Szemerédi's theorem and fractal dimensions of sets avoiding (k, ϵ)-arithmetic progressions

Kota Saito

1 Arithmetic progressions of integers

We say that a real sequence $(a_i)_{i=0}^{k-1} \subset \mathbb{R}$ is an arithmetic progression (AP) of length k if there exists $\Delta > 0$ such that

$$a_i = a_0 + i\Delta$$

for $i = 0, 1, \ldots, k-1$. We say that Δ is the gap length of an AP. We say that a set P is an AP of length k if there exists an AP $(a_i)_{i=0}^{k-1}$ of length k in the sense of a sequence such that $P = \{a_i \colon i = 0, 1, \ldots, k-1\}$. It is a big problem to show the existence or non-existence of arithmetic progressions of a given set. One of main topics of this problem is to find long APs of a given 'large' set. For example, a subset A of positive integers with positive upper density *i.e.*

$$\limsup_{N\to\infty} \frac{|A \cap [1, N]|}{N} > 0$$

contains arbitrarily long APs. This statement is known as Szemerédi's theorem which was proved by Szemerédi in 1975 [14]. This theorem implies van der Waerden's theorem [15] which states that if we partition a set of all positive integers into finitely many sets, then we can find at least one set which contains arbitrarily long arithmetic progressions. Erdős and Turán conjectured the case when the length of APs is 3 of Szemerédi's theorem in [2]. In the same paper, they introduced the number $r_k(N)$ as the maximum cardinality of A such that $A \subseteq \{1, 2, \ldots, N\}$ does not contain any APs of length k. We can get the equivalent statement of Szemerédi's theorem in terms of $r_k(N)$ as follows: Szemerédi's theorem is equivalent to

$$\lim_{N\to\infty} r_k(N)/N = 0$$

holds for all $k \geq 3$. Furthermore, if

$$r_k(N) \leq \frac{N}{\log N (\log N)^2}$$

was true for all $k \geq 3$ and $N \geq N(k)$, then Erdős-Turán conjecture would be true. We can see a proof of this implication in [7]. Here Erdős-Turán conjecture states that a subset A of positive integers satisfying

$$\sum_{a \in A} \frac{1}{a} = \infty$$

would contain arbitrarily long arithmetic progressions. This conjecture is still open even if the length of APs is equal to 3. Note that the sum of reciprocals of prime numbers is divergent. Therefore, if Erdős-Turán conjecture is true, we can prove the Green-Tao theorem [8] which states that the set of all prime numbers contains arbitrarily long arithmetic progressions. Hence it is important to find non-trivial bounds for $r_k(N)$. The result by Gowers in the paper [6] is the best in known upper bounds for $r_k(N)$ for general length k, that is, for all $N \geq 3$ and $k \geq 3$,

$$r_k(N) \leq \frac{N}{(\log \log N)^{2^{-2^{k+9}}}}. \tag{1}$$

Remark that we know better upper bounds for $r_k(N)$ in the particular case when $k = 3$ or 4 (see [1, 9]). O'Bryant gave a lower bound for $r_k(N)$. For all $N \geq 2$ and $k \geq 3$, we have

$$r_k(N) \geq CN \exp\left((\log 2)\left(-n2^{(n-1)/2}\sqrt[n]{\log_2 N} + \frac{1}{2n}\log_2\log_2 N\right)\right), \tag{2}$$

for some $C > 0$, where $n = \lceil \log_2 k \rceil$. By combining (1) and (2), for every $k \geq 3$, there exists positive constants C_k and D_k such that for every $N > e^e$, we have

$$N \exp(-C_k(\log N)^{n/2}) \leq r_k(N) \leq N \exp(-D_k \log\log\log N).$$

2 Arithmetic progressions of real numbers

We firstly construct a Cantor set which does not contain any APs of length 3. Let $\psi_1(x) = x/6$ and $\psi_2(x) = x/6 + 5/6$. Let $I_0 = [0, 1]$, and define

$$I_n = \psi_1(I_{n-1}) \cup \psi_2(I_{n-1})$$

for all $n \geq 1$. Then we define $C = \bigcap_{n=0}^{\infty} I_n$. Note that C is the attractor of the iterated function system $\{\psi_1, \psi_2\}$ on $\mathbb{R}$. We show that C does not contain any APs of length 3 by contradiction. If C contained at least one

AP of length 3, then let P be such an AP, and $P \subseteq I_1$ by definition. The set I_1 is the union of two disjoint closed intervals. The middle hole of I_1 is enough large such that all of elements in P belong to the left or right interval. Therefore there exists $i_1 \in \{1, 2\}$ such that

$$P \subseteq \psi_{i_1}(I_0) \tag{3}$$

By $P \subseteq I_2$ and (3), there exists $i_2 \in \{1, 2\}$ such that $P \subset \psi_{i_1} \circ \psi_{i_2}(I_0)$. By iterating these steps, the diameter of P can be arbitrarily small. This is a contradiction.

The iterated function system $\{\psi_1, \psi_2\}$ clearly satisfies open set condition. Therefore the Hausdorff dimension of C is equal to real number s such that

$$6^{-s} + 6^{-s} = 1.$$

Thus we have $\dim_{\mathrm{H}} C = \frac{\log 2}{\log 6}$ by Hutchinson's theorem [10], where $\dim_{\mathrm{H}} F$ denotes the Hausdorff dimension of F. In particular, C is uncountable since the Hausdorff dimension of C is positive. Therefore, we can not characterized the existence of APs by using the cardinality of a given set.

Question 1. Is it true that a set of real numbers with Hausdorff dimension one contains arbitrarily long arithmetic progressions?

We may guess that the answer to Question 1 is "yes" because of Ramsey-type phenomena like Szemerédi's theorem and van der Waerden's theorem. Surprisingly, the answer is no. It is due to Keleti in 1998 [11]. He proved that there exists a compact set A of real numbers with Hausdorff dimension one such that the solutions $x, y, z, w \in A$ to the equation

$$x + y = z + w, \quad x < y, \quad z < w \tag{4}$$

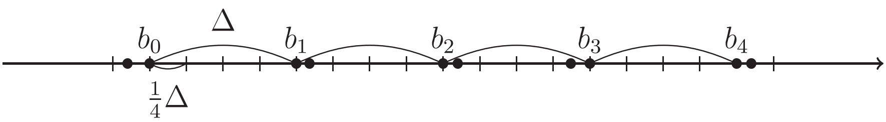
are only $x = z$ and $y = w$. If $y = z$ in the equation (4), sets $\{x, w, y\}$ are APs of length 3. Therefore A does not contain APs of length 3. Therefore, the existence of APs can not be characterized by the Hausdorff dimension.

3 Weak arithmetic progressions

For all $0 \leq \varepsilon < 1/2$, we say that a real sequence $(a_i)_{i=0}^{k-1}$ is a (k, ε)-AP if there exists an arithmetic progression $(b_i)_{i=0}^{k-1}$ of length k with gap length $\Delta > 0$ such that

$$|a_i - b_i| \leq \varepsilon\Delta$$

for all $i = 0, 1, \dots, k-1$. For example, when $k = 5$, $\varepsilon = 1/4$, the following sequence of points near to b_i's is an $(5, 1/4)$-AP:



The original notion of (k,ϵ)-APs can be seen in [5]. In the literature, the term "(k,ϵ)-APs" was firstly appeared in [4]. The existence of (k,ϵ)-APs can be caracterized by the Assouad dimension. Here for all bounded sets $E \subset \mathbb{R}^d$ and for all positive real number r, we define $N(E,r)$ as the minimum cardinality of a family of sets with diameter less than or equal to r which covers E. For all $F \subseteq \mathbb{R}^d$, we define

$$\dim_{\mathrm{A}} F = \inf \Bigg\{ \sigma \geq 0\colon \text{there exists } C > 0 \text{ such that for all } 0 < r < R \text{ and } x \in F, \text{ we have } N(B(x,R)\cap F, r) \leq C\left(\frac{R}{r}\right)^{\sigma} \Bigg\}$$

which is called the Assouad dimension of F. Here $B(x,R)$ denotes the closed ball of $\mathbb{R}^d$ with radius R centered at x. By the result of Fraser and Yu [5], a subset F of real numbers contains (k,ϵ)-APs for all $k \geq 3$ and $\epsilon \in (0,1/2)$ if and only if $\dim_{\mathrm{A}} F = 1$. Further, they proved that a subset of positive integers whose sum of reciprocals is divergent contains (k,ϵ)-APs for all $k \geq 3$ and $\epsilon \in (0,1/2)$ [5]. Fraser, the author, and Yu gave quantitative bounds for the Assouad (Hausdorff) dimension of a set which does not contain (k,ϵ)-APs for fixed k and ϵ. More precisely, for every $F \subset \mathbb{R}$, $k \geq 3$, and $\varepsilon \in (0,1/2)$, if F does not contain any (k,ε)-APs, then we have

$$\dim_{\mathrm{A}} F \leq 1 + \frac{\log(1-1/k)}{\log(k\lceil 1/(2\epsilon)\rceil)}. \tag{5}$$

Note that for every set $F \subseteq \mathbb{R}^d$, the following inequality holds:

$$\dim_{\mathrm{L}} F \leq \dim_{\mathrm{H}} F \leq \dim_{\mathrm{P}} F \leq \dim_{\mathrm{A}} F,$$

where $\dim_{\mathrm{L}} F$ denotes that the lower dimension of F, and $\dim_{\mathrm{P}} F$ denotes that the packing dimension of F. We can see this formula in [3]. Hence one can replace $\dim_{\mathrm{A}}$ in (5) by $\dim_{\mathrm{X}}$ for all $\mathrm{X} \in \{\mathrm{L},\mathrm{H},\mathrm{P}\}$. Remark that if $F \subset \mathbb{R}^d$ is bounded, then the upper box dimension of F is less than or equal to the Assouad dimension of it. Thus we can also replace the Assouad dimension by the upper box dimension in (5) if we restrict that F is bounded. Now define

$$D_{\mathrm{A}}(k,\varepsilon) := \sup\{\dim_{\mathrm{A}} F \mid F \subseteq \mathbb{R} \text{ does not contain } (k,\epsilon)\text{-APs}\}.$$

We can consider $D_{\mathrm{A}}(k,\epsilon)$ as an analogue of $r_k(N)$. The right hand side of (5) is an upper bound for $D_{\mathrm{A}}(k,\epsilon)$. In order to find effective lower bounds, we should construct a subset of real numbers which has large dimension and does not contain (k,ϵ)-APs. Fraser, the author, and Yu also gave a lower bound for $D_{\mathrm{A}}(k,\epsilon)$. For all $k \geq 3$ and $\epsilon \in (0,1)$ satisfying $\epsilon < (k-2)/4$, there exists a set $F \subset \mathbb{R}$ which does not contain any (k,ϵ)-APs and

$$\dim_{\mathrm{A}} F = \dim_{\mathrm{H}}\ F = \frac{\log 2}{\log \frac{2k-2-4\varepsilon}{k-2-4\varepsilon}}. \tag{6}$$

Now define $D_{\mathrm{H}}(k,\varepsilon) := \sup\{\dim_{\mathrm{H}} F \mid F \subseteq \mathbb{R} \text{ does not contain } (k,\varepsilon)\text{-APs}\}$. By combining (5), (6), and the inequality $\dim_{\mathrm{H}} F \leq \dim_{\mathrm{A}} F$, we have

$$\frac{\log 2}{\log \frac{2k-2-4\varepsilon}{k-2-4\varepsilon}} \leq D_{\mathrm{H}}(k,\varepsilon) \leq D_{\mathrm{A}}(k,\varepsilon) \leq 1 + \frac{\log(1-1/k)}{\log\left(k\lceil 1/(2\varepsilon)\rceil\right)}. \tag{7}$$

4 Main results

This section gives the results which the author states in the conference These results are one dimensional cases in the paper [13].

Theorem 2. For every $k \geq 3$ and $0 < \varepsilon < 1/2$, we have

$$D_{\mathrm{A}}(k,\varepsilon) \leq \frac{1}{2}\frac{\log(r_k(\lceil 1/\varepsilon\rceil)\lceil 1/\varepsilon\rceil)}{\log(\lceil 1/\varepsilon\rceil)}.$$

By Gowers' upper bounds , we have

$$D_{\mathrm{A}}(k,\epsilon) \leq 1 - \frac{1}{2^{1+2^{k+9}}}\frac{\log\log\log\lceil 1/\epsilon\rceil}{\log\lceil 1/\epsilon\rceil}. \tag{8}$$

This upper bound is better than (5) if ϵ is sufficiently small. Actually, if $0 < \epsilon < \exp(-\exp\exp\exp(2+2^{k+9}))$, then one has (8) < (5).

Theorem 3. Fix $k \geq 3$ and $0 < \epsilon < 1/16$. Then we can construct a compact set $F \subset \mathbb{R}$ which satisfies the following conditions:

(i) there exists a set of linear contractions $\{f_1,\ldots,f_m\}$ with open set condition such that F is the attractor of $\{f_1,\ldots,f_m\}$ *i.e.*
$$F = \bigcup_{i=1}^{m} f_i(F);$$

(ii) F does not contain any (k,ϵ)-APs;

(iii) $\dim_{\mathrm{H}} F \geq \dfrac{\log r_k(\lfloor 1/(8\epsilon)\rfloor)}{\log(1/\epsilon)}$.

Here $f : \mathbb{R} \to \mathbb{R}$ is called a linear contraction if $f(x) = ax + b$ for some $|a| < 1$. Define $D_{\mathrm{S}}(k, \varepsilon) = \sup\{\dim_{\mathrm{H}} F \mid F$ satisfies (i) and (ii) in Theorem 3$\}$. Then we have

$$\frac{\log r_k(\lfloor 1/(8\epsilon)\rfloor)}{\log(1/\epsilon)} \leq D_{\mathrm{S}}(k,\varepsilon) \leq D_{\mathrm{H}}(k,\varepsilon) \leq D_{\mathrm{A}}(k,\varepsilon) \leq \frac{1}{2}\frac{\log(r_k(\lceil 1/\varepsilon\rceil)\lceil 1/\varepsilon\rceil)}{\log(\lceil 1/\varepsilon\rceil)}.$$

By O'Bryant's lower bound for $r_k(N)$, we get a better lower bound for $D_{\mathrm{S}}(k, \epsilon)$ for sufficiently small $0 < \epsilon < \epsilon(k)$.

Corollary 4. There exists positive real numbers C and D such that For every $0 < \epsilon < 1/16$ and $k \geq 3$, we have

$$C\frac{r_k(\lfloor 1/(8\epsilon)\rfloor)}{\lfloor 1/(8\epsilon)\rfloor} \leq \epsilon^{1-D_{\mathrm{S}}(k,\epsilon)} \leq \epsilon^{1-D_{\mathrm{H}}(k,\epsilon)} \leq \epsilon^{1-D_{\mathrm{A}}(k,\epsilon)} \leq E\left(\frac{r_k(\lceil 1/\epsilon\rceil)}{\lceil 1/\epsilon\rceil}\right)^{1/2}.$$

Hence for fixed $k \geq 3$. It is known that a subset of positive integers with positive upper density contains arithmetic progressions of length k if and only if $\lim_{N\to\infty} r_k(N)/N = 0$. Therefore from Corollary 4, the following are equivalent:

(i) a subset of positive integers with positive upper density contains arithmetic progressions of length k;

(ii) $\lim_{N\to\infty} r_k(N)/N = 0$;

(iii) $\lim_{\epsilon\to+0} \epsilon^{1-D_{\mathrm{X}}(k,\epsilon)} = 0$ for some $\mathrm{X} \in \{\mathrm{S}, \mathrm{H}, \mathrm{A}\}$;

(iv) $\lim_{\epsilon\to+0} \epsilon^{1-D_{\mathrm{X}}(k,\epsilon)} = 0$ for all $\mathrm{X} \in \{\mathrm{S}, \mathrm{H}, \mathrm{A}\}$.

Actually, by Szemerédi's theorem, (i),(ii),(iii), and (iv) are true. If one showed (iii) for all $k \geq 3$, then we would get another proof of Szemerédi's theorem. Remark that by (5), implies that

$$\epsilon^{1-D_{\mathrm{A}}(k,\epsilon)} \leq \epsilon^{-\frac{\log(1-1/k))}{\log k\lceil 1/(2\varepsilon)\rceil}} = (1-1/k)^{\frac{\log(1/\epsilon)}{\log k\lceil 1/(2\varepsilon)\rceil}} \to 1 - 1/k$$

as $\epsilon \to +0$. Thus the upper bound given by Fraser, the author, and Yu does not reach to prove Szemerédi's theorem.

5 Sketch of proofs

Sketch of Proof of Theorem 2. Fix any $k \geq 3$ and $\epsilon \in (0, 1/2)$. Take any $F \subseteq \mathbb{R}$ which does not contain (k, ϵ)-APs. Take any interval $I = [a, a + 2R]$ centered at a point in F. Fix any positive integer N. We partition I into N/ϵ small closed intervals A_j with diameter $2R\epsilon/N$ as follows:

$$A_j = [a + 2jR\epsilon/N,\ a + 2(j+1)R\epsilon/N]$$

for all $j = 0, 1, \ldots, N/\epsilon - 1$. For any $0 \leq i < 1/\epsilon$, let

$$T(i) = \{i + j/\epsilon \colon 0 \leq j \leq N - 1\}.$$

Fix $0 \leq i < 1/\epsilon$. Since F does not contain any arithmetic progressions of length k, the number of $j \in T(i)$ such that $A_j \cap F \neq \emptyset$ is less than or equal to $r_k(N)$. Hence the number of $j \in \{0, 1, \ldots, N/\epsilon - 1\}$ such that $F \cap A_j \neq \emptyset$ is less than or equal to $r(N)/\epsilon$. We iterate this argument t-times for each smaller itervals which intersect F, where t is a real number such that

$$2R\left(\frac{\epsilon}{N}\right)^t = r.$$

Therefore we obtain that

$$N(F \cap I, r) \leq \left(\frac{r_k(N)}{\epsilon}\right)^t \leq \left(\frac{2R}{r}\right)^{\frac{\log(r_k(N)/\epsilon)}{\log(N/\epsilon)}}.$$

Therefore we conclude that

$$\dim_{\mathrm{A}} F \leq \frac{\log(r_k(N)/\epsilon)}{\log(N/\epsilon)}.$$

□

Sketch of Proof of Theorem 3. Fix any $k \geq 3$. Let $N = \lfloor 1/(8\epsilon) \rfloor$. Let A be a subset of $\{0, 1, \ldots, N-1\}$ with $|A| = r_k(N)$ which does not contain arithmetic progressions of length k. Let $\delta = 1/16$ and define

$$\psi_a(x) = \frac{x}{N - 1 + \delta} + a \quad (a \in A), \tag{9}$$

$$I_0 = [0, N - 1 + \delta], \quad I_n = \bigcup_{a \in A} \psi_a(I_{n-1}) \quad (n \geq 1), \tag{10}$$

and define $F = \bigcap_{n=0}^{\infty} I_n$. Then F is the attractor of the iterated function sysytem $\{\psi_a \colon a \in A\}$ which satisfies open set condition. Thus one has

$$\dim_{\mathrm{H}} F \geq \frac{\log r_k(\lfloor 1/(8\epsilon) \rfloor)}{\log(1/\epsilon)}.$$

By contradiction, we can see that F does not contain (k, ϵ)-APs. □

We refer [13] to the readers who want to know more details.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers 19J20878.

References

[1] T. F. Bloom. A quantitative improvement for Roth's theorem on arithmetic progressions, *J. London Math. Soc.* **93** (2016), 643–663.

[2] P. Erdős and P. Turán. On some sequences of integers, *J. London Math. Soc.* **11** (1936), 261–264.

[3] J. M. Fraser. Assouad type dimensions and homogeneity of fractals, *Trans. Amer. Math. Soc.*, **366** (2014), 6687–6733.

[4] J. M. Fraser, K. Saito and H. Yu. Dimensions of sets which uniformly avoid arithmetic progressions. *Int. Math. Res. Not. IMRN.*, 2017, rnx261.

[5] J. M. Fraser and H. Yu. Arithmetic patches, weak tangents, and dimension, *Bull. Lond. Math. Soc.*,**50** (2018), 85-95.

[6] W.T. Gowers. A new proof of Szemerédi's theorem, *Geom. Funct. Anal.*, **11** (2001), 465—588.

[7] W.T. Gowers. Erdős and arithmetic progressions, Erdős Centennial, Bolyai Society Mathematical Studies, 25, L. Lovasz, I. Z. Ruzsa, V. T. Sos eds., Springer 2013, 265-287.

[8] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions, *Ann. of Math. (2)*, **167** (2008), 481– 547.

[9] B. Green and T. Tao. New bounds for Szemerédi's theorem, III: a polylogarithmic bound for $r_4(N)$, *Mathematika* **63** (2017), 944–1040.

[10] J. E. Hutchinson. Fractals and self-similarity, *Indiana Univ. Math. J.*, **30** (1981), 713–747.

[11] T. Keleti, A 1-dimensional subset of the reals that intersects each of its translates in at most a single point, *Real Anal. Exchange*, 24:2 (1998), 843–844.

[12] K. O'Bryant. Sets of integers that do not contain long arithmetic progressions, *Electron. J. Combin.*, **18** (2011), Paper 59, 15 pp.

[13] K. Saito. New bounds for dimensions of a set uniformly avoiding multi-dimensional arithmetic progressions, *preprint*, 2019, available at https://arxiv.org/abs/1910.13071.

[14] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression, Collection of articles in memory of Juriĭ Vladimirovič Linnik. *Acta Arith.*, **27** (1975), 199–245.

[15] B. L. van der Waerden. Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wisk.* **15** (1927), 212–216.