

Szemerédi's Thm & density thms

- stronger than vdW (if true): density version:

$$\forall k, r \exists N_0 \text{ s.t. } \forall N > N_0$$

$$A \subseteq [N], |A| > N/r \Rightarrow A \supseteq k\text{-t. AP}$$

- in hypergraph language \otimes :

$$\mathcal{H}_N = \{k\text{-t. AP's in } [N]\} \quad (V(\mathcal{H}_N) = [N])$$

$$\underline{\text{vdW}}: \chi(\mathcal{H}_N) \rightarrow \infty$$

$$\underline{\text{density}}: \alpha(\mathcal{H}_N) = o(N) \quad (?)$$

$$\otimes (\alpha: \text{ind. \# ; ind: } \neq \text{edge})$$

- analogue for Ramsey: $V(\mathcal{H}_n) = \binom{[n]}{2}$
 $\mathcal{H}_n = \{\text{triangles}\}$] (say) \rightarrow

$$\text{Ramsey: } \chi(\mathcal{H}_n) \rightarrow \infty$$

$$\underline{\text{density}}: \alpha(\mathcal{H}_n) = o(n^2) \quad (?)$$

$$\text{--- nonsense: } \alpha(\mathcal{H}_n) = \lceil n^2/4 \rceil$$

... But density vdW is true:

another form:

$$\text{upper density of } A \subseteq \mathbb{N} := \overline{\lim} \frac{|A \cap [N]|}{N}$$

► Sz's thm ('74):

if $A \subseteq \mathbb{N}$ has pos. upper density (p.u.d.)

then $A \supseteq k$ -term AP $\forall k$.

[why equiv? $\left\{ \begin{array}{l} \Rightarrow: \dots \\ \Leftarrow: \text{EX} \end{array} \right. \]$

some history:

- Conj'd: Erdős-Turán '36 (\rightsquigarrow vdw)
- Roth '52: $k=3$ (Fourier; see GRS)
- Szemerédi '69: $k=4$
- Furstenberg '77:
ergodic theory pf of Sz. (hint in GRS)
- Green-Tao '04: long AP's in $\{\text{primes}\}$

BTW:

$$r_k(N) := \max \{ |A| : A \subseteq [N], A \not\subseteq k\text{-t. AP} \} \rightarrow$$

Szemerédi: $r_k(N) = o(N)$ (Roth: $r_3(N) = o(N)$)

• Gowers '04: $r_k(N) < (\log \log N)^c N$ [$c = 2^{-2^{k+9}}$]

▶ much work on $r_3(N)$; e.g. (skipping some constants):

• Roth '52: $r_3(N) \lesssim \frac{N}{\log \log N}$

• Bloom-Sisask '20: $O\left(\frac{N}{(\log N)^{1+c}}\right)$ (!)

(breaks the "logarithmic barrier")

• Kelley-Meka '23: $N \exp[-\Omega(\log^c N)]$ (!!!)

vs. Behrend '46: $r_3(N) \gtrsim N \exp[-\sqrt{\log N}]$ ↓

Szemerédi's (beautiful) pf of Roth

prelim's (briefly)

① "k-cube": $M(a; d_1, \dots, d_k) = \{a + \sum_{i \in I} d_i : I \subseteq [k]\}$ $[a, d_i \in \mathbb{P}]$

Cube lemma $A \subseteq [n], |A| > cn \Rightarrow$

$A \supseteq k$ -cube \bar{w} $k > \log \log n - f(c)$ $[\log = \log_2 \text{ but unimp.}]$

semipf / ex: (i) choose $d_1 \in \mathbb{P}$ \bar{w}

$$|A_1 := \{x \in A : x + d_1 \in A\}| \text{ max'm } \rightarrow |A_1| \geq c^2 n / 2$$

(ii) iterate (on A_1, A_2, \dots) until $|A_i| = 1$

• k -cube in A_1 gives k -cube in A etc.

• # of iterations $\approx \log \log n - \log \log (2/c)$ \square

② $S(n) := \max \{|A| : A \subseteq [n], A \not\subseteq 3\text{-t. AP}\}$

(= $r_3(n)$; for this step 3 could be k)

EX: S is subadditive, i.e. $S(m+n) \leq S(m) + S(n)$

► Fekete's Lemma $S: \mathbb{N} \rightarrow \mathbb{R}^+$ subadditive \Rightarrow

• $\alpha := \lim \frac{S(n)}{n}$ exists and

• $\alpha \leq \frac{S(n)}{n} \quad \forall n$

Pf: EX (not hard $\&$ a basic fact)

(now Roth) setting up:

● to show: $\Sigma(n)/n \rightarrow 0$

● $c := \lim \Sigma(n)/n$ (Fekete: c exists)

assume (for $\rightarrow \leftarrow$) that $c > 0$. Then:

$$\forall \varepsilon > 0 \exists n_\varepsilon: n > n_\varepsilon \Rightarrow cn \leq \Sigma(n) < (c+\varepsilon)n$$

● choose:

● ε very small (e.g. $\varepsilon < .01 c^2$)

● n very large: $.01 c^2 \log \log n > n_\varepsilon$

[$\log \log \leftrightarrow$ Cube Lemma; don't worry about values: it will be clear \exists val's that work & we'll see what's needed]

● Let $A \subseteq [n]$, $|A| \geq cn$, $A \not\subseteq$ 3-t. AP \rightarrow for $\rightarrow \leftarrow$

MAIN IDEA: if $[n] \approx \cup$ (long AP's) then $\uparrow > n_\varepsilon$

$|A| \geq cn$ and AP densities (in A) $< c + \varepsilon \Rightarrow$

can't lose much

⊗ first use at ① but main use at end; see PLAN below

$$\textcircled{1} \quad I := (.49n, .5n) \rightarrow \frac{|A \cap I|}{.01n} > \frac{c}{2}$$

[ingredients highlighted thus as they appear]

Pf O'wise $|A| < (c+\varepsilon) \cdot .99n + \frac{c}{2} (.01n) < cn \quad (\rightarrow \leftarrow)$ $\textcircled{2}$
 (triv.)
 \Rightarrow

$$\textcircled{2} \quad \exists \text{ interval } \gamma \subseteq I, |\gamma| = \sqrt{n} \text{ (really } |\gamma| \approx \sqrt{n}) \text{ s.t.}$$

$$|A \cap \gamma| > \frac{c}{2} |\gamma|$$

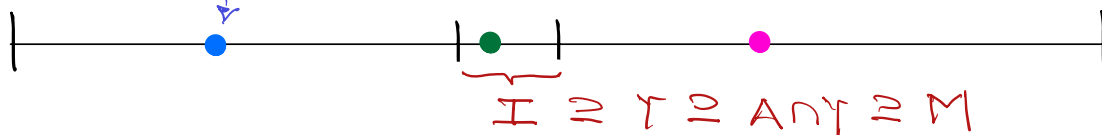
(Cube t.)
 \rightarrow

$$\textcircled{3} \quad A \cap \gamma \supseteq k\text{-cube } M = M(a; d_1, \dots, d_k)$$

$$(\because \{a + \sum_{i \in I} d_i : I \subseteq [k]\})$$

$$\bar{w} \quad k = \log \log n - o(1)$$

$$\textcircled{4} \quad X := A \cap [1, .49n]$$



PLAN: {3rd terms (•) of AP's \bar{w} }

first term (•) in X & second term (•) in M is a

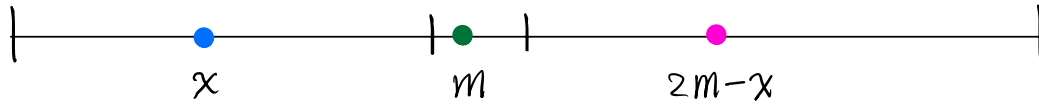
large subset of $(.49n, n] \setminus A$ with structure

\rightarrow part'n rest of this into (mostly) long AP's

\ni use "MAIN IDEA"

⑤ $M_i := M(a, d_1, \dots, d_i)$ ($M_0 := \{a\}$)

$N_i := \{zm - x : x \in X, m \in M_i\}$



• $A \cap N_i = \emptyset$ (or $A \cong 3\text{-t. AP}$)

• $N_i = N_{i-1} \cup (N_{i-1} + 2d_i)$

• $|N_0| = |X| > cn/4$ (say; actually $|X| \approx cn/2$)

• $|N_k| < n \rightarrow \exists i \quad |N_i \setminus N_{i-1}| < n/k \rightarrow$

⑥ $Z := (.49n, n] \setminus N_{i-1}$

• block := AP w diff. $2d_i$

• partition N_{i-1} into max'l blocks:

$< |N_i \setminus N_{i-1}| (< n/k)$

→ EX [easier: $|N_i \setminus N_{i-1}| + 2d_i$]

• partition Z into max'l blocks:

$< n/k + 2d_i < n/k + 2\sqrt{n}$
 ↓ why? ignore

done:

rough: $(A \cap N_{k-1}^c = \emptyset \rightarrow)$ lost $|N_{k-1}| = \Omega(n)$ \neq
 rest partitioned into mostly long AP's
 (which have densities $\leq c$) \rightarrow can't catch up.

precise: block is $\begin{cases} \text{short} & \text{if size} \leq n_\varepsilon \\ \text{long} & \text{otherwise} \end{cases}$

A has density:

- 0 on N_{k-1}^c
- $< c + \varepsilon$ on $[1, .49n) \cup U\{\text{long blocks}\}$
- (≤ 1 on short blocks)

\rightarrow

$$cn < |A|$$

$$< (c + \varepsilon) \left[.49n + \left(.51n - \frac{c}{4}n \right) \right] + \left(\frac{n}{k} + 2\sqrt{n} \right) n_\varepsilon$$

$$< cn \quad \longleftrightarrow$$

\square