Szemerédi's Thm & density thms

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

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

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

- in hypergraph language*:

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

  vdW: $\chi(\mathcal{H}_N) \longrightarrow \infty$

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

  * (α: ind. #; ind: $\not\supseteq$ edge)

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

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

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

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

... But density vdW is true:

another form:

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? $\begin{cases} \Rightarrow: & \cdots \\ \Leftarrow: & \text{EX} \end{cases}$]

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 {primes}

BTW:

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

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

- Gowers '01: $r_k(N) < (\log\log N)^{-c} N$ $[c = 2^{-2^{k+9}}]$

- 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: $r_3(N) \boxed{\gtrsim} N \exp[-\sqrt{\log N}]$

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 $A \subseteq [n]$, $|A| > cn \implies$

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

semipf/EX: (i) choose $d_1 \in \mathbb{P}$ s.t.

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

(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 $\approx \log\log n - \log\log(2/c)$ ∎

② $S(n) := \max\{|A| : A \subseteq [n], A \not\ni$ 3-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** $S : \mathbb{N} \to \mathbb{R}^+$ subadditive $\implies$

- $\alpha := \lim \dfrac{S(n)}{n}$ exists and
- $\alpha \leq \dfrac{S(n)}{n} \quad \forall n$

Pf: EX (not hard & a basic fact)

(now Roth) setting up:

- TO SHOW: $\boxed{S(n)/n \to 0}$
- $c := \lim S(n)/n$ (Fekete: c exists)

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

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

- choose:
  - ε very small (e.g. $\varepsilon < .01c^2$)
  - n very large: $.01c^2 \log\log n > n_\varepsilon$

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

- Let $A \subseteq [n]$, $|A| \geq cn$, $\underbrace{A \not\ni \text{3-t. AP}}_{\text{for } \to\leftarrow}$

MAIN IDEA(*): if $[n] \approx \dot{\bigcup}(\text{long AP's})$ then

long $> n_\varepsilon$

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

can't lose much

(*) first use at (I) but main use at end; see $\boxed{\text{PLAN}}$ below

① $I := (.49n, .5n) \to \dfrac{|A\cap I|}{.01n} > \dfrac{c}{2}$

[ingredients highlighted thus as they appear]

Pf O'wise $|A| < (c+\varepsilon).99n + \frac{c}{2}(.01n) < cn$ (→←) ☒

$\overset{\text{(triv.)}}{\Longrightarrow}$

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

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

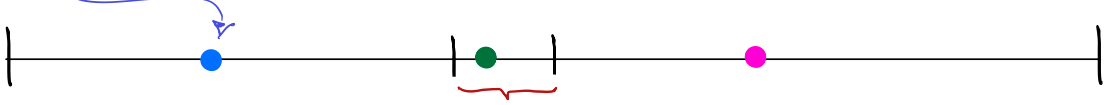
$\xrightarrow{\text{(Cube L.)}}$

③ $A\cap Y \supseteq k$-cube $M = M(a; d_1, \dots, d_k)$

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

w/ $k = \log\log n - O(1)$

④ $X := A\cap[1, .49n]$



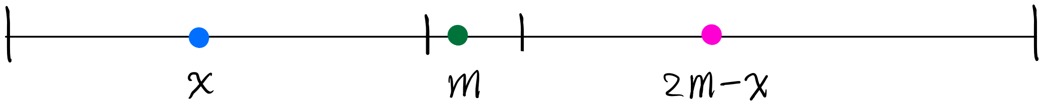
$I \supseteq Y \supseteq A\cap Y \supseteq M$

$\boxed{\text{PLAN:}}$ {3rd terms (•) of AP's w/ first term (•) in X & second term (•) in M} is a large subset of $(.49n, n] \setminus A$ with structure

$\longrightarrow$ part'n rest of this into (mostly) long AP's & use "MAIN IDEA"

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

$N_i := \{2m - x : x \in X, m \in M_i\}$



- $\boxed{A \cap N_i = \phi}$ (or $A \supseteq$ 3-t. AP)
- $N_i = N_{i-1} \cup (N_{i-1} + 2d_i)$
- $|N_0| = |X| > cn/4$ (say; actually $|X| \approx cn/2$)
- $|N_k| < n \longrightarrow \exists i \quad |N_i \setminus N_{i-1}| < n/k \longrightarrow$

⑥ $Z := (.49n, n] \setminus N_{i-1}$

- block := AP w/ diff. $2d_i$
- partition $\boxed{N_{i-1}}$ into max'l blocks:

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

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

- partition $\boxed{Z}$ into max'l blocks:

  \# $< n/k + 2d_i < n/k + \underbrace{2\sqrt{n}}_{\text{ignore}}$

  why?

& done: good EX for now & TB continued