Hausdorff dimension of the multiplicative golden mean shift

Richard Kenyon [a], Yuval Peres [b], Boris Solomyak [c]

[a] *Richard Kenyon, Brown University*
[b] *Yuval Peres, Microsoft Research*
[c] *Boris Solomyak, University of Washington*

Received *****; accepted after revision +++++

Presented by

We compute the Hausdorff dimension of the "multiplicative golden mean shift" defined as the set of all reals in $[0,1]$ whose binary expansion (x_k) satisfies $x_k x_{2k} = 0$ for all $k \geq 1$, and show that it is smaller than the Minkowski dimension.

Résumé

Dimension de Hausdorff du shift de Fibonacci multiplicatif. Nous calculons la dimension de Hausdorff du "shift de Fibonacci multiplicatif", c'est-à-dire l'ensemble des nombres réels dans $[0,1]$ dont le développement en binaire (x_k) satisfait $x_k x_{2k} = 0$ pour tout $k \geq 1$. Nous montrons que la dimension de Hausdorff est plus petite que la dimension de Minkowski.

1. Introduction

A classical result of Furstenberg [5] says that if X is a closed subset of $[0,1]$, invariant under the map $T_m:\ x \mapsto mx \pmod 1$, then its Hausdorff dimension equals the Minkowski (box-counting) dimension, which equals the topological entropy of $T_m|_X$ divided by $\log m$. A simple example is the set $\Psi_G := \left\{x = \sum_{k=1}^\infty x_k 2^{-k}:\ x_k \in \{0,1\},\ x_k x_{k+1} = 0 \text{ for all } k\right\}$ for which we have $\dim_H(\Psi_G) = \dim_M(\Psi_G) = \log_2\left(\frac{1+\sqrt{5}}{2}\right)$ (the subscript G here stands for the "Golden Ratio"). Instead, we consider the set

$$\Xi_G := \left\{x = \sum_{k=1}^\infty x_k 2^{-k}:\ x_k \in \{0,1\},\ x_k x_{2k} = 0 \text{ for all } k\right\}$$

Email addresses: `rkenyon@math.brown.edu` (Richard Kenyon), `peres@microsoft.com` (Yuval Peres), `solomyak@math.washington.edu` (Boris Solomyak).

which we call the "multiplicative golden mean shift." The reason for this term is that the set of binary sequences corresponding to the points of Ξ_G is invariant under the action of the semigroup of multiplicative positive integers $\mathbb{N}^*$: $M_r(x_k) = (x_{rk})$ for $r \in \mathbb{N}$. Fan, Liao, and Ma [4] showed that $\dim_M(\Xi_G) = \sum_{k=1}^{\infty} 2^{-k-1} \log_2 F_{k+1} = 0.82429\ldots$, where F_k is the k-th Fibonacci number: $F_1 = 1,\ F_2 = 2, F_{k+1} = F_{k-1} + F_k$, and raised the question of computing the Hausdorff dimension of Ξ_G.

Theorem 1.1 *We have* $\dim_H(\Xi_G) < \dim_M(\Xi_G)$. *In fact,*

$$\dim_H(\Xi_G) = -\log_2 p = 0.81137\ldots, \quad \text{where } p^3 = (1-p)^2, \quad 0 < p < 1. \tag{1}$$

Our manuscript [6] contains substantial generalizations of this result, extending it to a large class of "multiplicative subshifts." We state one of them at the end of the paper.

Although the set Ξ_G is on the real line, it appears to have a strong resemblance with a class of self-affine sets on the plane, namely, the Bedford-McMullen "carpets" [1,7], for which also the Hausdorff dimension is typically smaller than the Minkowski dimension. However, this seems to be more of an analogy than a direct link.

An additional motivation to study the multiplicative subshifts comes from questions on multifractal analysis of multiple ergodic averages raised in [4]. Perhaps, the simplest non-trivial case of such multifractal analysis is the study of the sets $A_\theta := \left\{ x = \sum_{k=1}^{\infty} x_k 2^{-k} :\ x_k \in \{0,1\},\ \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} x_k x_{2k} = \theta \right\}$. It is not hard to show that $\dim_H(A_0) = \dim_H(\Xi_G)$, which we compute in Theorem 1.1. With more work, our method can be used to compute the Hausdorff dimension of A_θ, but the details are beyond the scope of this note.

In this paper, we focus on Ξ_G to explain our ideas and methods in the simplest possible setting. To conclude the introduction, we should mention that the dimensions of some analogous sets, e.g., $\widetilde{\Xi} = \left\{ x = \sum_{k=1}^{\infty} x_k 2^{-k} :\ x_k \in \{0,1\},\ x_k x_{2k} x_{3k} = 0 \text{ for all } k \right\}$ are so far out of reach.

2. Proof of Theorem 1.1

It is more convenient to work in the symbolic space $\Sigma_2 = \{0,1\}^{\mathbb{N}}$, with the metric $\varrho((x_k),(y_k)) = 2^{-\min\{n:\ x_n \neq y_n\}}$. It is well-known that the dimensions of a compact subset of $[0,1]$ and the corresponding set of binary digit sequences in Σ_2 are equal (this is equivalent to replacing the covers by arbitrary interval with those by dyadic intervals). Thus, it suffices to determine the dimensions of the set X_G—the collection of all binary sequences (x_k) such that $x_k x_{2k} = 0$ for all k. Observe that

$$X_G = \left\{ \omega = (x_k)_{k=1}^{\infty} \in \Sigma_2 :\ (x_{i2^r})_{r=0}^{\infty} \in \Sigma_G \ \text{ for all } i \text{ odd} \right\} \tag{2}$$

where Σ_G is usual (additive) golden mean shift: $\Sigma_G := \{(x_k)_{k=1}^{\infty} \in \Sigma_2,\ x_k x_{k+1} = 0,\ \forall\, k \geq 1\}$.

We will use the following well-known result; it essentially goes back to Billingsley [2]. We state it in the symbolic space Σ_2 where $[u]$ denotes the cylinder set of sequences starting with a finite "word" u and $x_1^n = x_1 \ldots x_n$.

Proposition 1 (see [3]) *Let E be a Borel set in Σ_2 and let ν be a finite Borel measure on Σ_2.*
(i) *If* $\liminf_{n\to\infty}(-\frac{1}{n}) \log_2 \nu[x_1^n] \geq s$ *for ν-a.e. $x \in E$, then* $\dim_H(E) \geq s$.
(ii) *If* $\liminf_{n\to\infty}(-\frac{1}{n}) \log_2 \nu[x_1^n] \leq s$ *for all $x \in E$, then* $\dim_H(E) \leq s$.

Given a probability measure μ on Σ_G, we can define a probability measure on X_G by

$$\mathbb{P}_\mu[u] := \prod_{i\le n,\, i \text{ odd}} \mu[u|_{J(i)}], \quad \text{where } J(i) = \{2^r i\}_{r=0}^{\infty} \tag{3}$$

and $u|_{J(i)}$ denotes the "restriction" of the word u to the subsequence $J(i)$. It turns out that this class of measures is sufficiently rich to compute $\dim_H(X_G)$.

For $k \ge 1$ let α_k be the partition of Σ_G into cylinders of length k. For a measure μ on Σ_2 and a finite partition α, denote by $H^\mu(\alpha)$ the μ-entropy of the partition, with base 2 logarithms: $H^\mu(\alpha) = -\sum_{C\in\alpha} \mu(C)\log_2 \mu(C)$. Define

$$s(\mu) := \sum_{k=1}^{\infty} \frac{H^\mu(\alpha_k)}{2^{k+1}}\,. \tag{4}$$

Proposition 2 *Let μ be a probability measure on Σ_G. Then $\dim_H(X_G) \ge s(\mu)$.*

Proof. We are going to demonstrate that for every $\ell \in \mathbb{N}$,

$$\liminf_{n\to\infty} \frac{-\log_2 \mathbb{P}_\mu[x_1^n]}{n} \ge \sum_{k=1}^{\ell} \frac{H^\mu(\alpha_k)}{2^{k+1}} \quad \text{for } \mathbb{P}_\mu\text{-a.e. } x. \tag{5}$$

Then, letting $\ell \to \infty$ and using Proposition 1(i) will yield the desired inequality. Fix $\ell \in \mathbb{N}$. By a routine argument, to verify (5) we can restrict ourselves to $n = 2^\ell r,\ r \in \mathbb{N}$. In view of (3), we have

$$\mathbb{P}_\mu[x_1^n] \le \prod_{k=1}^{\ell} \prod_{\frac{n}{2^k} < i \le \frac{n}{2^{k-1}},\ i \text{ odd}} \mu[x_1^n|_{J(i)}]. \tag{6}$$

Note that $x_1^n|_{J(i)}$ is a word of length k for $i \in (n/2^k, n/2^{k-1}]$, with i odd, which is a beginning of a sequence in Σ_G. Thus, $[x_1^n|_{J(i)}]$ is an element of the partition α_k. The random variables $x \mapsto -\log_2 \mu[x_1^n|_{J(i)}]$ are i.i.d for $i \in (n/2^k, n/2^{k-1}]$, with i odd, and their expectation equals $H^\mu(\alpha_k)$, by the definition of entropy. Note that there are $n/2^{k+1}$ odds in $(n/2^k, n/2^{k-1}]$. Fixing k, ℓ with $k \le \ell$ and taking $n = 2^\ell r$, $r \to \infty$, we get an infinite sequence of i.i.d. random variables. Therefore, by a version of the Law of Large Numbers,

$$\forall\ k \le \ell, \sum_{\frac{n}{2^k} < i \le \frac{n}{2^{k-1}},\ i \text{ odd}} \frac{-\log_2 \mu[x_1^n|_{J(i)}]}{(n/2^{k+1})} \to H^\mu(\alpha_k) \ \text{ as } n = 2^\ell r \to \infty, \ \text{ for } \mathbb{P}_\mu\text{-a.e. } x. \tag{7}$$

By (6) and (7), for $\mathbb{P}_\mu$-a.e. x,

$$\frac{-\log_2 \mathbb{P}_\mu[x_1^n]}{n} \ge \sum_{k=1}^{\ell} \frac{1}{2^{k+1}} \sum_{\frac{n}{2^k} < i \le \frac{n}{2^{k-1}},\ i \text{ odd}} \frac{-\log_2 \mu[x_1^n|_{J(i)}]}{n/2^{k+1}} \to \sum_{k=1}^{\ell} \frac{H^\mu(\alpha_k)}{2^{k+1}}\,.$$

This confirms (5), so the proof is complete. □

Proof of the lower bound for the Hausdorff dimension in Theorem 1.1. Let $s := \sup\{s(\mu) :\ \mu$ is a probability measure on $\Sigma_G\}$. By Proposition 2, we have $\dim_H(X_G) \ge s$, and we will prove that this is actually an equality. To this end, we specify a measure which will turn out to be "optimal." This measure is Markov, but non-stationary. It could be "guessed" or derived by solving the optimization problem (which also yields that the optimal measure is unique). However, for the proof of dimension formula it suffices to produce the answer. Let μ be a Markov measure on Σ_G, with initial probabilities $(p, 1-p)$, and the matrix of transition probabilities $P = (P(i,j))_{i,j=0,1} = \begin{pmatrix} p & 1-p \\ 1 & 0 \end{pmatrix}$. Using elementary properties of entropy, it is not hard to see that $s(\mu) = \frac{H(p)}{2} + \frac{ps(\mu)}{2} + \frac{(1-p)s(\mu)}{4}$, whence $s(\mu) = \frac{2H(p)}{3-p}$. Maximizing over p yields $s(\mu) = 2\log_2 \frac{p}{1-p}$, and comparing this to $s(\mu) = \frac{2H(p)}{3-p}$ we get

$$p^3 = (1-p)^2, \quad s(\mu) = -\log_2 p. \tag{8}$$

Combined with Proposition 2, this proves the lower bound for the Hausdorff dimension in (1). □

Proof of the upper bound for the Hausdorff dimension in Theorem 1.1. Denote by $N_i(u)$ the number of symbols i in a word u. By the definition of the measure μ, we obtain for any $u = u_1 \dots u_k \in \{0,1\}^n$,

$$\mu[u] = p_{u_1} P(u_1, u_2) \cdot \ldots \cdot P(u_{k-1}, u_k) = (1-p)^{N_1(u_1 \ldots u_k)} p^{N_0(u_1 \ldots u_k) - N_1(u_1 \ldots u_{k-1})}. \tag{9}$$

Indeed, the probability of a 1 is always $1-p$, whereas the probability of a 0 is p, except in those cases when it follows a 1, which it must by the definition of Σ_G. In view of (9), by the definition of the measure $\mathbb{P}_\mu$ on X_G, we have $\mathbb{P}_\mu[x_1^n] = (1-p)^{N_1(x_1^n)} p^{N_0(x_1^n) - N_1(x_1^{n/2})}$ for any $x \in X_G$ and n even. Using that $(1-p)^2 = p^3$ and $N_0(x_1^n) = n - N_1(x_1^n)$, we obtain that

$$\mathbb{P}_\mu[x_1^n] = p^n p^{N_1(x_1^n)/2 - N_1(x_1^{n/2})}.$$

Let $a_\ell = -\frac{1}{n}\log_2 \mathbb{P}_\mu[x_1^n]$ for $n = 2^\ell$. Then $a_\ell = -\log_2 p \left(1 + \frac{1}{2}\left[\frac{N_1(x_1^n)}{n} - \frac{N_1(x_1^{n/2})}{n/2}\right]\right)$. Now we see that the average of a_ℓ's "telescopes":

$$\frac{a_1 + \cdots + a_\ell}{\ell} = -\log_2 p \left(1 + \frac{1}{2\ell}\left[\frac{N_1(x_1^{2^\ell})}{2^\ell} - N_1(x_1)\right]\right) \to -\log_2 p, \quad \text{as } \ell \to \infty.$$

It follows that

$$\liminf_{\ell\to\infty} a_\ell = \liminf_{\ell\to\infty} 2^{-\ell}(-\log_2 \mathbb{P}_\mu[x_1^{2^\ell}]) \le -\log_2 p = s,$$

for every $x \in X_G$, so $\dim_H(X_G) \le s$ by Proposition 1(ii). □

3. Generalization

Here we state a generalization of Theorem 1.1 to the case of arbitrary multiplicative subshifts of finite type; the proof can be found in [6].

Theorem 3.1 (i) *Let A be a 0-1 primitive $m \times m$ matrix (i.e. some power of A has only positive entries). Consider* $\Xi_A = \left\{x = \sum_{k=1}^\infty x_k m^{-k} : \ x_k \in \{0, \dots, m-1\},\ A(x_k, x_{2k}) = 1 \text{ for all } k\right\}$. *Then* $\dim_H(\Xi_A) = \frac{1}{2}\log_m \sum_{i=0}^{m-1} t_i$, *where* $(t_i)_{i=0}^{m-1}$ *is the unique vector satisfying* $t_i^2 = \sum_{j=0}^{m-1} A(i,j)t_j,\ \ t_i > 1,\ i = 0, \dots, m-1$.

(ii) *The Minkowski dimension of Ξ_A exists and equals* $\dim_M(\Xi_A) = \sum_{k=1}^\infty 2^{-k-1}\log_m(A^{k-1}\overline{1}, \overline{1})$ *where* $\overline{1} = (1, \dots, 1)^T \in \mathbb{R}^m$. *We have* $\dim_H(\Xi_A) = \dim_M(\Xi_A)$ *if and only if all row sums of A are equal.*

Acknowledgements

We are grateful to Jörg Schmeling for telling us about the problem, and to Aihua Fan, Lingmin Liao, and Jihua Ma for sending us their preprint [4] prior to publication. The research of R. K. and B. S. was supported in part by NSF.

References

[1] T. Bedford. *Crinkly curves, Markov partitions and box dimension in self-similar sets.* Ph.D. Thesis, University of Warwick, 1984.

[2] P. Billingsley. *Ergodic theory and information.* Wiley, New York, 1965.

[3] K. Falconer. *Techniques in fractal geometry.* John Wiley & Sons, Chichester, 1997.

[4] A. Fan, L. Liao, J. Ma, *Level sets of multiple ergodic averages*, preprint.

[5] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1** (1967), 1–49.

[6] R. Kenyon, Y. Peres, B. Solomyak, *Hausdorff dimension for fractals invariant under the multiplicative integers*, preprint, 2011.

[7] C. McMullen. The Hausdorff dimension of general Sierpinski carpets. *Nagoya Math. J.* **96** (1984), 1–9.