Density of arithmetic progressions

<u>Conjecture</u> (Erdős - Turán, 1936)

$\forall A \subset \mathbb{N}$: if $\limsup_n |A \cap [n]|/n > 0$ then $\forall k$: $A \supset AP_k$.

Proved by Roth for $k=3$ (53, Fields Medal)

by Szemeredi for $k=4$ (69, Abel Prize 2012)

& $\forall k$ (74)

Also by Furstenberg $\forall k$ (77) by ergodic theory.

<u>Cor</u>. $\forall r \forall \chi: \mathbb{N} \to [r]$: the most frequent color contains AP_k $\forall k$.

<u>Thm</u> (Furstenberg, Katznelson) $\forall t \geq 2 \, \forall \varepsilon > 0 \, \exists N: \forall n \geq N$

if $S \subseteq C_t^n$, $|S| \geq \varepsilon t^n$, then S contains a line.

(True for $t=2$ by Sperner's Thm; $\forall t \Rightarrow$ Szem. Thm.)

<u>Conjecture</u> (most valuable conj. of Erdős, \$3000)

If for $A \subseteq \mathbb{N}$: $\sum_{a \in A} \frac{1}{a} = \infty$ then $A \supset AP_k$ $\forall k$.

Converse ~~false~~

$1, 10, 11, 100-102$

$1000-1003, \ldots$

Proved by Green & Tao for primes (Tao - Fields m. 20~~00~~06)

Here we give Szemerédi's proof of Roth's Thm (S22)

<u>Thm</u> (Roth, 1953) If $\limsup_n |A \cap [n]|/n > 0$ then $A \supset AP_3$

<u>Proof</u> Let $S(n)$ = size of a largest subset of $[n]$ without AP_3.

The negation: $\limsup_n \frac{S(n)}{n} > 0$ (b/c Thm $\Leftrightarrow \frac{S(n)}{n} \to 0$)

Observe $S(n)$ is subadditive: $S(n_1+n_2) \leq S(n_1) + S(n_2)$

$\overset{L1}{\Rightarrow}$ $c = \lim_{n\to\infty} \frac{S(n)}{n}$ exists & $S(n) \geq cn$

<u>The idea</u>: Let $A \subset [n]$, $|A| = S(n)$, $A \not\supset AP_3$. We will show that $A \supset$ a special structure (k-cube) which will rule out many els. from $A \Rightarrow |A| < cn$ - contrad.

<u>L1</u> If $f: \mathbb{N} \to \mathbb{R}^+$ is subadditive then $c = \lim \frac{f(n)}{n}$ exists, $\frac{f(n)}{n} \geq c$.

<u>Def</u> k-cube $M(a, d_1, \dots, d_k)$, $a, d_1, \dots, d_k \in \mathbb{N}$, is the set

$$\left\{ a + \sum_{i=1}^{k} \varepsilon_i d_i : \varepsilon_i \in \{0,1\} \right\}$$

Equivalently, $M_0 = \{a\}$, $M_i = M_{i-1} \cup (M_{i-1} + d_i)$, $i = 1, \dots, k$

(Note that each M_i is an i-cube)

<u>Ex</u>. $a = 3$, $d_1 = 1$, $d_2 = 4$, $d_3 = 2$ $M_0 = \{3\}$, $M_0 + 1 = \{4\}$, $M_1 = \{3, 4\}$

$M_1 + d_2 = \{7, 8\}$, $M_2 = \{3, 4, 7, 8\}$, $M_2 + d_3 = \{5, 6, 9, 10\}$, $M_3 = \{3, 4, 5, 6, 7, 8, 9, 10\}$

(S23)

L2 (The cube lemma) $\forall n, \alpha \in \mathbb{N}$, define

$\alpha_0 = \alpha$, $\alpha_{i+1} = \binom{\alpha_i}{2} / (n-1)$, $i = 0, 1, 2, \ldots$. If $\alpha_k \geq 1$ then $\forall A \in \binom{[n]}{\alpha}$ contains a k-cube. $\alpha_k \approx \frac{c^{2^{k-1}}}{2^k} n \geq 1$

Cor Let $A^{(n)} \subset [n]$, $|A^{(n)}| \sim cn$, $n = 1, 2, \ldots$, $\mathbb{N} \ni k_n = \log\log n - x$, $x = x(c)$. Then for large enough n, $A^{(n)} \supset k_n$-cube.

Szemeredi's Proof: Let $\varepsilon = 10^{-10} c^2$. Choose n_0 so that $\forall n \geq n_0$: $c \leq \frac{S(n)}{n} < c + \varepsilon$.

$$|A \cap ([.49n] \cup \{\tfrac{n}{2}, \ldots, n\})| \leq S(.49n) + S(\tfrac{1}{2}n) < (c+\varepsilon)(.99n)$$

$$\Rightarrow |A \cap \{.49n, \ldots, \tfrac{n}{2}\}| > cn - (c+\varepsilon)(.99n) > \frac{.01}{2} cn$$

$\Rightarrow$ A has in $\{.49n, \ldots, \frac{n}{2}\}$ density $> \frac{c}{2}$

Split $\{.49n, \ldots, \frac{n}{2}\}$ into disjoint intervals of length $\sqrt{n}$
On one of these, A has density $> \frac{c}{2}$

By Cor, $A \cap \{.49n, \ldots, \frac{n}{2}\} \supset M(a, d_1, \ldots, d_k)$, $k = \log\log n + O(1)$, $d_i < \sqrt{n}$

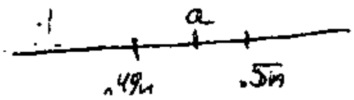
Set $M_0 = \{a\}$, $M_i = M(a, d_1, \ldots, d_i)$, $i = 1, \ldots, k$.

$N_i = \{2m - x : m \in M_i, x \in A, x < a\}$, $i = 0, \ldots, k$; clearly $N_i \subset [n]$

Since $(x, m, 2m-x)$ is an AP_3 & $\{m, x\} \subset A$, $A \cap N_i = \emptyset$.

S24

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



$\Rightarrow N_0 \subset N_1 \subset N_2 \subset \dots \subset N_k$

(<u>HW</u> $|N_i| \geq |N_0| = |A \cap [a-1]| \geq |A \cap [.49n]| = |A| - |A \cap \{.49n, \dots, n\}|$

$\geq cn - S(.51n) \geq cn - (c+\varepsilon)(.51)n = .49cn - .51\varepsilon n > \frac{c}{2}(.49)n$)

But $|N_k| \leq n \Rightarrow \exists i : |N_i - N_{i-1}| < \frac{n}{k}$. Fix that i.

Let us call ~~an~~ max. AP with diff. $2d_i$ a <u>block</u>

N_{i-1} can be split into $< \frac{n}{k}$ blocks. Indeed, to every block we may assign an el. of $N_{i+1} - N_i$ (by adding $2d_i$ to the last el. of the block — this number $\in N_i$ b/c $N_i \ni N_{i-1} + 2d_i$)

Split $[n]$ into $2d_i$ residue classes modulo $2d_i$:
$L_0, \dots, L_{2d_i - 1}$

The blocks in $L_j \cap N_{i-1}$ can't overlap — they appear sequentially & the gaps between them form blocks in ~~...~~
$L_j \cap ([n] - N_{i-1})$

Suppose there are t_j blocks in $N_{i-1} \cap L_j$, $\sum t_j < \frac{n}{k}$. Then there are at most
$\sum (t_j + 1) < \frac{n}{k} + 2d_i \sim \frac{n}{\log\log n}$ blocks in $[n] \setminus N_{i-1}$.

We call a block of $[n] - N_{i-1}$ <u>small</u> if $|\cdot| < 0.01c^2 \log\log n$

Small blocks have together $< .01c^2 n$ els.

Choose n so that $.01c^2 \log\log n > n_0$. Then $\forall$ <u>large</u> block B:

$|A \cap B| < (c+\varepsilon)|B|$ — B can be shrunk to $[|B|]$.

Finally, $|A| = |A \cap ([n] \setminus N_{i-1})| < (c+\varepsilon)(n - |N_{i-1}|) + 0.01c^2 n$

$< cn + \varepsilon n - c(\frac{c}{2}(.49)n) + .01cn^2 < cn$ b/c $\varepsilon < .235c^2$ $\square$

Twierdzenie Rotha (1953)

$$r_3(n) = o(n)$$

Dowód będzie oparty na następującym teoriografowym lemacie.

Lemat o Usuwaniu (Ruzsa i Szemerédi, 1978)

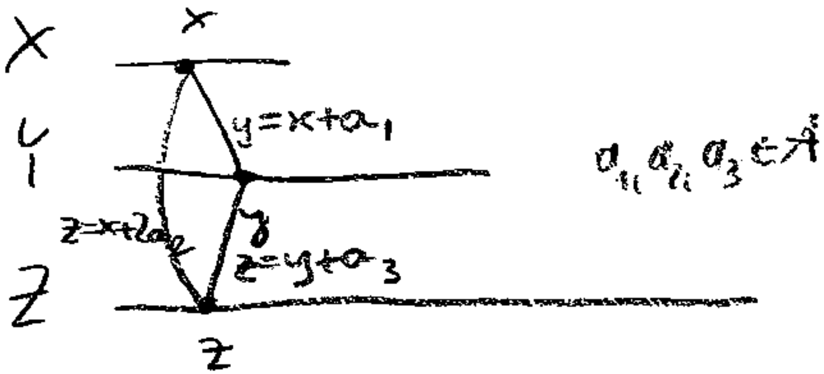
Dla każdego $c > 0$ istnieje $a > 0$ takie, że jeżeli graf G na n wierzchołkach ma co najwyżej an^3 trójkątów, to można je wszystkie zniszczyć usuwając co najwyżej cn^2 krawędzi.

Albo:

Lemat o Usuwaniu (Ruzsa i Szemerédi, 1978)

Dla każdego $c > 0$ istnieje $a > 0$ takie, że jeżeli graf G na n wierzchołkach ma co najmniej cn^2 krawędziowo rozłącznych trójkątów, to G ma co najmniej an^3 wszystkich trójkątów.

$A \subset [n]$



$\Downarrow$

$a_1 + a_3 = 2a_2 \Rightarrow a_1 = a_2 = a_3$

else $AP_3 \subset A$

Dowód Twierdzenia Rotha

Niech $A \subseteq \{1, 2, \ldots, n\}$, $|A| = r_3(n)$, będzie zbiorem niezawierającym ciągu arytmetycznego długości 3.

Pokażemy, że $|A| = o(n)$.

Niech $G = (V, E)$ będzie 3-dzielnym grafem na zbiorze wierzchołków:

$$V = X \cup Y \cup Z, \quad \text{gdzie } X = [n],\ Y = [2n],\ Z = [3n],$$

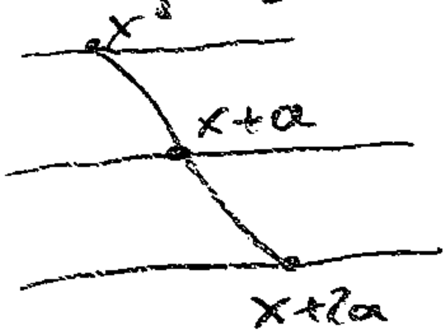
ze zbiorem krawędzi:

$$E = E_{XY} \cup E_{XZ} \cup E_{YZ},$$

gdzie

$$E_{XY} = \{\{x,y\} : x \in X,\ y \in Y,\ y = x + a \text{ dla pewnego } a \in A\},$$
$$E_{XZ} = \{\{x,z\} : x \in X,\ z \in Z,\ z = x + 2a \text{ dla pewnego } a \in A\},$$
$$E_{YZ} = \{\{y,z\} : y \in Y,\ z \in Z,\ z = y + a \text{ dla pewnego } a \in A\}.$$

The only K_3's in G are



$|X|\cdot|A|$ edge-disjoint K_3

If $|A| = \Omega(n)$ then, by RL,
G has $\Omega(n^3)$ triangles – contrad.

Dowód twierdzenia Rotha

Załóżmy, że $x \in X$, $y \in Y$ i $z \in Z$ jest trójkątem w G.

Wtedy dla pewnych $a_1, a_2, a_3 \in A$:

$$y = x + a_1, \quad z = x + 2a_2 \quad \text{i} \quad z = y + a_3.$$

Zatem,

$$a_1 + a_3 = 2a_2.$$

Stąd, $a_1 = a_2 = a_3$, bo inaczej (a_1, a_2, a_3) byłby ciągiem arytmetycznym długości 3 w A.

W konsekwencji każdy trójkąt jest postaci $\{x, x+a, x+2a\}$, gdzie $x \in X$ i $a \in A$.

Mamy więc dokładnie $|X| \cdot |A| = n \cdot r_3(n) \le n^2$ krawędziowo rozłącznych(!) trójkątów.

Przypuśćmy, nie wprost, że $r_3(n) = \Omega(n)$.

Wtedy, Lemat o Usuwaniu implikuje, że G zawiera $\Omega(n^3)$ trójkątów – sprzeczność! □