Maximal subsets avoiding arithmetical progressions in sets

Semchankau Aliaksei

Vilnius Conference on Combinatorics and Number Theory

21 July 2017

Definitions

Let us define a function on sets of integer:

$$f_k(B) = \max |A| : \{A \subseteq B, A \not\supseteq APk\}.$$

APk is a set of arithmetical progressions of length k.
It seems that is $f_k(B)$ is small, then B should be structured.
Let's consider extremal value of f_k:

$$\phi_k(n) = \min_{|B|=n} f_k(B).$$

Now consider a function f_k on the "most structured" set:

$$g_k(n) = f_k([1, 2, \dots, n])$$

and its "density":

$$\rho_k(n) = g_k(n)/n$$

For $\rho_k(n)$ next bounds are known

$$\frac{1}{e^{c_k\sqrt{\ln n}}} \ll \rho_k(n) \ll \frac{1}{(\ln \ln n)^{s_k}},$$

where are the c_k and s_k depend only only on k. The lower bound is due to Behrend and the upper bound is due to Gowers.

By definition, $\phi_k(n) \leqslant g_k(n)$ holds, but one can expect, that $\phi_k(n) = g_k(n)$, because, intuitively, $[1, 2, \ldots, n]$ is very structured and is full of arithmetic progressions.
Counterexample:

$$\phi_3(5) = f_3(\{1, 2, 3, 4, 7\}) = 3,$$

$$g_3(5) = f_3(\{1, 2, 3, 4, 5\}) = 4.$$

Previous results

So, one can suppose, that $\phi_k(n) \sim g_k(n)$.

- Komlós, J., Sulyok, M., Szemerédi, E. (1975):

$$\phi_3(n) \geqslant \left(\frac{1}{2^{15}} + o(1)\right) g_3(n), n \to \infty$$

- O'Bryant, with no proof:

$$\phi_3(n) \geqslant \left(\frac{1}{34} + o(1)\right) g_3(n), n \to \infty$$

Our result

For any $k \geq 3$ there exist such a sequence $n_1 < n_2 < \dots$ so that for any n in this sequence next inequality holds:

$$\phi_k(n) \geq \left(\frac{1}{4} + o(1)\right) g_k(n), n \to \infty,$$

and this sequence is dense enough: any segment of type $[n, ne^{(\ln n)^{\frac{1}{2}+o(1)}}]$ contains at least one number of this sequence.

Sets Compressions

Definition

Let $X = \{x_1, x_2, \dots, x_n\}$ be a set of integers. We call $Y = \{y_1, y_2, \dots, y_n\}$ a *compression* of X, if for any i, j, k such that $x_i - 2x_j + x_k = 0$ it is true, that $y_i - 2y_j + y_k = 0$. In other words, Y contains the arithmetical structure of X.

Now we can state a hypothesis:

Hypothesis:

For any $1 > \epsilon > 0$ there exists a function $h(n) = h_\epsilon(n)$ which grows slower than any power of n, so that can one take any set X of size n and remove at most ϵn integers, so that the rest set can be compressed into subset of $[1, 2, \dots, nh(n)]$.

At the current moment it is proved, that for $\epsilon > 3/4$ one can take $h_\epsilon(n) = C_\epsilon \ln n$.

Sketch of the Proof

Lemma 1:

Any set X of size n can be compressed into subset of segment $[1, \ldots, 4n^4 6^{n/2}]$.

Proof: some manipulations with matrices.

Sketch of the Proof

Lemma 2:

If set X of size n belongs to segment $[1, \cdots, M]$, where $M = 4n^4 6^{n/2}$, then $\exists X' \subset X, |X'| \geq |X|/2$ which can be compessed into subset of $[n^3]$.

Proof: one can find prime number $p \leq 2n^3$, so that p does not divide any difference $x_i - x_j$. It's easy to see now, that one of the sets

$$X_1 = X/\mathbb{Z}_p \cap [1, \ldots, p/2),$$

$$X_2 = X/\mathbb{Z}_p \cap (p/2, \ldots, p-1],$$

would satisfy our conditions.

Sketch of Proof

Lemma 3:

If set X of size n belongs to segment $[1, \cdots 8n^3]$, and $\epsilon > 0$ is given, then $\exists X' \subset X, |X'| \geq (1/2 - \epsilon)|X|$ which can be compressed into subset of segment $[1, \cdots, C_\epsilon n \ln n]$, where C_ϵ is a constant which depends only on ϵ.

Let's consider prime numbers in interval $[2n, \cdots, 2cn \ln n]$,where c is some positive constant. By Chebychev, one can find more then cn prime numbers there. Let's enumerate them as $p_1, p_2, \ldots, p_s$, $s > cn$.
Now we count triples (i, j, t), where i, j, t are such that $p_t|(x_i - x_j)$. One can notice, that one pair (i, j) may occur as first 2 arguments only twice because $|x_i - x_j| < 8n^3$ and can't be divisible on 3 different primes, greater than $2n$. Thus, there are no more than n^2 such triples. By Dirichlet, some p_t corresponds to less then $n^2/cn = n/c$ triples. Let's extract from X all x_i, x_j which a equal modulo p.
Now all x_i are different modulo p, and we can use the same approach, as in lemma 2.

Now we apply hypothesis to bound value of $\phi_k(n)$, so we will prove that for every set B of size n inequality $f_k(B) \geqslant (\frac{1}{4} + o(1)) f_k([1, 2, \dots, n])$ holds. One can find a subset in B with $\sim n/4$ elements, which can be compressed into subset B' of segment $[1, \dots, O(n \ln n)]$. Let A be a maximal set with no APk's in the greater segment $[1, \dots, O(n \ln n)]$ (i.e. $|A| = f_k([1, \dots, O(n \ln n)])$) One can demonstrate, using some combinatorical arguments, that we can find shift A, which has large intersection with B'.

Now we can just take a prototype of $B' \cap A$ in B, this prototype would be of size $\left(\frac{1}{4} + o(1)\right) g_k(n)$ and would not contain arithmetical progressions of length k.

Questions to solve

Improve constant:

It's expected, that inequality $\phi_k(n) \geq (1 + o(1))\, g_k(n)$ is true.

Hypothesis for all ϵ:

Prove the hypothesis for all $\epsilon > 0$.

Reverse problem:

What can we say about structure of B, if inequality $f_k(B) \leq C\phi_k(n)$ holds?

Thanks for attention!