Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
C. R. Acad. Sci. Paris, Ser. I 343 (2006) 129–133
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Dynamical Systems

Dimension of sets of sequences defined in terms of recurrence of their prefixes

Li Peng

Department of Mathematics, Wuhan University, Wuhan 430072, PR China

Received 21 February 2006; accepted 14 March 2006
Available online 27 June 2006
Presented by Jean-Pierre Kahane

Abstract

Let Σ be the set of sequences $x=(x_n)_{n\geqslant 1}$ of elements of $S=\{1,2,\dots,m\}$ endowed with the usual ultrametric $\mathrm{d}(x,y)=m^{-\inf\{k\geqslant 0:\ x_{k+1}\neq y_{k+1}\}}$. Let define

$$R_n(x)=\inf\{j>n:\ x_1x_2\cdots x_n=x_jx_{j+1}\cdots x_{j+n-1}\}.$$

We show that for any α and β such that $1\leqslant\alpha\leqslant\beta\leqslant\infty$ the Hausdorff dimension of the set

$$B_{\alpha,\beta}=\left\{x\in\Sigma:\ \varliminf_{n\to\infty}\frac{\log R_n(x)}{\log n}=\alpha \text{ and } \varlimsup_{n\to\infty}\frac{\log R_n(x)}{\log n}=\beta\right\}$$

is equal to 1. ***To cite this article: L. Peng, C. R. Acad. Sci. Paris, Ser. I 343 (2006).***

Résumé

Dimension d'ensembles de suites dont les préfixes réapparaissent en un temps prescrit. Soit Σ l'ensemble des suites $x=(x_n)_{n\geqslant 1}$ d'éléments de $S=\{1,2,\dots,m\}$ muni de l'ultramétrique usuelle $\mathrm{d}(x,y)=m^{-\inf\{k\geqslant 0:\ x_{k+1}\neq y_{k+1}\}}$. Posons

$$R_n(x)=\inf\{j>n:\ x_1x_2\cdots x_n=x_jx_{j+1}\cdots x_{j+n-1}\}.$$

Nous montrons que, quels que soient α et β tels que $1\leqslant\alpha\leqslant\beta\leqslant\infty$ l'ensemble

$$B_{\alpha,\beta}=\left\{x\in\Sigma:\ \varliminf_{n\to\infty}\frac{\log R_n(x)}{\log n}=\alpha \text{ et } \varlimsup_{n\to\infty}\frac{\log R_n(x)}{\log n}=\beta\right\}$$

a une dimension de Hausdorff égale à 1. ***Pour citer cet article : L. Peng, C. R. Acad. Sci. Paris, Ser. I 343 (2006).***

E-mail address: pengandtan@yahoo.com (L. Peng).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2006.05.005

Version française abrégée

Soit Σ l'ensemble des suites $x = (x_n)_{n \geqslant 1}$ d'éléments de l'alphabet $S = \{1, 2, \ldots, m\}$. Si I est un intervalle de $\mathbb{N}^*$, on note x_I le mot (suite finie ou non) composé des lettres x_j, avec $j \in I$. On munit Σ de la distance ultramétrique usuelle

$$\mathrm{d}(x, y) = m^{-\inf\{k \geqslant 0:\ x_{k+1} \neq y_{k+1}\}}$$

et l'on pose

$$R_n(x) = \inf\{j > n:\ x_{[j, j+n-1]} = x_{[1,n]}\}.$$

D'une part, Orstein et Weiss [2] on montré que, pour toute probabilité μ invariante par le décalage, le rapport $(\log R_n(x))/n$ tend pour μ-presque tout x vers l'entropie de μ par rapport au décalage. D'autre part, Feng et Wu [1] ont montré que, quels que soient α et β tels que $0 \leqslant \alpha \leqslant \beta \leqslant \infty$, l'ensemble des suites x telles que les limites inférieure et supérieure du rapport $(\log R_n(x))/n$ soient α et β a une dimension de Hausdorff égale à 1.

Il était dès lors naturel de s'intéresser aux suites pour lesquelles R_n se comporte comme une puissance. Voici notre résultat :

Théorème 0.1. *Si α et β vérifient $1 \leqslant \alpha \leqslant \beta \leqslant \infty$ on a*

$$\dim_H \left\{ x \in \Sigma:\ \varliminf_{n\to\infty} \frac{\log R_n(x)}{\log n} = \alpha,\ \varlimsup_{n\to\infty} \frac{\log R_n(x)}{\log n} = \beta \right\} = 1.$$

La démonstration de ce théorème repose sur le lemme suivant :

Lemme 0.2. *Soit $k = \{k_n\}_{n \geqslant 2}$ et $l = \{l_n\}_{n \geqslant 1}$ deux suites strictement croissantes de nombre entiers supérieurs ou égaux à 1 vérifiant les conditions*

$$l_{n+1} \geqslant l_n + k_{n+1} + 3 \quad \textit{pour}\ n \geqslant 1, \qquad l_n \geqslant k_{n+1} + 1 \quad \textit{et} \quad \lim_{n\to\infty} \frac{l_n}{\sum_{j=1}^n k_{j+1}} = \infty.$$

Alors l'ensemble

$$A^{k,l} = \bigl\{ x \in \Sigma:\ \exists N,\ \forall n \geqslant N,\ R_{k_n}(x) = \cdots = R_{k_{n+1}-1}(x) = l_n \bigr\}$$

a une dimension de Hausdorff égale à 1.

Démonstration. Soit q un entier supérieur à 3. Considérons l'ensemble

$$E_q = \bigl\{ x = (x_i)_{i \geqslant 1} \in \Sigma:\ x_j = m \text{ pour } 1 \leqslant j \leqslant q \text{ et } x_{kq+1} = x_{kq+q} = 1,\ \text{pour } k \geqslant 1 \bigr\}.$$

Sa dimension de Hausdorff est $(q-2)/q$.

Soit γ une bijection sans point fixe de l'ensemble S dans lui-même et n_0 un entier tel que $n_0 \geqslant 2$ et $k_{n_0+1} \geqslant q$.

Si $x \in E_q$ on définit par récurrence une suite $(x^n)_{n \geqslant n_0}$ de points de Σ :

$$x^{n_0} = x \quad \text{et,} \quad \text{pour } n \geqslant n_0, \quad x^{n+1} = x^n_{[1, l_n - 2]} 1 x^n_{[1, k_{n+1}-1]} \gamma\bigl(x^n_{k_{n+1}}\bigr) 1 x^n_{[l_n - 1, \infty)}.$$

La suite $(x^n)_{n \geqslant n_0}$ a une limite x^* appartenant à $A^{k,l}$. L'application $x \mapsto x^*$ est injective et son inverse est « presque lipschitzienne » (voir le sens exact de cette locution dans la version anglaise) de E_q dans $A^{k,l}$, ce qui montre que la dimension de $A^{k,l}$ est supérieure à $(q-2)/q$. □

Le théorème résulte alors de ce que, étant donnés α et β tels que $1 \leqslant \alpha \leqslant \beta \leqslant \infty$, on peut trouver deux suites k et l vérifiant outre les hypthèses du lemme précédent les conditions suivantes

$$\lim_{n\to\infty} \frac{\log k_n}{\log k_{n+1}} = 1, \qquad \varliminf_{n\to\infty} \frac{\log l_n}{\log k_n} = \alpha \quad \text{et} \quad \varlimsup_{n\to\infty} \frac{\log l_n}{\log k_n} = \beta.$$

1. Introduction

Let $S=\{1,2,\ldots,m\}$ be an m-letter alphabet ($m \geqslant 2$) and Σ be the set of one-sided sequences $x=(x_j)_{j\geqslant 1}$ over S. For such a sequence, if I is an interval of $\mathbb{N}^*$, we denote by x_I the word (finite or infinite) whose letters are the x_j with $j \in I$. Endowed with the following metric $\mathrm{d}(x,y)=m^{-\inf\{k\geqslant 0:\ x_{k+1}\neq y_{k+1}\}}$, Σ is a compact space.

For any $x=(x_i)_{i=1}^{\infty} \in \Sigma$ and positive integer n, define

$$R_n(x)=\inf\{j>n:\ x_j x_{j+1}\cdots x_{j+n-1}=x_1x_2\cdots x_n\},$$

that is, R_n is the first $j>n$ such that $\sigma^{j-1}(x)$ belongs to the n-cylinder $I_n(x)=\{y\in\Sigma:\ y_i=x_i \text{ for } 1\leqslant i\leqslant n\}$, where σ is the shift operator over Σ. We call $R_n(x)$ the nth recurrence time of x. Ornstein and Weiss [2] proved that, for each σ-invariant ergodic Borel probability measure μ on Σ,

$$\mu\left\{x\in\Sigma:\ \lim_{n\to\infty}\frac{\log R_n(x)}{n}=h_\mu(\sigma)\right\}=1,$$

where $h_\mu(\sigma)$ denotes the measure theoretic entropy of μ with respect to σ.

Then it is natural to inquire about the size of the exceptional set. Feng and Wu [1] considered the following set

$$A_{\alpha,\beta}=\left\{x\in\Sigma:\ \underline{\lim}_{n\to\infty}\frac{\log R_n(x)}{n}=\alpha \text{ and } \overline{\lim}_{n\to\infty}\frac{\log R_n(x)}{n}=\beta\right\}$$

and showed that, for $0\leqslant\alpha\leqslant\beta\leqslant\infty$, one has

$$\dim_H A_{\alpha,\beta}=1.$$

Thus there are many sequences which recur at a given exponential speed. What can we say about the polynomial speed? This is the subject of this Note.

2. Results and proofs

We consider the following sets:

$$B_{\alpha,\beta}=\left\{x\in\Sigma:\ \underline{\lim}_{n\to\infty}\frac{\log R_n(x)}{\log n}=\alpha \text{ and } \overline{\lim}_{n\to\infty}\frac{\log R_n(x)}{\log n}=\beta\right\}.$$

Our result is the following:

Theorem 1. *For any $\alpha,\beta\in[1,\infty]$ with $\alpha\leqslant\beta$, one has* $\dim_H B_{\alpha,\beta}=1$.

It should be pointed out that ideas and techniques to prove this theorem are reminiscent from [1]. In this latter work, the construction a big Cantor set contained in $A_{\alpha,\beta}$ relies on the rapid increase of $R_n(x)$: the letters between the places $R_n(x)+n$ and $R_{n+1}(x)$ can be chosen almost freely. However, in our case, $R_n(x)$ increases not so quickly (for example, if $\beta<2$, $R_{n+1}(x)-R_n(x)<n$ for any n in a subset of $\mathbb{N}$ of positive density). This causes difficulties, and we need more delicate constructions.

The proof of the above theorem strongly relies on the following lemma:

Lemma 2. *Let $k=\{k_n\}_{n\geqslant 2}$ and $l=\{l_n\}_{n\geqslant 1}$ be two increasing sequences of natural numbers satisfying the following conditions*:

(i) *For all $n\geqslant 1$, $l_{n+1}\geqslant l_n+k_{n+1}+3$,*
(ii) $l_n\geqslant k_{n+1}+1$,
(iii) $\lim_{n\to\infty}\dfrac{l_n}{\sum_{j=1}^{n}k_{j+1}}=\infty$.

Then the set

$$A^{k,l}=\left\{x\in\Sigma:\ \exists N,\ \forall n\geqslant N,\ R_{k_n}(x)=\cdots=R_{k_{n+1}-1}(x)=l_n\right\}$$

has Hausdorff dimension 1.

Proof. Denote $A = A^{k,l}$. Since $\dim_H \Sigma = 1$, it suffices to show that $\dim_H A \geqslant 1-\delta$, for any $\delta > 0$.

Fix $\delta > 0$. Choose an integer $q \geqslant 3$ such that $(q-2)/q > 1-\delta$, and define

$$E_q = \bigl\{x = (x_i)_{i\geqslant 1} \in \Sigma\colon\ x_j = m \text{ for } 1 \leqslant j \leqslant q \text{ and } x_{kq+1} = x_{kq+q} = 1,\ \text{for } k \geqslant 1\bigr\}.$$

Since the set $\{x = (x_i)_{i\geqslant 1} \in \Sigma\colon\ x_{kq+1} = x_{kq+q} = 1,\ \text{for}\ \ k \geqslant 0\}$ can be viewed as a self-similar set generated by m^{q-2} similitudes with ratio m^{-q}, its Hausdorff dimension is equal to $(\log m^{q-2})/(-\log m^{-q}) = (q-2)/q$. Therefore $\dim_H E_q = (q-2)/q > 1-\delta$.

In what follows, we will construct a one-to-one map φ from E_q into A satisfying the following condition: for any $\varepsilon > 0$, there exists a k_0 such that $\mathrm{d}(\varphi(x), \varphi(y)) < m^{-k}$ implies $\mathrm{d}(x, y) < m^{-k(1-\varepsilon)}$ for $k \geqslant k_0$. This means that $\dim_H \varphi(E_q) \geqslant \dim_H E_q$ and thus,

$$\dim_H A \geqslant \dim_H \varphi(E_q) \geqslant \dim_H E_q \geqslant 1-\delta,$$

as desired.

For $x = (x_i)_{i\geqslant 1} \in E_q$, we construct a sequence $\{x^n\}_{n\geqslant n_0}$ of points of Σ by induction, where n_0 is the smallest integer such that $n_0 \geqslant 2$ and $k_{n_0+1} > q$:

$$x^{n_0} = x,$$

and, for $n \geqslant n_0$,

$$x^{n+1} = x^n_{[1,l_n-2]} 1 x^n_{[1,k_{n+1}-1]} \gamma\bigl(x^n_{k_{n+1}}\bigr) 1 x^n_{[l_n-1,\infty)},$$

where γ is a permutation of S with no fixed point.

Since $l_{n+1} \geqslant l_n + k_{n+1} + 3$, the word $x^{n+1}_{[1,l_n+k_{n+1}+1]}$ is a prefix of x^{n+2}. Thus the sequence $\{x^n\}_{n\geqslant n_0}$ converges to a point $x^* \in \Sigma$.

One can easily check that blocks of q consecutive m's appear in x^* at positions 1 and l_j for $j \geqslant n_0$ only (this is the purpose of the presence of the 1's in the inserted words). Therefore, for any $\nu \leqslant q$ the word $x^*_{[1,\nu]}$ can only reappear at positions l_n (for $n \geqslant n_0$).

Consider $n \geqslant n_0$. One has $x^*_{[l_n,l_n+k_{n+1}-2]} = x^*_{[1,k_{n+1}-1]}$, because $k_{n+1} \leqslant l_n + 1$. We remark that, if $n > n_0$, the word $x^*_{[1,k_{n+1}]}$ cannot reappear before as $x^*_{[l_{n'},l_{n'}+k_{n+1}-1]}$ (with $n_0 \leqslant n' < n$) because $x^*_{[l_{n'},l_{n'}+k_{n'+1}}-1]} \neq x^*_{[1,k_{n'+1}]}$ (this is the purpose of inserting the letter $\gamma(x^n_{k_{n+1}})$).

Therefore, we have $R_\nu(x^*) = l_n$ for $n \geqslant n_0$ and $k_n \leqslant \nu < k_{n+1}$.

Now we define the map $\varphi : E_q \to A$ by $\varphi(x) = x^*$.

Obviously, φ is injective. Now, we are going to show that φ^{-1} is nearly Lipschitz, that is, for all $\varepsilon > 0$, there exists an integer k_0, such that $\mathrm{d}(x^*, y^*) \leqslant m^{-k}$ implies $\mathrm{d}(x, y) \leqslant m^{-k(1-\varepsilon)}$ for $k \geqslant k_0$.

In fact, since $\lim_{n\to\infty}(l_n/\sum_{j=1}^n k_{j+1}) = \infty$, for any $\varepsilon > 0$, there exists an integer $N \geqslant n_0$ such that $\sum_{j=1}^n (k_{j+1} + 2) < \varepsilon l_n$ for any $n \geqslant N$. Let $k_0 = l_N$, if $\mathrm{d}(x^*, y^*) \leqslant m^{-k}$ for some integer $k \geqslant k_0$, we have $x_1^* x_2^* \cdots x_k^* = y_1^* y_2^* \cdots y_k^*$. Let t be the integer such that $l_t \leqslant k < l_{t+1}$. Since $k \geqslant l_N$, it follows immediately $t \geqslant N \geqslant n_0$. By the construction of $x^*$and y^*, we have $x_1 x_2 \cdots x_{k'} = y_1 y_2 \cdots y_{k'}$ where $k' = k - \sum_{j=n_0}^{t-1}(k_{j+1}+2)$. Note that $k' > k - \sum_{j=1}^{t-1}(k_{j+1}+2) > k - \varepsilon l_t \geqslant k - \varepsilon k$. Then we have $\mathrm{d}(x, y) \leqslant m^{-k'} \leqslant m^{-(1-\varepsilon)k}$, as desired. □

To prove the theorem, it suffices to show that, given $\alpha, \beta \in [1,\infty]$ with $\alpha \leqslant \beta$, there exist sequences $\{k_n\}$ and $\{l_n\}$ meeting the conditions (i), (ii), and (iii) of Lemma 2 and satisfying

$$\lim_{n\to\infty} \frac{\log k_n}{\log k_{n+1}} = 1, \qquad \varliminf_{n\to\infty} \frac{\log l_n}{\log k_n} = \alpha, \quad \text{and} \quad \varlimsup_{n\to\infty} \frac{\log l_n}{\log k_n} = \beta. \tag{1}$$

In fact, due to the condition (iii) of Lemma 2 and the definition of the set $A^{k,l}$, the conditions (1) imply that, for any $x \in A^{k,l}$,

$$\varliminf_{n\to\infty} \frac{\log R_n(x)}{\log n} = \alpha \quad \text{and} \quad \varlimsup_{n\to\infty} \frac{\log R_n(x)}{\log n} = \beta.$$

Now, we give the definition of the sequences $\{k_n\}$ and $\{l_n\}$. Let $(N_i)_{i\geqslant 0}$ be an increasing sequence of integers such that $\lim_{i\to\infty}(N_i/N_{i+1}) = 0$. We need to consider several cases:

(i) $\alpha = \beta = 1$:

$$l_n = [n \exp n], \qquad k_n = [\exp n];$$

(ii) $1 = \alpha < \beta < \infty$:

$$u_n = \begin{cases} [\log n] & \text{if } N_{2i} \leqslant n < N_{2i+1} \text{ for some } i \in \mathbb{N}, \\ [n^{\beta-1}] & \text{otherwise,} \end{cases}$$

$$l_n = \sum_{i=1}^{k_n} u_i, \quad k_n = [e^n];$$

(iii) $1 < \alpha \leqslant \beta < \infty$:

$$u_n = \begin{cases} [n^{\alpha-1}] & \text{if } N_{2i} \leqslant n < N_{2i+1} \text{ for some } i \in \mathbb{N}, \\ [n^{\beta-1}] & \text{otherwise,} \end{cases}$$

$$l_n = \sum_{i=1}^{k_n} u_i, \quad k_n = [n^\gamma], \text{ with } \gamma \text{ satisfying } \alpha > 1 + \frac{1}{\gamma};$$

(iv) $1 < \alpha < \beta = \infty$:

$$u_n = \begin{cases} [n^{\alpha-1}] & \text{if } N_{2i} \leqslant n < N_{2i+1} \text{ for some } i \in \mathbb{N}, \\ [e^n] & \text{otherwise,} \end{cases}$$

$$l_n = \sum_{i=1}^{k_n} u_i, \quad k_n = [n^\gamma], \text{ with } \gamma \text{ satisfying } \alpha > 1 + \frac{1}{\gamma};$$

(v) $\alpha = \beta = \infty$:

$$l_n = [e^n], \qquad k_n = n.$$

Acknowledgements

The author thanks Prof. Jacques Peyrière for helpful discussions and careful modifications of the manuscript. She also thanks Prof. Jun Wu and Dr. Bo Tan for helpful discussions and thanks Prof. Zhi-ying Wen for inviting her to the MorningSide center of Mathematics (CAS) where the paper was conceived and finished.

References

[1] D.J. Feng, J. Wu, The Hausdorff dimension of recurrent sets in symbolic spaces, Nonlinearity 14 (1) (2001) 81–85.
[2] D.S. Ornstein, B. Weiss, Entropy and data compression schemes, IEEE Trans. Inform. Theory 39 (1) (1993) 78–83.