DIOPHANTINE PROPERTY OF MATRICES AND
ATTRACTORS OF PROJECTIVE ITERATED FUNCTION
SYSTEMS IN $\mathbb{RP}^1$

BORIS SOLOMYAK AND YUKI TAKAHASHI

Abstract. We prove that almost every finite collection of matrices in $GL_d(\mathbb{R})$ and $SL_d(\mathbb{R})$ with positive entries is Diophantine. Next we restrict ourselves to the case $d = 2$. A finite set of $SL_2(\mathbb{R})$ matrices induces a (generalized) iterated function system on the projective line $\mathbb{RP}^1$. Assuming uniform hyperbolicity and the Diophantine property, we show that the dimension of the attractor equals the minimum of 1 and the critical exponent.

1. Introduction and main results

1.1. Diophantine property of matrices. Recently there has been interest in Diophantine properties in non-Abelian groups. The following is a variant of [14, Definition 4.2].

Definition 1.1. Let $A = \{A_i\}_{i \in \Lambda}$ be a finite subset of a topological group $G$ equipped with a metric $\varrho$. Write $A_i = A_{i_1} \cdots A_{i_n}$ for $i = i_1 \cdots i_n$. We say that the set $A$ is Diophantine if there exists a constant $c > 0$ such that for every $n \in \mathbb{N}$, we have

\begin{equation}
    i, j \in \Lambda^n, \ A_i \neq A_j \Rightarrow \varrho(A_i, A_j) > c^n.
\end{equation}

The set $A$ is strongly Diophantine if there exists $c > 0$ such that for all $n \in \mathbb{N}$,

\begin{equation}
    i, j \in \Lambda^n, \ i \neq j \Rightarrow \varrho(A_i, A_j) > c^n.
\end{equation}

Clearly, $A$ is strongly Diophantine if and only if it is Diophantine and generates a free semigroup. Gamburd, Jacobson, and Sarnak [14, Definition 4.2] gave a definition of a Diophantine set, which is equivalent to ours, except that they always consider symmetric sets (that is, $g \in A \Rightarrow g^{-1} \in A$). Diophantine-type questions in groups arise in connection with spectral gap estimates, see [14, 8].

See [11, 2] for a recent discussion of Diophantine properties in groups and related problems. In [2] a Lie group $G$ is called Diophantine, if almost every $k$ elements of $G$, chosen independently at random according to the Haar measure, together with their...

Date: March 1, 2019.

Both authors were supported by the Israel Science Foundation grant 396/15 (PI. B. Solomyak).
inverses, form a Diophantine set in $G$. Gamburd et al. [14] conjectured that $SU_2(\mathbb{R})$ is Diophantine. More generally, it is conjectured that semi-simple Lie groups are Diophantine. Kaloshin and Rodnianski [18] proved a weaker Diophantine-type property: for a.e. $(A, B) \in SO_3(\mathbb{R}) \times SO_3(\mathbb{R})$, there exists $c > 0$ such that for any $n \geq 1$ and any two distinct words $W_1, W_2$ over the set $\mathcal{A} = \{A, B, A^{-1}, B^{-1}\}$ of length $n$,

$$\|W_1 - W_2\| \geq c n^2.$$ 

It is mentioned in [18] that their method is general, and applies to $SU_2(\mathbb{R})$ as well, and also to $m$-tuples of matrices for any $m \geq 2$.

Next we state our first result. For any collection of linearly independent vectors $v_1, \ldots, v_d$ in $\mathbb{R}^d$ consider the simplicial cone

$$(1.3) \quad \Sigma = \Sigma_{v_1, \ldots, v_d} = \{x_1 v_1 + \cdots + x_d v_d : x_1, \ldots, x_d \geq 0\}.$$ 

If a matrix $A \in GL_d(\mathbb{R})$ satisfies

$$A(\Sigma \setminus \{0\}) \subset \Sigma^o,$$

we say that $\Sigma$ is strictly invariant for $A$. Given a cone $\Sigma = \Sigma_{v_1, \ldots, v_d}$, denote by $\mathcal{X}_{\Sigma, m}$ (respectively, $\mathcal{Y}_{\Sigma, m}$) the set of all $GL_d(\mathbb{R})$ (respectively, $SL_d(\mathbb{R})$) $m$-tuples of matrices for which $\Sigma$ is strictly invariant. We consider $\mathcal{X}_{\Sigma, m}$ as an open subset of $\mathbb{R}^{d^2 m}$ and $\mathcal{Y}_{\Sigma, m}$ as a $(d^2 - 1)m$-dimensional manifold.

**Theorem 1.2.** Let $\Sigma = \Sigma_{v_1, \ldots, v_d}$ be a simplicial cone in $\mathbb{R}^d$ and $m \geq 2$.

(i) For a.e. $A \in \mathcal{X}_{\Sigma, m}$, the $m$-tuple $A$ is strongly Diophantine. In particular, a.e. $m$-tuple of positive $GL_d(\mathbb{R})$ matrices is strongly Diophantine.

(ii) For a.e. $A \in \mathcal{Y}_{\Sigma, m}$, the $m$-tuple $A$ is strongly Diophantine. In particular, a.e. $m$-tuple of positive $SL_d(\mathbb{R})$ matrices is strongly Diophantine.

**Remark 1.3.** 1. Unfortunately, our results do no cover any example of a symmetric set, since the strict invariance property cannot hold for a matrix $A$ and $A^{-1}$ simultaneously.

2. Every $m$-tuple of matrices with algebraic entries is Diophantine (but not necessarily strongly Diophantine), see, e.g., [14] Prop. 4.3.

3. It is well-known that Diophantine numbers in $\mathbb{R}$ form a set of full measure, which is, however, meagre in Baire category sense (its complement contains a dense $G_\delta$ set). Baire category genericity of non-Diophantine $m$-tuples in $SU_2(\mathbb{R})$ has been pointed out in [14]. In $G = SL_d(\mathbb{R})$ the situation is different, since there are, for example, open sets of $m$-tuples in $G \times G$ which satisfy (1.2). For instance, if $\mathbb{R}^d_+$ is mapped by $A, B$ into closed cones that are disjoint, except at the origin, then (1.2) holds for $\{A, B\}$. On the other hand, there are open sets in $(SL_d(\mathbb{R}))^m$ in which non-Diophantine pairs are dense. For instance, the set of elliptic matrices...
The scheme of the proof of Theorem 1.2 is as follows. We consider the induced action of the matrices on the projective space, and show that, given a non-degenerate family of $m$-tuples strictly preserving an open set, depending on a parameter real-analytically, for all parameters outside an exceptional set of zero Hausdorff dimension, the induced iterated function system (IFS) satisfies a version of the “exponential separation condition”. This property implies the strong Diophantine condition for the matrices. We then locally foliate the space of $m$-tuples of matrices and apply Fubini’s Theorem. The result on the zero-Hausdorff dimensional set of exceptions uses the notion of order-$k$ transversality, which is a modified version of that which appeared in the work of Hochman [15, 16]. The strict open set preservation property is needed to ensure that the induced IFS is contracting (uniformly hyperbolic).

1.2. Projective IFS and linear cocycles. Let $A = \{A_i\}_{i \in \Lambda}$ be a finite collection of $SL_2(\mathbb{R})$ matrices. The linear action of $SL_2(\mathbb{R})$ on $\mathbb{R}^2$ induces an action on the projective space $\mathbb{P}^1$, and thus $A$ defines an IFS $\Phi_A = \{\varphi_A\}_{A \in \mathcal{A}}$ on $\mathbb{P}^1$, called a (real) projective IFS. Such IFS were studied by Barnsley and Vince [4], and by De Leo [11, 10]. Following [4], we say that the IFS $\Phi_A$ has an attractor $K$ if for every nonempty compact set $B$ in a neighborhood of $K$, we have $\lim_{k \to \infty} \Phi_k(A)(B) = K$ in the Hausdorff metric, where $\Phi_k(A)(B) = \bigcup_{A \in \mathcal{A}} \varphi_A(B)$. Let $A = \{A_i\}_{i \in \Lambda}$ be a finite collection of $GL_2(\mathbb{R})$ matrices and $\Phi_A$ the associated IFS on $\mathbb{P}^1$.

An alternative, but closely related viewpoint, is to consider the linear cocycle $A : \Lambda^\mathbb{Z} \to SL_2(\mathbb{R})$ over the shift on $\Lambda^\mathbb{Z}$, defined by $A(i) = A_i$. Strict contractivity of the projective IFS turns out to be equivalent to uniform hyperbolicity of the cocycle [6]. Here we restrict ourselves to the case of $d = 2$, which was investigated in great detail by Avila, Bochi, and Yoccoz [3]. There is a natural identification between $[0, \pi)$ and the projective space $\mathbb{P}^1$. Below we use this identification freely, and whenever necessary we view $[0, \pi)$ as $\mathbb{R}/\pi\mathbb{Z}$. For $A \in GL_2(\mathbb{R})$ denote the action of $A$ on $[0, \pi) \cong \mathbb{R}/\pi\mathbb{Z}$ by the symbol $\varphi_A$.

Denote by $d_\theta$ the metric on $\mathbb{P}^1$ induced from the identification with $\mathbb{R}/\pi\mathbb{Z}$. Below we work with $m$-tuples of $SL_2(\mathbb{R})$-matrices, since the action of $GL_2(\mathbb{R})$ factors through the $SL_2(\mathbb{R})$ action in the obvious way, via $A \mapsto (\det A)^{-1}A$.

**Definition 1.4.** A multicone is a proper nonempty open subset $U$ of $\mathbb{P}^1$, having finitely many connected components with disjoint closures.

In the following theorem we extracted the results relevant for us from [3, 4] (note that [3] considers real projective IFS of any dimension).
Theorem 1.5 ([3, 4]). Let $\mathcal{A} = \{A_i\}_{i \in \Lambda}$ be a family of $SL_2(\mathbb{R})$ matrices and let $\Phi_\mathcal{A}$ be the associated IFS on $\mathbb{RP}^1$. The following are equivalent:

(i) the IFS $\Phi_\mathcal{A}$ has an attractor $K \neq \mathbb{RP}^1$; 
(ii) the associated linear cocycle over $\Lambda \mathbb{Z}$ is uniformly hyperbolic; 
(iii) there is a multicone $U$, such that $\Phi_\mathcal{A}(U) \subset U$; 
(iv) there is nonempty open set $V \subset \mathbb{RP}^1$ such that $\Phi_\mathcal{A}$ is contractive on $V$, with respect to a metric equivalent to $d_\mathcal{P}$.

Following [3], we will call a multicone satisfying $\Phi_\mathcal{A}(U) \subset U$, a strictly invariant multicone for the family of matrices and for the IFS. There are examples, see [3], which show that one may need a multicone having $k$ components, for any given $k \geq 2$, even for a pair of $SL_2(\mathbb{R})$ matrices $\{A_1, A_2\}$.

Our next result concerns the dimension of the attractor. Following De Leo [11], consider the $\zeta$-function

$$\zeta_\mathcal{A}(t) = \sum_{n \geq 1} \sum_{i \in \Lambda^n} \|A_i\|^{-t},$$

and define the critical exponent of $\mathcal{A}$ by

$$s_\mathcal{A} = \sup_{t \geq 0} \{t : \zeta_\mathcal{A}(t) = \infty\}.$$

(1.4)

Theorem 1.6. Let $\mathcal{A} = \{A_i\}_{i \in \Lambda}$ be a finite set of $SL_2(\mathbb{R})$ matrices which has a strictly invariant multicone (or satisfies any of the equivalent conditions from Theorem 1.5), and let $K$ be the attractor of the associated IFS $\Phi_\mathcal{A}$ on $\mathbb{RP}^1$. Assume that at least two of the maps $\varphi_{A_i}$ have distinct attracting fixed points. If $\mathcal{A}$ is strongly Diophantine, then $\dim_H(K) = \min\{1, \frac{1}{2}s_\mathcal{A}\}$, where $s_\mathcal{A}$ is the critical exponent (1.4).

In the special case when the IFS $\Phi_\mathcal{A}$ satisfies the Open Set Condition, this result is due to De Leo [11, Th.4]. Recall that the strong Diophantine condition holds, in particular, when $\mathcal{A}$ generates a free semigroup and all the entries of $A_i$ are algebraic.

Remark 1.7. It is further shown in [11] that for $\mathcal{A}$ hyperbolic (and in some parabolic cases),

$$s_\mathcal{A} = \lim_{r \to \infty} \frac{N_\mathcal{A}(r)}{\log r},$$

where $N_\mathcal{A}(r)$ is the number of elements of norm $\leq r$ of the semigroup generated by $\mathcal{A}$. An analogy is pointed out with the classical results on Kleinian and Fuchsian groups, see, e.g., [26].

Let $\Phi = \Phi_\mathcal{A}$. An alternative way to express the dimension, and one we actually use in the proof, is in terms of Bowen’s pressure formula

$$P_\Phi(s) = 0,$$

(1.5)
where $P_{\Phi}(\cdot)$ is the pressure function associated with the IFS $\Phi$. Throughout the paper we use the notation 

$$\varphi_i = \varphi_{i_1} \cdots \varphi_{i_n}.$$ 

The pressure is defined by

\begin{equation}
1.6 \quad P_{\Phi}(t) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{i \in \Lambda^n} \|\varphi'_i\|^{t},
\end{equation}

where $\|\cdot\|$ is the supremum norm on $U$. As will be clear from the Bounded Distortion Property, the definition of $P_{\Phi}(t)$ does not depend on the choice of strictly invariant multicone $U$, and moreover,

\begin{equation}
1.7 \quad 2s = s_A.
\end{equation}

It is a classical result, going back to Bowen [9] and Ruelle [22], see also [12], that if \{\varphi_i\}_{i \in \Lambda} is a hyperbolic IFS on $\mathbb{R}$ of smoothness $C^{1+\epsilon}$, satisfying the Open Set Condition, then the dimension of the attractor $K$ is given by the Bowen’s equation. In the case that the maps $\varphi_i$ are affine, $s > 0$ is the unique solution of

$$\sum_{i \in \Lambda} r_i^s = 1,$$

where $r_i \in (0,1)$ is the contraction ratio of $\varphi_i$. For an IFS with overlaps this is not necessarily true. In [24], Simon, Solomyak, and Urbański showed that for a one-parameter family of nonlinear IFS with overlaps (hyperbolic and some parabolic) satisfying the order-1 transversality condition, for Lebesgue-a.e. parameter the dimension of the attractor is given by

\begin{equation}
1.8 \quad \dim_H(K) = \min\{1, s\},
\end{equation}

where $s$ is from (1.5) and the pressure is given by (1.6).

**Definition 1.8.** Let $F = \{f_i\}_{i \in \Lambda}$ be an IFS on a metric space $(X, \rho)$, that is, $f_i: X \to X$. We say that $F$ satisfies the exponential separation condition on a set $J \subset X$ if there exists $c > 0$ such that for all $n \in \mathbb{N}$ we have

\begin{equation}
1.9 \quad \sup_{x \in J} \rho(f_i(x), f_j(x)) > c^n, \quad \text{for all } i, j \in \Lambda^n \text{ with } i_1 \neq j_1 \text{ and } f_i \neq f_j.
\end{equation}

If, in addition, the semigroup generated by $F$ is free, that is, $f_i \equiv f_j \iff i = j$, we say that $F$ satisfies the strong exponential separation condition. If these properties hold for infinitely many $n$, then we say that $F$ satisfies the (strong) exponential separation condition on $J$ along a subsequence.

It is rather straightforward to show that the (strong) Diophantine condition for an $m$-tuple in $SL_2(\mathbb{R})$ matrices is equivalent to the (strong) exponential separation condition for the associated projective IFS (see Lemma 3.6 below).
In [15 Cor. 1.2], Hochman proved (1.8) for an affine IFS $F = \{f_i\}_{i \in \Lambda}$ satisfying the exponential separation condition on $J = \{0\}$ along a subsequence. Thus our Theorem 1.6 is, in a sense, a generalization of Hochman’s result to the case of contractive projective IFS.

**Remark 1.9.** In fact, Hochman [15] used the condition (1.9) without the requirement $i_1 \neq j_1$. However, for an IFS $\{f_i\}_{i \in \Lambda}$ on an interval $J \subset \mathbb{R}$, such that

$$\inf_{x \in J, i \in \Lambda} |f'_i(x)| \geq r_{\min} > 0,$$

requiring $i_1 \neq j_1$ in (1.9) does not weaken the exponential separation condition — it only affects the constant $c$. This follows from the estimate

$$|f_i(x) - f_j(x)| = |f_{i \wedge j}(u(x)) - f_{i \wedge j}(v(x))| \geq r_{\min}^n |u(x) - v(x)|, \quad i, j \in \Lambda^n,$$

where $i \wedge j$ is the common initial segment of $i$ and $j$, so that $u_1 \neq v_1$.

1.3. **IFS of linear fractional transformations.** It is well-known that the action of $GL_2(\mathbb{R})$ on $\mathbb{R}P^1$ can be expressed in terms of linear fractional transformations. For

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{R}),$$

let $f_A(x) = (ax + b)/(cx + d)$, and define $\psi : [0, \pi) \to \mathbb{R}^*$ by $\psi(\theta) = \cos \theta / \sin \theta$, where $\mathbb{R}^* = \mathbb{R} \cup \{\infty\}$. It is easy to see that the following diagram commutes:

$$[0, \pi) \xrightarrow{\varphi_A} [0, \pi)$$

$$\psi \downarrow \quad \psi \downarrow$$

$$\mathbb{R}^* \xrightarrow{f_A} \mathbb{R}^*$$

Observe that $\psi$ is smooth, and on any compact subset of $(0, \pi)$ the derivatives of $\psi$ and $\psi^{-1}$ are bounded. The following is then an immediate corollary of Theorem 1.6.

**Corollary 1.10.** Let $F = \{f_i\}_{i \in \Lambda}$ be a finite collection of linear fractional transformations with real coefficients. Assume that there exists $U \subset \mathbb{R}$, a finite union of bounded open intervals with disjoint closures, such that $f_i(U) \subset U$ for all $i \in \Lambda$. If $F$ satisfies the strong exponential separation condition on $U$, then we have $\text{dim}_H(K) = \min\{1, s\}$, where $s > 0$ is the unique zero of the pressure function $P_F$.

1.4. **Furstenberg measure.** Let $A = \{A_i\}_{i \in \Lambda}$ be a finite collection of $SL_2(\mathbb{R})$ matrices, and let $p = (p_i)_{i \in \Lambda}$ be a probability vector. Assume that $p_i > 0$ for all
\( i \in \Lambda \) (we always assume this for any probability vector). We consider the finitely supported probability measure \( \mu \) on \( SL_2(\mathbb{R}) \):

\[
\mu = \sum_{i \in \Lambda} p_i \delta_{A_i}.
\]  

(1.10)

Our standing assumption is that \( A \) generates an unbounded and totally irreducible subgroup (i.e., does not preserve any finite set in \( \mathbb{R}P^1 \)). Then there exists a unique probability measure \( \nu \) on \( \mathbb{R}P^1 \) satisfying \( \mu \cdot \nu = \nu \), that is,

\[
\nu = \sum_{i \in \Lambda} p_i A_i \nu,
\]  

(1.11)

where \( A_i \nu \) is the push-forward of \( \nu \) under the action of \( A_i \), see [13]. The measure \( \nu \) is the stationary measure, or the Furstenberg measure, for the random matrix product \( A_{i_1} \cdots A_{i_n} \) where the matrices are chosen i.i.d. from \( A \) according to the probability vector \( p \).

The properties of the Furstenberg measure for \( SL_2(\mathbb{R}) \) random matrix products, such as absolute continuity, singularity, Hausdorff dimension, etc., were studied by many authors, including [19, 7]. In [21, 20, 25] this investigation was linked with the study of IFS consisting of linear fractional transformations. The reader is referred to [17] for a discussion of more recent applications. We will recall the main result of [17], since it will be the main tool in proving Theorem 1.6.

Let \( \chi_{A,p} \) be the Lyapunov exponent, which is the almost sure value of the limit

\[
\lim_{n \to \infty} \frac{1}{n} \log \| A_{i_1} \cdots A_{i_n} \|,
\]  

(1.12)

where \( i_1, i_2, \cdots \in \Lambda \) is a sequence chosen randomly according to the probability vector \( p = (p_i)_{i \in \Lambda} \). The Lyapunov exponent is usually defined as the almost sure value of the limit

\[
\lim_{n \to \infty} \frac{1}{n} \log \| A_{i_n} \cdots A_{i_1} \|,
\]  

(1.13)

but it is easy to see that (1.12) and (1.13) define the same value (e.g., by Egorov’s Theorem). Under the standing assumptions, the limit exists almost surely and is positive [13]. The Hausdorff dimension of a measure \( \nu \) is defined by

\[
\dim_H(\nu) = \inf \{ \dim_H(E) : \nu(E^c) = 0 \}.
\]

For a probability vector \( p = (p_i)_{i \in \Lambda} \), we denote the entropy \( H(p) \) by

\[
H(p) = -\sum_{i \in \Lambda} p_i \log p_i.
\]

**Theorem 1.11** ([17]). Let \( A = \{ A_i \}_{i \in \Lambda} \) be a finite collection of \( SL_2(\mathbb{R}) \) matrices. Assume that \( A \) is strongly Diophantine and generates an unbounded and totally
irreducible subgroup. Let \( p = (p_i)_{i \in \Lambda} \) be a probability vector, and let \( \nu \) be the associated Furstenberg measure. Then we have

\[
\dim_H(\nu) = \min \left\{ 1, \frac{H(p)}{2^\chi_{A,p}} \right\}.
\]

Theorem 1.2 implies, in particular, that the dimension formula (1.14) holds for the Furstenberg measure associated with a.e. finite family of positive matrices (independent of the probability vector).

Next we address the question: what is the Hausdorff dimension of the support of the Furstenberg measure? Sometimes, the support is all of \( \mathbb{R}P^1 \), in which case the answer is trivially one. The definition (1.14) implies that the support is invariant under the IFS \( \Phi \) induced by \( A \). Thus, Theorem 1.6 has the following immediate corollary:

**Corollary 1.12.** Let \( A = \{ A_i \}_{i \in \Lambda} \) be a Diophantine set of \( SL_2(\mathbb{R}) \) matrices which has a strictly invariant multicone, \( \mu \) a finitely supported measure defined by (1.14), and \( \nu \) the associated Furstenberg measure. Then \( \dim_H(\text{supp}\nu) = \min\{1, \frac{1}{2}s_A\} \), where \( s_A \) is the critical exponent of \( A \).

Denote by \( \mathcal{H}_m \) the set of \( m \)-tuples in \( SL_2(\mathbb{R}) \) which have a strictly invariant multicone. Avila (see [27, Prop.6]) proved that the interior of the complement of \( \mathcal{H}_m \) in \((SL_d(\mathbb{R}))^m\) is \( \mathcal{E}_m \), where \( \mathcal{E}_m \) is the set of \( m \)-tuples which generate a semigroup containing an elliptic matrix. Observe that if an elliptic matrix is conjugate to an irrational rotation, then certainly the invariant set (support of the Furstenberg measure) is all of \( \mathbb{R}P^1 \). On the other hand, if it is conjugate to a rational rotation, then the semigroup generated by \( A \) contains the identity and the strong Diophantine property fails. We expect that our methods can be extended to cover strongly Diophantine families on the boundary of \( \mathcal{H}_m \), which include parabolic systems.

1.5. **Structure of the paper.** The rest of the paper is organized as follows. In the next section we prove Theorem 1.2. In Section 3 we consider projective IFS and prove Theorem 1.6. Finally, in Section 4 we include proofs of some standard technical results for the reader’s convenience.

2. **Diophantine property of \( GL_{d+1}(\mathbb{R}) \) and \( SL_{d+1}(\mathbb{R}) \) matrices**

For notational reasons it is convenient to consider \( GL_{d+1}(\mathbb{R}) \) instead of \( GL_d(\mathbb{R}) \).

2.1. **\( GL_{d+1}(\mathbb{R}) \) actions.** Let \( A \in GL_{d+1}(\mathbb{R}) \) be a matrix that strictly preserves a cone \( \Sigma = \Sigma_{v_1, \ldots, v_{d+1}} \subseteq \mathbb{R}^{d+1} \). Without loss of generality, we can assume that \( \Sigma \setminus \{0\} \) is contained in the halfspace \( \{ x \in \mathbb{R}^{d+1} : x_{d+1} > 0 \} \). It is convenient to represent the induced action of \( A \) on \( \mathbb{R}^d \) on the affine hyperplane \( \{ x \in \mathbb{R}^{d+1} : x_{d+1} = 1 \} \), and
consider the corresponding action on \( \mathbb{R}^d \). To be precise, for \( x = (x_1, \ldots, x_d) \in \mathbb{R}^d \), we consider \((x, 1) = (x_1, \ldots, x_d, 1) \in \mathbb{R}^{d+1}\) and let
\[
f_A(x) = P_d \left( \frac{A(x, 1)}{A(x, 1)_{d+1}} \right), \quad \text{when} \quad A(x, 1)_{d+1} \neq 0,
\]
where \( P_d \) is the projection onto the first \( d \) coordinates. The components of \( f_A \) are rational functions, which are, of course, real-analytic on their domain. Consider
\[
\nabla := P_d(\Sigma \cap \{ x \in \mathbb{R}^{d+1} : x_{d+1} = 1 \}).
\]
By assumption, \( f_A \) is well-defined on \( \nabla \), and we have \( f_A(\nabla) \subset \nabla \).

We will also consider the action of \( A \) on the unit sphere, given by
\[
\varphi_A(x) := A \cdot x = \frac{Ax}{\|x\|},
\]
for a unit vector \( x \in \mathbb{S}^d \). Consider also \( \mathcal{U} \), the intersection of \( \Sigma \) with the upper hemisphere. We have \( \varphi_A(\mathcal{U}) \subset \mathcal{U} \). Lines through the origin provide a 1-to-1 correspondence between \( \mathcal{U} \) and \( \nabla \), which is bi-Lipschitz in view of the assumption \( \Sigma \setminus \{0\} \subset \{ x \in \mathbb{R}^{d+1} : x_{d+1} > 0 \} \).

It is well-known [3] (see also [4, Section 9]) that strictly preserving a cone implies that \( \varphi_A \) is a strict contraction in the Hilbert metric on \( \mathcal{U} \), which is by-Lipschitz with the round metric. We thus obtain the following:

**Lemma 2.1.** Suppose that the finite family \( A = \{A_i\}_{i \in \Lambda} \subset GL_d(\mathbb{R}) \) strictly preserves a simplicial cone \( \Sigma = \Sigma_{v_1, \ldots, v_{d+1}} \subset \{ x \in \mathbb{R}^{d+1} : x_{d+1} > 0 \} \cup \{0\} \). Then the associated IFS \( F_A = \{f_A\}_{A \in \Lambda} \) is real-analytic and uniformly hyperbolic on \( \nabla \subset \mathbb{R}^d \), in the sense that there exist \( C > 0 \) and \( \gamma \in (0, 1) \) such that
\[
\max_{x \in \nabla} \|f'_i(x)\| \leq C\gamma^n, \quad \text{for all} \quad i \in \Lambda^n,
\]
where \( f_i = f_{A_{i_1}} \circ \cdots \circ f_{A_{i_n}} \) and \( \|f'_i(x)\| \) is the operator norm of the differential at the point \( x \).

### 2.2. From exponential separation to the Diophantine property
Recall the strong exponential separation condition (Definition 1.8).

**Proposition 2.2.** Let \( A \) be a finite family of \( GL_{d+1}(\mathbb{R}) \) matrices, and let \( \Phi_A \) be the induced IFS on \( \mathbb{S}^d \). If \( \Phi_A \) satisfies the strong exponential separation on a nonempty set, then \( A \) is strongly Diophantine.

**Proof.** Let \( C_1 = \max_{i \in \Lambda} \{1, \|A_i\|\} \) and \( C_2 = \max_{i \in \Lambda} \{1, \|A_i^{-1}\|\} \). Suppose that \( i \neq j \) in \( \Lambda^n \). Let us write
\[
i = (i \wedge j)u, \quad j = (i \wedge j)v,
\]
where \( i \land j \) is the common initial segment of \( i \) and \( j \), so that \( u = u_1 \ldots u_k, \ v = v_1 \ldots v_k \) for some \( k \leq n \), with \( u_1 \neq v_1 \). We have

\[
(2.1) \quad \| A_i - A_j \| \geq \| A_{i \land j}^{-1} \| A_u - A_v \| \geq C_2^{-n} \| A_u - A_v \|.
\]

**Lemma 2.3.** For any \( A, B \in GL_{d+1}(\mathbb{R}) \) and any unit vector \( x \in \mathbb{R}^{d+1} \), we have

\[
\| A \cdot x - B \cdot x \| \leq \| A^{-1} \| (1 + \| B \| \| B^{-1} \|) \cdot \| A - B \|.
\]

**Proof.** We have

\[
\| A \cdot x - B \cdot x \| = \left\| \frac{Ax}{\| Ax \|} - \frac{Bx}{\| Bx \|} \right\| \leq \left\| \frac{Ax}{\| Ax \|} \right\| + \left\| \frac{Bx}{\| Bx \|} \right\| =: R_1 + R_2.
\]

Since

\[
1 = \| A^{-1} (Ax) \| \leq \| A^{-1} \| \| Ax \|,
\]
we have \( \| Ax \|^{-1} \leq \| A^{-1} \| \). Therefore,

\[
R_1 \leq \| A - B \| \cdot \| A^{-1} \|.
\]

Similarly,

\[
R_2 \leq \| B \| \cdot \| Ax \|^{-1} \| Bx \| \cdot \| Ax \|^{-1} \| Bx \|^{-1} \leq \| B \| \cdot \| A - B \| \cdot \| A^{-1} \| \| B^{-1} \|,
\]
and the desired estimate follows. \( \square \)

Applying the lemma to \( A_u \) and \( A_v \) yields, in view of \( \| A_w \| \leq C_1^n, \| A_w^{-1} \| \leq C_2^n \) for any \( w \in A^k, k \leq n \):

\[
(2.2) \quad \| A_u - A_v \| \geq 2^{-n} C_1^{-n} C_2^{-2n} \| A_u \cdot x - A_v \cdot x \|.
\]

Now we continue with the proof of the lemma. By assumption, \( \Phi_A \) satisfies the exponential separation condition on a nonempty set. Let \( c \in (0, 1) \) be the constant from the definition (1.9). It follows that there exists \( x \in S^d \) such that

\[
\| A_u \cdot x - A_v \cdot x \| \geq c^k \geq c^n.
\]
Combining this inequality with (2.2) and (2.1) yields

\[
\| A_i - A_j \| \geq 2^{-n} C_1^{-n} C_2^{-3n} c^n,
\]
confirming the strong Diophantine property. \( \square \)
2.3. Dimension of exceptions for one-parameter families. We consider a one-parameter family of real-analytic IFS on a compact subset of $\mathbb{R}^d$, and show that under some mild assumptions it satisfies the exponential separation condition outside of a Hausdorff dimension zero set. This section is based on [15, Section 5.4] and [16, Section 6.6], but we had to make a substantial number of modifications in the definitions and proofs.

Let $J$ be a compact interval in $\mathbb{R}$ and $V$ an open set in $\mathbb{R}^d$. For each $i \in \Lambda$ and $t \in J$, we assume that $f_{i,t} : V \to V$ is a continuously differentiable function. Assume also that $t \mapsto f_{i,t}(x)$ is real-analytic on a neighborhood of $J$ for any $i \in \Lambda$ and $x \in V$. Denote $F_t = \{f_{i,t}\}_{i \in \Lambda}$.

Further, assume that the IFS is uniformly hyperbolic in the following sense: there exist $C > 0$ and $0 < \gamma < 1$, such that

\begin{equation}
\|f'_{i,t}(x)\| \leq C\gamma^n, \quad \text{for all } i \in \Lambda^n, \ x \in V, \ t \in J.
\end{equation}

Fix $x_0 \in V$. For any finite sequence $i \in \Lambda^n$ we define $F_i(t) = f_{i,t}(x_0)$.

Of course, this depends on $x_0$, but we suppress it from notation. For $i \in \Lambda^N$ we have

\begin{equation}
\Pi_t(i) = F_i(t) := \lim_{n \to \infty} F_{i|_n}(t),
\end{equation}

where $\Pi_t : \Lambda^N \to \mathbb{R}^d$ is the natural projection corresponding to the IFS $F_t$ and $i|_n = i_1 \ldots i_n$. Notice that this is already independent of $x_0$. The proof of the next Claim is standard and follows from uniform hyperbolicity (2.3).

**Claim.** If $i^{(k)} \in \Lambda^{n(k)}$ is a sequence of words, such that $i^{(k)} \to i \in \Lambda^N$, in the sense that for any $N \in \mathbb{N}$ we have $i^{(k)}|_N = i|_N$ for $k$ sufficiently large, then

$F_{i^{(k)}}(\cdot) \to F_i(\cdot)$

uniformly on $J$. In particular, $F_{i|_n}(\cdot) \to F_i(\cdot)$ uniformly on $J$ for all $i \in \Lambda^N$. Thus $F_i(\cdot)$ is real-analytic on $J$, for any $i \in \Lambda^N$.

Next, for $i,j \in \bigcup_{n=1}^{\infty} \Lambda^n \cup \Lambda^N$, let

$\Delta_{ij}(t) = F_i(t) - F_j(t) \in \mathbb{R}^d$.

For any $\varepsilon > 0$, let

$E_\varepsilon = \bigcap_{N=1}^{\infty} \bigcup_{n>N} \left( \bigcup_{i,j \in \Lambda^n, i_1 \neq j_1} \left( \Delta_{ij} \right)^{-1} B_{\varepsilon^n} \right)$

and

\begin{equation}
E = \bigcap_{\varepsilon > 0} E_\varepsilon,
\end{equation}

where \( B_{\varepsilon^n} = \{ x \in \mathbb{R}^d : \| x \| \leq \varepsilon^n \} \). It is easy to see that if \( t \notin E \) then \( \mathcal{F}_t \) satisfies the strong exponential separation condition.

**Remark 2.4.** In [15, 16] Hochman considered the case where \( \mathcal{F}_t \) is an affine IFS. He defined the sets \( E'_\varepsilon \) and \( E' \) as follows:

\[
E'_\varepsilon = \bigcup_{N=1}^{\infty} \bigcap_{n>N} \left( \bigcup_{i,j \in \Lambda^n, i \neq j} (\Delta_{i,j}^{-1})B_{\varepsilon^n} \right)
\]

and

\[
E' = \bigcap_{\varepsilon > 0} E'_\varepsilon.
\]

If \( t \notin E' \) then \( \mathcal{F}_t \) satisfies the strong exponential separation condition along a subsequence.

**Definition 2.5.** For a family of IFS \( \mathcal{F}_t, t \in \mathcal{J} \), as above, and for \( i,j \in \Lambda^N \) let \( \Delta_{i,j}(t) = F_i(t) - F_j(t) \). We say that the family is non-degenerate if

\[
\Delta_{i,j}(\cdot) \equiv 0 \iff i = j \quad \text{for all } i,j \in \Lambda^N.
\]

We next prove the following:

**Theorem 2.6.** Suppose that the family of IFS \( \mathcal{F}_t, t \in \mathcal{J} \), is non-degenerate. Then the set \( E \) from (2.5) has Hausdorff dimension zero, and therefore, \( \mathcal{F}_t \) satisfies the strong exponential separation condition on \( \mathcal{J} \), outside of a set of zero Hausdorff dimension.

**Corollary 2.7.** For a family of IFS \( \mathcal{F}_t, t \in \mathcal{J} \), as above, assume that there exists \( t_0 \in \mathcal{J} \) such that the sets \( \{ f_{i,t_0}(V) \}_{i \in \Lambda} \) are pairwise disjoint. Then (2.7) holds, and hence the set \( E \) from (2.5) has Hausdorff dimension zero.

Hochman [15, 16] proved, for a non-degenerate family of affine IFS, with a real-analytic dependence on parameter, that the set \( E' \) from (2.6) has packing dimension zero.

For any smooth function \( F : \mathcal{J} \to \mathbb{R}^d \), denote \( F^{(p)}(t) = \frac{d^p}{dt^p} F(t) \).

**Definition 2.8.** The family \( \{ \mathcal{F}_t \}_{t \in \mathcal{J}} \) is said to be transverse of order \( k \) if there exists \( c > 0 \) such that for all \( n \in \mathbb{N} \) and \( i,j \in \Lambda^n \), with \( i_1 \neq j_1 \), we have

\[
\forall t \in \mathcal{J} \quad \exists p \in \{ 0, \cdots, k \} \quad \text{s.t.} \quad \| \Delta_{i,j}^{(p)}(t) \| > c.
\]

Here the norm \( \| \cdot \| \) is simply the Euclidean norm in \( \mathbb{R}^d \).

**Remark 2.9.** The above definition is different from [15] and it simplifies the proof of Theorem 2.6.
Lemma 2.10. Suppose that the non-degeneracy condition (2.7) holds. Then \( \{F_t\}_{t \in \mathcal{J}} \) is transverse of order \( k \) for some \( k \in \mathbb{N} \).

Proof. Suppose that for all \( k \in \mathbb{N} \) the family \( \{F_t\}_{t \in \mathcal{J}} \) is not transverse of order \( k \). Then by assumption, for \( \{c_k\} \) with \( c_k < 1/k \), we can choose \( n(k), i^{(k)}, j^{(k)} \in \Lambda^{n(k)} \) with \( i^{(k)}_1 \neq j^{(k)}_1 \) and a point \( t_k \in \mathcal{J} \) such that
\[
\|\Delta^{(p)}_{i^{(k)}_1,j^{(k)}_1}(t_k)\| < c_k
\]
for \( 0 \leq p \leq k \). Passing to a subsequence \( \{k_i\} \), we can assume that \( t_{k_i} \to t_0 \in \mathcal{J} \), \( i^{(k_i)} \to i \in \Lambda^\mathbb{N} \) and \( j^{(k_i)} \to j \in \Lambda^\mathbb{N} \), with \( i_1 \neq j_1 \). Note that \( \Delta^{(p)}_{i^{(k_i)}_1,j^{(k_i)}_1} \to \Delta_{i,j} \) uniformly on \( \mathcal{J} \) and the same holds for \( p \)-th derivatives by real-analyticity. Hence for all \( p \geq 0 \), we have
\[
\|\Delta^{(p)}_{i,j}(t_0)\| = \lim_{i \to \infty} \|\Delta^{(p)}_{i^{(k_i)}_1,j^{(k_i)}_1}(t_{k_i})\| = 0.
\]
Since \( \Delta_{i,j} \) is real-analytic, the vanishing of its derivatives implies \( \Delta_{i,j} \equiv 0 \) on \( \mathcal{J} \), contradicting (2.7), since \( i \neq j \) by construction. \( \square \)

For a \( C^k \)-smooth function \( F : \mathcal{J} \to \mathbb{R} \), write
\[
\|F\|_{\mathcal{J},k} = \max_{p \in \{0, \ldots, k\}} \max_{t \in \mathcal{J}} |F^{(p)}(t)|.
\]

Lemma 2.11 (Lemma 5.8 in [15]). Let \( k \in \mathbb{N} \) and let \( F : \mathcal{J} \to \mathbb{R} \) be a \( k \) times continuously differentiable function. Let \( M = \|F\|_{\mathcal{J},k} \), and let \( 0 < c < 1 \) be such that for every \( t \in \mathcal{J} \) there is \( p \in \{0, \ldots, k\} \) with \( |F^{(p)}(t)| > c \). Then there exists \( C = C_{c,M,|\mathcal{J}|} \geq 1 \) such that for every \( 0 < \rho < (c/2)^{2^k} \), the set \( F^{-1}(-\rho, \rho) \cap \mathcal{J} \) can be covered by \( C^k \) intervals of length \( \leq 2(\rho/c)^{1/2^k} \) each.

Lemma 2.12. If \( \{F_t\}_{t \in \mathcal{J}} \) is transverse of order \( k \geq 1 \) on the compact interval \( \mathcal{J} \), then the set \( E \) from (2.7) has Hausdorff dimension zero.

Proof. Extending the real-analytic functions to the complex plane, by Cauchy’s formula, since
\[
\sup_n \sup_{i,j \in \Lambda^n, i \neq j} \|\Delta_{i,j}\|_{\mathcal{V},0} < \infty
\]
on a neighborhood \( \mathcal{V} \) of \( \mathcal{J} \), and \( \Delta_{i,j}(\cdot) \) is real-analytic for all \( i, j \in \Lambda^n \), we have
\[
M := \sup_n \sup_{i,j \in \Lambda^n, i \neq j} \|\Delta_{i,j}\|_{\mathcal{J},k} < \infty.
\]
Let
\[
E_{\varepsilon,n} = \bigcup_{i,j \in \Lambda^n, i \neq j} (\Delta_{i,j})^{-1}(B_{\varepsilon^n}).
\]
Then
\[
E_{\varepsilon} = \bigcap_{N=1}^{\infty} \bigcup_{n>N} E_{\varepsilon,n}.
\]
Let \( i, j \in \Lambda^n \), with \( i_1 \neq j_1 \), and assume that \( \| \Delta_{i,j}(t) \| < \varepsilon^n \). By Lemma 2.11 applied to a component of \( \Delta_{i,j} \), for \( \varepsilon \) sufficiently small, the set

\[
(\Delta_{i,j})^{-1}(B_{\varepsilon^n})
\]

may be covered by \( C^k \) intervals of length \( \leq 2(\varepsilon^n \cdot c^{-1})^{1/2^k} \). It follows that the set \( E_{\varepsilon,n} \) from (2.9) may be covered by \( O(|\Lambda|^2n \cdot C^k) \) intervals of length \( \leq (\varepsilon^n \cdot c^{-1})^{1/2^k} \).

Fix \( s > 0 \) and write \( H_s \) for the \( s \)-dimensional Hausdorff measure. We obtain from (2.10) that

\[
H^s(E_{\varepsilon}) \leq O(1) \cdot \sum_{n \geq 1} |\Lambda|2n C^k (\varepsilon^n \cdot c^{-1})^{s/2^k} < \infty
\]

for \( \varepsilon \) sufficiently small. It follows that \( H^s(E) = 0 \). □

Proof of Theorem 2.6. This is now immediate from Lemmas 2.10 and 2.12. □

2.4. Proof of Theorem 1.2. The next lemma follows by an application of Fubini’s Theorem.

Lemma 2.13. Let \( F \subset \mathbb{R}^n \) and let \( v \in \mathbb{R}^n \) be a nonzero vector. Assume that for every \( x_0 \in \mathbb{R}^n \), the set \( \{x_0 + tv : t \in \mathbb{R}\} \cap F \) has 1-dimensional Lebesgue measure 0. Then the set \( F \) has \( n \)-dimensional Lebesgue measure 0.

Proof of Theorem 1.2. (i) Let \( \Sigma = \Sigma_{v_1, \ldots, v_{d+1}} \) be a simplicial cone in \( \mathbb{R}^{d+1} \). Let \( \mathcal{U} \subset X_{\Sigma,m} \) be a small open set in \( (GL_{d+1}(\mathbb{R}))^m \) of \( m \)-tuples of matrices for which \( \Sigma \) is strictly invariant. Choose vectors \( w_i \in \mathbb{R}^{d+1} (i \in \Lambda) \), with distinct directions, in such a way that

\[
(2.11) \quad w_i \in \Sigma_{A_{i}v_1, \ldots, A_{i}v_{d+1}} \quad \text{for all} \quad (A_i)_{i \in \Lambda} \in \mathcal{U}.
\]

This is possible when \( \mathcal{U} \) is sufficiently small. Let \( (A_i)_{i \in \Lambda} \in \mathcal{U} \), and for each \( t \geq 0 \) and \( i \in \Lambda \) let \( A_{i,t} \) be such that

\[
A_{i,t}v_j = A_i v_j + tw_i, \quad j = 1, \ldots, d + 1.
\]

Condition (2.11) guarantees that \( \{A_{i,t}v_j\}_{i \in \Lambda, j = 1}^{d+1} \) is linearly independent, and hence \( A_{i,t} \in GL_{d+1}(\mathbb{R}) \) for all \( t > 0 \). This is a consequence of the following elementary claim.

Claim. Let \( y_1, \ldots, y_{d+1} \in \mathbb{R}^{d+1} \) be linearly independent, and suppose that \( w = \sum_{k=1}^{d+1} a_k y_k \) for some \( a_k \geq 0 \). Then the family \( \{y_1 + w, \ldots, y_{d+1} + w\} \) is linearly independent as well.

Proof of the Claim. We have

\[
\sum_{j=1}^{d+1} c_j \left( y_j + \sum_{k=1}^{d+1} a_k y_k \right) = 0 \quad \Rightarrow \quad \sum_{j=1}^{d+1} \left( c_j + a_j \sum_{k=1}^{d+1} c_k \right) y_j = 0,
\]

where \( c_j \) are coefficients.
observe that, given \( \varepsilon > 0 \), the IFS is given by rational functions. Condition (2.3) holds if and only if \( \sum_{k=1}^{d+1} c_k \neq 0 \), we obtain a contradiction, in view of \( a_j \geq 0, j = 1, \ldots, d + 1 \); thus \( c_j = 0, j = 1, \ldots, d + 1 \), as claimed. 

Let \( A_t = \{ A_{i,t} \}_{i \in \Lambda} \) be the family of matrices defined above, for \( t \geq 0 \), and let \( \mathcal{F}_t = \mathcal{F}_{A_t} \) be the corresponding one-parameter family of IFS on the set \( \overline{V} \subset \mathbb{R}^d \) obtained by projection of \( \Sigma \cap \{ x \in \mathbb{R}^{d+1} : x_d = 1 \} \) onto \( \mathbb{R}^d \). Notice that the cone \( \Sigma \) is strictly preserved by all \( A_t, t \geq 0 \), by construction, hence by Lemma 2.1 these IFS are all uniformly hyperbolic. It is easy to see that the dependence on \( t \) is real-analytic, since the IFS is given by rational functions. Condition (2.3) holds for \( t \in [0, M] \), for any \( M < \infty \), by uniform hyperbolicity and compactness. Finally, observe that, given \( \varepsilon > 0 \), for \( t \) sufficiently large, we have

\[
\bigcap_{x \in \mathbb{R}^d} \bigcap_{x_d = 1} \{ x \}
\]

where \( \Sigma_c(w_i) \) is the cone of vectors \( \varepsilon \)-close to \( w_i \) in direction. By construction, \( w_i \) are all distinct, hence Corollary 2.1 applies. We obtain that for all \( t \in [0, \infty) \) outside a set of Hausdorff dimension zero, the IFS \( \mathcal{F}_t \) satisfies the exponential separation condition, and then Proposition 2.2 implies that the \( m \)-tuple of matrices \( (A_{i,t})_{i \in \Lambda} \) is Diophantine for all \( t \) outside of a zero-dimensional set, so certainly for Lebesgue-a.e. \( t \). Now Lemma 2.3 yields the desired claim.

(ii) We consider \((\text{SL}_{d+1}(\mathbb{R}))^m\) as a codimension-\( m \) submanifold of \((\text{GL}_{d+1}(\mathbb{R}))^m \subset \mathbb{R}^{(d+1)^2m}\). In the proof of part (i) we showed that for a.e. \( (A_i)_{i \in \Lambda} \in \mathcal{X}_{\Sigma,m} \), the induced IFS on a subset of \( \mathbb{R}^d \) satisfies the strong exponential separation condition. Suppose that there is a positive measure subset \( \mathcal{E} \subset \mathcal{Y}_{\Sigma,m} \) for which the strong Diophantine condition is violated. Then for every \( (A_i)_{i \in \Lambda} \in \mathcal{E} \), the induced IFS \( \Phi \) does not have strong exponential separation, by another application of Proposition 2.2. However, \( (A_i)_{i \in \Lambda} \in \mathcal{Y}_{\Sigma,m} \) and \( (c_i A_i)_{i \in \Lambda} \in \mathcal{X}_{\Sigma,m} \), for any \( c_i > 0 \), induce the same IFS on the projective space, and we get a set of positive measure in \( \mathcal{X}_{\Sigma,m} \) for which the strong Diophantine condition does not hold. This is a contradiction, and the theorem is proved completely. 

3. Dimension of the attractor

Let \( A \in \text{SL}_2(\mathbb{R}) \). It is easy to see that \( A^* A \) has eigenvalues \( \|A\|^2, \|A\|^{-2} \). Let \( (\cos t_A, \sin t_A)^t \) be the unit eigenvector corresponding to the eigenvalue \( \|A\|^{-2} \), where \( t_A \in [0, \pi) \). We recall some basic properties of the map \( \varphi_A \). For more details see sections 2.2, 2.3 and 2.4 in [17]. The following simple lemma is [17, Section 2.4].

**Lemma 3.1.** Let \( A \in \text{SL}_2(\mathbb{R}) \). Then the induced map \( \varphi_A \) expands by at most \( \|A\|^2 \) and contracts by at most \( \|A\|^{-2} \). Furthermore, for any \( \varepsilon > 0 \) there exists \( C_\varepsilon > 1 \) such that \( \|A\|^{-2} \leq |\varphi'_A(x)| < C_\varepsilon \|A\|^{-2} \) for all \( x \in [0, \pi) \setminus (t_A - \varepsilon, t_A + \varepsilon) \).

The following lemma is now immediate.
Lemma 3.2. Let $U \subset (0, \pi)$ be an open set. Then, for every $\varepsilon > 0$ there exists $C_\varepsilon > 1$ such that for any $A \in SL_2(\mathbb{R})$ with $(t_A - \varepsilon, t_A + \varepsilon) \subset U$, we have
\[ \pi - C_\varepsilon \|A\|^{-2} < |\varphi_A(U)| < \pi. \]

Lemma 3.2 implies the following:

Lemma 3.3. Let $U \subset (0, \pi)$ be an open set. Then, for every $\varepsilon > 0$ there exists $M = M(\varepsilon) > 0$ such that the following holds: for any $A \in SL_2(\mathbb{R})$ that satisfies $\varphi_A(U) \subset U$ and $\|A\| > M$, we have $(t_A - \varepsilon, t_A + \varepsilon) \not\subset U$.

Let $A = \{A_i\}_{i \in \Lambda}$ be a finite collection of $SL_2(\mathbb{R})$ matrices and let $\Phi = \{\varphi_A\}_{A \in \mathcal{A}}$ be the corresponding IFS on $[0, \pi) \cong \mathbb{RP}^1$. Recall the notation:
\[ \Phi(E) = \bigcup_{A \in \mathcal{A}} \varphi_A(E). \]

Assume that there is a strictly invariant multicone $U \subset [0, \pi)$, that is, a nonempty open set having finitely many connected components with disjoint closures, such that $\overline{U} \neq \mathbb{RP}^1$ and $\Phi(\overline{U}) \subset U$.

By Theorem 1.5, the associated cocycle is uniformly hyperbolic, which implies that there exist $c > 0$ and $\lambda > 1$ such that
\[ \|A_i\| \geq c \lambda^n \quad \text{for all } i \in \Lambda^n, \ n \in \mathbb{N}, \quad (3.1) \]
see [27] and [3, Theorem 2.2]. Fix $\varepsilon > 0$ such that the $(2\varepsilon)$-neighborhood of $\Phi(\overline{U})$ is contained in $U$, and let $M = M(\varepsilon)$ from Lemma 3.3. By (3.1), there exists $n_0 \in \mathbb{N}$ such that $\|A_i\| > M$ for $i \in \Lambda^n$, $n \geq n_0$. Lemma 3.3 implies that $(t_{A_i} - \varepsilon, t_{A_i} + \varepsilon) \not\subset U$, hence
\[ (t_{A_i} - \varepsilon, t_{A_i} + \varepsilon) \cap \Phi(\overline{U}) = \emptyset, \quad \text{for all } i \in \Lambda^n, \ n \geq n_0. \]

Hence, by Lemmas 3.1 and 3.2 we obtain
\[ \|A_i\|^{-2} \leq |\varphi'_i(x)| \leq C_\varepsilon \|A_i\|^{-2}, \quad \text{for all } x \in \overline{U}, \ i \in \Lambda^n, \ n \geq n_0. \quad (3.2) \]

Thus we obtain

Lemma 3.4. (i) The Bounded Distortion Property holds for $\Phi$ on $U$: there exists $C' > 1$ such that
\[ \frac{1}{C'} \leq \frac{|\varphi'_i(x)|}{|\varphi'_i(y)|} \leq C' \quad \text{for all } x, y \in \overline{U}, \ i \in \Lambda^n, \ n \in \mathbb{N}. \quad (3.3) \]

(ii) The IFS $\Phi^k$ is contractive on $U$ in the metric $d_\mathcal{P}$ for sufficiently large $k$. More precisely, there exists $C'' > 0$ such that
\[ \|\varphi'_i\|_{\overline{U}} \leq C'' \lambda^{-2n}, \quad (3.4) \]
where $\lambda > 1$ is from (3.1).
(iii) We have
\[ s = s_A/2, \]
where \( s \) is the solution of the Bowen’s equation \( P_\Phi(s) = 0 \), with the pressure given by (1.7) and \( s_A \) is the critical exponent, given by (1.4).

Let \( p = (p_i)_{i \in \Lambda} \) be a probability vector, and let \( x_0 \in U \). Let \( \chi_{\Phi,p} \) be the almost sure value of the limit
\[ \lim_{n \to \infty} -\frac{1}{n} \log |(\varphi_{i_1} \cdots i_n)'(x_0)|, \]
where \( i_1, i_2, \cdots \in \Lambda \) is a sequence chosen randomly according to the probability vector \( p = (p_i)_{i \in \Lambda} \). The equations (3.2) and (1.12) imply

Lemma 3.5. We have \( \chi_{\Phi,p} = 2\chi_{A,p} \).

By the Birkhoff Ergodic Theorem, applied to the shift transformation on \( \Lambda^N \) with the measure \( \mu = p^N \), in view of the bounded distortion (3.3), we have

(3.6) \[ \chi_{\Phi,p} = \lim_{n \to \infty} -\frac{1}{n} \log |(\varphi_{i_1} \cdots i_n)'(\Pi(i))| = -\int_{\Lambda^N} \log |\varphi_{i_1}'(\Pi(i))| \, d\mu(i), \]
where \( \Pi : \Lambda^N \to \mathbb{RP}^1 \) is the natural projection corresponding to \( \Phi \).

Lemma 3.6. (i) Let \( A \) be a finite set of matrices in \( GL_2(\mathbb{R}) \), and let \( \Phi \) be the IFS induced by \( A \) on the projective line \( \mathbb{RP}^1 \). If \( \Phi \) satisfies the strong exponential separation condition on a nonempty set, then \( A \) is strongly Diophantine.

(ii) Let \( A \) be a finite set of matrices in \( SL_2(\mathbb{R}) \), and let \( \Phi \) be the IFS induced by \( A \) on the projective line \( \mathbb{RP}^1 \). Then \( \Phi \) satisfies the strong exponential separation condition on a set containing at least three points if and only if \( A \) is strongly Diophantine.

Proof. (i) This is a special case of Proposition 2.2 since exponential separation for the induced action on a subset of \( \mathbb{RP}^1 \) is equivalent to that for the induced action on a subset of the circle.

(ii) One direction, that the strong exponential separation for \( \Phi \) implies the strong Diophantine property for \( A \), follows from (i). For the converse, we refer to [17, Lemma 2.5], which says that \( SL_2(\mathbb{R}) \) is quantitatively separated by the action on three points of \( \mathbb{RP}^1 \).

Proof of Theorem 1.6. Recall that \( A \) is a finite set of \( SL_2(\mathbb{R}) \) matrices satisfying the strong Diophantine condition and having a strictly invariant multicone \( U \), and \( \Phi = \Phi_A \) is the associated IFS on \( U \). Then \( \Phi \) has a compact attractor \( K \), and our
goal is to show that \( \dim_H(K) = s \), where \( P_\Phi(s) = 0 \) and \( P_\Phi \) is given by (1.6). It is known that
\[
(3.7) \quad \dim_H(K) \leq s,
\]
see the appendix for a short proof. Let us show the opposite inequality.

Let \( d_n > 0 \) be the solution of the equation
\[
\sum_{i \in \Lambda^n} |U_i|^{d_n} = 1,
\]
where \( U_i = \varphi_i(U) \) and \( |\cdot| \) denotes the Lebesgue measure on \([0, \pi) \cong \mathbb{RP}^1\). It is not hard to see that
\[
(3.8) \quad \lim_{n \to \infty} d_n = s.
\]
For convenience of the reader, we include the proof in the appendix, following [23].

Let \( p(n) = (p_i^{(n)})_{i \in \Lambda} \) be the probability vector such that \( p_i^{(n)} = |U_i|^{d_n} \). Let \( \eta^{(n)} \) be the invariant probability measure for the IFS \( \Phi^n \) on \( U \), corresponding to \( p^{(n)} \). Since \( \eta^{(n)} \) is supported on \( K \), we have \( \dim \eta^{(n)} \leq \dim_H(K) \).

We claim that \( A \) satisfies the assumptions of Theorem 1.11. Indeed, the existence of a strictly invariant multicone is known to imply that all the matrices in \( A \) are hyperbolic, hence the group generated by \( A \) is unbounded. Further, we assumed that not all attracting fixed points of \( A \) are the same, hence this group is totally irreducible. Thus the Furstenberg measure for \( (\mathcal{A}^n, p^{(n)}) \) is unique, and it coincides with \( \eta^{(n)} \). Since \( A \) is Diophantine, we have that \( \mathcal{A}^n \) is Diophantine as well. Now, by Theorem 1.11 and Lemma 3.5, we have
\[
\frac{H(p^{(n)})}{\chi_{\mathcal{A}^n, p^{(n)}}} \leq \dim_H(K).
\]

We claim that there exists \( C > 0 \) such that
\[
(3.9) \quad \chi_{\mathcal{A}^n, p^{(n)}} \leq -\sum_{i \in \Lambda^n} |U_i|^{d_n} \log |U_i| + C \quad \text{for all} \quad n \in \mathbb{N}.
\]
Indeed, by [310], we have
\[
\chi_{\mathcal{A}^n, p^{(n)}} \leq \sum_{i \in \Lambda^n} \mu([i]) \cdot \log \left( \min_{x \in U} |\varphi_i'(x)| \right)^{-1},
\]
where \([i]\) is the cylinder set of sequences starting with \( i \). Now \( \mu([i]) = |U|^{d_n} \), and
\[
\min_{x \in U} |\varphi_i'(x)| \geq \frac{|U_i|}{C' |U|},
\]
by the Bounded Distortion Property (3.3). Therefore,
\[
\chi_{\Phi_n, p(n)} \leq \sum_{i \in \Lambda_n} |U_i|^{d_n} \log \frac{C'|U|}{|U_i|} = - \sum_{i \in \Lambda_n} |U_i|^{d_n} \log |U_i| + \log C'|U|.
\]

By (3.9), we have
\[
H(p(n)) \chi_{\Phi_n, p(n)} > - \sum_{i \in \Lambda_n} |U_i|^{d_n} \log |U_i|^{d_n} - \sum_{i \in \Lambda_n} |U_i|^{d_n} \log |U_i| + C_n = d_n \left( 1 + \sum_{i \in \Lambda_n} |U_i|^{d_n} \log |U_i| \right)^{-1}.
\]

Since \( \lim_{n \to \infty} d_n = s \) and \( \lim_{n \to \infty} - \sum_{i \in \Lambda_n} |U_i|^{d_n} \log |U_i| = \infty \), we obtain \( s \leq \dim_H(K) \), as desired. Finally, \( s = s_A \) by Lemma 3.4(iii).

\section*{4. Appendix: the proof of (3.7) and (3.8)}

\subsection*{4.1. Proof of (3.8) [23]}
We have a projective IFS \( \Phi = \{ \varphi \}_{i \in \Lambda} \) on a strictly invariant multicone \( U \). Observe that
\[
P_\Phi(t) = \lim_{n \to \infty} \frac{1}{n} \frac{\log}{n} \sum_{i \in \Lambda_n} \| \varphi_i^{[t]} \|^{[t]} = \lim_{n \to \infty} \frac{1}{n} \log \sum_{i \in \Lambda_n} |U_i|^{[t]},
\]
by the Bounded Distortion Property (3.3). Let
\[
Q_n = \frac{1}{n} \log \sum_{i \in \Lambda_n} |U_i|^{[s]}
\]
Since \( P_\Phi(s) = 0 \), we have \( \lim_{n \to \infty} Q_n = 0 \). Let \( r_1 > 0 \) be such that \( r_1 \leq |\varphi_i(x)| \) for all \( i \in \Lambda \) and \( x \in U \). Recall (3.4), which says that
\[
\| \varphi_i^{[t]} \|^{[t]} \leq C'' \lambda^{-2n},
\]
for \( C'' > 0 \) and \( \lambda > 1 \). Then \( r_1^t |U| \leq |U_i| < C'' \lambda^{-2n} |U| \) for \( i \in \Lambda_n \), and hence we have
\[
(r_1^t |U|)^{s-d_n} : |U_i|^{d_n} < |U_i| < C''(\lambda^{-2n} |U|)^{s-d_n} : |U_i|^{d_n}.
\]

In view of \( \sum_{i \in \Lambda_n} |U_i|^{d_n} = 1 \), we have
\[
\frac{1}{n} \log (r_1^t |U|)^{s-d_n} < Q_n < \frac{1}{n} \log C'' + (s-d_n)[\log |U| - 2n \log \lambda],
\]
and it follows that
\[
\frac{Q_n - (\log C'')/n}{-2 \log \lambda + (\log |U|)/n} < s - d_n < \frac{Q_n}{\log r_1 + (\log |U|)/n},
\]
which implies \( d_n \to s \), as desired. \( \Box \)
4.2. Proof of (3.7). Fix $\varepsilon > 0$. Then for sufficiently large $n$ we have $d_n < s + \varepsilon/2$. Thus
\[ \sum_{i \in \Lambda^n} |U_i|^{s+\varepsilon} < \sum_{i \in \Lambda^n} |U_i|^{d_n+\varepsilon/2} < (r_n^s |U|)^{\varepsilon/2} \to 0, \text{ as } n \to \infty. \]
Therefore, the $(s+\varepsilon)$-dimensional Hausdorff measure of $K$ is zero. By the definition of the Hausdorff dimension, this proves (3.7).

Acknowledgements

The authors would like to acknowledge Balazs Bara’ny for helpful comments and for telling us about the papers [3, 6].

References

[1] Menny Aka, Emmanuel Breuillard, Lior Rosenzweig, and Nicolas de Saxcé. Diophantine properties of nilpotent Lie groups. Compos. Math., 151(6):1157–1188, 2015.
[2] Menny Aka, Emmanuel Breuillard, Lior Rosenzweig, and Nicolas de Saxcé. Diophantine approximation on matrices and Lie groups. Geom. Funct. Anal., 28(1):1–57, 2018.
[3] Artur Avila, Jairo Bochi, and Jean-Christophe Yoccoz. Uniformly hyperbolic finite-valued SL(2, $\mathbb{R}$)-cocycles. Comment. Math. Helvet., 85(4):813–884, 2010.
[4] Michael F. Barnsley and Andrew Vince. Real projective iterated function systems. J. Geom. Anal., 22(4):1137–1172, 2012.
[5] Garrett Birkhoff. Extensions of Jentzsch’s theorem. Trans. Amer. Math. Soc., 85:219–227, 1957.
[6] Jairo Bochi and Nicolas Gourmelon. Some characterizations of domination. Math. Z., 263(1):221–231, 2009.
[7] Philippe Bougerol and Jean Lacroix. Products of random matrices with applications to Schrödinger operators, volume 8 of Progress in Probability and Statistics. Birkhäuser Boston, Inc., Boston, MA, 1985.
[8] Jean Bourgain. An application of group expansion to the Anderson-Bernoulli model. Geom. Funct. Anal., 24(1):49–62, 2014.
[9] Rufus Bowen. Hausdorff dimension of quasicircles. Inst. Hautes Études Sci. Publ. Math., (50):11–25, 1979.
[10] Roberto De Leo. A conjecture on the Hausdorff dimension of attractors of real self-projective iterated function systems. Exp. Math., 24(3):270–288, 2015.
[11] Roberto De Leo. On the exponential growth of norms in semigroups of linear endomorphisms and the Hausdorff dimension of attractors of projective iterated function systems. J. Geom. Anal., 25(3):1798–1827, 2015.
[12] Kenneth Falconer. Techniques in fractal geometry. John Wiley & Sons, Ltd., Chichester, 1997.
[13] Harry Furstenberg. Noncommuting random products. Trans. Amer. Math. Soc., 108:377–428, 1963.
[14] Alex Gamburd, Dmitry Jakobson, and Peter Sarnak. Spectra of elements in the group ring of SU(2). J. Eur. Math. Soc. (JEMS), 1(1):51–85, 1999.
[15] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy. Ann. of Math. (2), 180(2):773–822, 2014.
[16] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy in $\mathbb{R}^d$. 
*To appear in Memoirs of the Amer. Math. Soc.*, 2015. http://arxiv.org/abs/1503.09043.

[17] Michael Hochman and Boris Solomyak. On the dimension of Furstenberg measure for $SL_2(\mathbb{R})$ random matrix products. *Invent. Math.*, 210(3):815–875, 2017.

[18] Vadim Kaloshin and Igor Rodnianski. Diophantine properties of elements of $SO(3)$. *Geom. Funct. Anal.*, 11(5):953–970, 2001.

[19] François Ledrappier. Une relation entre entropie, dimension et exposant pour certaines marches aléatoires. *C. R. Acad. Sci. Paris Sér. I Math.*, 296(8):369–372, 1983.

[20] Russell Lyons. Singularity of some random continued fractions. *J. Theoret. Probab.*, 13(2):535–545, 2000.

[21] Steve Pincus. Singular stationary measures are not always fractal. *J. Theoret. Probab.*, 7(1):199–208, 1994.

[22] David Ruelle. Repellers for real analytic maps. *Ergodic Theory Dynam. Systems*, 2(1):99–107, 1982.

[23] Károly Simon and Boris Solomyak. Iterated function systems with overlaps, *unpublished manuscript*.

[24] Károly Simon, Boris Solomyak, and Mariusz Urbański. Hausdorff dimension of limit sets for parabolic IFS with overlaps. *Pacific J. Math.*, 201(2):441–478, 2001.

[25] Károly Simon, Boris Solomyak, and Mariusz Urbański. Invariant measures for parabolic IFS with overlaps and random continued fractions. *Trans. Amer. Math. Soc.*, 353(12):5145–5164, 2001.

[26] Dennis Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Math.*, 153(3-4):259–277, 1984.

[27] Jean-Christophe Yoccoz. Some questions and remarks about $SL(2, \mathbb{R})$ cocycles. In *Modern dynamical systems and applications*, pages 447–458. Cambridge Univ. Press, Cambridge, 2004.

Yuki Takahashi, Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel

E-mail address: takahashi@math.biu.ac.il

Boris Solomyak, Department of Mathematics, Bar-Ilan University, Ramat Gan, Israel

E-mail address: bsolom3@gmail.com