

Pattern Complexity of Aperiodic Substitutive Subshifts

Etienne Moutot¹ and Coline Petit-Jean²

¹etienne.moutot@lis-lab.fr, Aix-Marseille Université, CNRS, LIS, France

²coline.petit-jean@ens-lyon.fr, ENS de Lyon, France

Abstract

This paper aims to better understand the link between aperiodicity in subshifts and pattern complexity. Our main contribution deals with substitutive subshifts, an equivalent to substitutive tilings in the context of symbolic dynamics. For a class of substitutive subshifts, we prove a quadratic lower bound on their pattern complexity. Together with an already known upper bound, this shows that this class of substitutive subshifts has a pattern complexity in $\Theta(n^2)$. We also prove that the recent bound of Kari and Moutot, showing that any aperiodic subshift has pattern complexity at least $mn + 1$, is optimal for fixed m and n .

1 Introduction

One of the most fascinating aspects of tilings of the plane is the existence of *aperiodic tilesets*. That is, a set of tiles that tessellates the plane but only in a non-periodic manner. Aperiodic tilesets even exist in the specific setting of Wang tilings, where tiles are square with colored edges and cannot be rotated nor reflected. Two tiles can then be placed next to each other if the colors of the matching edges are the same. Interestingly, aperiodic tilesets do not exist for a similar model in dimension one. If one considers tilings of the infinite line \mathbb{Z} by bi-color Wang dominoes, then any set of dominoes tiling the line also tiles it periodically. This has a consequence on the decidability of the most fundamental problem about tilings: the domino problem. This problem just asks if a given tileset can or cannot tile the plane. Considering this problem on a line or a plan changes everything: the domino problem for bi-infinite line is decidable in polynomial time, as it is equivalent to the existence of a cycle in a finite graph; whereas it is undecidable if one considers tilings of the plane. From now on, “tiling” will always mean “Wang tiling”, except stated otherwise.

The first aperiodic tileset is due to Berger in 1966 [3], who disproved Wang’s conjecture stating that aperiodic tilesets of the plane could not exist. His tileset was expressed as a set of Wang tiles, a model introduced by Wang to study fragments of first order logic [26]. Each tile is a unit square with a color on each edge. Two tiles can then be placed next to each other if the colors of the matching edges are the same. Then from a finite set of tiles, one tries to tile the plane by copies of these tiles without rotations. Berger set of tiles was initially made of 20426 tiles, reduced to 104 in his PhD thesis [2]. This started the search for the smallest aperiodic set of Wang tiles. Among many, one can cite Robinson’s tileset [24] which has only 56 tiles. With a radically different idea, Kari built a 14 tiles aperiodic tileset [14]. His construction was quickly improved by Culik

to 13 tiles [6], which held the record for many years. Finally, in 2015 Jeandel and Rao found an 11 tiles aperiodic tileset, and proved that it was the smallest possible [13].

If it is interesting to have aperiodic tilesets with few tiles, the number of tiles is not a good estimation of how “complex” the resulting tilings can be. Indeed, with the same number of tiles one can end up with very simple tilings as well as extremely complex (encoding for example the space-time diagram of a Turing machine). A better way of estimating how complex a tiling is by its *pattern complexity*. For a tileset T , we denote by $\mathcal{L}_{m,n}(T)$ the set of all $m \times n$ rectangular patterns appearing in at least one valid tiling by T . The pattern complexity of T is then the cardinal of this set. Besides to carrying more information about how “complicated” a tiling is, this definition of complexity can be seen as a generalization of the number of tiles, as the latter is simply the number of patterns of size 1×1 . Computing the exact complexity of a tileset is often complicated, and a more studied quantity is the topological entropy of a tileset. However, topological entropy is a very rough estimation of the complexity, and provide only an asymptotic estimation of the growth of the complexity function. In this paper, we are interested in bounding the exact complexity of aperiodic tilesets. More precisely, we want to understand what is the minimal complexity that can be achieved by an aperiodic tileset; or equivalently, how “simple” can an aperiodic tileset be. In [15], it is shown that any aperiodic tileset have complexity at least $mn + 1$. The first result of the current paper shows that this bound is optimal for fixed m and n .

One of the most common techniques to build an aperiodic tileset is to use two dimensional substitutions. In most of the cases, the resulting tilings are close to being fixpoints of well-chosen two-dimensional substitutions, leading to their aperiodicity. This is the case of Berger’s tileset, Robinson’s, and even Jeandel and Rao’s tileset was recently found to have a substitutive structure [16]. Their substitutive nature also ensures that they have zero topological entropy. Studying the complexity of substitutive tilesets is therefore a way of understanding the complexity of many aperiodic tilesets. One of the rare exceptions is Kari’s aperiodic tileset, which is known not to be substitutive, as it was shown to have positive topological entropy [7]. This also shows that it has an exponential pattern complexity.

More generally, *substitutive subshifts* have been extensively studied in dimension one [8, 10]. A substitutive subshift is a set of colorings of the infinite line \mathbb{Z} generated by infinite iterations of a one-dimensional substitution. In dimension one, Pansiot fully classified substitutions in terms of factor complexity [21].

In the two-dimensional case, substitutions have been mostly studied in the context of geometrical tilings [1, 9, 12, 22, 25]. In this paper, we focus on discrete substitutions over two-dimensional words. Our second main result is a quadratic lower bound on the pattern complexity of a large class of two dimensional substitutive subshifts. Together with Robinson’s upper bound on the complexity of two-dimensional substitutions [23, Theorem 7.16], this shows that the pattern complexity of subshifts from this class is in $\Theta(n^2)$. Such a bound paves the way towards a classification of two-dimensional substitutions in terms of pattern complexity, as Pansiot did for one-dimensional substitutions.

After introducing definitions and basic properties in Section 2, we show in Section 3 that the $mn + 1$ complexity bound is optimal for fixed m, n . Then we prove our lower bound on our class of substitutive subshifts in Section 4. Finally we conclude by some remarks on possible directions for future work in Section 5.

2 Preliminaries

In this section, we introduce all the useful definitions and properties.

We will use the following conventions for our notations. Capital A designates a finite alphabet and a and b elements of it. Capital X is a subshift, whose configurations are usually denoted by the letter c , and patterns by p . Vectors of \mathbb{Z}^2 are bold $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^2$. σ is a substitution, and i, j, k, l, m, n designate integers.

2.1 Subshifts

Tilings have an equivalent definition as subshifts of finite type. Let A be a finite alphabet. A *pattern* of support $D \subset \mathbb{Z}^2$ is a coloring of D with colors of A , i.e. $p \in A^D$, and a *configuration* $c \in A^{\mathbb{Z}^2}$ is a coloring of \mathbb{Z}^2 . We say that a pattern p of support D *appears* in a pattern or a configuration p' if there exists $\mathbf{u} \in \mathbb{Z}^2$ such that for all $\mathbf{v} \in D$, $p_{\mathbf{v}} = p'_{\mathbf{v}-\mathbf{u}}$, in this case we write $p \sqsubseteq p'$ (and $p \sqsubset p'$ when $p \neq p'$). Let F be a set of patterns, a *subshift* X_F is a set of configurations in which none of the patterns of F appear:

$$X_F := \{c \in A^{\mathbb{Z}^2} \mid \forall p \in F, p \not\sqsubseteq c\}.$$

If F is finite, X_F is called a *subshift of finite type* (SFT for short). Subshifts have an equivalent definition as subsets of $A^{\mathbb{Z}^2}$ that are shift-invariant and topologically closed. In other words, any shift of a configuration of a subshift X is still in X , and the limit of a sequence of shifts of configurations of X is in X .

Let $\tau_{\mathbf{u}}$ denote the shift action by the vector $\mathbf{u} \in \mathbb{Z}^2$: let p be a pattern or configuration of support D (which is \mathbb{Z}^2 if p is a configuration), then for all $\mathbf{v} \in D + \mathbf{u}$, $\tau_{\mathbf{u}}(p)_{\mathbf{v}} = p_{\mathbf{v}-\mathbf{u}}$. A configuration c is said to be *periodic* if there exists $\mathbf{u} \in \mathbb{Z}^2 \setminus \{(0, 0)\}$ such that $c = \tau_{\mathbf{u}}(c)$. A subshift is *aperiodic* if it is not empty and contains no periodic configuration.

The set of valid tilings by a set of Wang tiles mentioned in the introduction is an SFT. These models are actually equivalent, as any SFT can be made into a set of Wang tilings by application of a single morphism. As subshifts are the most natural model to deal with substitutions, we chose to use this formalism in this paper instead of Wang tilesets.

2.2 Pattern complexity

The set of all patterns appearing in a configuration c is called the language of c , $\mathcal{L}(c)$; and we denote by $\mathcal{L}_{m,n}(c) := \mathcal{L}(c) \cap A^{\{0, \dots, m-1\} \times \{0, \dots, n-1\}}$ the set of all $m \times n$ rectangular patterns appearing in c . Similarly, $\mathcal{L}(X) = \bigcup_{c \in X} \mathcal{L}(c)$ is the set of all patterns of the subshift X and $\mathcal{L}_{m,n}(X) := \bigcup_{c \in X} \mathcal{L}_{m,n}(c)$ is the set of $m \times n$ patterns of X . The *pattern complexity* (or just *complexity*) of a configuration (resp. a subshift) is the number of such rectangular patterns $|\mathcal{L}_{m,n}(c)|$ (resp. $|\mathcal{L}_{m,n}(X)|$).

For configurations of dimension one, periodicity and complexity are linked: the Morse-Hedlund theorem states that a one-dimensional configuration $w \in A^{\mathbb{Z}}$ is periodic if and only if there exists n such that $P_c(n) \leq n$ [18]. For one-dimensional SFTs, the question is trivial as every non-empty SFT contains a periodic configuration (see for example [17]), and thus the complexity of the subshift cannot have any impact on its aperiodicity.

In dimension two, the link between complexity and periodicity becomes much more involved. Nivat conjectured in 1997 an implication similar to Morse and Hedlund's: any two-dimensional configuration with $P_c(m, n) \leq m, n$ for some m, n must be periodic

[20]. The other direction does not hold, as there exists periodic configurations with exponential complexity [4]. Since two dimensional aperiodic SFTs do exist, it is natural to try to link their complexity with their aperiodicity, and a bound relating the two has been recently found:

Theorem (Kari, Moutot [15]). *Let X be a subshift such that $\exists c \in X, \exists m, n, |\mathcal{L}_{m,n}(c)| \leq mn$. Then there exists $d \in X$ that is periodic.*

Which can be rephrased as a lower bound on the complexity of aperiodic subshifts.

Corollary 1. *Let X be an aperiodic subshift. Then for all $c \in X$ and all m, n ,*

$$|\mathcal{L}_{m,n}(c)| \geq mn + 1.$$

2.3 Substitutions

Substitutions on words have been extensively studied, see for example [10] for a good reference about them. In dimension two, they have been studied mostly in the context of substitution of \mathbb{R}^2 , that is, substitutions producing geometrical tilings. In the current paper, we restrict ourselves to discrete uniform two-dimensional substitutions. In this context, a *substitution* of size $m \times n$ (or support $\{0, \dots, m-1\} \times \{0, \dots, n-1\}$) is a map $\sigma : A \rightarrow A^{\{0, \dots, m-1\} \times \{0, \dots, n-1\}}$. It is said to be *primitive* if there exists k such that for all $a, b \in A$, we have $a \sqsubseteq \sigma^k(b)$, and *invertible* if no two letters have the same image. The *subshift of a substitution*, also called a substitutive subshift, is the set of configurations that can be made from applying the substitution to letters of A :

$$X^\sigma := \left\{ c \in A^{\mathbb{Z}^2} \mid \forall p \sqsubset c, \exists a \in A, \exists k \in \mathbb{N}, p \sqsubseteq \sigma^k(a) \right\}.$$

In general this subshift is only a sofic subshift and not an SFT, as shown by Mozes [19].

Substitutions are a very convenient way of generating aperiodicity, by example by considering fixpoints configurations. In the context of geometrical tilings, Solomyak characterized when the subshift of a substitution is aperiodic or not [25]. We recall here his result in the specific case of uniform rectangular substitutions.

Definition 1 (Unique desubstitution). A \mathcal{D} -pattern $p \in A^{\mathcal{D}}$ can be *uniquely desubstituted* by a substitution σ of size (k, l) if there exists $\mathbf{t} \in \mathbb{Z}^2$ which is unique modulo (k, l) such that there exists $c \in X^\sigma, \tau_{\mathbf{t}}(\sigma(c))|_{\mathcal{D}} = p$.

Theorem 2 (Solomyak [25]). *Let σ be a primitive substitution. Then X^σ is aperiodic if and only if there exists $\rho > 0$ such that all patterns p whose support fits in $\{1, \dots, \rho\}^2$ can be uniquely desubstituted by σ .*

A quadratic upper bound is already known for substitutive subshifts, due to Hansen and Robinson.

Theorem (Hansen, Robinson [23, Theorem 7.17]). *Let σ be a primitive and invertible substitution with square support. Then there exists $K > 0$ such that*

$$|\mathcal{L}_{n,n}(X)| \leq Kn^2.$$

Our lower bound will not deal with all invertible substitutions, but only with the ones having what we call a *determining position*, which is a position in the image that allows us to uniquely recover the antecedent.

Definition 2 (Determining position). A substitution σ has a *determining position* (i, j) if for all $a, b \in A$:

$$\sigma(a)_{i,j} = \sigma(b)_{i,j} \Rightarrow a = b.$$

Equivalently, this means that $\{\sigma(a)_{i,j} \mid a \in A\} = A$. See Fig. 1 for an illustration of this property.

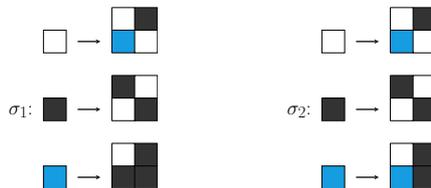


Figure 1: Illustration of Definition 2: σ_1 has a determining position $(0, 0)$ whereas σ_2 has no determining position.

Remark 1. If $|A| = 2$, all non-constant substitutions have a determining position. Indeed, as the images of the two letters must be different, they must differ in at most one position, which is then a determining position.

3 Optimality of the $mn + 1$ bound

Corollary 1 shows that all configurations in aperiodic subshifts must have complexity at least $mn + 1$ for all m, n .

It is not known whether this bound is optimal in general, however if m and n are fixed, it is possible to prove its optimality.

Theorem 3. *For all $m, n \in \mathbb{N}$, there exists an aperiodic SFT X such that for all $c \in X$, we have $|\mathcal{L}_{m,n}(c)| = mn + 1$.*

Proof. Let $m, n \in \mathbb{N}$ be fixed and $Y \subseteq A^{\mathbb{Z}^2}$ be an aperiodic SFT on some alphabet A (for example the set of all Robinson tilings [24]). We will “blow up” Y by encoding each of its colors into a large rectangle over alphabet $\{0, 1\}$.

Let k be such that $2^{k^2} \geq |A|$. Then, for each $a \in A$, it is easy to build a $k \times k$ square $S(a) \in \{0, 1\}^{\{0, \dots, k-1\}^2}$ encoding a in such a way that S is an isomorphism. S can be naturally extended to patterns or configurations of $A^{\mathbb{Z}^2}$.

Let R_0 and R_1 be the $m \times n$ rectangles with respectively a 0 and a 1 in position $(0, 0)$ and 0s elsewhere. We define the substitution

$$\sigma : \begin{cases} 0 \mapsto R_0 \\ 1 \mapsto R_1 \end{cases}.$$

Let $X := \sigma(S(Y)) = \{\sigma(S(c)) \mid c \in Y\}$. To prove that X is an SFT, let F_0 be the finite set of forbidden patterns of Y . Let

$$F_1 = \left\{ p \in \{0, 1\}^{\{0, \dots, 2km-1\} \times \{0, \dots, 2kn-1\}} \mid \forall q \in A^{\{0, \dots, 3\} \times \{0, \dots, 3\}}, p \not\subseteq \sigma(S(q)) \right\}$$

be the set of $2km \times 2kn$ patterns not appearing in any $2km \times 2kn$ patterns of the form $\sigma(S(q))$. Let $F = F_1 \cup \sigma(S(F_0))$. Then, $X = X_F$. Indeed, F_1 ensures that all configurations of X_F are desubstitutable by $\sigma(S)$, and together with $\sigma(S(F_0))$, we have

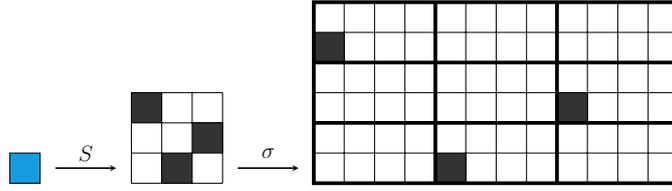


Figure 2: Illustration of the sparse encoding of Theorem 3 with $k = 3$, $m = 4$ and $n = 2$. We first apply S to some $s \in A$, and then apply σ to this image.

$c \in X_F \Leftrightarrow c \in \sigma(S(Y))$ thanks to the fact that $\sigma(S)$ is one-to-one. As F is finite, X is an SFT.

Now, since every $c \in X$ can be written $\sigma(d)$ with $d \in S(X_0)$, every $m \times n$ rectangular pattern of c contains at most one 1, therefore $|\mathcal{L}_{m,n}(c)| \leq mn + 1$. But because X is aperiodic, Corollary 1 ensures that $|\mathcal{L}_{m,n}(c)| \geq mn + 1$, so $|\mathcal{L}_{m,n}(c)| = mn + 1$. \square

One might hope to improve the $mn + 1$ bound of Corollary 1 by using a result of Cassaigne:

Theorem (Cassaigne, direct consequence of [5]). *Let $c \in A^{\mathbb{Z}^2}$ be such that for all m, n , $|\mathcal{L}_{m,n}(c)| = mn + 1$. then there exists $d \in \mathcal{O}(c)$ which is uniform. In particular, the subshift $\mathcal{O}(c)$ is not periodic.*

However, it gives an information about configurations with $|\mathcal{L}_{m,n}(c)| = mn + 1$ for all m, n . In order to improve Corollary 1 to $mn + 2$, one would need a similar result with the hypothesis of the existence of m, n such that $|\mathcal{L}_{m,n}(c)| = mn + 1$. Together with Theorem 3, this emphasizes the importance of the quantifiers and their order when studying the complexity of subshifts.

The SFT built in Theorem 3 depends on m and n , so it does not rule out the possibility of improving the uniform bound in m and n . In the next section we show that the uniform lower bound can indeed be improved for a large class of substitutive subshifts.

4 Lower bound for substitutive subshifts

In this section, we show our lower bound for a class of two-dimensional substitutive subshifts. First, we need a classical lemma on substitutive subshifts, whose (short) proof is included for sake of completeness.

Lemma 4. *Let σ be a primitive substitution and $c \in X^\sigma$. Then for all $p \sqsubset c$, $\sigma(p) \sqsubset c$.*

Proof. Let $p \sqsubset c$. By definition of X^σ , there exists $a \in A$, $k \in \mathbb{N}$ such that $p \sqsubset \sigma^k(a)$. This implies that $\sigma(p) \sqsubset \sigma^{k+1}(a)$. Moreover, σ is primitive: there exists some $l \in \mathbb{N}$ such that for all $b \in A$, $a \sqsubset \sigma^l(b)$ and consequently $\sigma^{k+1}(a) \sqsubset \sigma^{l+k+1}(b)$.

By definition of X^σ , for a big enough pattern p' appearing in c , there is some b such that $\sigma^{l+k+1}(b) \sqsubset p'$. Eventually, we have:

$$\sigma(p) \sqsubset \sigma^{k+1}(a) \sqsubset \sigma^{l+k+1}(b) \sqsubset p' \sqsubset c.$$

\square

Then, we take a look at the determining position property for composition of substitutions.

Lemma 5. *If σ and σ' have determining positions, then $\sigma \circ \sigma'$ has a determining position.*

Proof. Let σ and σ' be two substitutions of size (m, n) having determining positions (k, l) and (k', l') respectively. Let $a, a' \in A$. Then,

$$\begin{aligned} (\sigma \circ \sigma')(a)_{(mk+k', nl+l')} &= (\sigma \circ \sigma')(a')_{(mk+k', nl+l')} \\ \Rightarrow \sigma \left(\sigma'(a)_{(k,l)} \right)_{(k',l')} &= \sigma \left(\sigma'(a')_{(k,l)} \right)_{(k',l')} \\ \Rightarrow \sigma'(a)_{(k,l)} &= \sigma'(a')_{(k,l)} \\ \Rightarrow a &= a'. \end{aligned}$$

□

An immediate consequence is the following.

Corollary 6. *If σ has a determining position, then for all $k \geq 1$, σ^k has a determining position.*

The next technical lemma is the key of our lower bound. Essentially, it provides us a lower bound on the complexity based on the number of times k that a patterns can be uniquely desubstituted.

Lemma 7. *Let σ be a substitution of size (M, M) having a determining position and $k \in \mathbb{N}$. For all $n \in \mathbb{N}$, if all the patterns of size (n, n) of X^σ have a unique desubstitution by σ^k , then for all $c \in X^\sigma$:*

$$|\mathcal{L}_{n,n}(c)| \geq n^2 \left(1 + \frac{1}{\lceil \frac{n}{M^k} \rceil^2} \right).$$

Proof. Let us denote by $\mathcal{P}_c(k, l)$ the elements of $\mathcal{L}_{k,l}(c)$ for this proof for better reading of the equations. Using the fact that σ^k also has a determining position we build an injective function f from the disjoint union $\bigsqcup_{0 \leq i, j \leq M^k - 1} \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil)$ to $\mathcal{P}_c(n, n)$. The existence of such an injective function, put together with the result of Kari and Moutot (Corollary 1), yields the result.

Let $(d_1, d_2) \in \{0, \dots, M^k - 1\}$ be the determining position of σ^k obtained in Corollary 6, $i, j \in \{0, \dots, M^k - 1\}$ and $p \in \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil)$. The fact that p appears in c implies that it can be extended to bigger patterns also appearing in c . Then, for each such p , we fix one $p^{ext} \in A^{\{-1, \dots, \lceil \frac{n-i}{M^k} \rceil + 1\} \times \{-1, \dots, \lceil \frac{n-j}{M^k} \rceil + 1\}}$ such that $p^{ext}|_{\{0, \dots, \lceil \frac{n-i}{M^k} \rceil - 1\} \times \{0, \dots, \lceil \frac{n-j}{M^k} \rceil - 1\}} = p$ and $p^{ext} \sqsubset c$. We define the following pattern:

$$f_{i,j}(p) = \tau_{i-d_1, j-d_2}(\sigma^k(p^{ext}))|_{\{0, \dots, n-1\}^2}$$

This is well defined because the fact that $i, j, d_1, d_2 \in \{0, \dots, M^k - 1\}$ implies that

$$\{0, \dots, n-1\}^2 \subset (i-d_1, j-d_2) + \{-1, \dots, \lceil \frac{n-i}{M^k} \rceil + 1\} \times \{-1, \dots, \lceil \frac{n-j}{M^k} \rceil + 1\}.$$

By Lemma 4 it holds that $\sigma^k(p^{ext}) \sqsubset c$ and consequently that $f_{i,j}(p) \in \mathcal{P}_c(n, n)$.

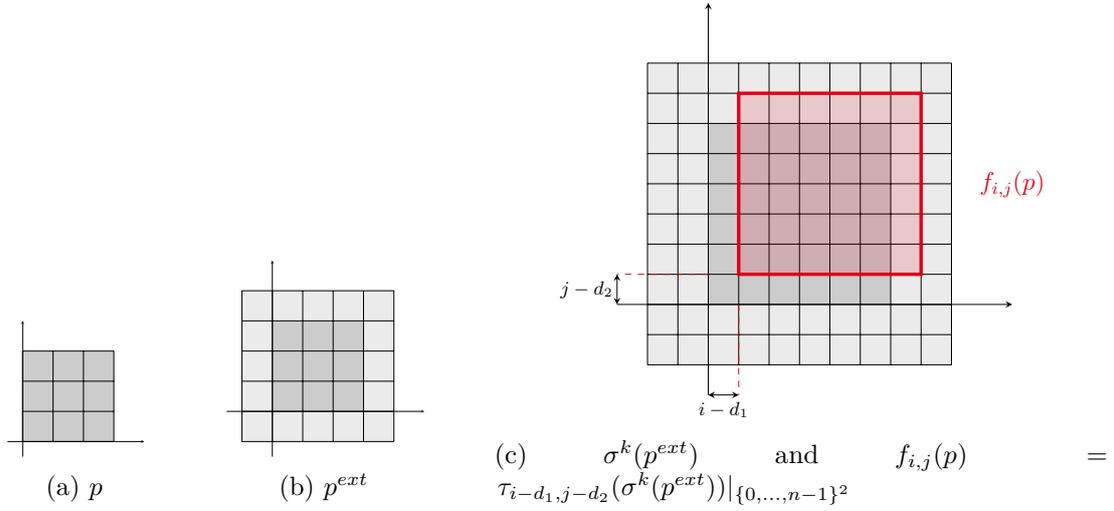


Figure 3: Illustration of the definition of $f_{i,j}(p)$ with $n = 6, i = j = 1, M = 2, k = 1$ and $(d_1, d_2) = (0, 0)$. Figure (a) shows $p \in \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil)$, (b) shows the completion to p^{ext} and (c) the image of p^{ext} by σ^k and the pattern $f_{i,j}(p)$ extracted from it.

Then, the following function:

$$f : \bigsqcup_{0 \leq i, j \leq M^k - 1} \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil) \rightarrow \mathcal{P}_c(n, n)$$

$$p \in \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil) \mapsto f_{i,j}(p)$$

is well defined. Now we show that it is injective.

Let $p \in \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil)$, $p' \in \mathcal{P}_c(\lceil \frac{n-i'}{M^k} \rceil, \lceil \frac{n-j'}{M^k} \rceil)$ and suppose that $f(p) = f(p')$. By definition of f , there are some patterns $p^{ext}, p'^{ext} \sqsubseteq c$, such that

$$p^{ext}|_{\{0, \dots, \lceil \frac{n-i}{M^k} \rceil - 1\} \times \{0, \dots, \lceil \frac{n-j}{M^k} \rceil - 1\}} = p \quad \text{and} \quad p'^{ext}|_{\{0, \dots, \lceil \frac{n-i'}{M^k} \rceil - 1\} \times \{0, \dots, \lceil \frac{n-j'}{M^k} \rceil - 1\}} = p'$$

and that:

$$f(p) = \tau_{i-d_1, j-d_2}(\sigma^k(p^{ext}))|_{\{0, \dots, n-1\}^2} = \tau_{i'-d_1, j'-d_2}(\sigma^k(p'^{ext}))|_{\{0, \dots, n-1\}^2} = f(p') \quad (1)$$

Moreover as $f(p) = f(p')$ is a pattern of size (n, n) , it has a unique desubstitution by σ^k by assumption. Therefore, $(i - d_1, j - d_2) \equiv (i' - d_1, j' - d_2) \pmod{(M^k, M^k)}$. As $i, j, i', j' \in \{0, \dots, M^k - 1\}$, it implies that $(i, j) = (i', j')$. Consequently, Eq. (1) can be rewritten as

$$\tau_{i-d_1, j-d_2}(\sigma^k(p^{ext}))|_{\{0, \dots, n-1\}^2} = \tau_{i-d_1, j-d_2}(\sigma^k(p'^{ext}))|_{\{0, \dots, n-1\}^2}.$$

That is,

$$\sigma^k(p^{ext})|_{\{0, \dots, n-1\}^2 - (i-d_1, j-d_2)} = \sigma^k(p'^{ext})|_{\{0, \dots, n-1\}^2 - (i-d_1, j-d_2)}.$$

And as $p^{ext} \sqsubseteq p$,

$$\sigma^k(p)|_{\{0, \dots, n-1\}^2 - (i-d_1, j-d_2)} = \sigma^k(p')|_{\{0, \dots, n-1\}^2 - (i-d_1, j-d_2)}, \quad (2)$$

on the domain where $\sigma^k(p)$ is defined.

Now we show that p and p' are the same on all their domain. Let $\mathbf{u} \in \{0, \dots, \lceil \frac{n-i}{M^k} \rceil - 1\} \times \{0, \dots, \lceil \frac{n-j}{M^k} \rceil - 1\}$. First, we have:

$$M^k \mathbf{u} + (d_1, d_2) \in \{0, \dots, n-1\}^2 - (i-d_1, j-d_2).$$

Therefore by using the fact that (d_1, d_2) is a determining position, Eq. (2) gives:

$$\begin{aligned} \sigma^k(p)_{M^k \mathbf{u} + (d_1, d_2)} &= \sigma^k(p')_{M^k \mathbf{u} + (d_1, d_2)}, \\ \Rightarrow \sigma^k(p_{\mathbf{u}})_{d_1, d_2} &= \sigma^k(p'_{\mathbf{u}})_{d_1, d_2}, \\ \Rightarrow p_{\mathbf{u}} &= p'_{\mathbf{u}}. \end{aligned}$$

This being true for all $\mathbf{u} \in \{0, \dots, \lceil \frac{n-i}{M^k} \rceil - 1\} \times \{0, \dots, \lceil \frac{n-j}{M^k} \rceil - 1\}$, it eventually holds that $p = p'$. This concludes the proof of the injectivity of f .

The injectivity of f yields the following inequality:

$$|\mathcal{L}_{n,n}(c)| \geq \left| \bigsqcup_{0 \leq i, j \leq M^k - 1} \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil) \right| = \sum_{0 \leq i, j \leq M^k - 1} \left| \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil) \right|.$$

Since c is aperiodic and by Corollary 1, for all $0 \leq i, j \leq M^k - 1$, it holds that:

$$\begin{aligned} \left| \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil) \right| &\geq \lceil \frac{n-i}{M^k} \rceil \times \lceil \frac{n-j}{M^k} \rceil + 1 \\ &\geq \lceil \frac{n-i}{M^k} \rceil \times \lceil \frac{n-j}{M^k} \rceil \left(1 + \frac{1}{\lceil \frac{n-i}{M^k} \rceil \times \lceil \frac{n-j}{M^k} \rceil} \right) \\ &\geq \lceil \frac{n-i}{M^k} \rceil \times \lceil \frac{n-j}{M^k} \rceil \left(1 + \frac{1}{\lceil \frac{n}{M^k} \rceil^2} \right) \end{aligned}$$

Eventually, taking the sum and using a classical property of the ceiling function [11, p85] gives:

$$\begin{aligned} |\mathcal{L}_{n,n}(c)| &\geq \sum_{0 \leq i, j \leq M^k - 1} \left| \mathcal{P}_c(\lceil \frac{n-i}{M^k} \rceil, \lceil \frac{n-j}{M^k} \rceil) \right| \\ &\geq \left(1 + \frac{1}{\lceil \frac{n}{M^k} \rceil^2} \right) \sum_{0 \leq i, j \leq M^k - 1} \lceil \frac{n-i}{M^k} \rceil \times \lceil \frac{n-j}{M^k} \rceil \\ &= \left(1 + \frac{1}{\lceil \frac{n}{M^k} \rceil^2} \right) n^2. \end{aligned}$$

□

Our last lemma gives us a bound on how many times patterns from a substitutive subshift can be desubstituted thanks to Solomyak result's (Theorem 2).

Lemma 8. *Let σ be an invertible substitution. If all patterns $p \in \mathcal{L}(X^\sigma)$ whose support contains $\{0, \dots, \rho - 1\}^2$ can be uniquely desubstituted by σ ; then for all k , all patterns in $\mathcal{L}(X^\sigma)$ whose support contains $\{0, \dots, (\rho + 1)M^{k-1} - 2\}^2$ can be uniquely desubstituted by σ^k .*

Proof. We show the lemma by induction on k . The result is true for $k = 1$. Suppose that the result is true for some $k \geq 1$ fixed. Let p be a pattern in $\mathcal{L}(X^\sigma)$ with support \mathcal{D} such that $\{0, \dots, (\rho + 1)M^k - 2\}^2 \subset \mathcal{D}$. Let $\mathbf{t}, \mathbf{t}' \in \mathbb{Z}^2$ be such that $p = \tau_{\mathbf{t}}(\sigma^{k+1}(c))|_{\mathcal{D}} = \tau_{\mathbf{t}'}(\sigma^{k+1}(c'))|_{\mathcal{D}}$ for some $c, c' \in X^\sigma$. By supposition, and because $(\rho + 1)M^k \geq \rho$, the pattern p can be uniquely desubstituted by σ . This, and the fact that $p = \tau_{\mathbf{t}}(\sigma(\sigma^k(c)))|_{\mathcal{D}} = \tau_{\mathbf{t}'}(\sigma(\sigma^k(c')))|_{\mathcal{D}}$ implies that $\mathbf{t} \equiv \mathbf{t}' \pmod{(M, M)}$. Therefore $\mathbf{t}' = \mathbf{t} + M\mathbf{u}$ for some $\mathbf{u} \in \mathbb{Z}^2$ and then:

$$\begin{aligned} \tau_{\mathbf{t}}(\sigma(\sigma^k(c)))|_{\mathcal{D}} &= \tau_{\mathbf{t}+M\mathbf{u}}(\sigma(\sigma^k(c')))|_{\mathcal{D}}, \\ \Rightarrow \sigma(\sigma^k(c))|_{\mathcal{D}+\mathbf{t}} &= \tau_{a\mathbf{u}}(\sigma(\sigma^k(c')))|_{\mathcal{D}+\mathbf{t}} \\ \Rightarrow \sigma(\sigma^k(c))|_{\mathcal{D}+\mathbf{t}} &= \sigma(\tau_{\mathbf{u}}(\sigma^k(c')))|_{\mathcal{D}+\mathbf{t}}. \end{aligned}$$

Let $\mathcal{R} = \{0, \dots, (\rho + 1)M^{k-1} - 2\}^2$. For some $\mathbf{v} \in \mathbb{Z}^2$, we have the following inclusion (see Fig. 4):

$$M(\mathcal{R} + \mathbf{v}) + \{0, \dots, M - 1\}^2 \subset \{0, \dots, (\rho + 1)M^k - 2\}^2 + \mathbf{t} \subset \mathcal{D} + \mathbf{t}. \quad (3)$$

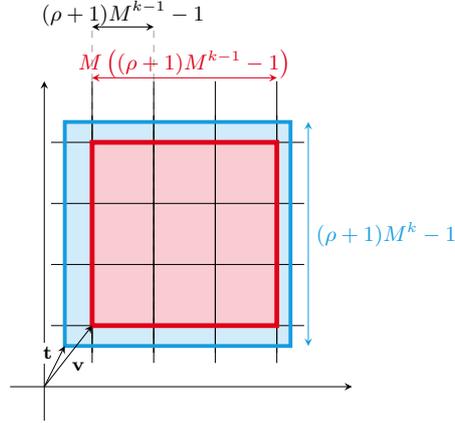


Figure 4: Inclusion of the sets of Eq. (3).

Then, because σ is invertible, we have:

$$\begin{aligned} \sigma(\sigma^k(c))|_{\mathcal{D}+\mathbf{t}} &= \sigma(\tau_{\mathbf{u}}(\sigma^k(c')))|_{\mathcal{D}+\mathbf{t}} \\ \Rightarrow \sigma(\sigma^k(c))|_{M(\mathcal{R}+\mathbf{v})+\{0,\dots,M-1\}^2} &= \sigma(\tau_{\mathbf{u}}(\sigma^k(c')))|_{M(\mathcal{R}+\mathbf{v})+\{0,\dots,M-1\}^2} \\ \Rightarrow \sigma^k(c)|_{\mathcal{R}+\mathbf{v}} &= \tau_{\mathbf{u}}(\sigma^k(c'))|_{\mathcal{R}+\mathbf{v}} \end{aligned}$$

By induction hypothesis, this pattern can be uniquely desubstituted by σ^k , therefore it holds that $\mathbf{u} \equiv 0 \pmod{M^k}$, which implies that $\mathbf{t}' = \mathbf{t} + M\mathbf{u} \equiv \mathbf{t} \pmod{M^{k+1}}$ and that p can be uniquely desubstituted by σ^{k+1} . \square

One can then desubstitute patterns from a substitutive subshifts as many times as possible, and bringing the two previous lemmas together gives the lower bound:

Theorem 9. *Let σ be a primitive square aperiodic substitution having a determining position. Then there exists $K > 1$ such that*

$$\forall c \in X^\sigma, \forall n \in \mathbb{N}, |\mathcal{L}_{n,n}(c)| \geq Kn^2$$

Proof. By Theorem 2, there exists ρ such that all patterns p appearing in the configurations of X^σ in which a square of size ρ fits can be uniquely desubstituted by σ . Let $c \in X^\sigma$ and $n \in \mathbb{N}$. Let $k(n) = \max\{k \in \mathbb{N} \mid \forall p \in \mathcal{L}_{n,n}(c), p \text{ can be uniquely desubstituted by } \sigma^k\}$. By Lemma 7 it holds that:

$$|\mathcal{L}_{n,n}(c)| \geq n^2 \left(1 + \frac{1}{\lceil \frac{n}{M^{k(n)}} \rceil^2} \right)$$

Additionally, $(\rho + 1)M^{k(n)} - 1 > n$ otherwise Lemma 8 would contradict the maximality of $k(n)$. Therefore, $\rho + 1 > \frac{n}{M^{k(n)}}$ and $\rho + 1 > \lceil \frac{n}{M^{k(n)}} \rceil$. This turns our bound into:

$$|\mathcal{L}_{n,n}(c)| \geq n^2 \left(1 + \frac{1}{(\rho + 1)^2} \right)$$

which gives a constant $K := 1 + \frac{1}{(\rho+1)^2} > 1$ which does not depend neither on c nor on n . \square

Remark 2. A similar bound can be found for non-square uniform substitutions. However, if one wants to know the pattern complexity $|\mathcal{L}_{n,n}(c)|$ for some m, n , the bound depends on how close the shape of the substitution is with respect to the rectangle (m, n) . More precisely, let σ be a substitution of size (M, N) . Then for any $R > 0$ such that

$$|\log_M(m) - \log_N(n)| \leq R,$$

there exists $K_R > 1$ such that for all $c \in X^\sigma$,

$$|\mathcal{L}_{n,n}(c)| \geq K_R mn.$$

To conclude this section, let us apply our result to Robinson's tiling. Let $X^R \subseteq A^{\mathbb{Z}^2}$ be the SFT constituted of all valid tilings by Robinson's tiling set A [24].

A natural way of constructing a substitution on A is to "grow" the pattern of each tile to a 2×2 pattern. Each tile is composed of a background marking (in light-blue and gray on our drawings) and of a main marking (in black). We generate the rules of our 2×2 substitution σ in the following way:

1. The image of each tile has the same background marking,
2. the upper right tile of the image of a tile t has the same main marking than t ,
3. the lower left tile of the image of t has a main marking with four outgoing arrows having its blue parts oriented as the background marking of t ,
4. the two other tiles are chosen to have one main arrow and to incoming arrows on each side so that the main arrow extends the upper right tile.

Fig. 5 illustrates the construction of $\sigma(t)$ for one tile $t \in A$ and Fig. 6 its iteration. For space considerations, we do not write the full substitution obtained by this process.

Even though X^R is not a substitutive subshift by itself, as substitutive patterns of its configurations can be separated by "fracture lines", one can show that $X^\sigma \subset X^R$, and therefore for all n , $|\mathcal{L}_{n,n}(X^\sigma)| \leq |\mathcal{L}_{n,n}(X^R)|$.

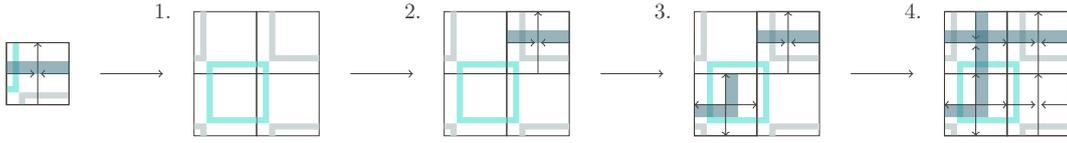


Figure 5: The four steps of construction of the substitution σ .

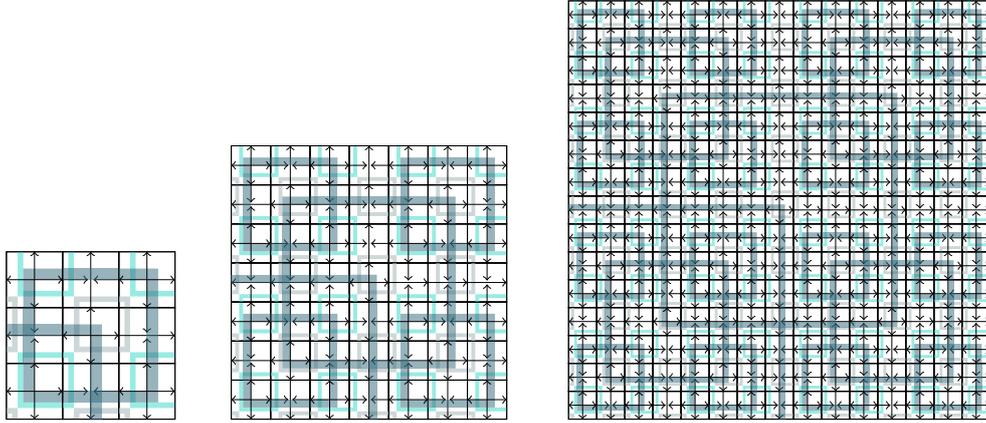


Figure 6: Portions of patterns generated by successive applications of σ .

Now, we partition A into two sets A_1 and A_2 . A_1 contains all the tiles with four outgoing arrows, of the form , and A_2 all the other tiles. Then, we define

$$f : \begin{cases} A & \rightarrow \{0, 1\} \\ t & \mapsto \begin{cases} 1 & \text{if } t \in A_1 \\ 0 & \text{if } t \in A_2 \end{cases} \end{cases}$$

f is compatible with σ in the sense that if $f(t_1) = f(t_2)$ for two tiles in A , then $f(\sigma(t_1)) = f(\sigma(t_2))$. Therefore the following substitution is well defined:

$$\sigma' : \begin{cases} \{0, 1\} & \rightarrow \{0, 1\}^{\{0,1\}^2} \\ a & \mapsto f((\sigma(t)) \text{ with } \sigma(t) = a \end{cases}$$

Fig. 7 illustrate σ' the kind of patters it generates.

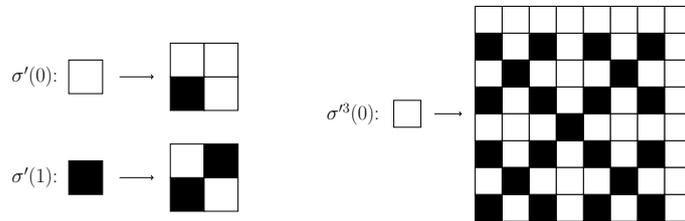


Figure 7: The substitution σ' and an example of pattern generated by it.

Let $X^{\sigma'}$ be the subshift of σ' . We have that $\mathcal{L}(X^{\sigma'}) = \{\sigma'^k(0), \sigma'^k(1) \mid k \in \mathbb{N}\} = \{f(\sigma^k(t)) \mid t \in A, k \in \mathbb{N}\}$. Therefore for all n , it holds that $\mathcal{L}_{n,n}(X^{\sigma'}) = f(\mathcal{L}_{n,n}(X^{\sigma}))$. This implies that $|\mathcal{L}_{n,n}(X^{\sigma'})| \leq |\mathcal{L}_{n,n}(X^{\sigma})|$. Then, σ' is primitive with a determining

position and $X^{\sigma'}$ is aperiodic since every 1 corresponds to one corner of a square in Robinson's tiling, therefore Theorem 9 gives that there exists $K > 1$ such that $|\mathcal{L}_{n,n}(X^{\sigma'})| \geq Kn^2$, and then:

$$|\mathcal{L}_{n,n}(X^R)| \geq Kn^2.$$

5 Perspectives

We believe that Theorem 9 can be extended to any square uniform primitive substitution, without the assumption of determining position. In the current proof, this condition allows us to “recover” information on patterns when de-substituting in Lemma 7, even when the pattern does not contain a full $\sigma^k(a)$ for some $a \in A$. For now, we do not see how to obtain a quadratic lower bound without this information, but we think that it is possible to do. Such a generalization would show that all substitutive subshifts of square primitive substitutions have complexity $|\mathcal{L}_{n,n}(X^\sigma)| \in \Theta(n^2)$. In addition to a better understanding of aperiodic subshifts, we hope that this is the first step towards a classification of two-dimensional substitutions in terms of pattern complexity, similar to Pansiot's [21] in the one dimensional case.

It is also unknown if the $mn + 1$ complexity bound of Corollary 1 can be improved for non-substitutive aperiodic subshifts.

Acknowledgements

The authors want to thank Nathalie Aubrun for her guidance and lots of helpful advice, as well as Guilhem Gamard for his careful proofreading.

References

- [1] Pierre Arnoux, Valérie Berthé, Thomas Fernique, and Damien Jamet. Functional stepped surfaces, flips, and generalized substitutions. *Theoretical Computer Science*, 380(3):251–265, 2007. Combinatorics on Words. URL: <https://www.sciencedirect.com/science/article/pii/S0304397507001995>, doi:<https://doi.org/10.1016/j.tcs.2007.03.031>.
- [2] Robert Berger. *The Undecidability of the Domino Problem*. PhD thesis, Harvard University, 1964.
- [3] Robert Berger. *The Undecidability of the Domino Problem*. Memoirs of the American Mathematical Society. The American Mathematical Society, 1966.
- [4] Valérie Berthé and Laurent Vuillon. Tilings and rotations on the torus: a two-dimensional generalization of sturmian sequences. *Discrete Mathematics*, 223(1):27–53, 2000. doi:[10.1016/S0012-365X\(00\)00039-X](https://doi.org/10.1016/S0012-365X(00)00039-X).
- [5] Julien Cassaigne. Double sequences with complexity $mn + 1$. *Journal of Automata, Languages and Combinatorics*, 4:153–170, 01 1999.
- [6] Karel Culik. An aperiodic set of 13 wang tiles. *Discrete Mathematics*, 160(1):245–251, 1996. doi:[https://doi.org/10.1016/S0012-365X\(96\)00118-5](https://doi.org/10.1016/S0012-365X(96)00118-5).
- [7] Bruno Durand, Guilhem Gamard, and Anaël Grandjean. Aperiodic tilings and entropy. In Arseny M. Shur and Mikhail V. Volkov, editors, *Developments in Language Theory*, page 166–177. Springer International Publishing, 2014. doi:[10.1007/978-3-319-09698-8_15](https://doi.org/10.1007/978-3-319-09698-8_15).
- [8] Fabien Durand, Bernard Host, and Christian Skau. Substitutional dynamical systems, bratteli diagrams and dimension groups. *Ergodic Theory and Dynamical Systems*, 19:953 – 993, 08 1999. doi:[10.1017/S0143385799133947](https://doi.org/10.1017/S0143385799133947).
- [9] Thomas Fernique. Multidimensional sturmian sequences and generalized substitutions. *International Journal of Foundations of Computer Science*, 17(03):575–599, 2006. arXiv:<https://doi.org/10.1142/S0129054106004005>, doi:[10.1142/S0129054106004005](https://doi.org/10.1142/S0129054106004005).
- [10] N P Fogg, V Berthé, S Ferenczi, C Mauduit, and A Siegel. *Substitutions in Dynamics, Arithmetics and Combinatorics*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2003. URL: <https://books.google.fr/books?id=U1tuCQAAQBAJ>.
- [11] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, 1989.
- [12] Branko Grünbaum and G. C. Shephard. *Tilings and patterns*, 1987.
- [13] Emmanuel Jeandel and Michaël Rao. An aperiodic set of 11 wang tiles. *Advances in Combinatorics*, January 2021. doi:[10.19086/aic.18614](https://doi.org/10.19086/aic.18614).
- [14] Jarkko Kari. A small aperiodic set of wang tiles. *Discrete Mathematics*, 160(1–3):259–264, 1996.

- [15] Jarkko Kari and Etienne Moutot. Decidability and periodicity of low complexity tilings. In Christophe Paul and Markus Bläser, editors, *37th International Symposium on Theoretical Aspects of Computer Science (STACS 2020)*, volume 154 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14:1–14:12. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2020. URL: <https://drops.dagstuhl.de/opus/volltexte/2020/11875>, doi:10.4230/LIPIcs.STACS.2020.14.
- [16] Sébastien Labbé. Substitutive structure of jeandel–rao aperiodic tilings. *Discrete & Computational Geometry*, 65(3):800–855, Apr 2021. doi:10.1007/s00454-019-00153-3.
- [17] Douglas A Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995.
- [18] Marston Morse and Gustav A. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60(4):815–866, 1938.
- [19] Shahar Mozes. Tilings, substitutions systems and dynamical systems generated by them. *Journal d’analyse mathématique*, 53:139–186, 1988. doi:10.1007/BF02793412.
- [20] Maurice Nivat. Keynote address at the 25th anniversary of eatcs, during icalp 1997, bologna. 1997.
- [21] Jean-Jacques Pansiot. Complexité des facteurs des mots infinis engendrés par morphismes itérés. In Jan Paredaens, editor, *Automata, Languages and Programming*, page 380–389. Springer Berlin Heidelberg, 1984.
- [22] R. Penrose. Pentaplexity a class of non-periodic tilings of the plane. *The Mathematical Intelligencer*, 2(1):32–37, Mar 1979. doi:10.1007/BF03024384.
- [23] E. Arthur Robinson. *Symbolic dynamics and tilings of \mathbb{R}^d* , volume 60. 2004. doi:10.1090/psapm/060/2078847.
- [24] Raphael M. Robinson. Undecidability and nonperiodicity for tilings of the plane. *Inventiones mathematicae*, 12(3):177–209, Sep 1971. doi:10.1007/BF01418780.
- [25] Boris Solomyak. Nonperiodicity implies unique composition for self-similar translationally finite tilings. *Discrete & Computational Geometry*, 20(2):265–279, 1998. Publisher: Springer.
- [26] Hao Wang. Proving theorems by pattern recognition – ii. *The Bell System Technical Journal*, 40(1):1–41, 1961. doi:10.1002/j.1538-7305.1961.tb03975.x.

A The Robinson tileset

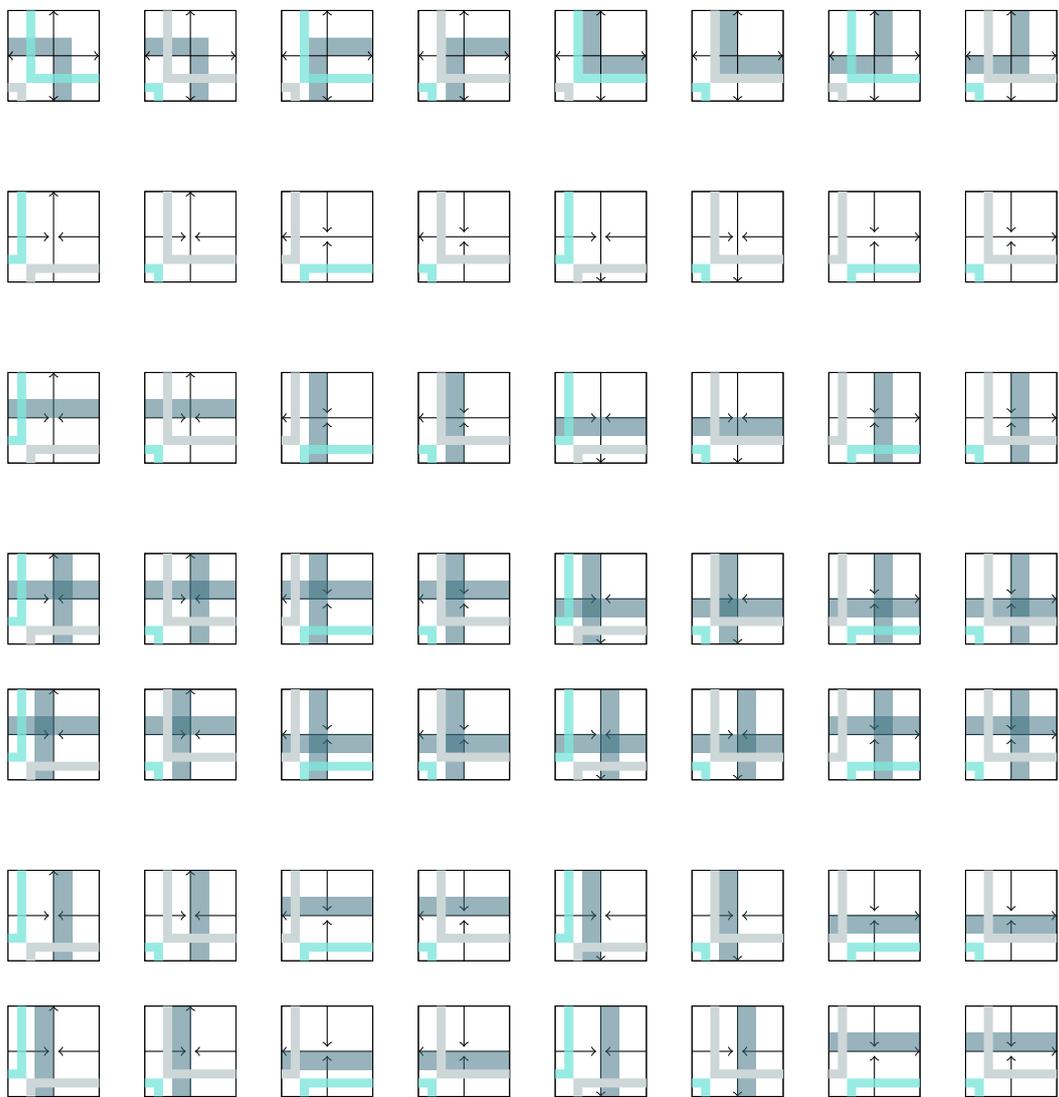


Figure 8: Robinson's tileset.