Pi01 sets and tilings

Emmanuel Jeandel, Pascal Vanier

Π_1^0 SETS AND TILINGS

EMMANUEL JEANDEL AND PASCAL VANIER

Abstract. In this paper, we prove that given any Π_1^0 subset P of $\{0,1\}^{\mathbb{N}}$ there is a tileset τ with a set of configurations C such that $P \times \mathbb{Z}^2$ is recursively homeomorphic to $C \setminus U$ where U is a computable set of configurations. As a consequence, if P is countable, this tileset has the exact same set of Turing degrees.

Introduction

Wang tiles have been introduced by Wang [17] to study fragments of first order logic. Knowing whether a tileset can tile the plane with a given tile at the origin (also known as the origin constrained domino problem) was proved undecidable also by Wang [18]. Knowing whether a tileset can tile the plane in the general case was proved undecidable by Berger [2, 3].

Understanding how complex, in the sense of recursion theory, the tilings of a given tileset can be is a question that was first studied by Myers [13] in 1974. Building on the work of Hanf [10], he gave a tileset with no recursive tilings. Durand/Levin/Shen [9] showed, 40 years later, how to build a tileset for which all tilings have high Kolmogorov complexity.

A Π_1^0-set is an effectively closed subset of $\{0,1\}^{\mathbb{N}}$, or equivalently the set of oracles on which a given Turing machine halts. Π_1^0-sets occur naturally in various areas in computer science and recursive mathematics, see e.g. [6, 15] and the upcoming book [7]. It is easy to see that the set of tilings of a given tileset is a Π_1^0-set (up to a recursive coding of $Q^{\mathbb{Z}^2}$ into $\{0,1\}^{\mathbb{N}}$). This has various consequences. As an example, every non-empty tileset contains a tiling which is not Turing-hard (see Durand/Levin/Shen [9] for a self-contained proof). The main question is how different the sets of tilings are from Π_1^0-sets. In the context of one-dimensional symbolic dynamics, some answers to these questions were given by Cenzer/Dashti/King/Tosca/Wyman [4, 5, 8].

The main result in this direction was obtained by Simpson [16], building on the work of Hanf and Myers: for every Π_1^0-set S, there exists a tileset whose set of tilings have the same *Medvedev* degree as S. The Medvedev degree roughly relates to the "easiest" Turing degree of S. What we are interested in is a stronger result: *can we find for every Π_1^0-set S a tileset whose set of tilings have the same Turing degrees ?* We prove in this article that this is true if S contains a recursive point. More exactly we build (theorem 4.1) for every Π_1^0-set S a set of tilings for which the set of Turing degrees is exactly the same as for S, possibly with the additional Turing degree of recursive points. In particular, as every *countable* Π_1^0-set contains a recursive point, the question is completely solved for countable sets: the sets of Turing degrees of countable Π_1^0-sets are the same as the sets of Turing degrees of countable sets of tilings. In particular, there exist countable sets of tilings with non-recursive points. This can be thought as a two-dimensional version of theorem 8 in [5].

This paper is organized as follows. After some preliminary definitions, we start with a quick proof of a generalization of Hanf, already implicit in Simpson [16]. We then build a very specific tileset, which forms a grid-like structure while having only countably many tilings. This tileset will then serve as the main ingredient in the theorem in the last section.

1. Preliminaries

1.1. **Π_1^0 sets and degrees.** A Π_1^0 *set* $P \subseteq \{0,1\}^{\mathbb{N}}$ is a set for which there exists a Turing machine that given $x \in \{0,1\}^{\mathbb{N}}$ as an oracle halts if and only if $x \notin P$. Equivalently, a subset $S \subseteq \{0,1\}^{\mathbb{N}}$ is Π_1^0 if there exists a recursive set L so that $w \in S$ if no prefix of w is in L.

We say that two sets S, S' are *recursively homeomorphic* if there exists a bijective recursive function $f : S \to S'$.

A point x of a set $S \subseteq \{0,1\}^{\mathbb{N}}$ is *isolated* if it has a prefix that no other point of S has. The *Cantor-Bendixson derivative* $D(S)$ of S is the set S without its isolated points. We define inductively $S^{(\lambda)}$ for any ordinal λ:

- $S^{(0)} = S$
- $S^{(\lambda+1)} = D\left(S^{(\lambda)}\right)$
- $S^{(\lambda)} = \bigcap_{\gamma<\lambda} S^{(\gamma)}$ when λ is limit.

The *Cantor-Bendixson rank* of S, noted $CB(S)$, is defined as the first ordinal λ such that $S^{(\lambda)} = S^{(\lambda+1)}$. An element x is of rank λ in S if λ is the least ordinal such that $x \notin S^{(\lambda)}$.

See Cenzer/Remmel [6] for Π_1^0 sets and Kechris [11] for Cantor-Bendixson rank and derivative.

For $x, y \in \{0,1\}^{\mathbb{N}}$ we say that x is *Turing-reducible* to y if y is computable by a Turing machine using x as an oracle and we write $y \leq_T x$. If $x \leq_T y$ and $y \leq_T x$, we say that x and y are *Turing-equivalent* and we write $x \equiv_T y$. The *Turing degree* of $x \in \{0,1\}^{\mathbb{N}}$ is its equivalence class under the relation $\equiv_T$.

1.2. **Tilings and SFTs.** *Wang tiles* are unit squares with colored edges which may not be flipped or rotated. A *tileset* T is a finite set of Wang tiles. A *configuration* is a mapping $c : \mathbb{Z}^2 \to T$ assigning a Wang tile to each point of the plane. If all adjacent tiles of a configuration have matching edges, the configuration is called a tiling. The set of all tilings of T is noted $\mathcal{T}(T)$. We say a tileset is *origin constrained* when the tile at position $(0, 0)$ is forced, that is to say, we only look at the valid tilings having a given tile t at the origin.

A *Shift of Finite Type (SFT)* $X \subseteq \Sigma^{\mathbb{Z}^2}$ is defined by (Σ, F) where Σ is a finite alphabet and F a finite set of *forbidden patterns.* A *pattern* is a coloring of a finite portion $P \subset \mathbb{Z}^2$ of the plane. A point x is in X if and only if it does not contain any forbidden pattern of F anywhere. In particular, the set of tilings of a Wang tileset is a SFT. Conversely, any SFT is recursively homeomorphic to a Wang tileset. More information on SFTs may be found in Lind and Markus' book [12].

A set of configurations $X \subseteq \Sigma_X^{\mathbb{Z}^2}$ is a *sofic shift* iff there exists a SFT $Y \subseteq \Sigma_Y^{\mathbb{Z}^2}$ and a local map $f : \Sigma_Y \to \Sigma_X$ such that for any point $x \in X$, there exists a point $y \in Y$ such that for all $z \in \mathbb{Z}^2, x(z) = f(y(z))$.

The notion of *Cantor-Bendixson derivative* is defined on configurations in a similar way as with Π_1^0 sets. This notion was introduced for tilings by Ballier/Durand/Jeandel [1]. A configuration c is said to be *isolated* in a set of configurations C if there exists a pattern P such that c is the only configuration of C containing P. The Cantor-Bendixson derivative of C is noted $D(C)$ and consists of all configurations of C except the isolated ones. We define $C^{(\lambda)}$ inductively for any ordinal λ as above.

2. Π_1^0 sets and origin constrained tilings

A straighforward corollary of Hanf [10] is that Π_1^0 subsets of $\{0,1\}^{\mathbb{N}}$ and origin constrained tilings are recursively isomorphic. This is stated explicitely in Simpson [16].

Theorem 2.1. *Given any Π_1^0 subset P of $\{0,1\}^{\mathbb{N}}$, there exists a tileset and a tile t such that each origin constrained tiling with this tileset describes an element of P.*

Proof. We take the basic encoding of Turing machines as stated in Robinson [14] for instance. We modify the bottom tiles, ie the tiles containing the initial tape, such that instead of being able to contain only the blank symbol, they can contain only 0s or 1s on the right of the starting head. The Turing machine we encode is the one that given $x \in \{0,1\}^{\mathbb{N}}$ as an input halts if and only if $x \notin P$. Then the constrained tilings, having at the origin the tile with the starting head of the Turing machine, are exactly the runs of the Turing machine on the members of P. □

Corollary 2.2. *Any Π_1^0 subset P of $\{0,1\}^{\mathbb{N}}$ is recursively homeomorphic to an origin constrained tileset.*

3. The tileset

The main problem in the construction of Hanf is that tilings which do not have the given tile at the origin can be very wild : they may correspond to configurations with no computation (no head of the Turing Machine) or computations starting from an arbitrary (not initial) configuration. A way to solve this problem is described in [13] but is unsuitable for our purposes.

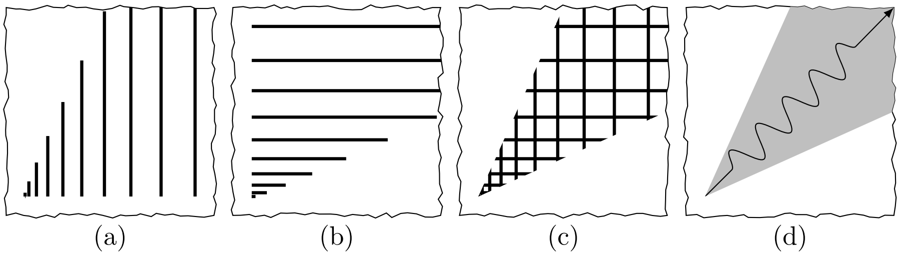

FIGURE 1. The tiling in which to encode the Turing machines

Our idea is as follows: We build a tileset which will contain, among others, the *sparse grid* of figure 1c. The main point is that all others tilings of the tileset will have at most one intersection point of two black lines. This means that if we put computation cells of a given Turing machine in the intersection points, every tiling which is not of the form of figure 1c will contain at most one cell of the Turing machine, thus will contain no computation.

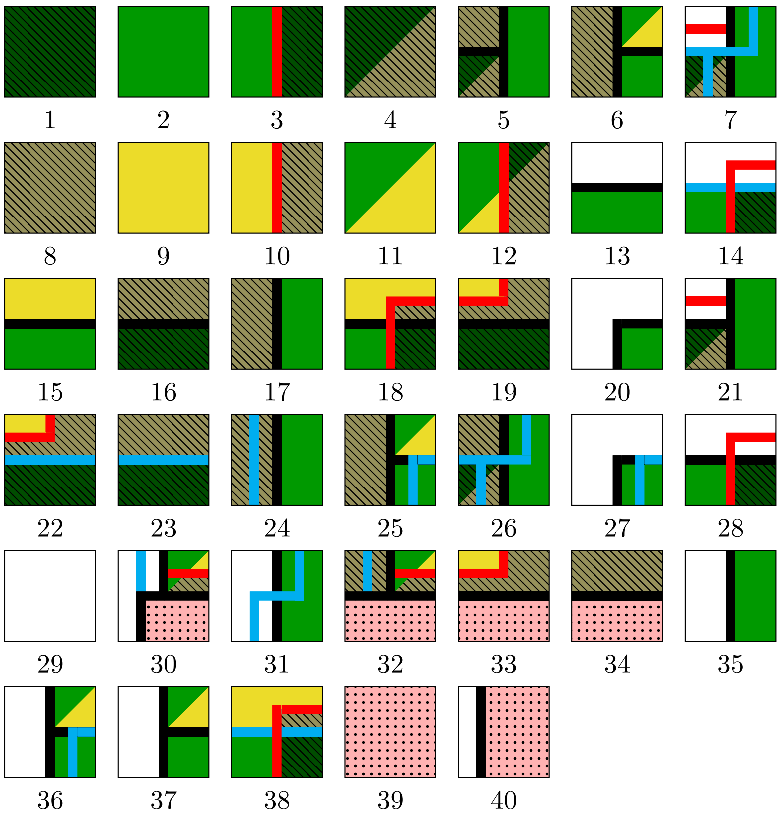

FIGURE 2. Our set of Wang tiles T.

To do this construction, we will first draw increasingly big and distant columns as in figure 1a and then superimposing the same construction for rows as in figure 1b, leading to the grid of figure 1c.

It is then fairly straightforward to see how we can encode a Turing machine inside a configuration having the skeleton of figure 1c by looking at it diagonally: time increases going to the north-east and the tape is written on the north west - south east diagonals[1].

Our set of tiles T of figure 2 gives the skeleton of figure 1a when forgetting everything but the black vertical borders. We will prove in this section that it is countable. We set here the vocabulary:

- a vertical line is formed of a vertical succession of tiles containing a vertical black line (tiles 5, 6, 17, 21, 24, 25, 26, 27, 31, 35, 36, 37).
- a horizontal line is formed of a horizontal succession of tiles containing a horizontal black line (tiles 13, 14, 15, 16, 22, 23, 38) or a bottom signal,
- the bottom signal is formed by a connected path of tiles among (30, 31, 27, 14, 7, 36, 38)
- the red signal is formed by a connected path of tiles containing a red line (tiles among 3 ,7, 10, 12, 14, 19, 22, 32, 33, 38).
- tile 30 is the corner tile
- tiles 30, 32, 33, 34 are the bottom tiles

Lemma 3.1. *The tileset T admits at most one tiling with two or more vertical lines.*

Proof. The idea of the construction is to force that whenever there are two vertical lines, then the only possible tiling is the one of figure 3. Note that whenever the corner tile appears in a tiling, it is necessarily a shifted version of the tiling on figure 3.

Suppose that we have a tiling in which two vertical lines appear. Suppose they are at distance $k+1$. Necessarily there must be horizontal lines between them forming squares. Inside these squares there must be a red signal: inside each square, this red signal is vertical, it is shifted to the right each time it crosses a horizontal line. This ensures that there are exactly k squares in this column. Furthermore, the bottom square has necessarily a bottom signal going through its top horizontal line. The bottom signal forces the square of the column before to be of size $k-1$ and the square of the column after to be of size exactly $k+1$.

□

Lemma 3.2. *The tileset T admits a countable number of tilings.*

Proof. Lemma 3.1 states that there is only one tiling that has more than 2 vertical lines. This means that the other tilings have at most one such line.

- If a tiling has exactly one vertical line, then it can have at most two horizontal lines: one on the left of the vertical one and one on the right. A red signal can then appear on the left or the right of the vertical line arbitrary far from it. There is a countable number of such tilings.
- If a tiling has no vertical line, then it has at most one horizontal line. A red signal can then appear only once. There is a finite number of such tilings.

[1] Note that we will have to skip one diagonal out of two in our construction, in order for the tape to increase at the same rate as the time.

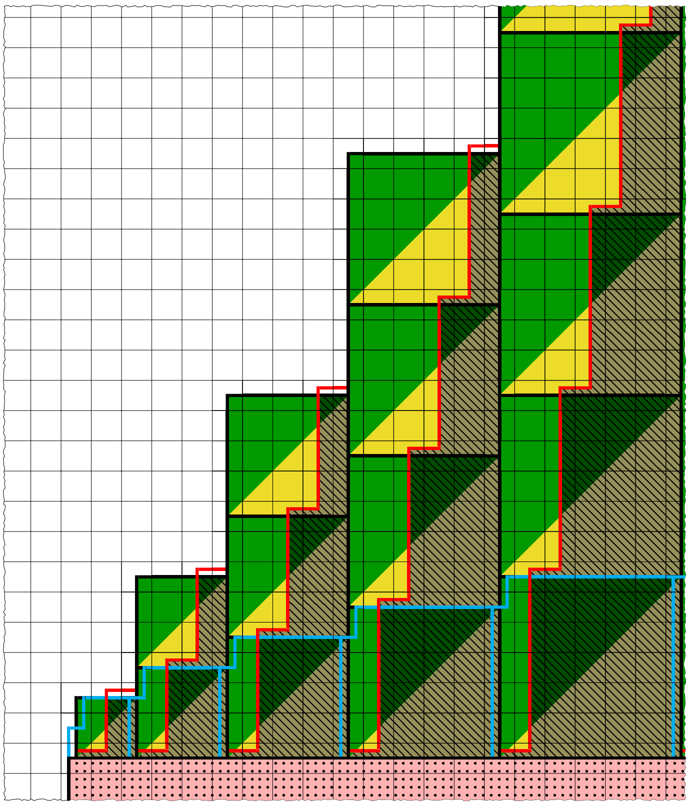

FIGURE 3. Tiling α: the unique valid tiling of T in which there are 2 or more vertical lines.

There is a countable number of tilings that can be obtained with the tileset T. All obtainable tilings are shown in figure 4 and 3.

□

By taking our tileset $T = \{1, \dots, 40\}$ and mirroring all the tiles along the south west-north east diagonal, we obtain a tileset $T' = \{1', \dots, 40'\}$ with the exact same properties, except it enforces the squeleton of figure 1b. Remember that whenever the corner tile appeared in a tiling, then necessarily this tiling was α. The same goes for T' and its corner tile. We hence construct a third tileset $\tau = (T \setminus \{30\} \times T' \setminus \{30'\}) \cup \{(30, 30')\}$. The corner tile $(30, 30')$ of τ has the property that whenever it appears, the tiling is the superimposition of the skeletons of figures 1a and 1b with the corner tiles at the same place: there is only one such tiling, call it β.

The skeleton of figure 1c is obtained if we forget about the parts of the lines of the T layer (resp. T') that are superimposed to white tiles, 29' (resp. 29), of T' (resp. T).

As a consequence of lemma 3.2, τ is countable. And as a consequence of lemma 3.1, the only tiling by τ in which computation can be embedded is β. The shape of β is the one of figure 1c, the coordinates of the points of the grid are the

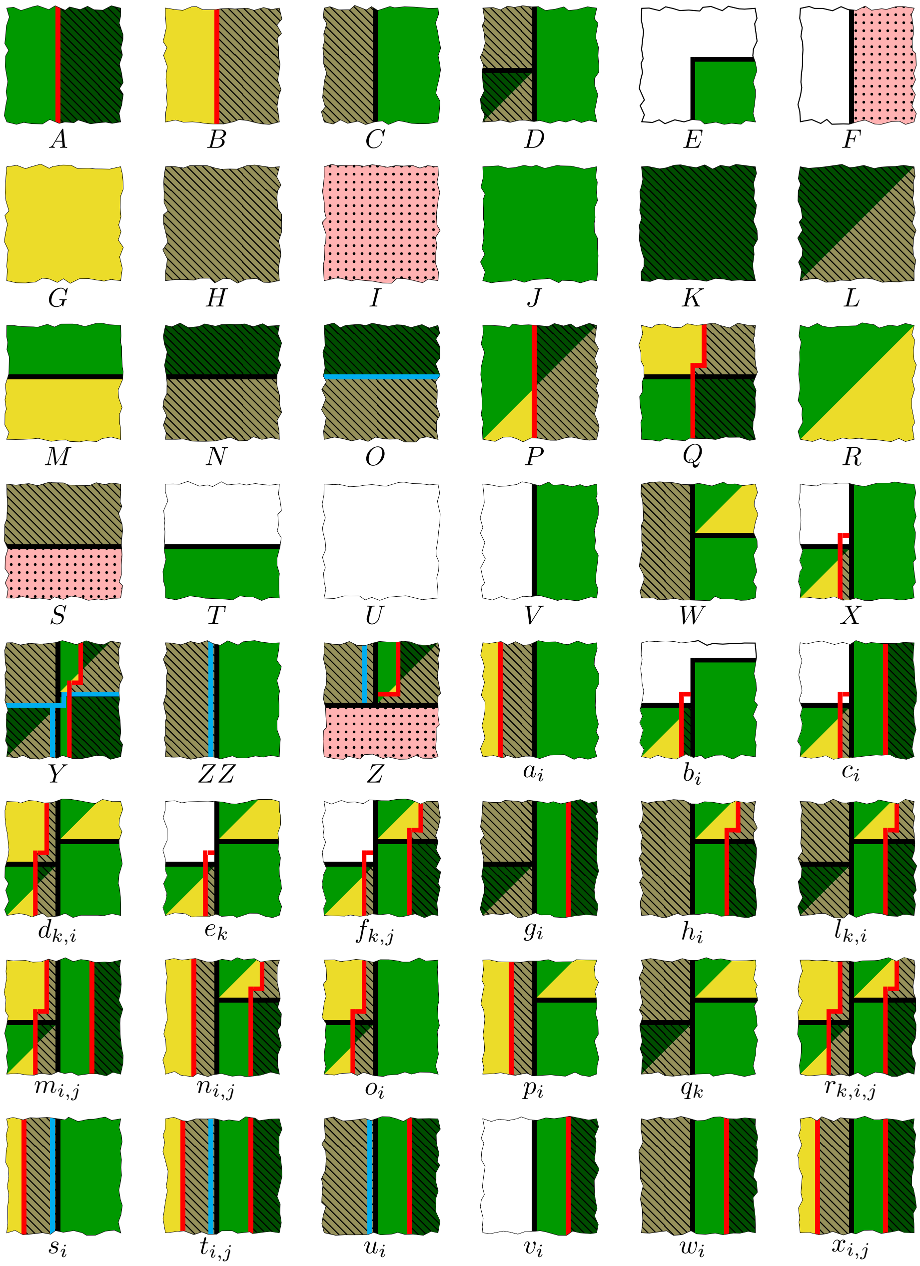

FIGURE 4. The other configurations: the $A - ZZ$ configurations are unique (up to shift), and the configurations with subscripts $i, j \in \mathbb{N}, k \in \mathbb{Z}^2$ represent the fact that distances between some of the lines can vary. Note that configuration ZZ cannot have a red signal on its left, because it would force another vertical line.

following (supposing tile $(30, 30')$ is at the center of the grid):

$$\{(f(n), f(m)) \mid f(m)/4 \leq f(n) \leq 4f(m)\}$$

$$\{(f(n), f(m)) \mid m/2 \leq n \leq 2m\}$$

where $f(n) = (n+1)(n+2)/2 - 1$.

Lemma 3.3. *The Cantor-Bendixson rank of $\mathcal{T}(\tau)$ is 12.*

Proof. The Cantor-Bendixson rank of $\mathcal{T}(T) \setminus \{\alpha\}$ is 6, see figure 4, thus the rank of $\mathcal{T}(T) \setminus \{\alpha\} \times \mathcal{T}(T') \setminus \{\alpha'\}$ is 11. Adding the configurations corresponding to the superimposition of α and α', τ is of rank 12.

□

4. Π_1^0 sets and tilings

Theorem 4.1. *For any Π_1^0 subset S of $\{0,1\}^{\mathbb{N}}$ there exists a tileset τ_S such that $S \times \mathbb{Z}^2$ is recursively homeomorphic to $\mathcal{T}(\tau_S) \setminus O$ where O is a computable set of configurations.*

Proof. This proof uses the construction of section 3. Let M be a Turing machine such that M halts with x as an oracle iff $x \notin S$. Take the tileset τ of section 3 and encode in it the Turing machine M having as an oracle x on an unmodifiable second tape. This gives us τ_M, O is the set all tilings except the β ones. To each $(x, p) \in S \times \mathbb{Z}^2$ we associate the β tiling having a corner at position p and having x on its oracle tape. It follows from lemma 3.2 that O is clearly computable.

□

Corollary 4.2. *For any countable Π_1^0 subset S of $\{0,1\}^{\mathbb{N}}$, there exists a tileset τ having exactly the same Turing degrees.*

Proof. We know, from Cenzer/Remmel [6], that countable Π_1^0 sets have $\mathbf{0}$ (computable elements) in their set of Turing degrees, thus the tileset τ_M described in the proof of theorem 4.1 has exactly the same Turing degrees as S.

□

Theorem 4.3. *For any countable Π_1^0 subset S of $\{0,1\}^{\mathbb{N}}$ there exists a tileset τ_S such that $CB(\mathcal{T}(\tau_S)) = CB(S) + 11$.*

Proof. Lemma 3.3 states that $\mathcal{T}(\tau)$ is of Cantor-Bendixson rank 12, 11 without α. In the tileset τ_M of the previous proof, the Cantor-Bendixson rank of the contents of the tape is exactly $CB(S)$, hence $CB(\mathcal{T}(\tau_S)) = CB(S) + 11$.

□

From Ballier/Durand/Jeandel [1] we know that for any tileset X, if $CB(\mathcal{T}(X)) \geq 2$, then X has only recursive points. Thus an optimal construction improves the Cantor-Bendixson rank by at least 2.

Corollary 4.4. *For any countable Π_1^0 subset S of $\{0,1\}^{\mathbb{N}}$ there exists a sofic subshift X such that $CB(X) = CB(S) + 2$.*

Proof. Take a projection that just keeps the symbols of the Turing machine tape τ_M of the proof of theorem 4.1 and maps everything else to a blank symbol. Recall the Turing machine tape cells are the intersections of the vertical lines and horizontal lines. This projection leads to 3 possible configurations :

- a completely blank configuration,
- a completely blank configuration with only one symbol somewhere,
- a configuration with a white background and points corresponding to the intersections in the sparse grid of figure 1c.

□

References

[1] Alexis Ballier, Bruno Durand, and Emmanuel Jeandel. Structural aspects of tilings. *25th International Symposium on Theoretical Aspects of Computer Science (STACS).*, 2008.

[2] Robert Berger. *The Undecidability of the Domino Problem*. PhD thesis, Harvard University, 1964.

[3] Robert Berger. *The Undecidability of the Domino Problem*. Number 66 in Memoirs of the American Mathematical Society. The American Mathematical Society, 1966.

[4] Douglas Cenzer, Ali Dashti, and Jonathan L. F. King. Computable symbolic dynamics. *Mathematical Logic Quarterly*, 54(5):460–469, 2008.

[5] Douglas Cenzer, Ali Dashti, Ferit Toska, and Sebastian Wyman. Computability of Countable Subshifts. In *Computability in Europe (CiE)*, volume 6158 of *Lecture Notes in Computer Science*, pages 88–97, 2010.

[6] Douglas Cenzer and J.B. Remmel. Π_1^0 classes in mathematics. In *Handbook of Recursive Mathematics - Volume 2: Recursive Algebra, Analysis and Combinatorics*, volume 139 of *Studies in Logic and the Foundations of Mathematics*, chapter 13, pages 623–821. Elsevier, 1998.

[7] Douglas Cenzer and Jeffrey Remmel. *Effectively Closed Sets*. ASL Lecture Notes in Logic, 2011. in preparation.

[8] Ali Dashti. *Effective Symbolic Dynamics*. PhD thesis, University of Florida, 2008.

[9] Bruno Durand, Leonid A. Levin, and Alexander Shen. Complex tilings. *Journal of Symbolic Logic*, 73(2):593–613, 2008.

[10] William Hanf. Non Recursive Tilings of the Plane I. *Journal of Symbolic Logic*, 39(2):283–285, June 1974.

[11] Alexander S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[12] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, New York, NY, USA, 1995.

[13] Dale Myers. Non Recursive Tilings of the Plane II. *Journal of Symbolic Logic*, 39(2):286–294, June 1974.

[14] Raphael M. Robinson. Undecidability and Nonperiodicity for Tilings of the Plane. *Inventiones Math.*, 12, 1971.

[15] Stephen Simpson. Mass Problems Associated with Effectively Closed Sets. in preparation.

[16] Stephen G. Simpson. Medvedev Degrees of 2-Dimensional Subshifts of Finite Type. *Ergodic Theory and Dynamical Systems*, 2011.

[17] Hao Wang. Proving theorems by Pattern Recognition II. *Bell Systems technical journal*, 40:1–41, 1961.

[18] Hao Wang. Dominoes and the ∀∃∀ case of the decision problem. *Mathematical Theory of Automata*, pages 23–55, 1963.

E-mail address: emmanuel.jeandel@lif.univ-mrs.fr

E-mail address: pascal.vanier@lif.univ-mrs.fr

LABORATOIRE D'INFORMATIQUE FONDAMENTALE DE MARSEILLE