p)$, $\delta \in (0, 1/2]$, $A \subseteq [0,1]^d$ measurable, $|A| \geq \delta$. Then $\ell^p\text{-gaps}_n(A)$ contains an interval of length at least

$$\begin{cases} \big(\exp(\exp(\delta^{-C(n,p,d)}))\big)^{-1} & \text{when } 3 \leq n \leq 4 \\ \big(\exp(\exp(\exp(\delta^{-C(n,p,d)})))\big)^{-1} & \text{when } n \geq 5 \end{cases}$$

Modifying Bourgain's example $\longrightarrow$ sharp regarding the values of p

One can take $D(n,p) = 2^{n+3}(n+p)$ $\longrightarrow$ certainly not sharp

Quantities that detect progressions

$$\sigma(x) = \delta(\|x\|_{\ell^p}^p - 1)$$

a measure supported on the unit sphere in the ℓ^p-norm

$$\mathcal{N}_\lambda^0(A) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \prod_{i=0}^{n-1} \mathbb{1}_A(x + iy)\, \mathrm{d}\sigma_\lambda(y)\, \mathrm{d}x$$

$$\mathcal{N}_\lambda^0(A) > 0 \implies (\exists x, y)\big(x, x+y, \ldots, x+(n-1)y \in A,\ \|y\|_{\ell^p} = \lambda\big)$$

$$\mathcal{N}_\lambda^\varepsilon(A) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \prod_{i=0}^{n-1} \mathbb{1}_A(x + iy)(\sigma_\lambda * \varphi_{\varepsilon\lambda})(y)\, \mathrm{d}y\, \mathrm{d}x$$

for a smooth $\varphi \geq 0$ with $\int_{\mathbb{R}^d} \varphi = 1$

Quantities that detect progressions

"The largeness/smoothness multiscale approach"

Essentially introduced by Cook, Magyar, and Pramanik (2017)

$$\mathcal{N}_{\lambda}^{\varepsilon}(A) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \prod_{i=0}^{n-1} \mathbb{1}_A(x + iy)(\sigma_\lambda * \varphi_{\varepsilon\lambda})(y) \, dy \, dx$$

$\lambda \in (0, \infty)$, scale of largeness $\longrightarrow$ detect APs with $\|y\|_{\ell^p} = \lambda$

$\varepsilon \in [0, 1]$, scale of smoothness $\longrightarrow$ the picture is blurred up to scale ε

Proof strategy

$$\mathcal{N}_\lambda^0(A) = \mathcal{N}_\lambda^1(A) + \left(\mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^1(A)\right) + \left(\mathcal{N}_\lambda^0(A) - \mathcal{N}_\lambda^\varepsilon(A)\right)$$

- $\mathcal{N}_\lambda^1(A) =$ the structured part ("the main term")
  Controlled uniformly in λ using Szemerédi's theorem (with the best known bounds) as a black box
- $\mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^1(A) =$ the error part
  Certain pigeonholing in λ is needed
  Leads to some multilinear singular integrals
- $\mathcal{N}_\lambda^0(A) - \mathcal{N}_\lambda^\varepsilon(A) =$ the uniform part
  Controlled uniformly in λ using Gowers uniformity norms
  Leads to some oscillatory integrals

Three propositions

Proposition handling the structured part

If $\lambda \in (0,1],\ \delta \in (0,1/2],\ A \subseteq [0,1]^d,\ |A| \geq \delta$, then

$$\mathcal{N}_\lambda^1(A) \geq \begin{cases} \big(\exp(\delta^{-E})\big)^{-1} & \text{when } 3 \leq n \leq 4 \\ \big(\exp(\exp(\delta^{-E}))\big)^{-1} & \text{when } n \geq 5 \end{cases}$$

This is essentially the analytical reformulation of Szemerédi's theorem

Three propositions

Proposition handling the error part

IF $J \in \mathbb{N}$, $\lambda_j \in (2^{-j}, 2^{-j+1}]$ for $j = 1, 2, \ldots, J$, $\varepsilon \in (0, 1/2]$, $A \subseteq [0, 1]^d$ measurable, then

$$\sum_{j=1}^{J} \left| \mathcal{N}^{\varepsilon}_{\lambda_j}(A) - \mathcal{N}^{1}_{\lambda_j}(A) \right| \leq \varepsilon^{-F} J^{1-2^{-n+2}}$$

Note the gain of $J^{-2^{-n+2}}$ over the trivial estimate for each fixed λ_j

Three propositions

Proposition handling the uniform part

If $d \geq D(n, p),\ \lambda, \varepsilon \in (0, 1],\ A \subseteq [0, 1]^d$ measurable, then

$$\left|\mathcal{N}_\lambda^{\,0}(A) - \mathcal{N}_\lambda^{\,\varepsilon}(A)\right| \leq G\varepsilon^{1/3}$$

Consequently, the uniform part can be made arbitrarily small by choosing a sufficiently small ε

How to finish the proof?

Denote

$$\vartheta := \begin{cases} \big(\exp(\delta^{-E})\big)^{-1} & \text{when } 3 \le n \le 4 \\ \big(\exp(\exp(\delta^{-E}))\big)^{-1} & \text{when } n \ge 5 \end{cases}$$

and choose

$$\varepsilon := \left(\frac{\vartheta}{3G}\right)^3, \quad J := \left\lfloor \left(3\vartheta^{-1}\varepsilon^{-F}\right)^{2^{n-2}} \right\rfloor + 1$$

Observe

$$2^{-J} \ge \begin{cases} \big(\exp(\exp(\delta^{-C(n,p,d)}))\big)^{-1} & \text{when } 3 \le n \le 4 \\ \big(\exp(\exp(\exp(\delta^{-C(n,p,d)})))\big)^{-1} & \text{when } n \ge 5 \end{cases}$$

How to finish the proof?

Take $A \subseteq [0,1]^d$ such that $|A| \geq \delta$

By pigeonholing we find $j \in \{1, 2, \dots, J\}$ such that for every $\lambda \in (2^{-j}, 2^{-j+1}]$ we have

$$\left|\mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^1(A)\right| \leq \varepsilon^{-F} J^{-2^{-n+2}}$$

Now, $I = (2^{-j}, 2^{-j+1}]$ is the desired interval!

Indeed, for $\lambda \in I$ we can estimate

$$\begin{aligned}\mathcal{N}_\lambda^0(A) &\geq \mathcal{N}_\lambda^1(A) - \left|\mathcal{N}_\lambda^\varepsilon(A) - \mathcal{N}_\lambda^1(A)\right| - \left|\mathcal{N}_\lambda^0(A) - \mathcal{N}_\lambda^\varepsilon(A)\right| \\ &\geq \vartheta - \varepsilon^{-F} J^{-2^{-n+2}} - G\varepsilon^{1/3} \\ &\geq \vartheta - \vartheta/3 - \vartheta/3 > 0\end{aligned}$$

The error part

The most interesting part for us is the error part

We need to estimate

$$\sum_{j=1}^{J} \kappa_j \Big(\mathcal{N}_{\lambda_j}^{\varepsilon}(A) - \mathcal{N}_{\lambda_j}^{1}(A) \Big)$$

for arbitrary scales $\lambda_j \in (2^{-j}, 2^{-j+1}]$ and arbitrary complex signs κ_j, with a bound that is sub-linear in J

The error part

It can be expanded as

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \prod_{i=0}^{n-1} \mathbb{1}_A(x + iy) K(y) \,\mathrm{d}y \,\mathrm{d}x,$$

where

$$K(y) := \sum_{j=1}^{J} \kappa_j \big((\sigma_{\lambda_j} * \varphi_{\varepsilon\lambda_j})(y) - (\sigma_{\lambda_j} * \varphi_{\lambda_j})(y) \big)$$

is a translation-invariant Calderón–Zygmund kernel

The error part

If $d = 1$ and $K(y)$ is a truncation of $1/y$, then this becomes the (dualized and truncated) multilinear Hilbert transform,

$$\int_{\mathbb{R}} \int_{[-R,-r]\cup[r,R]} \prod_{i=0}^{n-1} f_i(x+iy) \frac{\mathrm{d}y}{y} \, \mathrm{d}x$$

- When $n \geq 4$, no L^p-bounds uniform in r, R are known
- Tao (2016) showed a bound of the form $o(J)$
- Durcik, K., and Thiele (2019) showed a bound of the form $O(J^{1-\varepsilon})$

A stronger property

Most papers on the Euclidean density theorem simultaneously also establish a stronger property for subsets $A \subseteq \mathbb{R}^d$ of positive upper Banach density

Upper Banach density of a measurable set $A \subseteq \mathbb{R}^d$:

$$\overline{\delta}(A) := \limsup_{R \to \infty} \sup_{x \in \mathbb{R}^d} \frac{|A \cap (x + [0, R]^d)|}{R^d} \in [0, 1]$$

An open problem

Problem

Prove or disprove: if $n \geq 4$, $p \neq 1, 2, \ldots, n-1$, $d \geq D'(n,p)$, $A \subseteq \mathbb{R}^d$ measurable, $\overline{\delta}(A) > 0$, then $\exists \lambda_0 = \lambda_0(n,p,d,A) \in (0,\infty)$ such that for every $\lambda \geq \lambda_0$ one can find $x, y \in \mathbb{R}^d$ satisfying $x, x+y, \ldots, x+(n-1)y \in A$ and $\|y\|_{\ell^p} = \lambda$

Its difficulty $\longleftrightarrow$ lack of our understanding of boundedness properties of multilinear Hilbert transforms

Thank you for your attention!