

How many points contain arithmetic progressions in their continued fraction expansion?

by

XIN TONG and BAOWEI WANG (Wuhan)

1. Introduction. Continued fraction expansion is induced by the Gauss transformation $T : [0, 1) \rightarrow [0, 1)$ given by

$$T(0) := 0, \quad T(x) := \frac{1}{x} \pmod{1}, \quad 0 < x < 1,$$

in the sense that every irrational $x \in [0, 1)$ has a unique infinite expansion

$$(1.1) \quad x = \frac{1}{a_1(x) + \frac{1}{a_2(x) + \frac{1}{a_3(x) + \ddots}}}} = [a_1(x), a_2(x), a_3(x), \dots],$$

where $a_n(x) \in \mathbb{N}$, for all $n \geq 1$, are called the partial quotients of x .

Namely, the Gauss transformation supplies a mechanism attaching to each irrational $x \in [0, 1)$ an infinite integer sequence $\{a_n(x)\}_{n=1}^{\infty}$. Furthermore, if the sequence $\{a_n(x)\}_{n=1}^{\infty}$ is strictly increasing, there is a one-to-one correspondence between such irrational numbers and an integer subset. Szemerédi [6] showed that an integer subset contains arbitrarily long arithmetic progressions if it is of positive density in \mathbb{N} . Inspired by Szemerédi's theorem, we are led to ask naturally whether a point is a Szemerédi point (see definition below) and how large the set of such points is.

To make the statement precise, we introduce the following definitions.

DEFINITION 1.1. An irrational $x \in [0, 1)$ is called a *Szemerédi point* if the sequence of its partial quotients is strictly increasing and contains arbitrarily long arithmetic progressions.

2010 *Mathematics Subject Classification*: 11A55, 11B25, 28A80.

Key words and phrases: Szemerédi point, arithmetic progressions, continued fractions, Hausdorff dimension.

B. W. Wang is the corresponding author.

DEFINITION 1.2. An irrational $x \in [0, 1)$ is called an *absolute Szemerédi point* if it is a Szemerédi point and the sequence of its partial quotients contains arbitrarily long arithmetic progressions with arbitrary common difference.

Denote by Sz and ASz the sets of Szemerédi points and absolute Szemerédi points respectively. In this note, we show

THEOREM 1.3. $\dim_{\mathbb{H}} \text{Sz} = \dim_{\mathbb{H}} \text{ASz} = 1/2$, where $\dim_{\mathbb{H}}$ denotes the Hausdorff dimension.

2. Preliminaries. This section is devoted to collecting some elementary properties enjoyed by continued fractions, and recalling some known results, which will be used later.

We begin with some notation. For any $n \geq 1$ and $(a_1, \dots, a_n) \in \mathbb{N}^n$, call

$$I_n(a_1, \dots, a_n) = \begin{cases} \left[\frac{p_n}{q_n}, \frac{p_n + p_{n-1}}{q_n + q_{n-1}} \right) & \text{when } n \text{ is even,} \\ \left(\frac{p_n + p_{n-1}}{q_n + q_{n-1}}, \frac{p_n}{q_n} \right] & \text{when } n \text{ is odd,} \end{cases}$$

an *n*th order interval, where p_n, q_n are defined recursively by

$$(2.1) \quad p_{-1} = 1, \quad p_0 = 0, \quad p_k = a_k p_{k-1} + p_{k-2}, \quad 1 \leq k \leq n,$$

$$(2.2) \quad q_{-1} = 0, \quad q_0 = 1, \quad q_k = a_k q_{k-1} + q_{k-2}, \quad 1 \leq k \leq n.$$

In fact, $I_n(a_1, \dots, a_n)$ just represents the set of numbers in $[0, 1)$ which have a continued fraction expansion beginning with a_1, \dots, a_n , i.e.

$$I_n(a_1, \dots, a_n) = \{x \in [0, 1) : a_1(x) = a_1, \dots, a_n(x) = a_n\}.$$

For each irrational $x \in [0, 1)$ and $n \geq 1$, let $p_n(x) = p_n, q_n(x) = q_n$ be defined recursively by (2.1) and (2.2). Then $p_n(x)/q_n(x)$ is called the *n*th convergent of x .

PROPOSITION 2.1 ([4, 5]). For any $n \geq 1$ and $(a_1, \dots, a_n) \in \mathbb{N}^n$, let q_n be given by (2.2). Then

$$q_n \geq 2^{(n-1)/2}, \quad \prod_{k=1}^n a_k \leq q_n \leq \prod_{k=1}^n (a_k + 1), \quad |I_n(a_1, \dots, a_n)| = \frac{1}{q_n(q_n + q_{n-1})}.$$

For any $n \geq 1$ and $1 \leq k \leq n$,

$$1 \leq \frac{q_n(a_1, \dots, a_k, a_{k+1}, \dots, a_n)}{q_k(a_1, \dots, a_k) \cdot q_{n-k}(a_{k+1}, \dots, a_n)} \leq 2.$$

LEMMA 2.2 ([7]). For any $n \geq 1$ and $1 \leq k \leq n$,

$$\frac{a_k + 1}{2} \leq \frac{q_n(a_1, \dots, a_k, a_{k+1}, \dots, a_n)}{q_{n-1}(a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_n)} \leq a_k + 1.$$

LEMMA 2.3 ([3]). $\dim_{\mathbb{H}}\{x \in [0, 1] : a_n(x) \rightarrow \infty \text{ as } n \rightarrow \infty\} = 1/2$.

To end this section, we state a mass distribution principle (also known as the Billingsley Theorem), which is a classical tool to obtain a lower bound on the Hausdorff dimension of a fractal set.

THEOREM 2.4 ([1, 2]). *Let $E \subset [0, 1]$ be a Borel set and μ a measure with $\mu(E) > 0$. If for any $x \in E$,*

$$\lim_{r \rightarrow \infty} \frac{\log \mu(B(x, r))}{\log r} = s,$$

where $B(x, r)$ is the ball with center x and radius r , then $\dim_{\mathbb{H}} E = s$.

3. Proof of Theorem 1.3. Lemma 2.3 supplies the desired upper bound on $\dim_{\mathbb{H}} \text{Sz}$ and $\dim_{\mathbb{H}} \text{ASz}$, i.e.,

$$\dim_{\mathbb{H}} \text{Sz} \leq 1/2, \quad \dim_{\mathbb{H}} \text{ASz} \leq 1/2.$$

Consequently, we only need to prove the lower bound. Our strategy is:

STEP 1. Construct a subset E_{ASz} of ASz .

STEP 2. Establish a connection between E_{ASz} and some regularly defined set E by means of a $(1 - \epsilon)$ -Hölder function.

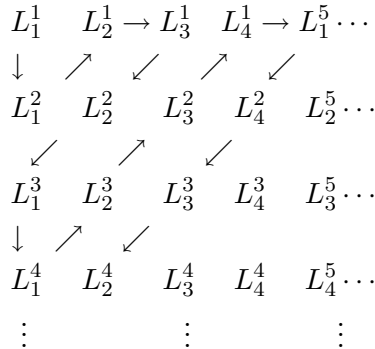
If these two procedures are done properly, then

$$\dim_{\mathbb{H}} \text{ASz} \geq (1 - \epsilon) \dim_{\mathbb{H}} E.$$

So, we are led to the last step.

STEP 3. Estimate $\dim_{\mathbb{H}} E$.

3.1. Subset of ASz . For any $k \geq 1$ and $d \geq 1$, let $L_k^d = \{d, 2d, \dots, kd\}$ be an arithmetic progression with length k and common difference d . We arrange $\{L_k^d : k, d \geq 1\}$ in the following order.



We denote the new sequence by $\{L_k : k \geq 1\}$. It is easy to observe that, for any $k \geq 1$,

$$(3.1) \quad |L_k| \leq k \quad \text{and} \quad \max L_k \leq k^2,$$

where $|L|$ denotes the cardinality of L . Without any confusion, sometimes L_k is considered as a vector with its elements arranged in increasing order.

Let $\alpha \geq 2$ be some integer and set

$$E(\alpha) = \{x \in [0, 1) : (2n)^\alpha \leq a_n(x) < (2n + 1)^\alpha \text{ for all } n \geq 1\}.$$

For every $x \in E(\alpha)$, we will construct an absolute Szemerédi point y . This is achieved by inserting the vectors $\{L_k : k \geq 1\}$ in the continued fraction expansion of x at appropriate positions.

Let $\{n_k : k \geq 0\}$ be a sequence such that $n_0 = 0$ and for all $k \geq 1$,

$$(3.2) \quad (2n_k + 1)^\alpha + k^2 < (2n_k + 2)^\alpha \quad \text{and} \quad \lim_{k \rightarrow \infty} \frac{k^2 \log n_k}{n_k} = 0.$$

For each $x \in E(\alpha)$, an absolute Szemerédi point y is constructed as follows. For $n \leq n_1 + |L_1|$, set

$$a_n(y) = a_n(x), \quad 1 \leq n \leq n_1, \quad (a_{n_1+1}(y), \dots, a_{n_1+|L_1|}(y)) = a_{n_1}(x) + L_1.$$

For $n > n_1 + |L_1|$, let $k \geq 1$ be the integer such that $n_k + |L_1| + \dots + |L_k| < n \leq n_{k+1} + |L_1| + \dots + |L_k| + |L_{k+1}|$.

(i) When $n_k + |L_1| + \dots + |L_k| < n \leq n_{k+1} + |L_1| + \dots + |L_k|$, set

$$a_n(y) = a_{n-(|L_1|+\dots+|L_k|)}(x).$$

(ii) When $n_{k+1} + |L_1| + \dots + |L_k| < n \leq n_{k+1} + |L_1| + \dots + |L_k| + |L_{k+1}|$, set

$$(a_{n_{k+1}+|L_1|+\dots+|L_k|+1}(y), \dots, a_{n_{k+1}+|L_1|+\dots+|L_k|+|L_{k+1}|}(y)) = a_{n_{k+1}}(x) + L_{k+1}.$$

By the above construction, we conclude that $y \in \text{ASz}$. This is because, on the one hand, all arithmetic progressions occur in the sequence $\{a_n(y)\}_{n \geq 1}$, and on the other hand, the large gaps between $a_{n_k}(x)$ and $a_{n_{k+1}}(x)$ (see formula (3.2)), for each $k \geq 1$, guarantee that the sequence $\{a_n(y)\}_{n \geq 1}$ is still strictly increasing.

We call x the *generator* of y and denote by $E_{\text{ASz}}(\alpha)$ the collection of y 's given by the above procedure. Evidently, we have

FACT 1. $\dim_{\mathbb{H}} \text{ASz} \geq \dim_{\mathbb{H}} E_{\text{ASz}}(\alpha)$.

3.2. $(1 - \epsilon)$ -Hölder function. For any $\epsilon > 0$, choose n_{k_0} large enough such that for every $n \geq n_{k_0}$ and $k \geq k_0$,

$$(3.3) \quad 32(2n + 3)^{2\alpha} \leq 2^{(n-1)\epsilon/2}, \quad 2^{2k+1} \prod_{j=1}^k ((2n_j + 1)^\alpha + j^2)^j \leq 2^{(n_k-1)\epsilon/2}.$$

Fix $a_1, \dots, a_{n_{k_0}}$ with $(2j)^\alpha \leq a_j < (2j + 1)^\alpha$ for $1 \leq j \leq n_{k_0}$. Define

$$E(a_1, \dots, a_{n_{k_0}}) := E(\alpha) \cap I(a_1, \dots, a_{n_{k_0}})$$

and denote by $E_{ASz}(a_1, \dots, a_{n_{k_0}})$ the corresponding set of absolute Sze-merédi points.

Now we define a map

$$f : E_{ASz}(a_1, \dots, a_{n_{k_0}}) \rightarrow E(a_1, \dots, a_{n_{k_0}}), \quad y \mapsto x,$$

where x is the generator of y .

CLAIM. f is $1/(1 + \epsilon)^2$ -Hölder.

For any pair $y_1, y_2 \in E_{ASz}(a_1, \dots, a_{n_{k_0}})$, let x_1, x_2 be the generators of y_1 and y_2 respectively. Denote by n the smallest integer such that $a_{n+1}(y_1) \neq a_{n+1}(y_2)$. Then $n \geq n_{k_0} + |L_1| + \dots + |L_{k_0}|$. Assume that

$$n_k + |L_1| + \dots + |L_k| \leq n < n_{k+1} + |L_1| + \dots + |L_k| + |L_{k+1}|$$

for some $k \geq k_0$.

First, we establish a relationship between $q_{n_k+|L_1|+\dots+|L_k|}(y_1)$ and $q_{n_k}(x_1)$. By Proposition 2.1 and Lemma 2.2, we can see that

$$\begin{aligned} (3.4) \quad q_{n_k+|L_1|+\dots+|L_k|}(y_1) &\leq 2q_{|L_k|}(a_{n_k}(x_1) + L_k)q_{n_k+|L_1|+\dots+|L_{k-1}|}(y_1) \\ &\leq 2((2n_k + 1)^\alpha + k^2)^k q_{n_k+|L_1|+\dots+|L_{k-1}|}(y_1) \\ &\leq 2^2((2n_k + 1)^\alpha + k^2)^k q_{n_{k-1}+|L_1|+\dots+|L_{k-1}|}(y_1) \\ &\quad \times q_{n_k-n_{k-1}}(a_{n_{k-1}+1}(x_1), \dots, a_{n_k}(x_1)) \\ &\quad \vdots \\ &\leq 2^{2k} \prod_{j=1}^k ((2n_j + 1)^\alpha + j^2)^j q_{n_j-n_{j-1}}(a_{n_{j-1}+1}(x_1), \dots, a_{n_j}(x_1)) \\ &\leq 2^{2k} \prod_{j=1}^k ((2n_j + 1)^\alpha + j^2)^j q_{n_k}(x_1) \leq \frac{1}{2} q_{n_k}^{1+\epsilon}(x_1). \end{aligned}$$

Secondly, we estimate the gap between y_1 and y_2 . Note that

$$a_{n+2}(y_1) \geq a_{n_k+|L_1|+\dots+|L_k|}(y_1) \geq a_{n_k}(x_1) \geq (2n_k)^\alpha,$$

and the same for y_2 . So,

$$y_1 \in \bigcup_{a_{n+2} \geq (2n_k)^\alpha} I_{n+2}(a_1(y_1), \dots, a_{n+1}(y_1), a_{n+2}) \subset I_{n+1}(a_1(y_1), \dots, a_{n+1}(y_1)),$$

$$y_2 \in \bigcup_{a_{n+2} \geq (2n_k)^\alpha} I_{n+2}(a_1(y_2), \dots, a_{n+1}(y_2), a_{n+2}) \subset I_{n+1}(a_1(y_2), \dots, a_{n+1}(y_2)).$$

Assume that n is even and $a_{n+1}(y_1) < a_{n+1}(y_2)$ (the other cases can be handled in the same manner). So, the gap between y_1 and y_2 is no less than the distance between the left endpoint of $I_{n+1}(a_1(y_1), \dots, a_{n+1}(y_1))$ and the

left endpoint of $I_{n+2}(a_1(y_1), \dots, a_{n+1}(y_1), (2n_k)^\alpha)$. Hence,

$$\begin{aligned} |y_1 - y_2| &\geq \frac{(2n_k)^\alpha p_{n+1}(y_1) + p_n(y_1)}{(2n_k)^\alpha q_{n+1}(y_1) + q_n(y_1)} - \frac{p_{n+1}(y_1) + p_n(y_1)}{q_{n+1}(y_1) + q_n(y_1)} \\ &= \frac{(2n_k)^\alpha - 1}{((2n_k)^\alpha q_{n+1}(y_1) + q_n(y_1))(q_{n+1}(y_1) + q_n(y_1))} \\ &\geq \frac{1}{32a_{n+1}^2(y_1)q_n^2(y_1)}. \end{aligned}$$

Since $a_{n+1}(y_1) \leq a_{n+1}(x_1) < (2n + 3)^\alpha$, by (3.3) we get

$$(3.5) \quad |y_1 - y_2| \geq \frac{1}{q_n^{2(1+\epsilon)}(y_1)}.$$

Thirdly, we estimate the gap between x_1 and x_2 . We will distinguish two cases.

- (i) $n_k + |L_1| + \dots + |L_k| \leq n < n_{k+1} + |L_1| + \dots + |L_k|$. In this case,

$$a_t(x_1) = a_t(x_2) \quad \text{for } 1 \leq t \leq t_0 = n - (|L_1| + \dots + |L_k|).$$

So,

$$(3.6) \quad |x_1 - x_2| \leq \frac{1}{q_{t_0}^2}.$$

Note that, by (3.4),

$$(3.7) \quad \begin{aligned} q_n(y_1) &\leq 2q_{t_0-n_k}(a_{n_k+1}(x_1), \dots, a_{t_0}(x_1))q_{n_k+|L_1|+\dots+|L_k|}(y_1) \\ &\leq q_{t_0-n_k}(a_{n_k+1}(x_1), \dots, a_{t_0}(x_1))q_{n_k}^{1+\epsilon}(x_1) \leq q_{t_0}^{1+\epsilon}(x_1). \end{aligned}$$

Combining (3.5), (3.6) and (3.7), we have

$$(3.8) \quad |f(y_1) - f(y_2)| = |x_1 - x_2| \leq |y_1 - y_2|^{1/(1+\epsilon)^2}.$$

- (ii) $n_{k+1} + |L_1| + \dots + |L_k| \leq n < n_{k+1} + |L_1| + \dots + |L_k| + |L_{k+1}|$. In this case,

$$a_t(x_1) = a_t(x_2) \quad \text{for } 1 \leq t \leq n_{k+1}.$$

So,

$$(3.9) \quad |x_1 - x_2| \leq \frac{1}{q_{n_{k+1}}^2}.$$

On the other hand,

$$q_n(y_1) \leq q_{n_{k+1}+|L_1|+\dots+|L_{k+1}|}(y_1) \leq q_{n_{k+1}}^{1+\epsilon}(x_1).$$

Thus,

$$(3.10) \quad |f(y_1) - f(y_2)| = |x_1 - x_2| \leq |y_1 - y_2|^{1/(1+\epsilon)^2}.$$

So, as f is $1/(1 + \epsilon)^2$ -Hölder and ϵ is arbitrary, we get

FACT 2. $\dim_{\mathbb{H}} E_{ASz}(\alpha) \geq \dim_{\mathbb{H}} E(\alpha)$.

3.3. Hausdorff dimension of $E(\alpha)$. For any $n \geq 1$, define

$$D_n = \{(a_1, \dots, a_n) \in \mathbb{N}^n : (2k)^\alpha \leq a_k < (2k + 1)^\alpha, 1 \leq k \leq n\}.$$

For each $n \geq 1$ and $(a_1, \dots, a_n) \in D_n$, we call $I_n(a_1, \dots, a_n)$ an *admissible cylinder* of order n with respect to $E(\alpha)$. Then it is evident that

$$E(\alpha) = \bigcap_{n=1}^{\infty} \bigcup_{(a_1, \dots, a_n) \in D_n} \text{cl } I_n(a_1, \dots, a_n),$$

where “cl” denotes closure. Now we define a set function on all admissible cylinders by setting

$$\mu(I_n(a_1, \dots, a_n)) = \frac{1}{\#D_n} \quad \text{for all } n \geq 1.$$

By the Carathéodory extension theorem, μ can be extended to a probability measure supported on $E(\alpha)$.

We will apply the Billingsley Theorem 2.4 to estimate $\dim_{\mathbb{H}} E(\alpha)$. So we are led to estimate the μ -measure of arbitrary balls $B(x, r)$ with center $x \in E(\alpha)$ and radius $r > 0$ small enough.

For each $x \in E(\alpha)$, there exists a sequence (a_1, a_2, \dots) such that for each $n \geq 1$, $x \in I_n(a_1, \dots, a_n)$ and $(a_1, \dots, a_n) \in D_n$. For any $0 < r < 1/9^\alpha$, there exists $n \geq 1$ such that

$$|I_{n+1}(a_1, \dots, a_{n+1})| < r \leq |I_n(a_1, \dots, a_n)|.$$

Then $B(x, r)$ can intersect at most five n th order admissible cylinders and at least one $(n + 1)$ th order admissible cylinder. So

$$\frac{\log 5\mu(I_n(a_1, \dots, a_n))}{\log |I_{n+1}(a_1, \dots, a_{n+1})|} \leq \frac{\log \mu(B(x, r))}{\log r} \leq \frac{\log \mu(I_{n+1}(a_1, \dots, a_{n+1}))}{\log |I_n(a_1, \dots, a_n)|}.$$

Since

$$\lim_{n \rightarrow \infty} \frac{\log 5\mu(I_n(a_1, \dots, a_n))}{\log |I_{n+1}(a_1, \dots, a_{n+1})|} = \lim_{n \rightarrow \infty} \frac{\log \mu(I_{n+1}(a_1, \dots, a_{n+1}))}{\log |I_n(a_1, \dots, a_n)|} = \frac{\alpha - 1}{2\alpha},$$

it follows that, for each $x \in E(\alpha)$,

$$(3.11) \quad \lim_{r \rightarrow 0} \frac{\log \mu(B(x, r))}{\log r} = \frac{\alpha - 1}{2\alpha}.$$

So, by the Billingsley Theorem 2.4, we obtain

FACT 3. $\dim_{\mathbb{H}} E(\alpha) = \frac{\alpha-1}{2\alpha}$.

Proof of Theorem 1.3. Facts 1–3 give directly $\dim_{\mathbb{H}} \text{ASz} \geq \frac{\alpha-1}{2\alpha}$. Letting $\alpha \rightarrow \infty$, we get $\dim_{\mathbb{H}} \text{ASz} \geq 1/2$. ■

Acknowledgements. The authors express their sincere thanks to Professor J. Wu and the referee for many helpful suggestions.

References

- [1] P. Billingsley, *Ergodic Theory and Information*, Wiley, New York, 1965.
- [2] K. J. Falconer, *Fractal Geometry. Mathematical Foundations and Applications*, Wiley, 1990.
- [3] I. J. Good, *The fractional dimensional theory of continued fractions*, Proc. Cambridge Philos. Soc. 37 (1941), 199–228.
- [4] M. Iosifescu and C. Kraaikamp, *Metrical Theory of Continued Fractions*, Math. Appl. 547, Kluwer, Dordrecht, 2002.
- [5] A. Ya. Khintchine, *Continued Fractions*, Noordhoff, Groningen, 1963.
- [6] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.
- [7] J. Wu, *A remark on the growth of the denominators of convergents*, Monatsh. Math. 147 (2006), 259–264.

Department of Mathematics
Huazhong University of Science and Technology
430074 Wuhan, China
E-mail: tongxinxin@126.com
bwei_wang@yahoo.com.cn

*Received on 3.10.2008
and in revised form on 20.1.2009*

(5816)