

# Undecidable properties of self-affine sets and multi-tape automata

TIMO JOLIVET<sup>1,2</sup> AND JARKKO KARI<sup>1</sup>

<sup>1</sup>*Department of Mathematics, University of Turku, Finland*

<sup>2</sup>*LIAFA, Université Paris Diderot, France*

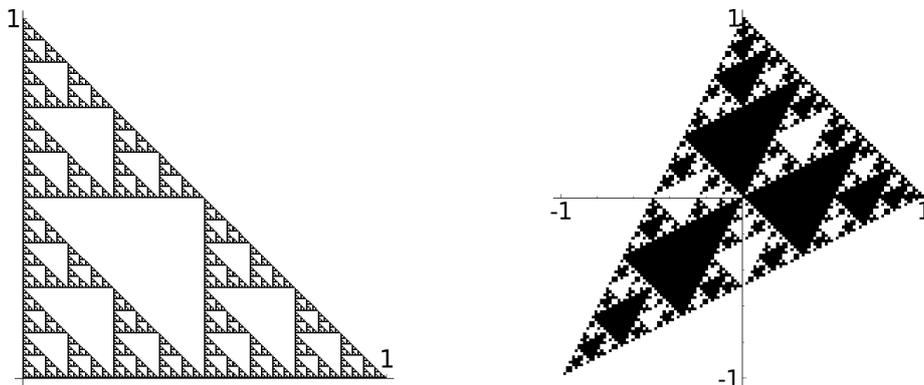
## Abstract

We study the decidability of the topological properties of some objects coming from fractal geometry. We prove that having empty interior is undecidable for the sets defined by two-dimensional graph-directed iterated function systems. These results are obtained by studying a particular class of self-affine sets associated with multi-tape automata. We first establish the undecidability of some language-theoretical properties of such automata, which then translate into undecidability results about their associated self-affine sets.

## 1 Introduction

A classical way to define fractals is to use an *iterated function system (IFS)*, specified by a finite collection of maps  $f_1, \dots, f_n : \mathbb{R}^d \rightarrow \mathbb{R}^d$  which are all *contracting*: there exists  $0 \leq c < 1$  such that  $\|f_i(x) - f_i(y)\| \leq c\|x - y\|$  for all  $x, y \in \mathbb{R}^d$ . The associated fractal set, called the *attractor* of the IFS, is the unique nonempty compact set  $X$  such that  $X = \bigcup_{i=1}^n f_i(X)$ . Such a set  $X$  always exists and is unique thanks to a famous result of Hutchinson [Hut81], based on an application of the Banach fixed-point theorem; see also [Fal03] or [Bar93]. For example, the classical Cantor set can be defined as the unique compact set  $X \subseteq \mathbb{R}$  satisfying the set equation  $X = \frac{1}{3}X \cup (\frac{1}{3}X + \frac{2}{3})$ . Two other examples are given in Figure 1.

A question of interest is to determine when the fractal set  $X$  has nonempty interior. This question arises in several areas, including tiling theory, dynamical systems, number theory and Fourier analysis (see [Wan99, LLR13] and references therein). A well studied case is when the



**Figure 1:** Two self-affine sets defined by  $X = \bigcup_{v \in \mathcal{D}} M^{-1}(X + v)$ , where  $M = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$  and  $\mathcal{D} = \{(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix})\}$  (left), and  $\mathcal{D} = \{(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} -1 \\ -1 \end{smallmatrix})\}$  (right). The set on the left is the Sierpiński triangle and has empty interior. The set on the right is an example of a self-affine tile with nonempty interior (see [LW96]).

contracting maps are affine mappings of the form  $f_i(x) = M^{-1}(x + v_i)$  where  $v_i \in \mathbb{Z}^d$  and  $M$  is an integer expanding matrix which is common to all the  $f_i$ , like in the examples of Figure 1. In this case, having nonempty interior is equivalent with having nonzero Lebesgue measure, and there are efficient algorithms to decide this [GY06, BK11].

Much less is known in the more general case of *self-affine attractors*, where the maps  $f_i$  are only restricted to be affine (of the form  $f_i = M_i x + v_i$  where the  $M_i$  are real matrices and  $v_i \in \mathbb{R}^d$ ). No algorithm is known to decide nonempty interior in this case, and specific results such as computation of Hausdorff dimension are known only for some very specific families of self-affine sets [Bed84, McM84, Fra13].

**Our results** We are interested in the following question: to what extent can we decide if a self-affine set has nonempty interior?

We will answer this question by an undecidability result for a natural generalization of iterated function systems, which consist of a *finite system* of equations instead of just one, hence defining a finite number of attractors. This is formalized in the following definition: a *d-dimensional graph-directed iterated function system (GIFS)* is a directed graph in which each edge  $e$  is labelled by a contracting mapping  $f_e : \mathbb{R}^d \rightarrow \mathbb{R}^d$ . The *attractors* of the GIFS are the unique nonempty compact sets  $\{X_q\}_{q \in \mathcal{Q}}$  such that

$$X_q = \bigcup_{r \in \mathcal{Q}} \bigcup_{e \in E_{q,r}} f_e(X_r),$$

where  $\mathcal{Q}$  is the set of vertices of the directed graph defining the GIFS, and  $E_{q,r}$  denote the set of edges from vertex  $q$  to vertex  $r$ . Again, such a collection of compact sets  $\{X_q\}_{q \in \mathcal{Q}}$  exists and is unique [Fal97]. Fractals defined by GIFS are widely used to define various self-similar tilings of the plane, the study of which have applications in physics, dynamics and number theory. Note that the case of single-vertex graphs corresponds to classical iterated function systems.

Our main result, Theorem 4.2, states that it is undecidable if the attractors of a 2-dimensional, 3-state affine GIFS have empty interior. We follow an approach initiated by Dube [Dub93] by associating self-affine sets with finite multi-tape automata. Then we relate some properties of the automaton with topological properties of its associated attractor, and we obtain the undecidability of the latter by proving the undecidability of the former. The original motivation of [Dub93] is to prove that it is undecidable if a if the attractor of a rational 2-dimensional affine IFS intersects the diagonal  $\{(x, x) : x \in [0, 1]\}$ .

In Section 2 we define multi-tape automata and we consider a variant of the Post correspondence problem in Section 2.2, which we use in Section 2.3 to prove undecidability results about multi-tape automata. We then relate some language-theoretical properties of an automaton with some topological properties of its attractor in Section 3. The main results are stated in Section 4.

**Acknowledgements.** Research supported by the Academy of Finland Grant 131558 and by project Fractals and Numeration ANR-12-IS01-0002.

## 2 Multi-tape automata

### 2.1 Definitions

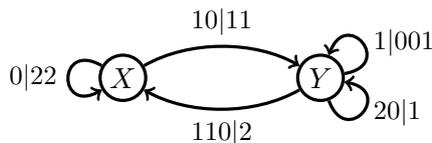
A *d-tape automaton*  $\mathcal{M}$  on alphabet  $\mathcal{A} = A_1 \times \cdots \times A_d$  is defined by a finite set of *states*  $\mathcal{Q}$ , and a finite set of *transitions*  $\mathcal{R} \subseteq \mathcal{Q} \times \mathcal{Q} \times (A_1^+ \times \cdots \times A_d^+)$ . A *d-tape automaton* on state  $\mathcal{Q}$  is conveniently represented by a directed graph with vertex set  $\mathcal{Q}$  and an edge  $(q, r)$  labelled by  $w_1 | \cdots | w_d$  for every transition  $(q, r, (w_1, \dots, w_d))$ . This is illustrated in Example 2.1.

A *configuration* is an infinite sequence  $c \in \mathcal{A}^{\mathbb{N}} = (A_1 \times \cdots \times A_d)^{\mathbb{N}}$ . For  $k \in \{1, \dots, d\}$ , the *kth tape* of  $c$  refers to the infinite sequence  $((c_n)_k)_{n \in \mathbb{N}}$ , which is an infinite concatenation of

words in  $A_k^*$ . For convenience, configurations will be denoted by writing the tape components separated by the symbol “|”. For example,  $00\dots | 11\dots | 00\dots$  denotes the 3-tape configuration  $(0, 1, 0), (0, 1, 0), \dots \in (\{0, 1\} \times \{0, 1\} \times \{0, 1\})^{\mathbb{N}}$ .

Let  $q$  be a state of  $\mathcal{M}$ . A configuration  $c \in A^{\mathbb{N}}$  is  **$q$ -accepted** by  $\mathcal{M}$  if there exists an infinite sequence of transitions  $((q_n, r_n, (w_{n,1}, \dots, w_{n,d})))_{n \geq 1}$  such that  $q_1 = q$ ,  $r_n = q_{n+1}$  for all  $n \geq 1$ , and for every  $k \in \{1, \dots, d\}$ , the infinite word  $w_{1,k}w_{2,k}\dots$  is equal to the  $k$ th tape of  $c$  (that is,  $w_{1,k}w_{2,k}\dots = (c_1)_k(c_2)_k\dots$ ). Such an infinite sequence of transitions will sometimes be referred to as a **run of  $\mathcal{M}$  starting at  $q$** . Note that we also forbid  $\varepsilon$ -transitions as the words  $w_1, \dots, w_d$  used in transitions are nonempty, so that each infinite run provides an infinite word on every tape.

**Example 2.1.** Consider the following 2-tape, 2-state automaton on alphabet  $\mathcal{A} = \{0, 1\} \times \{0, 1, 2\}$ , with state set  $\mathcal{Q} = \{X, Y\}$  and transitions given by the following.



It is easy to check that the configuration  $00\dots | 22\dots$  is not  $Y$ -accepted but is  $X$ -accepted by  $\mathcal{M}$  (by repeatedly using the transition  $(X, X, (0, 22))$ ). However, giving a precise description of the set of configurations which are accepted by  $\mathcal{M}$  seems difficult.

**Remark 2.2.** Multi-tape automata are very powerful computational devices because of the fact that the words  $w_1, \dots, w_d$  in a transition are allowed to have different lengths. This is the fundamental feature that will allow us to establish several undecidability results about multi-tape automata later in this section. On the other hand, if the words  $w_1, \dots, w_d$  all have the same length in every transition, then it is easy to see that this model is not more powerful than a classical finite automaton on a product alphabet.

## 2.2 Post correspondence problems

The undecidability results of this article are all derived from the undecidability of the following decision problems. The **Post correspondence problem (PCP)** is: given  $n$  pairs of nonempty words  $(u_1, v_1), \dots, (u_n, v_n)$ , decide if there exist  $m \geq 1$  and a word  $i_1 \dots i_m$  such that  $u_{i_1} \dots u_{i_m} = v_{i_1} \dots v_{i_m}$ . This is a well-known undecidable problem [Pos46].

We will need a slight variant of PCP, the **prefix Post correspondence problem (prefix-PCP)**: given  $n$  pairs of nonempty words  $(u_1, v_1), \dots, (u_n, v_n)$ , decide if there exist  $m, m' \geq 1$  and two words  $i_1 \dots i_m$  and  $i_1 \dots i_{m'}$  such that  $u_{i_1} \dots u_{i_m} = v_{i_1} \dots v_{i_{m'}}$  and one of the two words  $i_1 \dots i_m$  and  $i_1 \dots i_{m'}$  is a prefix of the other.

A positive PCP always yields a positive prefix-PCP instance (by taking  $m = m'$ ), but the converse is not always true. For example, the instance  $(u_1, v_1) = (a, abb)$ ,  $(u_2, v_2) = (bb, aa)$  admits the prefix-PCP solution given by  $u_1 u_2 u_1 u_1 = v_1 v_2 = abbaa$ , that is,  $m = 4$ ,  $m' = 2$  and the two words  $i_1 i_2 i_3 i_4 = 1211$  and  $i_1 i_2 = 12$ . However, this instance cannot admit any PCP solution because no pair of words ends by the same symbol.

**Lemma 2.3.** *Prefix-PCP is undecidable.*

*Proof.* We reduce PCP to prefix-PCP. Let  $(u_1, v_1), \dots, (u_n, v_n)$  be an instance of PCP on alphabet  $\mathcal{A}$ . Let  $\mathcal{B} = \mathcal{A} \cup \{\#, *\}$  be a new alphabet, where  $\#$  and  $*$  are two new symbols not contained in  $\mathcal{A}$ . We construct a prefix-PCP instance  $(A_1, B_1), \dots, (A_n, B_n), (U_1, V_1), \dots, (U_n, V_n), (Y, Z)$  on the new alphabet  $\mathcal{B}$ , defined by

$$\begin{aligned} A_i &= \#x_1*x_2*\dots*x_k & U_i &= *x_1*x_2*\dots*x_k & Y &= *\# \\ B_i &= \#y_1*y_2*\dots*y_\ell* & V_i &= y_1*y_2*\dots*y_\ell* & Z &= \# \end{aligned}$$

for all  $i \in \{1, \dots, n\}$ , where  $u_i = x_1 \cdots x_n$  and  $v_i = y_1 \cdots y_\ell$  and the  $x_j, y_j$  are in  $\mathcal{A}$ . We now prove that the PCP instance has a solution if and only if the prefix-PCP instance has a solution. Suppose that there exists a solution  $i_1 \cdots i_m$  to the PCP instance, that is  $u_{i_1} \cdots u_{i_m} = v_{i_1} \cdots v_{i_m}$ . Then clearly the prefix-PCP also has a solution, given by  $A_{i_1} U_{i_2} \cdots U_{i_m} Y = B_{i_1} V_{i_2} \cdots V_{i_m} Z$ .

Conversely, suppose that the prefix-PCP instance has a solution. By construction, because of  $\#$  and  $*$ , there must exist a prefix-PCP solution of the form  $A_{i_1} U_{i_2} \cdots U_{i_m} Y = B_{i_1} V_{i_2} \cdots V_{i_{m'}} Z$ , where  $i_1 \cdots i_m$  is a prefix of  $i_1 \cdots i_{m'}$  or vice-versa. But the pairs  $(U_i, V_i)$  do not contain any  $\#$ , so the pair  $(Y, Z)$  is used exactly once, both after  $m$ th pair and the  $m'$ th pair, so  $m = m'$  and the PCP instance has a solution.  $\square$

### 2.3 Undecidable properties of multi-tape automata

Let  $\mathcal{M}$  be a  $d$ -tape automaton on alphabet  $\mathcal{A}$ , and let  $q$  be a state of  $\mathcal{M}$ . State  $q$  is **universal** if every sequence in  $\mathcal{A}^{\mathbb{N}}$  is  $q$ -accepted by  $\mathcal{M}$ . A finite sequence  $x \in \mathcal{A}^*$  is a **universal prefix** for state  $q$  if for every infinite sequence  $y \in \mathcal{A}^{\mathbb{N}}$ , the infinite sequence  $xy$  is  $q$ -accepted by  $\mathcal{M}$ .

**Example 2.4.** Let  $\mathcal{M}$  be a 1-tape, 1-state automaton on alphabet  $\{0, 1\}$  with three transitions labelled by 1, 10 and 00. The single state of  $\mathcal{M}$  is not universal because every sequence starting with 01 is rejected, but the word 1 is a universal prefix: any sequence starting with 1 is accepted, because any finite segment  $10^n 1$  is accepted by transitions 1,  $00 \times k$ , 1 if  $n = 2k$  or 10,  $00 \times k$ , 1 if  $n = 2k + 1$ , and any infinite tail of 0's or 1's is obviously accepted. Hence there exist some multi-tape automata without universal states but that admit universal prefixes. The self-affine set associated with this automaton is discussed in Example 3.5.

**Theorem 2.5.** *It is undecidable whether a given state of a given  $d$ -tape automaton is universal. This problem remains undecidable if we restrict to 2-tape automata with 3 states.*

*Proof.* We reduce prefix-PCP, which is undecidable thanks to Lemma 2.3. Let  $(u_1, v_1), \dots, (u_n, v_n)$  be an instance of prefix-PCP where the  $u_i, v_i$  are words over  $\mathcal{B}$ . We define a 2-tape automaton  $\mathcal{M}$  on 3 states (denoted by  $X, U, V$ ). The alphabet of  $\mathcal{M}$  is  $A_1 \times A_2$ , with  $A_1 = \{1, \dots, n\}$  and  $A_2 = \mathcal{B} \cup \{\#\}$ , where  $n$  is the size of the prefix-PCP instance,  $\mathcal{B}$  is the alphabet of words  $u_i, v_i$  and  $\#$  is a new symbol not in  $\mathcal{B}$ . The transitions of  $\mathcal{M}$  are

- (1)  $X \xrightarrow{i|u_i} U$  and  $U \xrightarrow{i|u_i} U$  for every  $i \in A_1$ ;
- (2)  $X \xrightarrow{i|v_i} V$  and  $V \xrightarrow{i|v_i} V$  for every  $i \in A_1$ ;
- (3)  $U \xrightarrow{i|u} X$  for every  $i \in A_1$  and  $u \in A_2^+$  such that
  - (i)  $|u| \leq |u_i|$ ,
  - (ii)  $u$  is not a prefix of  $u_i$ ,
  - (iii)  $u$  does not begin with  $\#$ ;
- (4)  $X \xrightarrow{i|u} X$  for every  $i \in A_1$  and  $u \in A_2^+$  such that (i) and (ii) above hold;
- (5)  $V \xrightarrow{i|v} X$  for every  $i \in A_1$  and  $v \in A_2^+$  such that
  - (i)  $|v| \leq |v_i|$ ,
  - (ii)  $v$  is not a prefix of  $v_i$ ,
  - (iii)  $v$  does not begin with  $\#$ ;
- (6)  $X \xrightarrow{i|v} X$  for every  $i \in A_1$  and  $v \in A_2^+$  such that (i) and (ii) above hold.

We now prove that the prefix-PCP instance  $(u_1, v_1), \dots, (u_n, v_n)$  has a solution if and only if state  $X$  is **not** universal in  $\mathcal{M}$ .

$(\Rightarrow)$  Suppose that the prefix-PCP instance admits a solution: there exist  $m, m' \geq 1$  and two words  $i_1 \cdots i_m$  and  $i_1 \cdots i_{m'}$  such that  $u_{i_1} \cdots u_{i_m} = v_{i_1} \cdots v_{i_{m'}}$  and one of the two words  $i_1 \cdots i_m$  and  $i_1 \cdots i_{m'}$  is a prefix of the other. Without loss of generality we can assume that  $m \geq m'$

and  $i_1 \cdots i_{m'}$  is a prefix of  $i_1 \cdots i_m$ . We prove that  $\mathcal{M}$  cannot accept any infinite sequence in  $(A_1 \times A_2)^\mathbb{N}$  beginning with

$$i_1 \cdots i_m \mid u_{i_1} \cdots u_{i_m} \#$$

when starting from state  $X$ , so  $\mathcal{M}$  is not universal. Indeed, let us describe the evolution of  $\mathcal{M}$  when reading such a sequence.

- We start from  $X$ , so  $\mathcal{M}$  necessarily uses a transition defined in (1) and (2) and moves to state  $U$  or  $V$  after having read  $i_1|u_{i_1}$  or  $i_1|v_{i_1}$ , respectively. (The other transitions (4) and (6) cannot be used because of the conditions (i) and (ii).) Note that both  $u_{i_1}$  and  $v_{i_1}$  are prefixes of the content of the second tape.
- Now if  $\mathcal{M}$  is in state  $U$ , the remaining input starts with some  $i$  on the first tape and starts with  $u_i$  on the second tape. So  $\mathcal{M}$  must use transition (1): stay in state  $U$  and read  $i|u_i$ . (Transition (3) cannot be used because of the conditions (i) and (ii).) The same holds if  $\mathcal{M}$  is in state  $V$ .

It follows that when  $\mathcal{M}$  reads  $i_1, \dots, i_{m'}$  on the first tape, then it is either in state  $U$  and has read  $u_{i_1} \cdots u_{i_{m'}}$  on the second tape, or it is in state  $V$  and has read  $v_{i_1} \cdots v_{i_{m'}} = u_{i_1} \cdots u_{i_m}$  on the second tape. In the second case, the next symbol on the second tape is  $\#$ , so  $\mathcal{M}$  is “blocked” on this input (there is no suitable transition for this sequence because of (iii)). In the first case, the computation must continue in the same way as before, so eventually  $\mathcal{M}$  is still in state  $U$  and has read  $i_1 \cdots i_m | u_{i_1} \cdots u_{i_m}$ , and again,  $\mathcal{M}$  is blocked because the next symbol on the second tape is  $\#$ .

( $\Leftarrow$ ) Suppose that no solution exists for the prefix-PCP instance. The following strategy shows that a move by the automaton can always be made, whatever its tape contents is. If  $\mathcal{M}$  is in state  $U$  or  $V$ , make any available move. In state  $X$ , if no loop in  $X$  is possible, then in the current configuration  $(i_1 i_2 \cdots | w)$ , both  $u_{i_1}$  and  $v_{i_1}$  must be prefixes of  $w$ , otherwise (4) or (6) could have been used. Write  $w = u_{i_1} w' = v_{i_1} w''$ . Then:

- (a) if  $u_{i_1} \cdots u_{i_k} \#$  is a prefix of  $w$  for some  $k$ , do not go to  $U$  by reading  $i_1|u_{i_1}$ ;
- (b) if  $v_{i_1} \cdots v_{i_k} \#$  is a prefix of  $w$  for some  $k$ , do not go to  $V$  by reading  $i_1|v_{i_1}$ ;

The only possible ways to be stuck at this point are:

- $\mathcal{M}$  is in state  $U$  or  $V$  and the next symbol on the second tape is  $\#$ ;
- $\mathcal{M}$  is in state  $X$  and (a), (b) prevent from moving to  $U$  or  $V$ .

The second case cannot happen because it implies the existence of a prefix-PCP solution. If we are in the first case, we can assume by symmetry that we are in state  $U$ . In the last step where  $\mathcal{M}$  went from  $X$  to  $U$ , the configuration must start with  $i_1 i_2 \cdots | u_{i_1} u_{i_2} \cdots u_{i_k} \# \cdots$  for some  $k$ , because this is the only way to get stuck in  $U$  some  $k$  steps later. However, this contradicts the choice made in (a) above, because  $\mathcal{M}$  should have moved to  $V$  instead of state  $U$ .  $\square$

**Theorem 2.6.** *It is undecidable whether a given state of a given  $d$ -tape automaton admits a universal prefix. This problem remains undecidable if we restrict to 2-tape automata with 3 states.*

*Proof.* We modify the prefix-PCP reduction made in the proof of Theorem 2.5. Let  $(u_1, v_1), \dots, (u_n, v_n)$  be an instance of prefix-PCP where the  $u_i, v_i$  are words over  $\mathcal{B}^*$ . First we modify the  $u_i, v_i$  by adding a new symbol  $*$  not in  $\mathcal{B}$  after each letter of each  $u_i$  and each  $v_i$  (a word  $x_1 x_2 \cdots x_k$  becomes  $x_1 * x_2 * \cdots * x_k *$ ). This modified instance is clearly equivalent to the original one, so we denote it again by  $(u_1, v_1), \dots, (u_n, v_n)$ .

We now define a 2-tape automaton  $\mathcal{M}$  on 3 states  $X, U, V$ . We take the same alphabet  $A_1 \times A_2$  as in the other reduction, with a new symbol  $\&$  for both  $A_1$  and  $A_2$ , and the symbol  $*$  for  $A_2$ . This gives  $A_1 = \{1, \dots, n\} \cup \{\&\}$  and  $A_2 = \mathcal{B} \cup \{\#, \&, *\}$ , where  $n$  is the size of the prefix-PCP instance,  $\mathcal{B}$  is the alphabet of the words  $u_i, v_i$  and  $\#, \&, *$  are new symbol not in  $\mathcal{B}$ . The transitions of  $\mathcal{M}$  consist of

- (1) and (2) like in the proof of Theorem 2.5, without allowing any symbol  $\&$  or  $*$ ;
- (3), (4), (5), (6) like in the proof of Theorem 2.5, where symbols  $\&$  or  $*$  are allowed, except in the first letter of  $u$  or  $v$ ;

plus the following transitions:

- (7)  $X \xrightarrow{a|\&} X$ ,  $U \xrightarrow{a|\&} X$  and  $V \xrightarrow{a|\&} X$  for every  $a \in A_1$ ;
- (8)  $X \xrightarrow{\&|a} X$ ,  $U \xrightarrow{\&|a} X$  and  $V \xrightarrow{\&|a} X$  for every  $a \in A_2 \setminus \{*\}$ ;
- (9)  $X \xrightarrow{a|*b} X$ ,  $U \xrightarrow{a|*b} X$  and  $V \xrightarrow{a|*b} X$  for every  $a \in A_1$  and  $b \in A_2$ .

We now prove that the prefix-PCP instance  $(u_1, v_1), \dots, (u_n, v_n)$  has a solution if and only if state  $X$  does **not** have any universal prefix.

( $\Rightarrow$ ) Suppose that the prefix-PCP instance has a solution: there exist  $m, m' \geq 1$  and two words  $i_1 \dots i_m$  and  $i_1 \dots i_{m'}$  such that  $u_{i_1} \dots u_{i_m} = v_{i_1} \dots v_{i_{m'}}$  and one of the two words  $i_1 \dots i_m$  and  $i_1 \dots i_{m'}$  is a prefix of the other. Consider the following claim.

**Claim.** Let  $x \in A_1^*$  and  $y \in A_2^*$  be such that  $x\&\&\dots|y\&\&\dots$  is  $X$ -accepted by at most  $k \geq 1$  different runs of  $\mathcal{M}$ . Then there exist  $x' \in A_1^*$  and  $y' \in A_2^*$  such that  $xx'\&\&\dots|yy'\&\&\dots$  is  $X$ -accepted by at most  $k - 1$  different runs.

This claim implies that  $X$  does not have any universal prefix, *i.e.*, that for every finite words  $x \in A_1^*$  and  $y \in A_2^*$ , there exists a configuration starting with  $x|y$  that is not  $X$ -accepted. Indeed, for every such  $x, y$ , there can be only finitely many different accepting runs (say  $k$ ) for  $x\&\&\dots|y\&\&\dots$ , because  $\mathcal{M}$  eventually loops on state  $X$  with transition  $\&|\&$ . So it suffices to apply the claim  $k$  times to obtain a configuration starting with  $x|y$  which is not  $X$ -accepted.

We now prove the claim, using the prefix-PCP solution. Let  $x \in A_1^*$  and  $y \in A_2^*$  be such that  $x\&\&\dots|y\&\&\dots$  is  $X$ -accepted by  $k$  different runs. Denote by  $R_1, \dots, R_k$  the finite prefixes of the  $k$  runs, each cut when  $\mathcal{M}$  reaches the  $\&\&\dots|\&\&\dots$  part. Let  $s = i_1 \dots i_m \in A_1^*$  and let  $t = u_{i_1} \dots u_{i_m}$ , which can be written in the form  $t = a_1 * a_2 * \dots * a_{|t|-1} * \in A_2^*$ , where each  $a_i$  is in  $A_2 \setminus \{\#, \&, *\}$ , thanks to the modification made to the instance.

Let  $\ell$  be the distance between the two tapes heads when  $\mathcal{M}$  has completed the finite run  $R_1$ . (Note that the first head is always behind the second one because it can only move by one cell at a time.) Without loss of generality we can assume that  $R_1$  is the run for which such an  $\ell$  is minimal. We now construct a configuration  $c$  which will “block” any run starting with  $R_1$ , without giving the other runs any possibilities for new nondeterministic branching. Let  $L, L' \geq 0$  such that  $s$  (on the first tape) begins  $\ell$  positions behind  $t$  (on the second tape) in the configuration

$$c = x\&^L s\&\&\dots | y\&^{L'} t\#\&\&\dots,$$

so that during any run starting with  $R_1$ ,  $\mathcal{M}$  starts reading  $s$  and  $t\#$  exactly at the same time. It follows that  $R_1$  cannot be extended to an accepting run for  $c$ , because  $s, t$  corresponds to a prefix-PCP solution, similarly as in the proof of Theorem 2.5. The same is true for any other run  $R_i$  for which such an  $\ell$  is the same as  $R_1$ .

Let us now consider another accepting run  $R_i$ . By minimality of  $\ell$ , the distance between the two tapes heads when  $\mathcal{M}$  first reaches  $\&\&\dots|\&\&\dots$  during run  $R_i$  is strictly larger than  $\ell$ . We now prove that  $R_i$  can be extended in a unique way to an accepting run for  $c$ . Indeed, any run of  $\mathcal{M}$  starting with  $R_i$  must evolve in the following way:

- when  $t$  starts being read the second tape,  $s$  is not yet being read on the first tape, so at this time  $\mathcal{M}$  is reading  $\&$  on the first tape and  $a_1$  on the second tape;
- the only possible transition is (8), so  $\mathcal{M}$  moves one step on both tapes, and is now reading  $*$  on the second tape;
- the only possible transition is (9), so  $\mathcal{M}$  moves one step on the first tape and two steps on the second, and is again reading  $*$  on the second tape;

- this continues until the whole  $t = a_1 * a_2 * \dots * a_{|t|-1} *$  has been read on the second tape, and  $\mathcal{M}$  is deterministically looping on  $\&|\&$ .

From this analysis, it follows that  $R_i$  can be extended in a *unique* way to an accepting run for  $c$ . Hence  $c$  is a configuration starting with  $x|y$  with at most  $k - 1$  accepting runs, because every accepting run for  $c$  must start with an  $R_i$ , each of which can be extended in at most one way if  $i \in \{2, \dots, k\}$ , or in no way at all if  $i = 1$ . Thus the claim is proved by taking  $x' = \&^L s$  and  $y' = \&^{L'} t \#$ .

( $\Leftarrow$ ) Suppose that no solution exists for the prefix-PCP instance. The strategy described in the “ $\Leftarrow$ ” direction of the proof of Theorem 2.5 can be applied to prove that state  $X$  is universal, with the additional case that if the tape begins by  $\&$  or  $*$ , then the transition (7), (8) or (9) can always be used.  $\square$

**Remark 2.7.** In the reduction made in the above proof of Theorem 2.6, if state  $X$  has a universal prefix, then in fact  $X$  is universal. Also, in this case, it is easy to see that any finite word satisfying (i), (ii) and (iii) of transition (3) is a universal prefix for  $U$  (and  $V$ ), so  $X$ ,  $U$  (and  $V$ ) have a common universal prefix. Hence we have the following: given a 2-tape automaton  $\mathcal{M}$  on 3 states and two states  $q, r$  of  $\mathcal{M}$ , it is undecidable if  $q$  and  $r$  have a common universal prefix.

### 3 Affine GIFS associated with multi-tape automata

Let  $\mathcal{M}$  be a  $d$ -tape automaton on alphabet  $\mathcal{A} = A_1 \times \dots \times A_d$ . We want to give a “numerical interpretation” to a finite word  $u \in \mathcal{A}^*$  or to an infinite configuration  $c \in \mathcal{A}^{\mathbb{N}}$ . We must first specify, for each  $k \in \{1, \dots, n\}$ , a numerical interpretation of the letters of  $A_k$  by choosing a bijection  $\delta_k : A_k \rightarrow \{0, \dots, |A_k| - 1\}$ . We then define  $\Delta_k : \mathcal{A}_k^* \rightarrow \mathbb{R}$  by

$$\Delta_k(u) = \sum_{1 \leq i \leq |u|} \delta_k(u_i) |A_k|^{-i}.$$

Equivalently, for  $u = u_1 \dots u_n \in A_k^n$ , the number  $\Delta_k(u)$  is represented by  $0.\delta_k(u_1) \dots \delta_k(u_n)$  in base  $|A_k|$ . Finally, let  $\Delta : A_1^+ \times \dots \times A_d^+ \rightarrow \mathbb{R}^d$  be defined by  $\Delta(w_1, \dots, w_d) = (\Delta_1(w_1), \dots, \Delta_d(w_d))$ . The domains of  $\Delta_k$  and  $\Delta$  can naturally be extended to  $A_k^{\mathbb{N}}$  and  $\mathcal{A}^{\mathbb{N}}$ , respectively.

In the examples that will follow, if the alphabets  $A_k$  are all of the form  $\{0, \dots, |A_k| - 1\}$  and the maps  $\delta_k : A_k \rightarrow \{0, \dots, |A_k| - 1\}$  are not specified, we will assume for convenience that they are identity mappings.

**Definition 3.1.** Let  $\mathcal{M}$  be a  $d$ -tape automaton on state set  $\mathcal{Q}$  and alphabet  $\mathcal{A} = A_1 \times \dots \times A_n$ . The **GIFS associated with  $\mathcal{M}$**  is the GIFS defined by the graph  $G$  with vertex set  $\mathcal{Q}$  and, for every transition  $R = (q, r, (w_1, \dots, w_d))$  of  $\mathcal{M}$ , an edge  $(q, r)$  labelled by the map  $f_R : [0, 1]^d \rightarrow [0, 1]^d$  defined by

$$f_R(x) = \begin{pmatrix} |A_1|^{-|w_1|} & & 0 \\ & \ddots & \\ 0 & & |A_d|^{-|w_d|} \end{pmatrix} x + \Delta(w_1, \dots, w_d).$$

**Example 3.2.** Let  $\mathcal{M}$  be a 2-tape automaton on alphabet  $\mathcal{A} = \{0, 1\} \times \{0, 1\}$ , and let  $c \in \mathcal{A}^{\mathbb{N}}$  be configuration. If  $\mathcal{M}$  contains a transition  $R = (q, r, (1011, 11))$ , then applying the contracting map  $f_R$  on  $\Delta(c) = (0.x_1x_2\dots, 0.y_1y_2\dots) \in [0, 1]^2$  has the following effect:

$$\begin{aligned} f_R(\Delta(c)) &= \begin{pmatrix} 1/16 & 0 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} 0.x_1x_2\dots \\ 0.y_1y_2\dots \end{pmatrix} + \Delta(1011, 11) \\ &= \begin{pmatrix} 0.0000x_1x_2\dots \\ 0.00y_1y_2\dots \end{pmatrix} + \begin{pmatrix} 0.1011 \\ 0.11 \end{pmatrix} = \begin{pmatrix} 0.1011x_1x_2\dots \\ 0.11y_1y_2\dots \end{pmatrix}. \end{aligned}$$

This suggests that applying a sequence of mappings  $f_{R_1} \cdots f_{R_n}(\Delta(c))$  corresponds to concatenating the words associated with the transitions  $R_n$  in the numerical interpretation  $\Delta(c)$  of a configuration  $c$ . This is the key idea to establish a correspondence between the GIFS of an automaton and its accepted configurations. This is formalized in the next proposition.

**Proposition 3.3.** *Let  $\mathcal{M}$  be a 2-tape automaton and let  $q$  be a state of  $\mathcal{M}$ . The GIFS attractor of  $\mathcal{M}$  associated with  $q$  is equal to the set  $\{\Delta(c) \in \mathbb{R}^d : c \in \mathcal{A}^{\mathbb{N}} \text{ is } q\text{-accepted by } \mathcal{M}\}$ .*

*Proof.* Let  $x \in [0, 1]^d$ . It follows from a standard fact in the theory of iterated function systems [Fal03, Chapter 9] that  $x \in X_q$  if and only if there is an infinite run  $(R_n)_{n \geq 1}$  starting at  $q$  such that  $x = \bigcap_{n \geq 1} f_{R_1} \cdots f_{R_n}([0, 1]^d)$ , where  $f_{R_n}$  is the mapping of the GIFS of  $\mathcal{M}$  associated with run  $R_n$ . Moreover, by definition of the GIFS of  $\mathcal{M}$ , for every such run  $(R_n)_{n \geq 1}$ , the configuration  $c = w_{1,1}w_{2,1} \cdots \mid \cdots \mid w_{1,d}w_{2,d} \cdots$  is such that  $x = \Delta(c)$ , where the  $w_{n,k}$  are given by the transitions  $(q_n, r_n, (w_{n,1}, \dots, w_{n,d}))$  for all  $n \geq 1$ , so the proposition is proved because  $c$  is a  $q$ -accepted configuration.  $\square$

**Example 3.4.** Let  $\mathcal{M}$  be the 1-state, 2-tape automaton on alphabet  $\{0, 1\}$  with transitions  $0|0, 0|1, 1|0$ . The iterated function system associated with  $\mathcal{M}$  consists of the maps  $x \mapsto \frac{x}{2}$ ,  $x \mapsto \frac{x}{2} + (\frac{1}{2}, 0)$ ,  $x \mapsto \frac{x}{2} + (0, \frac{1}{2})$  and it can easily be seen that the associated attractor is the Sierpiński triangle (see Figure 1).

**Example 3.5.** The 1-tape, 1-state automaton  $\mathcal{M}$  on alphabet  $\{0, 1\}$  with three transitions 1, 10 and 00 (given in Example 2.4) is an example of a non-universal automaton which admits universal prefixes. This reflects in the attractor associated with  $\mathcal{M}$  in the following way: it is not equal to  $[0, 1]$  but it has nonempty interior. This can be proved either by Proposition 3.7, or by proving directly that a configuration  $x \in \{0, 1\}^{\mathbb{N}}$  is accepted by  $\mathcal{M}$  if and only if it does not start with  $0^{2k+1}1$  for some  $k \geq 0$ , which implies that the attractor is equal to  $\bigcup_{k \geq 0} [2^{-2k-1}, 2^{-2k}]$ .

**Remark 3.6.** Given a  $d$ -tape automaton and a point  $x \in [0, 1]^d$ , if there exists two configurations  $c, c'$  such that  $x = \Delta(c) = \Delta(c')$  and such that the tapes components  $c_k \in A_k^{\mathbb{N}}$  and  $c'_k \in A_k^{\mathbb{N}}$  differ for some  $k \in \{1, \dots, d\}$ , then  $c_k$  and  $c'_k$  are both stationary, ending with  $0^\omega$  or  $(|A_k| - 1)^\omega$ . In particular,  $\Delta : \mathcal{A}^{\mathbb{N}} \rightarrow \mathbb{R}^d$  is finite-to-one.

The next proposition establishes the desired correspondence between word-theoretical properties of multi-tape automata and topological properties of the associated self-affine attractors.

**Proposition 3.7.** *Let  $\mathcal{M}$  be a  $d$ -tape automaton on alphabet  $\mathcal{A}$ , let  $q$  be a state of  $\mathcal{M}$ , and let  $X_q$  be the associated GIFS attractor. We have:*

- (1)  $q$  is universal if and only if  $X_q = [0, 1]^d$ ,
- (2)  $q$  has a universal prefix if and only if  $X_q$  has nonempty interior.

*Proof.* (1) If state  $q$  is universal the expansion of every element of  $[0, 1]^d$  is  $q$ -accepted so  $X_q = [0, 1]^d$  thanks to Proposition 3.3. Conversely, suppose that there exists an infinite sequence  $c$  that is not  $q$ -accepted. By a compactness argument, there must exist a prefix  $w$  of  $c$  such that  $wc'$  is not  $q$ -accepted for any infinite sequence  $c'$ . Thanks to Remark 3.6, by choosing  $c'$  with no tape components ending by  $0^\omega$  or  $(|A_k| - 1)^\omega$ , the sequence  $wc'$  is the *only* sequence such that  $x = \Delta(wc')$ , so  $\Delta(wc') \notin X_q$  because otherwise  $wc'$  would be  $q$ -accepted. It follows that  $X_q \neq [0, 1]^d$ .

(2) For a finite word  $w \in \mathcal{A}^*$ , define the *cylinder*  $[w]$  to be equal to the set of configurations that start with  $w$ . If  $q$  admits a universal prefix  $w$ , then  $\Delta([w]) \subseteq X_q$  by Proposition 3.3, so  $X_q$  has nonempty interior. Conversely, suppose that there exists a nonempty open set  $U \subseteq X_q$ , and let  $w \in \mathcal{A}^*$  be a finite word such that  $\Delta([w]) \subseteq U$ . By a reasoning similar as in the proof of (1), we can prove that  $w$  is a universal prefix for  $q$ .  $\square$

## 4 Undecidability results for self-affine sets

Thanks to the undecidability results obtained for multi-tape automata in Theorem 2.5 and to the correspondence between word-theoretical and topological properties in Proposition 3.7, we obtain the following undecidability results about topological properties of self-affine attractors.

The first result below states that it is undecidable if an attractor “takes up the whole space”, that is, equals  $[0, 1]^d$ . It follows directly from Theorem 2.5 and Proposition 3.7, (1).

**Theorem 4.1.** *The following problem is undecidable. Instance: a  $d$ -dimensional affine GIFS  $\mathcal{G}$  specified by maps with rational coefficients, and a state  $q$  of  $\mathcal{G}$ . Question: is  $X_q = [0, 1]^d$ ? This problem remains undecidable if we restrict to 2-dimensional GIFS with 3 states.*

The next result states the undecidability of a fundamental topological property for self-affine sets: having empty interior. It is a direct corollary of Theorem 2.6 and Proposition 3.7, (2).

**Theorem 4.2.** *The following problem is undecidable. Instance: a  $d$ -dimensional affine GIFS  $\mathcal{G}$  specified by maps with rational coefficients, and a state  $q$  of  $\mathcal{G}$ . Question: does  $X_q$  have empty interior? This problem remains undecidable if we restrict to 2-dimensional GIFS with 3 states.*

**Remark 4.3.** All the undecidability results above have been obtained via a reduction using affine GIFS associated with a multi-tape automaton. Hence it follows that undecidability holds even if we restrict to affine GIFS in which the linear part of the contractions  $f_i$  are diagonal matrices whose entries are negative powers of integers. By adding dummy duplicate symbols, undecidability holds even if the entries are negative powers of two.

**Remark 4.4.** We can deduce from Remark 2.7 that the following problem is undecidable. Instance: a  $d$ -dimensional affine GIFS  $\mathcal{G}$  specified by maps with rational coefficients, and two states  $q, r$  of  $\mathcal{G}$ . Question: does  $X_q \cap X_r$  have empty interior? Indeed, it can be shown that  $q$  and  $r$  have a common universal prefix if and only if  $X_q \cap X_r$  has nonempty interior, similarly as in Proposition 3.7.

## 5 Conclusion

We conclude this article by some questions and perspectives for further work. Is nonempty interior decidable for 1-state GIFS? (That is, for classical affine IFS.) What about the 1-dimensional case? Using multi-tape automata may lead to an undecidability result for the 1-state case, but not for the 1-dimensional case. Indeed, 1-tape automata are not more powerful than classical finite automata, for which the properties we used in this article are all decidable. Note that for 1-state multi-tape automata, universality is trivially decidable, but the status of prefix-universality is not known in this case.

Also, let us note that having nonempty interior is equivalent to having nonzero Lebesgue measure in the case of integer self-affine tiles (as mentioned in the introduction), but not in the more general setting of self-affine (G)IFS (see for example [CJP<sup>+</sup>06]). How do these properties relate in the case of self-affine sets arising from multi-tape automata?

Another interesting aspect is the computability of fractal dimension (such as Hausdorff dimension). For example, can we decide if the Hausdorff dimension of a 2-dimensional self-affine set is equal to 2? And in the case of a self-affine set with nonempty interior, can we compute the Hausdorff dimension of its boundary? Very few results are known in this direction, apart from some very specific families such as Bedford-McMullen carpets [Bed84, McM84, Fra13]. A possible approach towards undecidability would be to adapt the reductions of this article in such a way that the Hausdorff dimension can be controlled in the reductions, or to relate the entropy of the automaton language with the Hausdorff dimension of its attractor and prove that entropy is uncomputable.

## References

- [**Bar93**] Michael F. Barnsley, *Fractals everywhere*, second ed., Academic Press Professional, Boston, MA, 1993.
- [**Bed84**] Tim Bedford, *Crinkly curves, Markov partitions and box dimensions in self-similar sets*, Ph.D. thesis, University of Warwick, 1984.
- [**BK11**] Ievgen V. Bondarenko and Rostyslav V. Kravchenko, *On Lebesgue measure of integral self-affine sets*, *Discrete Comput. Geom.* **46** (2011), no. 2, 389–393.
- [**CJP<sup>+</sup>06**] Marianna Csörnyei, Thomas Jordan, Mark Pollicott, David Preiss, and Boris Solomyak, *Positive-measure self-similar sets without interior*, *Ergodic Theory Dynam. Systems* **26** (2006), no. 3, 755–758.
- [**Dub93**] Simant Dube, *Undecidable problems in fractal geometry*, *Complex Systems* **7** (1993), no. 6, 423–444.
- [**Fal97**] Kenneth Falconer, *Techniques in fractal geometry*, John Wiley & Sons Ltd., Chichester, 1997.
- [**Fal03**] ———, *Fractal geometry*, second ed., John Wiley & Sons Inc., Hoboken, NJ, 2003, Mathematical foundations and applications.
- [**Fra13**] Jonathan M. Fraser, *Dimension theory and fractal constructions based on self-affine carpets*, Ph.D. thesis, The University of St Andrews, 2013.
- [**GY06**] Jean-Pierre Gabardo and Xiaojiang Yu, *Natural tiling, lattice tiling and Lebesgue measure of integral self-affine tiles*, *J. London Math. Soc. (2)* **74** (2006), no. 1, 184–204.
- [**Hut81**] John E. Hutchinson, *Fractals and self-similarity*, *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–747.
- [**LW96**] Jeffrey C. Lagarias and Yang Wang, *Self-affine tiles in  $\mathbf{R}^n$* , *Adv. Math.* **121** (1996), no. 1, 21–49.
- [**LLR13**] Chun-Kit Lai, Ka-Sing Lau, and Hui Rao, *Spectral structure of digit sets of self-similar tiles on  $\mathbb{R}^1$* , *Trans. Amer. Math. Soc.* **365** (2013), no. 7, 3831–3850.
- [**McM84**] Curt McMullen, *The Hausdorff dimension of general Sierpiński carpets*, *Nagoya Math. J.* **96** (1984), 1–9.
- [**Pos46**] Emil L. Post, *A variant of a recursively unsolvable problem*, *Bull. Amer. Math. Soc.* **52** (1946), 264–268.
- [**Wan99**] Yang Wang, *Self-affine tiles*, *Advances in wavelets (Hong Kong, 1997)*, Springer, Singapore, 1999, pp. 261–282.