Natural coding of minimal rotations of the torus, induction and exduction

Mélodie Andrieu

[PhD thesis version, March 2021]

Abstract

We discuss a topological definition of natural coding of a minimal rotation on the d-dimensional torus, inspired by the seminal works of Rauzy on the Tribonacci word. In particular, we show that under the axiom of choice, it is possible to wisely complete the pseudo-fundamental domain of the torus into a fundamental domain, while preserving the property of piecewise translation and a weak form of sequential continuity. We prove then that if w is a natural coding of a minimal rotation of the d-torus, admitting furthermore $d+1$ return words to a letter a, then its derivated word to the letter a is still a natural coding of a minimal rotation of the d-torus, that we fully describe. In particular, this result completes an argument of Cassaigne, Ferenczi and Zamboni: under this assumption, and if furthermore the fundamental domain associated with the natural coding is bounded, then the cylinder $[a]$ is a bounded remainder set for w (i.e. the empiric frequency with which the symbolic trajectory w visits the set $[a]$ tends to its expected value at speed at least $1/n$), which is equivalent to finite imbalance on the letter a. As a consequence, no Arnoux-Rauzy word with infinite imbalance is a natural coding of a minimal rotation of the 2-dimensional torus, with bounded fundamental domain. The same holds for primitive C-adic words and, more generally, uniformly recurrent tree words.

Besides, we prove that to any natural coding of a minimal rotation of the d-torus we can associate other natural codings constructed by a reverse induction process, that we call *exduction*. We study the return words of Arnoux-Rauzy and primitive C-adic words within the S-adic framework and obtain that, for these two classes of words, being a natural coding of a minimal rotation of the 2-torus is a property that only depends on the asymptotic behavior of the directive sequence.

1 Introduction

Backgrounds

In [Rau82], Rauzy undertakes the study of symbolic systems associated with minimal rotations of the 2-dimensional torus by a remarkable example: the Tribonacci word. The "canonical association" he obtains, through a well-chosen partition, was refered to (though not written at that time) under the name of "natural coding". Later, when it appeared that the Tribonacci word was a remarkable element of a wide class of words generalizing Sturmian words on a 3-letter alphabet [AR91], it was believed that the canonical association property would extend to the whole class of words (now known as Arnoux-Rauzy words).

In [CFZ00], Cassaigne, Ferenczi and Zamboni disproved this belief by exhibiting an Arnoux-Rauzy word satisfying a remarkable combinatorial property: infinite imbalance (see Definition 1). The first and main part of their paper is devoted to the construction of this unsuspected object; in the second part, relying on a theorem of Rauzy on bounded remainder sets (see [Rau84] or Theorem B below), they state that if w_∞ is an Arnoux-Rauzy word with infinite imbalance, then either w_∞ or one of its induced words (which are still Arnoux-Rauzy words with infinite imbalance) is not a natural coding of a rotation of the 2-torus (the definition of natural coding is discussed in Section 3). Even if the proof is incorrect, their result is true, under the additional assumptions of boundedness of the fundamental domain and minimality of the rotation (it is a consequence of Corollary 44). In Section 4 of the present document, we rectify and complete the proof that Cassaigne, Ferenczi and Zamboni sketched to achieve a more significant result: Theorem A. Besides, the existence of non-coding Arnoux-Rauzy words was established by other techniques in [CFM08].

Since then, substantial advances have been made in the counterpart. Under a measure theory definition, [AI01], [BJS12] and [BŠW13] show that purely substitutive Arnoux-Rauzy words are natural codings of rotation of the 2-torus; [BST19] extends this positive result in the generic S-adic case to a large subclass of Arnoux-Rauzy words.

Our work

First, we propose a topological definition of natural coding of a minimal rotation on the d-dimensional torus, inspired by the seminal works of Rauzy [Rau82] (Definition 2). Under this framework, we cannot elude the question of borders.

We show that if w is a natural coding of a minimal rotation of the d-torus, then: 1) w is written with d+1 letters and is uniformly recurrent (see immediate Lemma 3); 2) under the axiom of choice, it is possible to wisely complete the pseudo-fundamental domain of the torus into a fundamental domain, while preserving the property of piecewise translation as well as a the continuity, in an of course weak sense, of the coding function (see Proposition 9 and Lemma 16). If furthermore w admits $d+1$ return words to a factor v, then its derivated word to v (see Definition 27) is still a natural coding of a minimal rotation of the d-torus, that we fully describe (see Theorem A).

Fulfilling an argument of [CFZ00], we prove then that the cylinder $[v]$ is a bounded remainder set for w, which is equivalent to finite imbalance on w for the factor v (see Proposition 37). As a consequence, no Arnoux-Rauzy word with infinite imbalance is a natural coding of a minimal rotation of the 2-torus. This consequence holds for primitive C-adic words as well, and more generally for tree words (see Corollary 44).

On another hand, we show that the property of being a natural coding of a minimal rotation also passes through a reverse induction operation that we call *exduction* (Theorem G).

In the case of Arnoux-Rauzy and primitive C-adic words, we link the induction and exduction processes to the action of a multidimensional continued fraction algorithm on the letter frequencies vector of w, through the S-adic expression of return words (see Theorem D for Arnoux-Rauzy words, which is a restatement of [JV00], and Theorem E for primitive C-adic words).

Finally, we show that for Arnoux-Rauzy and primitive C-adic subshifts, being a natural coding of a minimal rotation on the 2-torus only depends on the asymptotic behavior of the directive sequence (Theorem H).

Contents

2 Preliminaries

2.1 Finite and infinite words

An *alphabet* $\mathcal{A}$ is a finite set; its elements are called *letters*. For instance, in what follows, we work with the $(d+1)$-letter alphabet $\mathcal{I} = \{1, ..., d+1\}$. A *finite word of length* n, where n is a nonnegative integer, is the concatenation of n letters: $u = a_0 \cdot a_1 \cdot ... \cdot a_{n-1} \in \mathcal{A}^n$. As soon as there is no ambiguity, the concatenation symbol $\cdot$ will be omitted. We denote by $\mathcal{A}^* = \cup_{n\in\mathbb{N}} \mathcal{A}^n$ the set of finite words; an infinite word is an element $w = a_0 a_1 ... \in \mathcal{A}^{\mathbb{N}}$. Following `Python`, we denote by $w[k]$, for $k \in \mathbb{N}$, the $(k+1)$-th letter of a nonempty (finite or infinite) word w.

A finite word u is a *factor of length* n of a (finite or infinite) word w if there exists a nonnegative integer i such that for all $k \in \{0, ..., n-1\}$, $w[i+k] = u[k]$; in the particular case $i = 0$, we say that u is the *prefix of length* n of w, and denote it by $u = \mathrm{pref}_n(w)$. We denote by $\mathcal{F}_n(w)$ the set of factors of w of length n and by $\mathcal{F}(w)$ the set of factors of all lengths. An infinite word is said *recurrent* if every factor occurs infinitely often.

We endow the set $\mathcal{A}^{\mathbb{N}}$ with the product topology, for which it is compact. Given a finite word $u \in \mathcal{A}^*$, we denote by $[u]$ the set of words in $\mathcal{A}^{\mathbb{N}}$ which admit u as prefix. The sets $[u]$ are called *cylinders*; they are clopen and form a neighborhood basis for the topology.

2.2 Symbolic dynamics

We denote by S the *shift map*, which acts on infinite words by 'erasing' the first letter: if w is an infinite word, for all $k \in \mathbb{N}, S(w)[k] = w[k+1]$. If w_0 is an infinite word, we call *subshift associated with* w_0, denoted by X_0, the closure set (for the product topology) of the trajectory of w_0 under the shift action: $X_0 = \overline{\{S^k(w_0)|k \in \mathbb{N}\}}$.

The shift map is at the core of symbolic dynamics. Given a dynamical system, one can choose to partition the space into a finite number of regions $A_1, ..., A_d$, and study the possible sequences of regions crossed over time (for a general introduction to symbolic dynamics, the reader should refer to [LM95]). The difficulty and the interest of the discrete dynamical system thus obtained highly depends on the choice made for the partition. In this paper, we study the behavior, under the induction and the exduction operations, of a family of remarkable partitions for the discrete flow of a minimal rotation on a d-dimensional torus (called hereafter d-torus).

2.3 Imbalance

Definition 1. *The* imbalance *of an infinite word w is the quantity (possibly infinite):*

$$\text{imb}(w) = \sup_{n \in \mathbb{N}} \sup_{u,u' \in \mathcal{F}_n(w)} \max_{a \in \mathcal{A}} ||u|_a - |u'|_a|,$$

where $|u|_a$ denotes the number of times the letter a appears in the word u.

The imbalance measures iniquities in the distribution of letters in a given word. This notion appeared for the first time in the works of Morse and Hedlund ([MH38] and [MH40]); in [CH73] Coven and Hedlund showed that this quantity characterizes Sturmian words: a binary word is Sturmian if and only if it is aperiodic and its imbalance equals 1.

This quantity has been much studied since, through the notions of C-balancedness (a word is *C-balanced* if and only if its imbalance is lower than C), balancedness (originally, a word is *balanced* if and only if its imbalance is lower than 1; in recent papers, balanced words tend to denote words with finite imbalance). In [Ada03], Adamczewski introduced the *balance function* of an infinite word w: $B_w(n) = \max_{a \in \mathcal{A}} \max_{u,v \in \mathcal{F}_n(w)} ||u|_a - |v|_a|$. The imbalance of w is the smallest upper bound for this balance function.

For instance, the imbalance of the Tribonacci word is 2 (see [RSZ09] for a proof, but this fact was mentioned before). Because they were constructed as a generalization of Sturmian words from the combinatorial viewpoint, it was expected that Arnoux-Rauzy words would have bounded imbalance. This is not the case: Cassaigne, Ferenczi and Zamboni exhibited in [CFZ00] families of Arnoux-Rauzy words with arbitrary high imbalance, and even families of words with infinite imbalance (see also [And21] for an alternative construction).

One extend the notion of imbalance by considering factors instead of letters. For $u, v \in \mathcal{F}(w)$, denote by $|u|_v$ the number of occurrences of v in u, i.e., the number of indices $i \in \{0, ..., |u|-1\}$ such that $u[i]..u.[|v|-1+i] = v$ (for instance, $|131313|_{1313} = 2$.) The imbalance of w *on the factor* v is:

$$\text{imb}_v(w) = \sup_{n \in \mathbb{N}} \sup_{u,u' \in \mathcal{F}_n(w)} ||u|_v - |u'|_v|.$$

This notion appears in Section 4.3.

3 Natural coding of minimal rotations

3.1 A topological definition and its consequences

Let d be a positive integer. Recall that L is a lattice of $\mathbb{R}^d$ if there exist $e_1, ..., e_d \in \mathbb{R}^d$, linearly independent, such that $L = \mathbb{Z}e_1 + ... + \mathbb{Z}e_d$. We denote by $\mathbb{T}_L := \mathbb{R}^d/L$ the d-torus associated with the lattice L, and by p_L the projection map on the torus. The torus is endowed with the quotient topology (consisting of all sets with an open preimage under p_L), which makes p_L open and continuous.

A set $\Omega \subset \mathbb{R}^d$ is *L-simple* if the map $p_L : \Omega \to \mathbb{T}_L$ is one-to-one; Ω is a *fundamental domain* of L if the map $p_L : \Omega \to \mathbb{T}_L$ is one-to-one and onto. As soon as Ω is L-simple, we introduce the cover map $r_{\Omega,L}$ which maps each point in $p_L(\Omega)$ to its unique preimage in Ω. If the set Ω is open, the cover map $r_{\Omega,L}$ is open and continuous for the topology on $\mathbb{R}^d$. Remark: in [Rau84], L-simple sets are furthermore assumed to be bounded - in our work we will explicitly mention this assumption each time it is required.

Now, given $\alpha \in \mathbb{R}^d$, the rotation of the torus $\mathbb{T}_L$ through the angle α is the map $R_{\alpha,L} : \mathbb{T}_L \to \mathbb{T}_L,\ x \mapsto x + p_L(\alpha)$ (possibly denoted by R_α if there is no ambiguity on the lattice). Following [Rau84], a pair (α, L) is said *minimal* if for all $\tilde{x} \in \mathbb{T}_L$, the sequence $(R^n_\alpha(\tilde{x}))_{n\in\mathbb{N}}$ is dense in $\mathbb{T}_L$ - or equivalently, if there exists one such $\tilde{x}$ in $\mathbb{T}_L$.

Definition 2. *A word $w_0 \in \mathcal{A}^{\mathbb{N}}$ is a* natural coding of a minimal rotation of the d-torus *if there exists a lattice $L \subset \mathbb{R}^d$ together with a vector $\alpha \in \mathbb{R}^d$ such that:*

- *(minimality) The pair (α, L) is minimal.*
- *(partition of a pseudo-fundamental domain) There exist Ω_1,...,Ω_{d+1} nonempty, open sets of $\mathbb{R}^d$ such that:*
  - *the sets Ω_1,...,Ω_{d+1} are pairwise disjoint;*
  - *the union set $\Omega = \cup_{i\in\{1,...,d+1\}}\Omega_i$ is L-simple;*
  - *the projection set $p_L(\Omega)$ is dense in the torus $\mathbb{T}_L$.*
- *(exchange of pieces) There exist $\alpha_1, ..., \alpha_{d+1} \in \mathbb{R}^d$ such that for all index $i \in \{1, ..., d+1\}$ and for all point $\tilde{x} \in p_L(\Omega_i) \cap R^{-1}_\alpha(p_L(\Omega))$, $r_{\Omega,L}(R_\alpha(\tilde{x})) = r_{\Omega,L}(\tilde{x}) + \alpha_i$.*
- *(a coding trajectory) There exists $\tilde{x}_0$ in $p_L(\Omega)$ such that, for all $n \in \mathbb{N}$, $R^n_\alpha(\tilde{x}_0) \in p_L(\Omega_{w_0[n]})$, where $w_0[n]$ denotes the (n+1)-th letter of w_0.*

  We set $x_0 = r_{\Omega,L}(\tilde{x}_0)$ and we say that $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$ are elements *of the natural coding w_0.*

Lemma 3. *If w_0 is a natural coding of a minimal rotation of the d-torus, then w is written with exactly $d + 1$ letters and is uniformly recurrent.*

Proof. By minimality of the rotation, the trajectory of the point $\tilde{x}_0$ is dense in the torus and, thus, visits each open set $p_L(\Omega_i)$ - so w_0 contains each letter i in $\{1, ..., d + 1\}$.

Let u be a factor of w_0. Then there exists a nonnegative integer n such that $S^n(w_0) \in [u]$. Observe that the set $\tilde{\Omega}_u := \cap_{l=0}^{|u|-1} R^{-l}_\alpha(p_L(\Omega_{u[l]}))$ is nonempty (it contains indeed the point $R^n_\alpha(\tilde{x}_0)$) and open (the Ω_i are open, the projection p_L is open and the rotation is continuous). Therefore, by minimality of the pair (α, L), we obtain a cover of the torus by a countable family of open sets: $\mathbb{T}_L = \cup_{n\in\mathbb{N}} R^{-n}_\alpha(\tilde{\Omega}_u)$, from which we extract, by compacity of $\mathbb{T}_L$, a finite cover. We conclude that there exists a nonnegative integer m such that $\mathbb{T}_L = \cup_{n=0}^{m} R^{-n}_\alpha(\tilde{\Omega}_u)$ and, finally, that w_0 is uniformly recurrent. □

At last, we say that a *subshift* is a *natural coding of a minimal rotation of the d-torus* if it is minimal and if one of its elements is a natural coding of a minimal rotation of the d-torus.

Hereafter, we denote by $\mathcal{I} = \{1, ..., d+1\}$ the alphabet.

Notation 4. *In this document, we shall work with a second lattice, called M. To avoid confusion, we will use the symbol $\tilde{.}$ (tilda) to refer to points or sets in the torus $\mathbb{T}_L$, whereas the symbol $\bar{.}$ (bar) will be devoted to elements in $\mathbb{T}_M$- the absence of symbol referring by default to the covering space $\mathbb{R}^d$. From now on, we denote $\tilde{\Omega} = p_L(\Omega)$ and $\tilde{\Omega}_i = p_L(\Omega_i)$.*

Example 5. *A Sturmian subshift with slope α is a natural coding of the minimal rotation of the torus $\mathbb{R}/\mathbb{Z}$ through the angle α. One can take the pseudo-fundamental domain $\Omega =]0, 1[$ together with the partition $\Omega_1 =]0, 1-\alpha[$, $\Omega_2 =]1-\alpha, 1[$.*

The Tribonacci subshift is a natural coding of the minimal rotation of the torus $\mathbb{R}^2/\mathbb{Z}^2$ through the angle (ζ, ζ^2), where ζ is is the unique real root of the polynomial $x^3 + x^2 + x - 1$. Furthermore, the Rauzy fractal gives a pseudo-fundamental domain for which the pieces of the partition are furthermore bounded and simply connected (see [Rau82]).

Given a natural coding w_0 with elements $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$, we introduce the numbering function $\nu : \tilde{\Omega} \to \mathcal{I}$, which maps all elements of $\tilde{\Omega}_i$ to the letter i; and we consider the coding function f given by $f(x) = (\nu(R_\alpha^n(p_L(x))))_{n\in\mathbb{N}}$ which makes sense each time the trajectory of $p_L(x)$ for the rotation action is included in $\tilde{\Omega}$. Let $\mathcal{D}$ denote the maximal subset of Ω on which the coding function f is defined. The covering rotation $T = r_{\Omega,L} \circ R_\alpha \circ p_L$ is well-defined on $\mathcal{D}$ and satisfies $T(\mathcal{D}) \subset \mathcal{D}$. Following Notation 4, we denote $\tilde{\mathcal{D}} = p_L(\mathcal{D})$.

Lemma 6. *The trajectory of x_0 under the action of T is included in $\mathcal{D}$ and dense in Ω.*

Proof. By definition of natural coding, the trajectory of $\tilde{x}_0$ under R_α is included in $\tilde{\mathcal{D}}$ (we even know that $f(x_0) = w_0$). The rotation being minimal, the trajectory of $\tilde{x}_0$ is dense in the torus $\mathbb{T}_L$ and in particular in $\tilde{\Omega}$. By continuity of the cover map $r_{\Omega,L}$ (Ω is open), we conclude that this property is preserved in the cover space. □

Proposition 7. *The coding function f is continuous for the induced topology on $\mathcal{D}$. Furthermore, the diagram below is commutative:*

$$\begin{array}{ccc} \mathcal{D} & \xrightarrow{\quad T \quad} & \mathcal{D} \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ \mathcal{I}^{\mathbb{N}} & \xrightarrow[\quad S \quad]{} & \mathcal{I}^{\mathbb{N}}, \end{array}$$

and the image set $f(\mathcal{D})$ is included in X_0, the subshift generated by the word w_0.

Proof. First, observe that $f(x)$ belongs to the cylinder $[i_0...i_{n-1}]$ if and only if $x \in \mathcal{D}$ and for all k in $\{0, ..., n-1\}$, $R_\alpha^k(p_L(x)) \in \tilde{\Omega}_{i_k}$; if and only if $x \in \mathcal{D} \cap \cap_{k=0}^{n-1} p_L^{-1}(R_\alpha^{-k}(\tilde{\Omega}_{i_k}))$, which is open for the induced topology on $\mathcal{D}$ - hence the continuity of f.

Secondly, the diagram is commutative by definition of f.

Thirdly, let $w \in f(\mathcal{D})$ and $x \in \mathcal{D} \subset \Omega$ be one of its preimages. By Lemma 6, there exists an extracted sequence $(T^{n_k}(x_0))_k$ in $\mathcal{D}^{\mathbb{N}}$ that tends to x. But f being continuous, the image sequence, which is $(S^{n_k}(w_0))_k$, tends to $w = f(x)$ - meaning that the word $w \in \mathcal{I}^{\mathbb{N}}$ actually belongs to X_0; we conclude that $f(\mathcal{D}) \subset X_0$. □

We will prove in Proposition 18 that the coding function f is one-to-one.

Lemma 8. *For all $x \in \mathcal{D}$, we have $T(x) = x + \alpha_i$, with $i = \nu(p_L(x))$.*

Proof. This is an immediate consequence of Item 2 (exchange of pieces) in Definition 2 (natural coding). □

3.2 Borders assignment

In this subsection, we show that, under the axiom of choice, it is possible to wisely assign borders to the pieces Ω_i, in order to complete the L-simple set Ω into a fundamental domain Ω' and enlarge the remarkable property of exchange of pieces, while keeping (under a weak form) the continuity of the coding function f.

Proposition 9. *Let L be a lattice of $\mathbb{R}^d$, and $\Omega_1, ..., \Omega_{d+1}$ nonempty, open, pairwise disjoint sets, such that moreover $\Omega = \cup_{i\in\mathcal{I}}\Omega_i$, where $\mathcal{I} = \{1, ..., d+1\}$, is L-simple (H1). Let $\alpha \in \mathbb{R}^d$ be such that (α, L) is minimal (H2). Assume there exists $x_0 \in \Omega$ such that for all nonnegative integer n, $R^n_\alpha(p_L(x_0)) \in \Omega$(H3), and denote by $(i_n)_n \in \mathcal{I}^{\mathbb{N}}$ the unique sequence satisfying: for all $n \in \mathbb{N}$, $R^n_\alpha(p_L(x_0)) \in \Omega_{i_n}$. At last, assume that there exist $\alpha_1, ..., \alpha_{d+1} \in \mathbb{R}^d$ such that for all nonnegative integer n, $T^{n+1}(x_0) = T^n(x_0) + \alpha_{i_n}$, where $T = r_{\Omega,L} \circ R_\alpha \circ p_L$ (H4).*

Then, under the axiom of choice, there exist $\Omega'_1, ..., \Omega'_{d+1} \subset \mathbb{R}^{d+1}$ such that:

- *(C1) for all $i \in \mathcal{I}$, $\Omega_i \subset \Omega'_i$;*
- *(C2) the union set $\Omega' = \cup_{i\in\mathcal{I}}\Omega'_i$ is a fundamental domain of L;*
- *(C3) the sets Ω'_i are pairwise disjoint.*

Furthermore, if T' denotes the covered rotation $T' = r_{\Omega',L} \circ R_\alpha \circ p_L$, then:

- *(C4) for all $x \in \Omega'_i$, $T'(x) = x + \alpha_i$;*
- *(C5) for all $x \in \Omega'$, for all $q \in \mathbb{N}$, there exists τ an extraction (i.e. an increasing map from $\mathbb{N}$ to $\mathbb{N}$) such that: (i) the sequence $(T^{\tau(m)}(x_0))_{m\in\mathbb{N}}$ converges to x; (ii) for all $n \in \{0, ..., q\}$, for all nonnegative integer m, $T^{\tau(m)+n}(x_0) \in \Omega_{\iota_n}$, where ι_n is given by $T'^n(x) \in \Omega'_{\iota_n}$.*

Proof. **General idea.** The proof consists of a lifting process, based on the axiom of choice. Initially, the sets $\Omega'_1, ..., \Omega'_{d+1}$ are empty. We browse each orbit for the action of the rotation R_α to the future and back to the past, from a well-chosen point, in order to assign to each visited point of the torus $\mathbb{T}_L$ a covering point in $\mathbb{R}^d$, that we furthermore stow in one of the $d+1$ sets $\Omega'_1, ..., \Omega'_{d+1}$. The pair (α, L) being minimal (H2), each point of the torus is visited exactly once by this process, and the sets $\Omega'_1, ..., \Omega'_{d+1} \subset \mathbb{R}^d$ form a partition of a fundamental domain of the torus.

Following Notation 4, we denote $\tilde{x}_0 = p_L(x_0)$, $\tilde{\Omega}_i = p_L(\Omega_i)$ for all $i \in \mathcal{I}$, and $\tilde{\Omega} = p_L(\Omega)$.

Method to lift one orbit. Let $(\tilde{y}_n)_{n\in\mathbb{Z}}$ be an orbit for the action of R_α on the torus. The pair (α, L) being minimal (H2), the pair $(-\alpha, L)$ is minimal as well, and there exists an increasing sequence of nonnegative indices $(n_k)_{k\in\mathbb{N}}$ such that for all $k \in \mathbb{N}$, $\tilde{y}_{-n_k}$ belongs to the nonempty open set $\tilde{\Omega}$ (H1). Without loss of generality, we assume $n_0 = 0$.

We now intend to construct, by a diagonal process, a lifted sequence $(y_n)_{n\in\mathbb{Z}}$ for the orbit $(\tilde{y}_n)_{n\in\mathbb{Z}}$, together with a numbering sequence $(j_n)_{n\in\mathbb{Z}}$, such that for all $n \in \mathbb{Z}$:

1. $p_L(y_n) = \tilde{y}_n$;
2. if moreover $\tilde{y}_n \in \tilde{\Omega}$, then $y_n = r_{\Omega,L}(\tilde{y}_n)$;

3. $y_{n+1} = y_n + \alpha_{j_n}$;
4. for all $q \in \mathbb{N}$, there exists an extraction τ such that $T^{\tau(m)}(x_0) \to_{m\to\infty} y_n$ and for all nonnegative integer m, $i_{\tau(m)}...i_{\tau(m)+q} = j_n...j_{n+q}$.

The lifted orbit $(y_n)_{n\in\mathbb{Z}}$ and its symbolic trajectory $(j_n)_{n\in\mathbb{Z}}$ will be obtained as the limit sequences, when k tends to infinity, of the finite sequences $(y_n)_{n\in\{-n_k,...,n_k\}}$ and $(j_n)_{n\in\{-n_k,...,n_k\}}$: we now give the details of the construction.

First, since $\tilde{y}_0 \in \tilde{\Omega}$, there exists a unique $i \in \mathcal{I}$ such that $\tilde{y}_0 \in \tilde{\Omega}_i$. We set $y_0^0 = r_{\Omega,L}(\tilde{y}_0)$ and $j_0^0 = i$. By minimality of (α, L) (H2), and since $\Omega_{j_0^0}$ is open (H1), there exists an extraction σ_0 such that the sequence $(T^{\sigma_0(m)}(x_0))_{m\in\mathbb{N}}$ is included in $\Omega_{j_0^0}$ and tends to y_0^0.

Now, let $k \in \mathbb{N}$, and assume there exist an extraction σ_k together with a finite word $j^k_{-n_k}...j^k_{n_k} \in \mathcal{I}^{2n_k+1}$, such that $T^{\sigma_k(m)}(x_0) \to_{m\to\infty} r_{\Omega,L}(\tilde{y}_{-n_k})$ and for all nonnegative integer m, $i_{\sigma_k(m)}...i_{\sigma_k(m)+2n_k} = j^k_{-n_k}...j^k_{n_k}$. We want to construct an extraction σ_{k+1} together with a finite word $j^{k+1}_{-n_{k+1}}...j^{k+1}_{n_{k+1}} \in \mathcal{I}^{2n_{k+1}+1}$, such that:

$$\begin{cases} T^{\sigma_{k+1}(m)}(x_0) \to_{m\to\infty} r_{\Omega,L}(\tilde{y}_{-n_{k+1}}), \\ \text{for all } m \in \mathbb{N}, \quad i_{\sigma_{k+1}(m)}...i_{\sigma_{k+1}(m)+2n_{k+1}} = j^{k+1}_{-n_{k+1}}...j^{k+1}_{n_{k+1}}. \end{cases}$$

Denote $l = n_{k+1} - n_k \in \mathbb{N}^*$, and let $m_0 \in \mathbb{N}$ be such that $\sigma_k(m_0) \geq l$. Denote $\gamma : m \mapsto m + m_0$. The sequence $(T^{\sigma_k\circ\gamma(m)-l}(x_0))_{m\in\mathbb{N}}$ is the image, by the continuous function $T^{-l} = r_{\Omega,L} \circ R_\alpha^{-l} \circ p_L$ (Ω is open by (H1)), of the convergent sequence $(T^{\sigma_k\circ\gamma(m)}(x_0))_{m\in\mathbb{N}}$; it thus admits a limit that we denote by $y^{k+1}_{-n_{k+1}}$. Furthermore, the possible values of the sequence $(i_{\sigma_k\circ\gamma(m)-l}...i_{\sigma_k\circ\gamma(m)-l+2n_{k+1}})_{m\in\mathbb{N}}$ belong to the finite set $\mathcal{I}^{2n_{k+1}+1}$, so there exists $j^{k+1}_{-n_{k+1}}...j^{k+1}_{n_{k+1}} \in \mathcal{I}^{2n_{k+1}+1}$ together with an extraction υ such that for all nonnegative integer m, $i_{\sigma_{k+1}(m)}...i_{\sigma_{k+1}(m)+2n_{k+1}} = j^{k+1}_{-n_{k+1}}...j^{k+1}_{n_{k+1}}$, where $\sigma_{k+1}(m) = \sigma_k \circ \gamma \circ \upsilon(m) - l$, for all nonnegative integer m. Finally, for $n \in \{-n_{k+1}, ..., n_{k+1}\}$, we define $y_n^{k+1} := \lim_{m\to\infty} T^{\sigma_{k+1}(m)+n+n_{k+1}}(x_0)$.

Observe that the sequence $(T^{\sigma_{k+1}(m)+l}(x_0))_m$ is a subsequence of $(T^{\sigma_k(m)}(x_0))_m$. Consequently, for all $n \in \{-n_k, ..., n_k\}$:

$$\begin{cases} y_n^{k+1} = \lim_m \ T^{\sigma_{k+1}(m)+l+n+n_k}(x_0) = \lim_m \ T^{\sigma_k(m)+n+n_k}(x_0) = y_n^k; \\ j_n^{k+1} = j_n^k. \end{cases}$$

Therefore, we can construct $(y_n)_{n\in\mathbb{Z}}$ and $(j_n)_{n\in\mathbb{Z}}$ as the biinfinite limits of the finite words $(y_n^k)_{n\in\{-n_k,...,n_k\}}$ and $(j_n^k)_{n\in\{-n_k,...,n_k\}}$, for $k \in \mathbb{N}$.

Properties of the lifted orbit.

Now we formally check that the lifted orbit $(y_n)_{n\in\mathbb{Z}}$ satisfies the properties (1), (2), (3) and (4).

Lemma 10. *For all $n \in \mathbb{Z}$, $p_L(y_n) = \tilde{y}_n$; if moreover $\tilde{y}_n \in \tilde{\Omega}$, then $y_n = r_{\Omega,L}(\tilde{y}_n)$.*

Proof. Let $n \in \mathbb{Z}$ and $k \in \mathbb{N}$ be such that $-n_k \leq n$. By continuity of p_L and R_α, we have:

$$\begin{aligned} p_L(y_n) &= \lim_m \ p_L \circ T^{\sigma_k(m)+n+n_k}(x_0) \\ &= \lim_m \ R_\alpha^{n+n_k} \circ p_L \circ T^{\sigma_k(m)}(x_0) \\ &= R_\alpha^{n+n_k}(\tilde{y}_{-n_k}) \\ &= \tilde{y}_n. \end{aligned}$$

Furthermore, if $\tilde{y}_n \in \tilde{\Omega}$, then there exists $\lambda \in L$ such that $y_n = r_{\Omega,L}(\tilde{y}_n) + \lambda$. Since the sequence $(T^{\sigma_k(m)+n+n_k}(x_0))_{m\in\mathbb{N}}$ of elements in $\Omega^{\mathbb{N}}$ tends to y_n, which belongs to the open set $\Omega + \lambda$, we must have $\Omega \cap \Omega + \lambda \neq \emptyset$. By L-simplicity of Ω (H1), this implies $\lambda = 0$, hence $y_n = r_{\Omega,L}(\tilde{y}_n)$. □

Lemma 11. *For all $n \in \mathbb{Z}$, $y_{n+1} = y_n + \alpha_{j_n}$.*

Proof. Let $n \in \mathbb{Z}$ and $k \in \mathbb{N}$ be such that $-n_k \leq n < n_k$. For all nonnegative integer m, $i_{\sigma_k(m)+n+n_k} = j_n$, hence $T^{\sigma_k(m)+n+n_k+1}(x_0) = T^{\sigma_k(m)+n+n_k}(x_0) + \alpha_{j_n}$ (H4). Taking the limit when $m \to \infty$ on both sides gives $y_{n+1} = y_n + \alpha_{j_n}$. □

Lemma 12. *For all $n \in \mathbb{Z}$ and for all $q \in \mathbb{N}$, there exists an extraction τ such that $T^{\tau(m)}(x_0) \to_{m\to\infty} y_n$ and for all nonnegative integer m, $i_{\tau(m)}...i_{\tau(m)+q} = j_n...j_{n+q}$.*

Proof. Let $n \in \mathbb{Z}$ and $q \in \mathbb{N}$. We choose $k \in \mathbb{N}$ such that $-n_k \leq n$ and $q \leq n_k - n$. By construction of σ_k, the function $\tau : m \mapsto \sigma_k(m) + n + n_k$ suits. □

Construction of a fundamental domain with a "good" partition.

Thanks to the axiom of choice, we run this process for each orbit in the action of the rotation R_α on the torus $\mathbb{T}_L$. We denote $\Omega' \subset \mathbb{R}^d$ the set of all the lifted points we obtain. Since the orbits form a partition of the torus, and by minimality of (α, L) (H2), each point of the torus is lifted exactly once, meaning that Ω' is a fundamental domain of the torus (C2). Hereafter we denote by $r_{\Omega',L}(\tilde{y})$ the covering point of $\tilde{y}$ given by the process.

Again, because each point in the torus is visited exactly once, we can define the numbering map $\nu' : \mathbb{T}_L \to \mathcal{I}$, which maps the n-th point of a given orbit $(\tilde{y}_n)_{n\in\mathbb{Z}}$ (with respect to the indexation used for the lifting process) to the n-th term of the associated numbering sequence $(j_n)_{n\in\mathbb{Z}}$. For all $i \in \mathcal{I}$, we set $\Omega'_i = \{r_{\Omega',L}(\tilde{y}) \mid \tilde{y} \in \mathbb{T}_L \text{ s.t. } \nu'(\tilde{y}) = i\}$. The sets Ω'_1,...,Ω'_{d+1} form a partition of Ω' (C3).

At last, Lemma 10 implies that for all $\tilde{y} \in p_L(\Omega)$, $r_{\Omega',L}(\tilde{y}) = r_{\Omega,L}(\tilde{y})$, hence the inclusion $\Omega_i \subset \Omega'_i$ for all i in $\mathcal{I}$. Lemmas 11 and 12 respectively ensure (C4) and (C5). □

Definition 13. *Let w_0 be a natural coding of a minimal rotation of the d-torus with elements $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$. We say that $(\Omega' : \Omega'_1, ..., \Omega'_{d+1})$ is a* borders assignment *of w if:*

1. *for all i in $\mathcal{I} = \{1, ..., d+1\}$, Ω_i is included in Ω'_i;*
2. *the sets $\Omega'_1, ..., \Omega'_{d+1}$ forms a partition of Ω';*
3. *the set Ω' is a fundamental domain of L;*
4. *for all i in $\mathcal{I}$, for all x in Ω'_i, $T'(x) = x + \alpha_i$, where T' denotes the covered map of the rotation to the fundamental domain Ω': $T' = r_{\Omega',L} \circ R_\alpha \circ p_L$;*
5. *for all $x \in \Omega'$, for all $q \in \mathbb{N}$, there exists an extraction τ such that: (i) $T^{\tau(m)}(x_0) \to_{m\to\infty} x$; (ii) for all $n \in \{0, ..., q\}$, for all nonnegative integer m, $T^{\tau(m)+n}(x_0) \in \Omega_{\iota_n}$, where ι_n is given by $T'^n(x) \in \Omega'_{\iota_n}$.*

Corollary 14. *If $(\Omega' : \Omega'_1, ..., \Omega'_{d+1})$ is a borders assignment of a natural coding with elements $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$, then for all $i \in \mathcal{I} := \{1, ..., d+1\}$, the set Ω_i is dense in Ω'_i.*

Proof. Let $i \in \mathcal{I}$ and $x \in \Omega'_i$. Denote by τ the extraction given by Definition 13 for x and $q = 0$. Then the sequence $(T^{\tau(n)}(x_0))_{n\in\mathbb{N}}$ belongs to $\Omega_i^{\mathbb{N}}$ and converges to x - hence the density of Ω_i in Ω'_i. □

Corollary 15 (Immediate consequence of Proposition 9)**.** *Under the axiom of choice, a natural coding of a minimal rotation of the d-torus admits borders assignments.*

Given a natural coding of a minimal rotation w_0 endowed with a borders assignment $(\Omega' : \Omega'_1, ..., \Omega'_{d+1})$, we extend the numbering and coding functions ν and f into $\nu' : \tilde{\Omega}' \mapsto \mathcal{I}$ and $f' : \Omega' \mapsto \mathcal{I}^{\mathbb{N}}$: for all $\tilde{x} \in \tilde{\Omega}'_i$, we set $\nu'(\tilde{x}) = i$; for all x in Ω', $f'(x) = (\nu'(R^n_\alpha(p_L(x))))_{n\in\mathbb{N}}$. The extended coding function is defined on the whole fundamental domain Ω' and coincides with f wherever f is defined, i.e., on the subset $\mathcal{D}$.

The following lemma, which is an immediate consequence of Definition 13, is the keystone of the paper.

Lemma 16 (Weak sequential continuity)**.** *For all $x \in \Omega'$, there exists a sequence $(y_n)_n \in \mathcal{D}^{\mathbb{N}}$ such that $y_n \to_{n\to\infty} x$ and $f'(y_n) = f(y_n) \to_{n\to\infty} f'(x)$.*

Proof. Let $x \in \Omega'$. For $n \in \mathbb{N}$, we set $y_n = T^{\tau_n(n)}(x_0)$, where τ_n is the extraction given for $q = n$ in Definition 13. The sequence $(y_n)_{n\in\mathbb{N}}$ belongs to $\mathcal{D}^{\mathbb{N}}$, converges to x and satisfies, for all nonnegative integer n, $f(y_n)[0...n] = f'(x)[0...n]$. □

This implies in particular that the image set of the extended coding function f' belongs to the subshift (which is a close set) generated by w_0: $f'(\Omega') \subset X_0$.

We finally show that the extended coding function f' is one-to-one. This results of the minimality of the covered dynamical system (Ω', T').

Lemma 17. *The nonnegative orbit of any x in Ω' under the action of the extended covered rotation T' is dense in Ω'.*

Proof. Let $x, z \in \Omega'$ and $\varepsilon > 0$. By density of Ω in Ω' (Corollary 14), one can pick y in Ω at distance less than $\varepsilon/2$ from z. Consider an open ball $\mathcal{B}$ with center y and diameter less than $\varepsilon/2$ included in the open set Ω. The projected set $p_L(\mathcal{B})$ is still a nonempty open set; by minimality of the pair (α, L), there exists $n \in \mathbb{N}$ such that $R^n_{\alpha,L}(p_L(x)) \in p_L(\mathcal{B})$. Back to the covering space, we have that ${T'}^n(x) \in \mathcal{B}$; the point ${T'}^n(x)$ is thus at distance less than $\varepsilon/2$ from y and less than ε from z. □

Proposition 18. *The extended coding function $f' : \Omega' \mapsto X_0$ is one-to-one.*

Proof. By contradiction, consider $x \neq y \in \Omega'$ such that $f'(x) = f'(y)$. Because of Item (4) in Definition 13, an easy induction argument shows that ${T'}^n(y) = {T'}^n(x) + y - x$ for any nonnegative integer n. Taking the closure set of their nonnegative orbit, we obtain that $\overline{\{{T'}^n(y) | n \in \mathbb{N}\}} = \overline{\{{T'}^n(x) | n \in \mathbb{N}\}} + y - x$. Since $\overline{\{{T'}^n(z) | n \in \mathbb{N}\}} = \bar{\Omega}' = \bar{\Omega}$ for any z in Ω' (immediate consequence of Lemma 17), this implies that the set $\bar{\Omega}$ is invariant under the translation by the vector $y - x$.

Now, consider $\mathcal{B}_0 \subset \Omega$ an open ball with diameter less than $y - x$. By compacity of $\mathbb{T}_L$, there exists a positive integer n such that the intersection $R^n_{y-x,L}(p_L(\mathcal{B}_0)) \cap p_L(\mathcal{B}_0)$ is nonempty. Denote $\mathcal{B}_1 = \mathcal{B}_0 + n(y - x)$. On one hand, due to their small diameter, the balls $\mathcal{B}_0$ and $\mathcal{B}_1$ are disjoint. On another hand, the translated ball $\mathcal{B}_1$ is still included in $\overline{\Omega}$. Thus, the intersection $\mathcal{B}_1 \cap \Omega$ is dense in $\mathcal{B}_1$ and its projected set $p_L(\mathcal{B}_1 \cap \Omega)$ is dense in $p_L(\mathcal{B}_1)$. In particular, since $p_L(\mathcal{B}_1) \cap p_L(\mathcal{B}_0)$ is open and nonempty (indeed, $p_L(\mathcal{B}_1) = R^n_{y-x,L}(p_L(\mathcal{B}_0)))$ by definition of n the intersection $p_L(\mathcal{B}_1 \cap \Omega) \cap p_L(\mathcal{B}_0)$ is also nonempty. Given that $\mathcal{B}_0 \subset \Omega$ and that the balls $\mathcal{B}_0$ and $\mathcal{B}_1$ are disjoint, this nonemptyness is conflicting with the L-simplicity of Ω. □

3.3 The underlying group of a natural coding

Let w_0 be a natural coding of a minimal rotation of the d-torus with elements $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$. We introduce the *underlying group* of the natural coding w_0:

$$G = \sum_{i\in\mathcal{I}} \alpha_i \mathbb{Z}.$$

We now state two general lemmas about the group G, that will be useful in Sections 4 and 6.

Lemma 19. *The group G is a free abelian group of rank $d+1$, which is dense in $\mathbb{R}^d$. The family $(\alpha_1, ..., \alpha_{d+1})$ forms a basis of G.*

Proof. We first show that the group G is dense in $\mathbb{R}^d$. By minimality of the pair (α, L), the orbit $(R^n_{\alpha,L}(p_L(x_0)))_{n\in\mathbb{N}}$, which is included in $p_L(\Omega)$, is dense in $\mathbb{T}_L$. The set Ω being open and L-simple, the covered map $r_{\Omega,L}$ is well-defined and continuous, and thus, the covered orbit $(r_{\Omega,L} \circ R^n_{\alpha,L} \circ p_L(x_0))_{n\in\mathbb{N}} \subset G + x_0$ is dense in Ω. Finally, the group G is dense in the nonempty open set $\Omega - x_0 \subset \mathbb{R}^d$, and thus, in the whole space $\mathbb{R}^d$.

Now, we show that the family $(\alpha_1, ..., \alpha_{d+1})$ is free over $\mathbb{Z}$, which proves that G is a free abelian group of rank $d+1$, with basis $(\alpha_1, ..., \alpha_{d+1})$. By contradiction, assume that there exist $n_1, ..., n_{d+1} \in \mathbb{Z}$, non simultaneously equal to zero, and such that $\sum_{i\in\mathcal{I}} n_i\alpha_i = 0$. Without loss of generality, assume that $n_{d+1} \in \mathbb{N}^*$; so we have $n_{d+1}\alpha_{d+1} = -\sum_{i=1}^{d} n_i\alpha_i$. Now, denote by $\mathcal{V}$ the vectorial space over $\mathbb{R}$ generated by the vectors $\alpha_1, ..., \alpha_d$. If $\mathcal{V}$ is a strict subspace of $\mathbb{R}^d$, then we can find a vector $e \in \mathbb{R}^d$ such that the distance between e and the subspace $\mathcal{V}$ is greater than 1 - which is impossible since G is included in $\mathcal{V}$ and dense in $\mathbb{R}^d$. Therefore, $\alpha_1, ..., \alpha_d$ form a basis of $\mathbb{R}^d$ and $N = \sum_{i=1}^{d} \alpha_i\mathbb{Z}$ is a lattice of $\mathbb{R}^d$. We now show that

$$G = \bigcup_{r\in\{0,...,n_{d+1}-1\}} N + r\alpha_{d+1}.$$

Let $g = \sum_{i\in\mathcal{I}} m_i\alpha_i$ be an element of G. Denote respectively by q and r the quotient and the rest of m_{d+1} in the euclidean division by n_{d+1}. Then we have $g = \sum_{i=1}^{d}(m_i - qn_i)\alpha_i + r\alpha_{d+1}$. We conclude that G is included in the union set $\cup_{r\in\{0,...,n_{d+1}-1\}} N + r\alpha_{d+1}$; since the converse inclusion is trivially true, we have the equality. Then, as the finite union of discrete sets, G is discrete - a contradiction. Finally, the vectors $\alpha_1, ..., \alpha_{d+1}$ form a basis of the group G. □

Lemma 20. *We have $G = L + \alpha\mathbb{Z}$.*

Proof. Since $\alpha_i = \alpha \mod L$ for all $i \in \mathcal{I}$, we immediately have $G \subset L + \alpha\mathbb{Z}$. Conversely, let $x \in L + \alpha\mathbb{Z} + x_0$. We are going to show that $x \in G + x_0$ - which will end the proof. By density of G and openness of Ω, there exists $g \in G$ such that $x - g \in \Omega$. Since $G \subset L + \alpha\mathbb{Z}$, we deduce that $p_L(x - g) = R^l_{\alpha,L}(\tilde{x}_0)$ for a certain $l \in \mathbb{Z}$. But then, there exists $l_1, ..., l_{d+1} \in \mathbb{Z}$ such that $x - g = x_0 + \sum l_i\alpha_i$; hence $x \in x_0 + G$. □

3.4 Discussion on the definition of natural coding

The notion of natural coding of rotation, sometimes better called *natural coding of translation of the torus*, goes back to the works of Morse and Hedlund [MH40], and to the seed paper of Rauzy [Rau82] (study of the Tribonacci word) in dimension 2. Nonetheless, the terminology appears later (for instance in [CFZ00] and [Fog02]). As far as we know, the terminology was introduced, through not written, by Rauzy [Arn20].

Roughly speaking, a "natural" coding of rotation denotes a word obtained as the coding trajectory of a point of the torus, under the action of a rotation, with respect to a remarkable partition that can be covered such that the induced rotation on the associated fundamental domain coincides, on each covered piece, with a translation. Of course, we are interested in partitions with as few pieces as possible; moreover we would appreciate coding words with minimal complexity (the complexity of an infinite word w is the function which maps each nonnegative integer n to the number of factors of length n in w). The study of dimension 1, and the results obtained in dimension 2 (for instance [Rau82], [AI01], [BB12]) lead us to hope for a generic coding strategy with classes of words of complexity $dn+1$ (see also the recent papers [BST20] and [Fog20]).

We start by discussing the definition proposed in the article [CFZ00] (which inspired the present work), where no assumption is made on the topological nature of the partition. This definition is still used, under a weaker form (pieces are assumed disjoint up to measure 0) in [BST19].

Definition 21. *([CFZ00] No topological assumption.) Let L be a lattice of $\mathbb{R}^d$. A word $w \in \{1,...,d+1\}^{\mathbb{N}}$ is a* natural coding "with no topological assumption" *of the rotation $R_{\alpha,L}$ if there exist a fundamental domain Ω of $\mathbb{T}_L$, together with a partition $\Omega = \Omega_1 \cup ... \cup \Omega_{d+1}$, such that on each piece Ω_i the covered rotation coincides with a translation by a vector α_i; and the sequence w is the symbolic coding of the orbit of a point $x \in \Omega$ with respect to the partition in Ω_i.*

This definition is not restrictive enough, as illustrated by the following proposition.

Proposition 22. *Under the axiom of choice, for any lattice $L \subset \mathbb{R}^d$ and any $\alpha \in \mathbb{R}^d$ such that (α, L) is minimal, any word in $\{1,...,d+1\}^{\mathbb{N}}$ is a natural coding "with no topological assumption" of the rotation $R_{\alpha,L}$.*

Proof: a stupid Cantor example. Consider a lattice $L \subset \mathbb{R}^d$, and $\alpha \in \mathbb{R}^d$ such that (α, L) is minimal. Let w be a word in $\{1,...,d+1\}^{\mathbb{N}}$. Thanks to the axiom of choice, we cover each orbit $(\tilde{y}_n)_{n\in\mathbb{Z}}$ for the action of R_α on $\mathbb{T}_L$ as follows. We choose y_0 in $p_L^{-1}(\tilde{y}_0)$ and set, for any integer n, $y_n = y_0 + n\alpha$. We thus have $p_L(y_n) = p_L(y_0) + np_L(\alpha) = R_\alpha^n(\tilde{y}_0) = \tilde{y}_n$. Furthermore, we put the point y_n into the set $\Omega_{w[n]}$ if $n \geq 0$, or Ω_1 otherwise. By minimality of (α, L), each point of the torus is visited exactly once by this process, and the sets $\Omega_1, ..., \Omega_{d+1}$ form a partition of a fundamental domain of L. At last, by construction, the covered rotation coincides with the translation by the vector α on each set Ω_i, and the point indexed by 0 on each orbit admits w as symbolic coding. □

We thus have to restrict what we accept for the partition. As evidenced by Proposition 33 further, natural codings are made to preserve rotations while inducing on pieces, so the first property should ensure that these inductions are well-defined: nonempty interior for a topological study. This is what Berthé, Steiner and Thuswaldner require in [BST20]. We state here their definition with our notations.

Definition 23. *([BST20] Topological and metric assumptions, eluding borders). Let $\alpha \in \mathbb{R}^d$ be such that $(\alpha, \mathbb{Z}^d)$ is minimal. A* measurable fundamental domain *of $\mathbb{R}^d/\mathbb{Z}^d$ is a set $\Omega \subset \mathbb{R}^d$ with Lebesgue measure 1 that satisfies $\Omega + \mathbb{Z}^d = \mathbb{R}^d$. A collection $\{\Omega_1, ..., \Omega_h\}$ is said to be a* natural measurable partition *of Ω with respect to $R_{\alpha,\mathbb{Z}^d}$ if the sets Ω_i are measurable, they are the closure of their interior and zero measure boundaries, $\cup_{i=1}^h \Omega_i = \Omega$, the (Lebesgue) measure of $\Omega_i \cap \Omega_j$ is 0 for all $i \neq j$, and moreover there exist vectors $\alpha_1, ..., \alpha_h$ in $\mathbb{R}^d$ such that $\alpha_i + \Omega_i \subset \Omega$ with $\alpha_i \equiv \alpha$* mod *$\mathbb{Z}^d$, $1 \leq i \leq h$. This allows to define a map T (which depends on the partition) as an exchange of domains defined a.e. on Ω as $T(x) = x + \alpha_i$ whenever $x \in \mathring{\Omega}_i$.*

A sequence $(i_n)_{n\in\mathbb{N}} \in \{1,...,h\}^{\mathbb{N}}$ is said to be a natural coding *of $(\mathbb{R}^d/\mathbb{Z}^d, R_\alpha)$ w.r.t. the natural measurable partition $\Omega_1, ..., \Omega_h$ if there exists $x \in \Omega$ such that $(i_n)_{n\in\mathbb{N}}$ codes the orbit of x under the action of T , i.e. $T^n(x) \in \Omega_{i_n}$ for all $n \in \mathbb{N}$.*

By introducing objects up to measure 0, they manage to elude the question of borders, for which several arbitrary choices are admissible. However, this definition leads to induce the covered rotation on the interior of the pieces (where it is defined) instead of on the whole pieces. In our mind, this induction does not behave as well as it could be, as evidenced in Example 24.

Example 24. *For d=1, consider $L = \mathbb{Z}$ and α an irrational number in $[0, 1/2]$. We introduce $\Omega = \Omega_0 \cup \Omega_1$, where $\Omega_0 = [0, 1-\alpha]$ and $\Omega_1 = [1-\alpha, 1]$, which turns to be a natural partition according to Definition 23. Let $T_{0,ind}$ denote the first return map to $\mathring{\Omega}_0$ of the covered rotation T. We have: $T_{0,ind}(x) = x+\alpha$ if $x \in A_0 =]0, 1-2\alpha[$ and $T_{0,ind}(x) = x+2\alpha-1$ if $x \in A_1 =]1-2\alpha, 1-\alpha[$. Hence, for all $x \in A_0 \cup A_1$, $T_{0,ind}(x)$ coincides with the rotation through the angle α modulo $(1-\alpha)$. But looking at the remaining point $x = 1-2\alpha$, we have $T_{0,ind}(x) = x + 3\alpha - 1 \not\equiv \alpha \mod (1-\alpha)$. In other words, the induced map of T on $\mathring{\Omega}_0$, which is defined everywhere, is a rotation almost everywhere, but* not *everywhere.*

This is why, following [Rau82] and [Rau84], we chose to work with an exclusively topological background. This requires to carefully examine what happens with the borders. By doing so, on one hand, we guarantee that the assignment we choose for the borders enjoys the weak continuity property (Lemma 16) which turns to be central; on the other hand, we open the definition to sets whose borders have positive Lebesgue measure. As far as we know, it is an open question to determinate if the (S-adic) Rauzy fractal of all Arnoux-Rauzy words (which are good candidates for coding of rotation of the 2-torus) have borders of measure zero.

Finally, the conditions we ask for the pieces Ω_i (to be open, and they union set to be dense) are inherited from [Rau82] (see the definition of 'morcellement'). They are all needed in our work. Note that they appear in recent articles as well ([BST20] and [Fog20] for instance), in addition to the usual metric assumptions. At last, let us highlight that we did not assume the pieces to be bounded. Indeed, this assumption is required only when resorting to Rauzy's theorem on bounded remainder sets (Theorem B further). It would be of high interest to know (1) what remains of Rauzy's theorem if we remove this assumption; (2) if a word with infinite imbalance could be a natural coding of a minimal rotation with an unbounded pseudo-fundamental domain.

4 Stability under induction

4.1 Main result for induction

Definition 25. *[Dur98] A finite word u is a* return word *to the factor v in the recurrent word w if $u = w[i]...w[j-1]$, where $i, j \in \mathbb{N}$ are two consecutive occurrences of v in w.*

Lemma 26. *[Dur98] Let w be a uniformly recurrent word and $\mathcal{U}$ the set of return words to the factor v in w. Then $\mathcal{U}$ is finite. Furthermore, if w' is an element of the subshift generated by w, then the set of return words to the letter v in the word w' is again $\mathcal{U}$. If furthermore w' starts with the factor v, then it can be written in a unique way as a concatenation of elements in $\mathcal{U}$.*

Definition 27. *[Dur98] Let w be a uniformly recurrent word and v one of its factors. Denote by $\mathcal{U}$ the set of return words to v in w, that we enumerate: $u_1, ..., u_n$. Let l denote the index of the first occurrence of v in w. The* derivated word of w to v, *with respect to the chosen numeration, is the unique word $D_v(w)$ in $\{1, ..., n\}^{\mathbb{N}}$ satisfying: $\sigma(D_v(w)) = S^l(w)$, where σ is the substitution that maps k to the word u_k, for all $k \in \{1, ..., n\}$.*

Remark 28. *The derivated word of w to a is unique up to the choice made for the numeration. Whenever this choice has no significance, we will talk about* the *derivated word $D_v(w)$. However, it sometimes happens that the choice of the numeration is of interest: see for instance the case of primitive C-adic word in [Section 5, Example 51].*

We know from Lemma 3 that a natural coding of a minimal rotation of the torus is uniformly recurrent. We deduce that if v is one of its factors, then it admits a finite number of return words to v. The following theorem states that, whenever this number is $d+1$, i.e., coincides with the size of the partition (or equivalently, the size of the alphabet), then its derivated word to v is still a natural coding of a minimal rotation of the torus.

Theorem A. *Let w_0 be a natural coding of a minimal rotation of a d-dimensional torus, and denote by $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$ its elements, and by $(\Omega' : \Omega'_1, ..., \Omega'_{d+1})$ a borders assignment. Assume that v is a factor of w_0 which admits $d+1$ return words $u_1, ..., u_{d+1}$, and denote:*

$$\Omega'_v := \bigcap_{k=0}^{|v|-1} T'^{-k}(\Omega'_{v[k]}).$$

Then there exist a second lattice M together with an angle $\beta \in \mathbb{R}^d$ such that:

1. *the pair (β, M) is minimal;*
2. *the set Ω'_v is a fundamental domain of M;*
3. *for all x in Ω'_v, $T_{ind,v}(x) = x + \beta \mod M$, where $T_{ind,v}$ denotes the first return map of the covered rotation $T' = r_{\Omega',L} \circ R_\alpha \circ p_L$ to the set Ω'_v;*
4. *the derivated word to the factor v, $D_v(w_0)$, is a natural coding of the rotation of $\mathbb{T}_M$ through the angle β, whose elements and borders assignment are explicit (they are given in Proposition 35).*

Furthermore, the return words $u_1, ..., u_{d+1}$ to the factor v form a basis of the abelian free group generated by $\mathcal{I} = \{1, ..., d+1\}$.

Remark 29. *The induced natural coding $D_v(w_0)$ inherits from the choice made for the borders assignment of the original natural coding w_0. There is no need to resort to the axiom of choice a second time.*

4.2 Proof of Theorem A

Let w_0 be a natural coding of a minimal rotation of the d-torus with elements $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$ and borders assignment $(\Omega' : \Omega'_1, ..., \Omega'_{d+1})$. Let $v \in \mathcal{F}(w_0)$ be such that w_0 admits $d+1$ return words $u_1, ..., u_{d+1}$ to the factor v. We keep the notations and tools developed in Section 3. In particular, we denote by f the coding function with respect to the partition $(\Omega_1, ..., \Omega_{d+1})$ of the L-simple set Ω, $\mathcal{D}$ the maximal set on which f is defined, and f' the extended coding function on Ω'. We still denote $T = r_{\Omega,L} \circ R_\alpha \circ p_L$ the covered rotation on the L-simple set Ω, which is defined on $r_{\Omega,L}(\tilde{\Omega} \cap R_\alpha^{-1}(\tilde{\Omega}))$ and its extension $T' = r_{\Omega',L} \circ R_\alpha \circ p_L$ defined on the whole fundamental domain Ω'.

We denote:

$$\Omega_v := \bigcap_{k=0}^{|v|-1} T'^{-k}(\Omega_{v[k]}) \quad \text{and} \quad \Omega'_v := \bigcap_{k=0}^{|v|-1} T'^{-k}(\Omega'_{v[k]}).$$

Lemma 30. *The set Ω_v is nonempty, open and included in Ω'_v.*

Proof. The set Ω_v is open as the preimage by the continuous map $r_{\Omega,L}$ of the open set $\cap_{k=0}^{|v|-1} R_\alpha^{-k}(\Omega_{v[k]})$. Furthermore, since v is a factor of w_0, there exists a nonnegative integer n such that $S^n(w_0) \in [v]$. The set Ω_v is nonempty since it contains the point $T^n(x_0)$. At last, the inclusion of Ω_v in Ω'_v comes from the inclusion of Ω_i in Ω'_i, for all $i \in \mathcal{I}$. □

The dynamical system (Ω', T') inherits from the minimality of the dynamical system $(\mathbb{T}_L, R_\alpha)$ (Lemma 17). Since the set Ω'_v has nonempty interior (Lemma 30), all trajectories end up with passing through it. Thus, we can define the first return map to Ω'_v:

$$\begin{array}{llll} T_{ind,v}: & \Omega'_v & \to & \Omega'_v \\ & x & \mapsto & T'^{n_0}(x), \qquad\qquad \text{where} \quad n_0 = \inf\{n \in \mathbb{N} | T'^n(x) \in \Omega'_v\} < \infty. \end{array}$$

For $i \in \mathcal{I}$, we introduce the sets:

$$A_i = \bigcap_{k=0}^{|u_iv|-1} T'^{-k}(\Omega_{u_iv[k]}) \quad \subset \Omega_v \qquad \text{and} \qquad A'_i = \bigcap_{k=0}^{|u_iv|-1} T'^{-k}(\Omega'_{u_iv[k]}) \quad \subset \Omega'_v.$$

Lemma 31. *[A subpartition]*

- *The sets A'_i, for i in $\mathcal{I}$, form a partition of Ω'_v.*
- *For all i in $\mathcal{I}$, A_i is a nonempty open set included and dense in A'_i.*

Proof. By definition of f', we have $A'_i = f'^{-1}([u_iv])$ for all $i \in \mathcal{I}$. Since the words u_iv contain exactly two occurrences of the factor v, with one as suffix, they cannot be strict prefix one another; this implies that the cylinders $[u_iv]$ are pairwise disjoint - hence the disjointedness of their preimage sets A'_i. Moreover, for all $y \in \Omega'_v$, $f'(y) \in X_0 \cap [v] \subset \cup_{i\in\mathcal{I}}[u_iv]$ by Lemmas 16 and 26; consequently $y \in \cup_{i\in\mathcal{I}} A'_i$. Finally, the sets A'_i, for $i \in \mathcal{I}$, form a partition of Ω'_v.

The set A_i is the preimage, by the continuous map $r_{\Omega,L}$, of the open set $\cap_{k=0}^{|u_iv|-1} R_\alpha^{-k}(\Omega_{u_iv[k]}) \subset \Omega$ - it is thus open. Since the word u_iv is a factor of w_0, there exists a nonnegative integer n such that $S^n(w_0) \in [u_iv]$. The set A_i is nonempty since it contains the set $f^{-1}([u_iv])$, which contains itself the point $T^n(x_0)$. The inclusion $A_i \subset A'_i$ is inherited from the inclusions $\Omega_j \subset \Omega'_j$ for j in $\mathcal{I}$. At last, for the density, consider $y \in A'_i$. Applying Lemma 16, we can find a sequence $(x_n)_n \in \mathcal{D}^{\mathbb{N}}$ such that $x_n \to y$ and $\text{pref}_{|u_iv|}(f(x_n)) = \text{pref}_{|u_iv|}(f'(y)) = u_iv$; in particular $(x_n)_n \subset f^{-1}([u_iv]) \subset A_i$. □

The following lemma states that the induced map $T_{ind,v}$ acts on the sets $A'_1, ..., A'_{d+1}$ as an exchange of pieces.

Lemma 32 (Exchange of pieces)**.** *Let i in $\mathcal{I}$. For all $x \in A'_i$, $T_{ind,v}(x) = x + \beta_i$, where $\beta_i = \sum_{j\in\mathcal{I}} |u_i|_j\, \alpha_j$.*

Proof. If x belongs to A'_i, then its coding word $f'(x)$ belongs to $[u_iv]$, meaning that the $|u_iv| - 1$ first steps of the trajectory of x are fully known. More precisely, starting from $\Omega'_v \subset \Omega'_{u_i[0]}$, x is translated by the vector $\alpha_{u_i[0]}$ and falls into $\Omega'_{u_i[1]}$; then it is translated by $\alpha_{u_i[1]}$ and falls into $\Omega'_{u_i[2]}$, and so on; until arriving into $\Omega'_{u_i[|u_i|-1]}$ from where it is translated by $\alpha_{u_i[|u_i|-1]}$ and falls at last - and for the first time - into Ω'_v. All in all, from $A'_i \subset \Omega'_v$ to its first return into Ω'_v, the point x was translated by the vector $\beta_i = \sum_{j\in\mathcal{I}} |u_i|_j\, \alpha_j$. □

We introduce the vectors of $\mathbb{R}^d$:

$$\mathrm{v}_k = \beta_k - \beta_{d+1} \qquad \text{for } k \in \{1, ..., d\}$$

and the subgroup $M = \sum_{k=1}^{d} \mathbb{Z}\mathrm{v}_k$.

Proposition 33 (A rotation on a new torus)**.** *The following assertions are true.*

(i) The subgroup M is a lattice of $\mathbb{R}^d$.

(ii) The vectors β_j, $j \in \mathcal{I}$, are equals modulo M.

From now on, we denote $\beta = \beta_{d+1}$.

(iii) The pair (β, M) is minimal.

(iv) For all x in Ω'_v, we have $T_{ind,v}(x) = x + \beta \mod M$.

Proof. The assertions (ii) and (iv) stem from the definition of M and Lemmas 30 and 32. We now propose to show that the subgroup $\sum_{i\in\mathcal{I}} \beta_i\mathbb{Z} = \beta\mathbb{Z} + \sum_{k=0}^{d} \mathrm{v}_k\mathbb{Z}$ is dense in $\mathbb{R}^d$. This fact implies that:

- the vectors v_k, $k \in \{1, ..., d\}$, are linearly independent over $\mathbb{R}$- thus proving (i);
- the trajectory of $p_M(0)$ is dense in the torus $\mathbb{T}_M$ for the action of the rotation $R_{\beta,M}$ - thus proving (iii).

Let l be such that $x := T^l(x_0) \in \Omega'_v$. The trajectory of x under the map T being dense in Ω', the sequence $(T^n_{ind,v}(x))_{n\in\mathbb{N}}$ - consisting of all the points falling into Ω'_v - is dense in the open subset Ω_v and is, by Lemma 32, included in $x + \sum_{i\in\mathcal{I}} \beta_i\mathbb{Z}$. We conclude that the subgroup $\sum_{i\in\mathcal{I}} \beta_i\mathbb{Z}$ is dense in $\Omega_v - x$, which is a nonempty open set of $\mathbb{R}^d$, and thus, is actually dense in $\mathbb{R}^d$ itself. □

Proposition 34. *The set Ω'_v is a fundamental domain of M.*

Proof. We successively prove that the projection map $p_M : \Omega'_v \to \mathbb{T}_M$ is one-to-one and onto.

Let $x, y \in \Omega'_v$ be such that $p_M(x) = p_M(y)$, i.e. $y = x + \sum_{j=1}^{d} b_j\mathrm{v}_j$ for some $b_j \in \mathbb{Z}$. Since each β_i is a linear combination of $\alpha_1, ..., \alpha_{d+1}$ (Lemma 32), which are all congruent to α modulo L, it comes that $\beta_i = k_i\alpha \mod L$, where k_i is the length of the associated return word u_i. The previous equality can then be rewritten $y = x + \sum_{j=1}^{d} b_j(k_j - k_{d+1})\alpha + l$, for some $l \in L$; hence $p_L(y) = R^n_{\alpha,L}(p_L(x))$, with $n = \sum_{k=1}^{d} b_j(k_j - k_{d+1}) \in \mathbb{N}$ (if needed, we swap x and y), and thus $y = T'^n(x)$. But, given that both x and y belong to Ω'_v, y is not only on the trajectory of x for the action of T', but also for the action of the first return map $T_{ind,v}$: there exists $m \in \mathbb{N}$ s.t. $y = T'^m_{ind,v}(x)$. Finally, we had $y = x \mod M$ and now, we have $y = x + m\beta \mod M$, meaning that either $m = 0$, or the trajectory of $p_M(x)$ under $R_{\beta,M}$ is periodic, which is forbidden by minimality of (β, M). It eventually comes that $m = 0$, and $x = y$ - hence the injectivity.

Now, let $\overline{y} \in \mathbb{T}_M$. By minimality of $(-\beta, M)$, and because $p_M(\Omega'_v)$ has nonempty interior (Lemma 30), there exist an element $\overline{x} \in p_M(\Omega'_v)$ and a nonnegative integer n such that $R^n_{\beta,M}(\overline{x}) = \overline{y}$. Denote $x = r_{\Omega'_v,M}(\overline{x}) \in \Omega'_v$ (the covering map $r_{\Omega'_v,M}$ is well-defined by the previous paragraph). The trajectory of x under the map $T_{ind,v}$ remains in Ω'_v; in particular, $y := T'^n_{ind,v}(x)$ belongs to Ω'_v. Then, $p_M(y) = p_M(T'^n_{ind,v}(x)) = R^n_\beta(p_M(x)) = R^n_\beta(\overline{x}) = \overline{y}$. We conclude that $\overline{y}$ admits a preimage by p_M in Ω'_v - hence the surjectivity. □

Proposition 35 (An induced natural coding.)**.** *The derivated word of w_0 to the factor v, $D_v(w_0)$, is a natural coding of the minimal rotation of $\mathbb{T}_M$ through the angle β, with elements $((\beta, M); (A : A_1, ..., A_{d+1});\ T^l(x_0); (\beta_1, ..., \beta_{d+1}))$, where l denotes the minimal nonnegative integer such that $S^l(w_0)$ starts with the factor v. Furthermore, $(\Omega'_v : A'_1, ..., A'_{d+1})$ is a borders assignment of this natural coding.*

Proof. Minimality. We know from Proposition 33 that the pair (β, M) is minimal.

Partition of a pseudo-fundamental domain. By Lemma 31, the sets $A_1, ..., A_{d+1}$ are nonempty, open and pairwise disjoint. Furthermore, their union set $A = \cup_{i\in\mathcal{I}} A_i$ inherits from the M-simplicity of the set Ω'_v it is included in (Proposition 34). We now show that the projection set $p_M(A)$ is dense in the torus $\mathbb{T}_M$. Denote $y_0 = T^l(x_0) \in \Omega_v$. The trajectory of y_0 under the action of T is included in $\mathcal{D}$; so its trajectory under the action of the induced map $T_{ind,v}$ is included in $\Omega_v \cap \mathcal{D} \subset A$. We deduce that the trajectory of $\overline{y}_0 := p_M(y_0)$ under the rotation $R_{\beta,M}$, which is dense in $\mathbb{T}_M$ by minimality of (β, M), is included in $p_M(A)$; this implies that the set $p_M(A)$ is dense in $\mathbb{T}_M$.

Exchange of pieces. By Lemma 32, for all $i \in \mathcal{I}$ and for all $\overline{x} \in p_M(A_i) \cap R^{-1}_{\beta,M}(p_M(A))$, we have $r_{A,M} \circ R_{\beta,M}(\overline{x}) = T_{ind,v}(r_{A,M}(\overline{x})) = r_{A,M}(\overline{x}) + \beta_i$.

A coding trajectory. By construction of the sets $A_1, ..., A_{d+1}$, for all nonnegative integer n, we have $R^n_{\beta,M}(\overline{y}_0) \in p_M(A_i)$ if and only if $T'^n_{ind,v}(y_0) \in A_i$ if and only if $f(T'^n_{ind,v}(y_0))$ starts with the word u_i if and only if the n-th letter of the derived word $D_v(w_0)$ is i.

A borders assignment. (1) & (2) By Lemma 31, for all $i \in \mathcal{I}$, A_i is included in A'_i and the sets A'_i form a partition of Ω'_v. (3) The set Ω'_v is a fundamental domain of M (Proposition 34). (4) For all $i \in \mathcal{I}$ and for all $x \in A'_i$, we have $r_{\Omega'_v,M} \circ R_{\beta,M} \circ p_M(x) = T_{ind,v}(x) = x + \beta_i$ (Proposition 33 and Lemma 32). (5) Let $x \in \Omega'_v$ and $q \in \mathbb{N}$. Denote $l = \max_{i\in\mathcal{I}} |u_i|$ and σ the extraction given by the definition of borders assignment associated with the natural coding w_0 for the point x (seen as an element of Ω') and the integer $(q+1)l + |v|$. For any nonnegative integer m, the prefix of length $q(l+1) + |v|$ of the words $f(T^{\sigma(m)}(x_0))$ and $f'(x)$ coincide. In particular, the sequence $(T^{\sigma(m)}(x_0))_{m\in\mathbb{N}}$ is included in Ω_v, so it is a subsequence of the trajectory of x_0 under the action of the first return map to Ω'_v. Denote by τ the extraction such that for all nonnegative integer m, $T^{\tau(m)}_{ind,v}(x_0) = T^{\sigma(m)}(x_0)$. We immediately have that $T^{\tau(m)}_{ind,v}(x_0) = T^{\sigma(m)}(x_0) \to_{m\to\infty} x$. Furthermore, by definition of l, the prefix of length $(q+1)l + |v|$ of $f'(x)$ contains at least $q+2$ occurrences of the factor v; we deduce that, for any $m \in \mathbb{N}$, the first $q+1$ return words to v of the symbolic trajectory of x and $T^{\tau(m)}_{ind,v}(x_0)$ for the action of $T_{ind,v}$ coincide, i.e.: for all $n \in \{0, ..., q\}$, for all nonnegative integer m, $T^{\tau(m)+n}_{ind,v}(x_0) \in A_{\iota_n}$, where the index ι_n is defined by $T^n_{ind,v}(x) \in A'_{\iota_n}$. □

Proposition 36. *The return words $u_1, ..., u_{d+1}$ to the factor v form a basis of the free abelian group generated by $\mathcal{I}$.*

Proof. We are going to show that $\mathcal{M} = (|u_j|_i)_{i,j} \in GL(\mathbb{Z})$. Since $(\beta_1, ..., \beta_{d+1}) = (\alpha_1, ..., \alpha_{d+1})\mathcal{M}$, and since the vectors $\alpha_1, ..., \alpha_{d+1}$ form a basis of the $\mathbb{Z}$-module G (Lemma 19), it is sufficient to show that the $d+1$ vectors $(\beta_1, ..., \beta_{d+1})$ are free over $\mathbb{Z}$. This is the case by Lemma 19 again, given that the word $D_v(w_0)$ is a natural coding of the minimal rotation with elements $((\beta, M); (A : A_1, ..., A_{d+1}); T^l(x_0); (\beta_1, ..., \beta_{d+1}))$ (Proposition 35). □

Propositions 33, 34, 35 and 36 prove Theorem A.

4.3 Correction of the proof of [CFZ00]

We now complete the idea of [CFZ00] to prove, resorting to Rauzy's theorem on remainder sets (Theorem B below), that being a natural coding of a minimal rotation, with a bounded fundamental domain, implies finite imbalance on all factors (see Section 2.3 for the definitions, and Proposition 37 below for a formal statement). For letters (i.e. factors of length 1), a direct and general proof of this fact can be found in the latest version of [Thu19].

Proposition 37. *Let w_0 be a natural coding of a minimal rotation of a d-dimensional torus, and denote by $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$ its elements, and by $(\Omega' : \Omega'_1, ..., \Omega'_{d+1})$ a borders assignment. Assume that w_0 admits $d+1$ return words to a finite word v. Assume furthermore* *that the pseudo-fundamental domain Ω is bounded. Then the set $p_L(\Omega'_v)$ is a bounded remainder set for for any trajectory $(R_\alpha(\tilde{x}))_{n\in\mathbb{N}}$, and the imbalance of w_0 on the factor v is finite.*

Definition 38 (Following [Rau84])**.** *A set A is a* bounded remainder set for a sequence $(u_n)_{n\in\mathbb{N}}$ *if there exist two real numbers (ν, C) such that, for all positive integer N:*

$$|\sum_{n=0}^{N-1} 1_A(u_n) - N\nu| < C.$$

The numbers ν and C can be understood as a frequency and a tolerance margin for the event 'falling into A'. So, A is a bounded remainder set for the sequence (u_n) means that (u_n) is well-distributed relatively to A: the observed frequency of visits to A converges to it expected value at speed $1/n$.

Theorem B. *[Rau84] Let d be a positive integer, L a lattice of $\mathbb{R}^d$, α an element of $\mathbb{R}^d$ such that the pair (α, L) is minimal. Let $A \subset \mathbb{R}^d$, L-simple, bounded, with nonempty interior. Let T denote the transformation on A induced by the rotation R_α.*

If there exist a lattice M of $\mathbb{R}^d$, together with an element $\beta \in \mathbb{R}^d$, such that:

(i) A is M-simple,

(ii) for all $x \in A$, $T(x) = x + \beta \mod M$,

then $p_L(A)$ is a bounded remainder set for all sequence $(R^n_\alpha(\tilde{x}))_n$, with $\tilde{x} \in \mathbb{T}_L$.

Remark 39. *In [Rau84], Rauzy integrates the assumption of boundedness of A in the definition of* L-simplicity*. This assumption is crucial at two stages in his proof.*

Remark 40. *Theorem B gives a sufficient condition for a set to be a bounded remainder set. A necessary and sufficient condition generalizing this criterion is given in [Fer92] under the framework of measurable dynamical systems. Though not mentioned, the assumption of boundedness is still required.*

Proof of Proposition 37. The pseudo-fundamental domain Ω being bounded, so are the fundamental domain Ω' and its subset Ω'_v. Therefore, by Theorems A and B, for all $\tilde{x} \in \mathbb{T}_L$, $p_L(\Omega'_v)$ is a bounded remainder set for the sequence $(R_\alpha(\tilde{x}))_{n\in\mathbb{N}}$.

On another hand, by Definition 2 (natural coding), for all nonnegative integer n, we have $R_\alpha(\tilde{x}_0) \in \Omega'_v$ if and only if and only if $S^n(w_0) \in [v]$. We deduce from this equivalence that the cylinder $[v]$ is also a bounded remainder set for the sequence $(S^n(w_0))_{n\in\mathbb{N}}$: there exist two real numbers ν and C such that for all positive integer N:

$$|\sum_{n=0}^{N-1} 1_{[v]}(S^n(w_0)) - N\nu| < C.$$

In other words, for all positive integer N, $|\text{pref}_{N+|v|-1} S^n(w_0)|_v = \sum_{n=0}^{N-1} 1_{[v]}(S^n(w_0)) \in]\nu N - C; \nu N + C[$, from which we deduce that for all factor $u \in \mathcal{F}(w_0)$, $|u|_v \in]\nu|u| - 2C; \nu|u| + 2C[$. This implies that the imbalance of w_0 on the factor v is lower than the constant $4C$. □

Remark 41. *Finite imbalance on a letter a is equivalent to the cylinder $[a]$ being a bounded remainder set for the sequence $(S^n(w_0))_n$ (see [Ada03]).*

Remark 42. *The main mistake in the original proof of [CFZ00] is that no information on the second lattice M is given, and thereby, one cannot guarantee that the set A is M-simple. This confusion is still present in the first versions of the lecture notes [Thu19].*

5 Applications

5.1 Tree words

Theorem A claims that being a natural coding of a minimal rotation of the d-torus is a property preserved by the derivation operation. This is why good candidates should be families of words stable under this operation. This is the case for the class of infinite words admitting d return words to any factor [BPS08]; this is also the case of its remarkable subclass comprised of tree words.

We recall that a finite word u is a *return word to the factor* v in the recurrent word w if $u = w[i]...w[j-1]$, where i and j are two consecutive occurrences of v (Definition 25); or equivalently, if $uv \in \mathcal{F}(w)$, v is a prefix of uv and if there are exactly two occurrences of v in uv [Dur98].

Let w be an infinite word over an alphabet A, and u one of its factor. Following [BFD$^+$15a], we denote $L(u)$ (resp. $R(u)$) the set of letters a in A such that au (resp. ua) is still a factor of w. The *extension graph* of u is the undirected graph whose vertices are the disjoint union of $L(u)$ and $R(u)$, and whose edges are the pairs $(a, b) \in L(u) \times R(u)$ such that aub is a factor of w. An infinite word w is a *tree word* (or a *dendric word*, in recent texts) if the extension graph of each of its factors is acyclic and connected (viz. a tree).

On the two-letter alphabet, the set of infinite words admitting two return words to any factor, the set of uniformly recurrent tree words, the set of Sturmian words and the set of words whose subshift is a natural coding of a minimal rotation of the circle coincide ([Vui01], [JV00]).

More generally:

- Uniformly recurrent tree words on the alphabet $\mathcal{I} = \{1, ..., d+1\}$ admit $d+1$ return words to any factor (so in particular to each letter), which moreover form a basis of the free group over $\mathcal{I}$; but when $d \geq 2$, we also have examples of infinite words admitting $d+1$ return words to any factors which are not tree words [BFD$^+$15a].

- Strict episturmian words are uniformly recurrent tree words, but on alphabets with three letters or more, there exists other families of words, such as primitive C-adic words (see Definition 47 below), that belong to this class too [CLL17].

The following proposition and corollary are immediate applications of Theorem A.

Proposition 43. *If a uniformly recurrent tree word w_0 on the alphabet $\mathcal{I} = \{1, ..., d+1\}$ is a natural coding of a minimal rotation of the d-torus, then its derivated sequences to any factors are also natural codings of a minimal rotation of the d-torus.*

Corollary 44. *No uniformly recurrent tree word with infinite imbalance on a factor is a natural coding of a minimal rotation of the 2-torus with a bounded pseudo-fundamental domain.*

In particular, no Arnoux-Rauzy word with infinite imbalance is a natural coding of a minimal rotation of the 2-torus. This result rectifies and strengthens the one stated in [CFZ00]. Constructions of Arnoux-Rauzy words with infinite imbalance are detailed in [CFZ00] and [And21]. Likewise, no primitive C-adic word with infinite imbalance is a natural coding of a minimal rotation of the 2-torus. Primitive C-adic words with infinite imbalance have been constructed in [And18].

On the counterpart, remember that a lot of Arnoux-Rauzy words and C-adic words are natural codings of rotations, under the definition of [BST20].

Once a uniformly recurrent tree word is a natural coding of a minimal rotation of the d-torus, then its derivated words to the d letters of the alphabet are again tree words (see [BFD$^+$15b]) and natural codings of minimal rotations (Theorem A) - in particular, they are again uniformly recurrent by Lemma 3. We can thus iterate the derivation, and study the trajectory of words under this operation. In the remarkable cases of Arnoux-Rauzy and primitive C-adic words, these trajectories are driven by generalized euclidean maps (often referred to as *multidimensional continued fraction algorithms*), as evidenced through the S-adic framework (see the book [Sch00] for a general introduction to multidimensional continued fractions, and for instance the surveys [Ber11] or [BD14] for their study from the symbolic dynamical standpoint).

5.2 Return words for Arnoux-Rauzy words (under the S-adic framework)

We recall that *Arnoux-Rauzy words* are infinite words on the alphabet $\mathcal{I} = \{1, 2, 3\}$ with complexity $p(n) = 2n + 1$, such that for each n there is exactly one right and one left special factor of length n [AR91]. By a result of Boshernitzan [Bos84], Arnoux-Rauzy words are uniquely ergodic; hence the existence of frequencies, which are positive, for each factor. We introduce the set $AR = \{\sigma_i | i \in \mathcal{I}\}$ of Arnoux-Rauzy substitutions:

$$\begin{aligned} \sigma_i : \ & \mathcal{I} \to \mathcal{I}^* \\ & i \mapsto i \\ & j \mapsto ij \text{ for } j \in \mathcal{I} \backslash \{i\}. \end{aligned}$$

The following theorem evidences the link between Arnoux-Rauzy words and one generalization of the Euclid map.

Theorem C ([AS13]). *Let w be an Arnoux-Rauzy word. Then there exists a unique sequence of substitutions (called* directive sequence*) $d = (\sigma_{i_n})_n$ in $AR^{\mathbb{N}}$, and a unique Arnoux-Rauzy word w' such that:*

1. *each prefix of w' is a left-special factor;*
2. *the sets of factors of w and w' are equal;*
3. *$w' = lim_{n\to\infty} \sigma_{i_0} \circ ... \circ \sigma_{i_{n-1}}(1)$.*

Furthermore:
- we have $w' = lim_{n\to\infty} \sigma_{i_0} \circ ... \circ \sigma_{i_{n-1}}(2)$ and $w' = lim_{n\to\infty} \sigma_{i_0} \circ ... \circ \sigma_{i_{n-1}}(3)$;
- each Arnoux-Rauzy substitution appears infinitely many times in d;
- d is uniquely defined by the frequencies of letters in w: the sequence $(i_n)_{n\in\mathbb{N}}$ is the symbolic trajectory of the letters frequency vector under the action of the generalized Euclid map:

$$F_{AR} : \quad (x, y, z) \quad \mapsto \quad \begin{cases} (x - y - z, y, z) & \text{if } x \geq y + z, \\ (x, y - x - z, z), & \text{if } y \geq x + z, \\ (x, y, z - x - y), & \text{if } z \geq x + y, \end{cases}$$

with respect to the partition given by its piecewise definition.

As evidenced by Lemma 32 and Proposition 33, the action of the induction/derivation operation of a natural coding on the lattice and the angle of the rotation is driven by the abelianized vectors of the return words to a letter.

We now describe how to obtain the three return words to a letter a for Arnoux-Rauzy words. This result comes from [JV00]; we just state it under the S-adic formalism, i.e., as a function of the sequence of substitutions (namely, the directive sequence) given by Theorem C.

Notation 45. *We denote by s the* circular shift on (nonempty) finite words*: $s(u) = a_2...a_na_1$, where $a_1, ..., a_n$ are letters and $u = a_1...a_n$. The map s is bijective.*

Theorem D ([JV00], under a slighly different formalism.)**.** *If w is an Arnoux-Rauzy word with directive sequence $d = (\sigma_{i_n})_n$, and $a \in \{1,2,3\}$ is a letter, then w admits three return words to a, namely: $s^{-1} \circ d_0 \circ ... \circ d_{n_0-1} \circ s \circ d_{n_0}(b)$, for $b \in \{1,2,3\}$, where s is the circular shift on finite words and $n_0 = min\{n \in \mathbb{N} | \, i_n = a\}$. Furthermore, the derivated word of w to a is an Arnoux-Rauzy word with directive sequence $d' = (\sigma_{i_n})_{n > n_0}$.*

Return words to any factor are described in [JV00].

At last, we denote by $M_\sigma = (|\sigma(j)|_i)_{(i,j)\in\mathcal{I}^2}$ the *incidence matrix* of a substitution σ.

Corollary 46. *Let w be an Arnoux-Rauzy word with directive sequence $d = (d_n)_n$, and $a \in \{1,2,3\}$ a letter. If w is a natural coding of a minimal rotation of the 2-torus, with elements $((\alpha, L); (\Omega : \Omega_1, \Omega_2, \Omega_3); x_0; (\alpha_1, \alpha_2, \alpha_3))$, then the vectors β_1, β_2 and β_3 describing the induced rotation on Ω_a are given by:*

$$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3) M_{d_0} ... M_{d_{n_0}},$$

where $n_0 = min\{n \in \mathbb{N} | \, i_n = a\}$.

Proof. We have $(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3)\mathcal{M}$, with $\mathcal{M} = (|u_j|_i)_{i,j}$. By theorem D, for all $i, j \in \{1,2,3\}$, $|u_j|_i = |d_0 \circ ... \circ d_{n_0}(j)|_i$; therefore $\mathcal{M}$ is the incidence matrix of the substitution $d_0 \circ ... \circ d_{n_0}$, and $(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3) M_{d_0} ... M_{d_{n_0}}$. □

5.3 Return words for primitive C-adic words

We now deal with primitive C-adic words. This class of words was introduced in [CLL17], emerging from the research of a generalized Euclid map defined on $(\mathbb{R}^+)^3$ defined for any projective direction -contrary to F_{AR} which is defined for almost none (see [AR91], [AS13] and [AHS13])- and producing words with the lowest possible complexity: $p(n) = 2n + 1$. This led to the map:

$$F_C : \quad (x, y, z) \quad \mapsto \quad \begin{cases} (x - z, z, y) & \text{if } x \geq z \\ (y, x, z - x) & \text{otherwise} \end{cases},$$

and to the associated substitutions $C = \{c_1, c_2\}$ given by:

$$c_1 : \begin{array}{lll} 1 & \mapsto & 1 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 2 \end{array} \qquad c_2 : \begin{array}{lll} 1 & \mapsto & 2 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 3. \end{array}$$

Definition 47 ([CLL17])**.** *An infinite word w is* C-adic *if there exist a directive sequence $d = (d_n) \in C^{\mathbb{N}}$, together with a letter $a \in \{1,2,3\}$, such that w can be written $w = lim_{n\to\infty} d_0 \circ ... \circ d_{n-1}(a)$.*

As long as d contains infinitely many occurrences of c_1 and c_2, the sequence of finite words $(d_0 \circ ... \circ d_{n-1}(a))_n$ converges to an infinite words w that, furthermore, does not depend on the letter a [CLL17].

Proposition 48 ([CLL17])**.** *Let w be a C-adic word with directive sequence d. If $d \notin C^*\{c_1^2, c_2^2\}^\omega$, then w is a uniformly recurrent tree word.*

A C-adic word whose directive sequence does not belong to $C^*\{c_1^2, c_2^2\}^\omega$ is said *primitive* [CLL17]. The primitivity condition is central in the study of return words: it guarantees the termination of the process described in the proof of Theorem E (below).

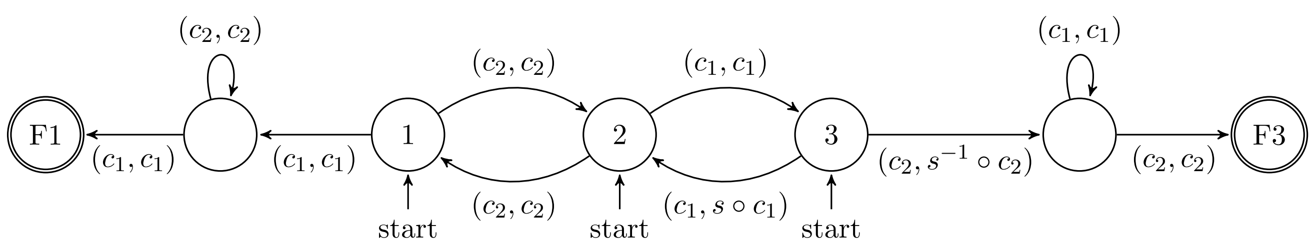
Lemma 49. *A primitive C-adic word w admits a unique directive sequence, which can be deduced from the knowledge of its set of factors $\mathcal{F}(w)$.*

Proof. Let $d = (d_n)_{n\geq 0} \in \{c_1, c_2\}^{\mathbb{N}}$ be such that $d \notin C^*\{c_1^2, c_2^2\}^\omega$, and w its (unique) associated C-adic word, which is primitive. Denote by w' and w'' the C-adic words obtained with the directive sequences $(d_n)_{n\geq 1}$ and $(d_n)_{n\geq 2}$ respectively, which are also primitive. We then have that 2 is factor of w'', which implies that 13 is factor of w', which implies in turn that 12 xor 23 is a factor of w. If $12 \in \mathcal{F}(w)$, then $d_0 = c_1$; if $23 \in \mathcal{F}(w)$, then $d_0 = c_2$. Furthermore, the word w' can be deduced from the knowledge of w and d_0. We iterate this process to determine the entire sequence d. □

Notation 50. *We denote by l (resp. r) the map which extracts the first (resp. the second) component x (resp. y) of a pair (x, y).*

Theorem E. *Let w be a primitive C-adic word with directive sequence $d = (d_n)_{n\in\mathbb{N}}$, and $a \in \{1, 2, 3\}$ a letter. The following assertions are true.*

1. *There exists in the automaton of partial quotients of C-adic words (Figure 5.3) a unique accepted path $e = (e_0, ..., e_{n_1})$ starting from the initial state a and such that the finite sequence $l_e = l(e_0)...l(e_{n_1}) \in \{c_1, c_2\}^*$ is a prefix of d.*
   We denote by n_2 the length of this prefix, and by w' the primitive C-adic word with directive sequence $(d_n)_{n\geq n_2}$.
2. *The set of return words to the letter a of w, denoted $\mathcal{U}$, is the image set of the alphabet $\{1, 2, 3\}$ by the application $r_e = r(e_0) \circ ... \circ r(e_{n_1})$.*
3. *The set $\mathcal{U}$ contains 3 elements, and if we denote them by $u_i = r_e(i)$ for i in $\{1, 2, 3\}$, then the derivated word of w to a, with respect to the numeration $u_i \to i$, is the word w' if the final state of e is F_1, and $S(w')$ (where S denotes the shift map) if the final state of e is F_3.*



with s the circular shift on finite words: $s(a_1...a_n) = a_2...a_na_1$.

Figure 1: Automaton of partial quotients for C-adic words.

Example 51. *We consider the primitive C-adic word w with directive sequence $d = c_1c_1c_2c_2(c_1c_2)^\omega$, and the letter $a = 3$.*

$$w = 113131213113121311313113121311313121311313113121311313121311...$$

Applying Theorem E, we obtain $l(e) = c_1c_1c_2c_2$ and $r_e = s \circ c_1 \circ c_1 \circ s^{-1} \circ c_2 \circ c_2$; hence $u_1 = 311$, $u_2 = 3121$, $u_3 = 31$ and w' is the (primitive) C-adic word with directive sequence $(c_1c_2)^\omega$. Since the path e leads to the final state F_3, the derivated word (with respect to the chosen numeration) is:

$$D_3(w) = S(w') = 32121312132131213212132131213212131213212132131213212131213213121321...$$

A final remark: if we had chosen another numeration, say $\tilde{u}_1 = 3121$, $\tilde{u}_2 = 311$ *and* $\tilde{u}_3 = 31$, *we would have obtained:*

$$\tilde{D}_3(w) = 31212321231232123121231232123121232123121231232123121232123123212312 1...$$

which is not is the subshift of a primitive C-adic word, since it does not contain the factor $13 = c_1(2) = c_2(2)$.

Proof of Theorem E. Let w be a primitive C-adic word with directive sequence $d = (d_n)_{n\geq 0}$. Since d contains infinitely many occurrences of c_1 and c_2, so does any sequence of the form $(d_n)_{n\geq n_0}$ for $n_0 \in \mathbb{N}$, so that the sequence of finite words $(d_{n_0} \circ ... \circ d_{n-1}(b))_{n\geq n_0}$ converges to an infinite word w', which does not depend on the letter b (but depends of course on n_0), and is again a primitive C-adic word.

1. If d starts with c_1, then the set of return words to 2 is the image set by c_1 of the return words to 3 in the word with directive sequence $(d_n)_{n\geq 1}$. Symmetrically, if d starts with c_2, then the set of return words to 2 is the image set by c_2 of the return words to 1 in the word with directive sequence $(d_n)_{n\geq 1}$.

2. If d starts with c_1, then there exists n_0 such that d starts with $c_1 \circ c_2^{n_0} \circ c_1$. If $n_0 = 2k+1$, the images of the letters 1, 2 and 3 by $c_1 \circ c_2^{n_0} \circ c_1$ are respectively $132^k, 132^{k+1}$ and 12^{k+1}; if $n_0 = 2k$, they are respectively $12^k, 12^{k+1}$ and 132^k. Since furthermore the word w' with directive sequence $(d_n)_{n\geq n_0+2}$ contains the three letters 1, 2 and 3, w contains three return words to 1, which are the images of the letters by the substitution $c_1 \circ c_2^{n_0} \circ c_1$. Otherwise, if d starts with c_2, the set of return words to 1 is the image set by c_2 of return words to 2 in the word with directive sequence $(d_n)_{n\geq 1}$.

3. Symmetrically, if d starts with c_2, then there exists n_0 such that d starts with $c_2 \circ c_1^{n_0} \circ c_2$. If $n_0 = 2k+1$, the images of the letters 1, 2 and 3 by $c_2 \circ c_1^{n_0} \circ c_2$ are respectively $2^{k+1}3, 2^{k+1}13$ and 2^k13; if $n_0 = 2k$, they are respectively $2^k13, 2^{k+1}3$ and 2^k3. Since furthermore the word w' with directive sequence $(d_n)_{n\geq n_0+2}$ contains the three letters 1, 2 and 3, w contains three return words to 3, which are the images of the letters by the application $s^{-1} \circ c_2 \circ c_1^{n_0} \circ c_2$. Otherwise, if d starts with c_1, the return words to 3 are the images by the map $s \circ c_1$ of the return words to 2 of the word with directive sequence $(d_n)_{n\geq 1}$.

We recursively combine the three situations above to obtain the return words to any letter a in w; indeed, the primitivity condition (i.e. $d \notin C^*\{c_1^2, c_2^2\}^\omega$) guarantees that this recursive process always halts. We conclude that each letter $a \in \{1, 2, 3\}$ admits three distinct return words by observing that the images by c_1 (resp. c_2) of two distinct finite words are again distinct. Indeed, two distinct words u and v can always be written $u = u's$ and $v = v's$ (resp. $u = pu'$ and $v = pv'$), where u' and v' end (resp. start) with distinct letters (one of them at most is allowed to be empty); then $c_1(u')$ and $c_1(v')$ also end (resp. start) with distinct letters, implying $c_1(u) \neq c_1(v)$ (resp. $c_2(u) \neq c_2(v)$). □

Corollary 52. *Let w be a primitive C-adic word with directive sequence $d = (d_n)_{n\geq 0}$, and $a \in \{1, 2, 3\}$ a letter. If w is a natural coding of a minimal rotation of the 2-torus, with elements $((\alpha, L); (\Omega : \Omega_1, \Omega_2, \Omega_3); x_0; (\alpha_1, \alpha_2, \alpha_3))$, then the vectors β_1, β_2 and β_3 describing the induced rotation on Ω_a are given by:*

$$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3) M_{d_0} ... M_{d_{n_2-1}},$$

where n_2 is the length of the unique prefix of d accepted by the partial quotients automaton for C-adic words from the initial state a.

Proof. The proof is identical to the proof of Corollary 46 for Arnoux-Rauzy words. □

At last, we deduce from Theorem E an algorithm which, given a primitive C-adic words w and v one of its factor, outputs the three return words to v of w.

Theorem F. *Let w be a primitive C-adic word with directive sequence d, and $v \in \mathcal{F}(w)\backslash\{1,2,3\}$ one of its factors of length at least 2. Let $n_0 = \min\{n \in \mathbb{N} | u \in \mathcal{F}(d_0 \circ ... \circ d_{n_0-1}(2))\}$. Let p and s be such that $(d_0 \circ ... \circ d_{n_0-1}(2)) = pvs$. At last, let $\mathcal{P} = \{p^{-1} d_0 \circ ... \circ d_{n_0-1}(u)p \;| u \in \mathcal{U}\}$, where $\mathcal{U}$ is the set of return words to the letter 2 of the C-adic word w' with directive sequence $(d_n)_{n \geq n_0}$. Then $\mathcal{P}$ contains three words, which start with v and pave a suffix of w.*

Proof. Since $v \in \mathcal{F}(w)$ and since the sequence $(d_0 \circ ... \circ d_{n-1}(1))_n$ shares a growing common prefix with w, we can define the nonnegative integer:

$$n_0 = \min\{n \in \mathbb{N} | u \in \mathcal{F}(d_0 \circ ... \circ d_{n-1}(1)) \cup \mathcal{F}(d_0 \circ ... \circ d_{n-1}(2)) \cup \mathcal{F}(d_0 \circ ... \circ d_{n-1}(3))\}.$$

Since v contains at least two letters, n_0 is actually positive. Observe that if $v \in \mathcal{F}(d_0 \circ ... \circ d_{n-1}(1))$ for $n \geq 1$, then $v \in \mathcal{F}(d_0 \circ ... \circ d_{n-2}(1))$ if $d_{n-1} = c_1$ and $v \in \mathcal{F}(d_0 \circ ... \circ d_{n-2}(2))$ otherwise. Symmetrically, if $v \in \mathcal{F}(d_0 \circ ... \circ d_{n-1}(3))$ for $n \geq 1$, then $v \in \mathcal{F}(d_0 \circ ... \circ d_{n-2}(2))$ if $d_{n-1} = c_1$ and $v \in \mathcal{F}(d_0 \circ ... \circ d_{n-2}(3))$ otherwise. We deduce from the minimality of n_0 that $v \in \mathcal{F}(d_0 \circ ... \circ d_{n_0-1}(2))$ and $v \notin \mathcal{F}(d_0 \circ ... \circ d_{n_0-1}(a))$ for $a \in \{1,3\}$.

The C-adic word w with directive sequence $(d_n)_{n \in \mathbb{N}}$ being primitive, so is the C-adic word w' with directive sequence $(d_n)_{n \geq n_0}$; the word w' thus admits three return words to the letter 2, whose set is denoted by $\mathcal{U}$. Let $k_0 = \min\{k \in \mathbb{N} | S^k(w') \in [2]\}$, where S denotes the shift map. Then the words in $\mathcal{U}$ pave the infinite word $S^{k_0}(w')$. Denote by k_1 the length of the image by the substitution $d_0 \circ ... \circ d_{n_0-1}$ of the prefix of length k_0 of w', and $k_2 = k_1 + |p|$, where p is such that $d_0 \circ ... \circ d_{n-1}(2) = pvs$. Then the set $\mathcal{P} = \{p^{-1} d_0 \circ ... \circ d_{n_0-1}(u)p \;| u \in \mathcal{U}\}$, which contains three elements (the images of distinct words by c_1 or c_2 remaining distinct - see the end of the proof of Theorem E) that start with v, pave the infite word $S^{k_2}(w)$. □

The set $\mathcal{P}$ is not always the set of return words to the factor v of w, as illustrated by Example 54. Nonetheless, the set of return words to the factor v of w is easily deduced from $\mathcal{P}$.

Corollary 53. *If we denote $\mathcal{P} = \{p_1, p_2, p_2\}$, the set of return words to the factor v in w is exactly the set of return words to v in the finite word $p_1 p_2 p_3 p_1$.*

Example 54. *We consider the primitive C-adic word w with directive sequence $d = c_2 c_2 c_1 c_2 c_1 c_1 (c_1 c_2)^\omega$, and the factor $v = 31 \in \mathcal{F}(w)$.*

$$w = 13313323133132313313313231331332313313313231331332313313231331332313 31...$$

Applying Theorem F, we obtain $n_0 = 6$, and $\sigma = d_0 \circ ... \circ d_5 = c_2 \circ c_2 \circ c_1 \circ c_2 \circ c_1 \circ c_1$ is given by:

$$\begin{array}{rrcl} \sigma : & 1 & \mapsto & 133 \\ & 2 & \mapsto & 13\underline{31}323 \\ & 3 & \mapsto & 13323, \end{array}$$

hence $p = 13$ and $s = 323$. By Theorem E, the three return words to the letter 2 of the primitive C-adic word w' with directive sequence $(d_n)_{n \geq n_0} = (c_1 c_2)^\omega$ are 21, 213 and 2131. We finally obtain the paving set: $\mathcal{P} = \{313231331332313313,\ 3132313313,\ 313231331332313\}$, from which we deduce, following Corollary 53, the three return words to 31 of w: 313, 3132 and 31332.

In this example, no element of $\mathcal{P}$ is a return word to the factor v. This is a consequence of v appearing in $w = \sigma(w')$ not only as factor of $\sigma(2)$, but also at each junction of images of letters by σ: indeed, here, all images by the substitution σ starts with 1 and ends with 3.

6 Stability under exduction (reverse induction)

We now prove that being a natural coding of a minimal rotation is a property which passes through the reverse operation of induction, that we call, following Rauzy (see [AI01]), *exduction*.

We start by an example in dimension 1 (Sturmian case).

Example 55. *For d=1, consider $M = \mathbb{Z}$ and β an irrational number. We introduce $A_1 =]0, 1-\alpha[$, $A_2 =]1-\alpha, 1[$, $A'_1 = [0, 1-\alpha[$ and $A'_2 = [1-\alpha, 1[$. Then the standard Sturmian word with slope β, that we denote w_{st}, is a natural coding with elements $((\beta, M), A : (A_1, A_2), \beta, (\beta_1, \beta_2))$, where $\beta_1 = \beta$ and $\beta_2 = \beta - 1$ and borders assignment $(A' : (A'_1, A'_2))$.*

Now, we consider the substitution σ given by $\sigma(1) = 1$ and $\sigma(2) = 12$, and its incidence matrix M_σ:

$$M_\sigma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We set $(\alpha_1, \alpha_2) = (\beta_1, \beta_2) M_\sigma^{-1}$, which gives: $\begin{cases} \alpha_1 = \beta \\ \alpha_2 = -1. \end{cases}$

Denote $\alpha = \alpha_1 = \beta$ and $L = (\alpha_2 - \alpha_1)\mathbb{Z} = (\beta + 1)\mathbb{Z}$. At last, we introduce $\Omega_1 = A =]0, 1[$, $\Omega'_1 = A' = [0, 1[$ $\Omega_2 =]1, 1+\beta[$ and $\Omega'_2 = [1, 1+\beta[$ (see Figure 2).

Observe that the pair (α, L) is minimal, that the sets $\{\Omega'_1, \Omega'_2\}$ form a partition of a fundamental domain for the lattice L, and that the rotation $R_{\alpha,L}$ acts on the piece Ω'_1 (resp. Ω'_2) of the fundamental domain as a translation by the vector α_1 (resp. α_2). In fact, the word $\sigma(w_{st})$ is a natural coding of a minimal rotation of the circle with elements $((\alpha, L), \Omega : (\Omega_1, \Omega_2), \beta, (\alpha_1, \alpha_2))$ and borders assignment $(\Omega' : (\Omega'_1, \Omega'_2))$.

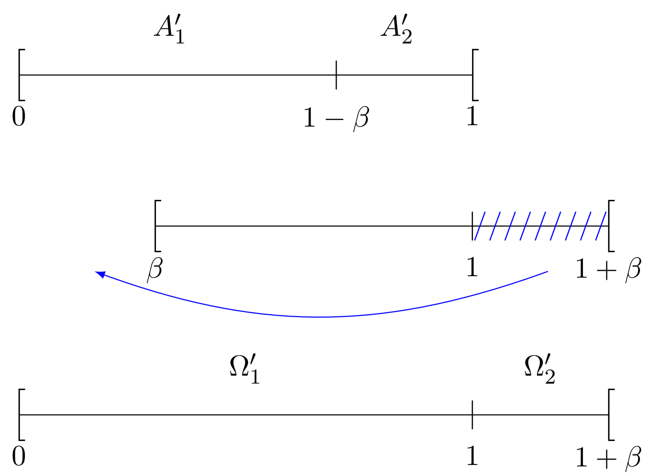


Figure 2: Example of exduction in dimension 1.

Our aim is to show that this construction is valid in a more general context.

6.1 Main result for exduction

Assumptions. Let w_0 be a natural coding of a minimal rotation of the d-torus, with elements $((\beta, M); (A : A_1, ..., A_{d+1}); x_0; (\beta_1, ..., \beta_{d+1}))$ and borders assignment $(A' : A'_1, ..., A'_{d+1})$. Let $a \in \mathcal{I} = \{1, ..., d+1\}$ a letter, and u_1,...,$u_{d+1} \in \mathcal{I}^*$ be such that:

1. For all i in $\mathcal{I}$, the word u_i starts with the letter a and admits no other occurrence of a;

2. the matrix $\mathcal{M} = (|u_j|_i)_{i,j} \in GL_{d+1}(\mathbb{Z})$.

CONSTRUCTION. Let $(\alpha_1, ..., \alpha_{d+1}) = (\beta_1, ..., \beta_{d+1})\mathcal{M}^{-1}$. We set $\alpha = \alpha_a$ and denote by L the additive subgroup of $\mathbb{R}^d$ given by: $L = \sum_{i=1, i\neq a}^{d+1}(\alpha_i - \alpha_a)\mathbb{Z}$. Furthermore, for $i \in \mathcal{I}$ and $k \in \{0, ..., |u_i|\}$, we introduce:

$$v_{i,k} = \sum_{l=0}^{k-1} \alpha_{u_i[l]} \in \mathbb{R}^d.$$

For $k \in K_i := \{0, ..., |u_i| - 1\}$, we set $A_{i,k} = A_i + v_{i,k}$ and $A'_{i,k} = A'_i + v_{i,k}$. Observe that for $k = |u_i|$ we have $v_{i,k} = \beta_i$, and that $A'_i + \beta_i \subset A'$. Now, let:

$$\Omega_j = \bigcup_{\substack{(i,k)\in\mathcal{I}\times K_i \\ u_i[k]=j}} A_{i,k} \qquad \text{and} \qquad \Omega'_j = \bigcup_{\substack{(i,k)\in\mathcal{I}\times K_i \\ u_i[k]=j}} A'_{i,k},$$

and set $\Omega := \cup_{j\in\mathcal{I}}\Omega_j$ and $\Omega' := \cup_{j\in\mathcal{I}}\Omega'_j$. At last, denote by σ the substitution given by $\sigma(i) = u_i$.

Theorem G. *The word $\sigma(w_0)$ is a natural coding of a minimal rotation of the d-torus, with elements $((\alpha, L); (\Omega : \Omega_1, ..., \Omega_{d+1}); x_0; (\alpha_1, ..., \alpha_{d+1}))$ and borders assignment $(\Omega' : \Omega'_1, ..., \Omega'_{d+1})$. Furthermore, we have $\Omega_a = A$, $\Omega'_a = A'$, and the induced map of $T_{sus} := r_{\Omega',L} \circ R_{\alpha,L} \circ p_L$ on the set with nonempty interior Ω_a is the map $T = r_{A',M} \circ R_{\beta,M} \circ p_M$.*

6.2 Proof of Theorem G

Let $G = \sum_{i\in\mathcal{I}} \beta_i\mathbb{Z}$ be the underlying group of the natural coding w_0.

Lemma 56. *The vectors $\alpha_1, ..., \alpha_{d+1}$ form a basis of G.*

Proof. By Lemma 19, the vectors $\beta_1, ..., \beta_{d+1}$ form a basis of the $\mathbb{Z}$-module G and, since $\mathcal{M} \in GL(\mathbb{Z})$, so do the vectors $\alpha_1, ..., \alpha_{d+1}$. □

Corollary 57. *The subgroup L is a lattice of $\mathbb{R}^d$ and the pair (α, L) is minimal.*

Proof. We write $G = \sum_{i\in\mathcal{I}} \alpha_i\mathbb{Z} = \sum_{i\in\mathcal{I}\backslash\{a\}}(\alpha_i - \alpha)\mathbb{Z} + \alpha\mathbb{Z}$. The group G being dense, with a similar argument than in Lemma 19, we show that the vectors $\alpha_i - \alpha$, for $i \in \mathcal{I}\backslash\{a\}$, form a basis of $\mathbb{R}^d$, and thus, that the group $L = \sum_{i\in\mathcal{I}\backslash\{a\}}(\alpha_i - \alpha)\mathbb{Z}$ is a lattice of $\mathbb{R}^d$. By density of G again, we obtain that the pair (α, L) is minimal. □

Lemma 58. *Let $x, y \in \mathbb{R}^d$. The three following assertions are equivalent:*

(i) the points x and y are equal modulo G;

(ii) there exists a unique $n \in \mathbb{Z}$ such that $p_L(y) = R^n_{\alpha,L}(p_L(x))$;

(iii) there exists a unique $m \in \mathbb{Z}$ such that $p_M(y) = R^m_{\beta,M}(p_M(x))$.

Proof. $(i \Rightarrow ii)$ Let $x, y \in \mathbb{R}^d$ be such that $y - x \in G$. Then, $\alpha_1, ...\alpha_{d+1}$ being a basis of G (Lemma 56), there exist $n_1, ..., n_{d+1} \in \mathbb{Z}$ such that $y - x = \sum_{i\in\mathcal{I}} n_i\alpha_i$. Therefore, $p_L(y) = R^n_{\alpha,L}(p_L(x))$, with $n = \sum_{i\in\mathcal{I}} n_i$. The pair (α, L) being minimal (57), the integer n is unique. $(ii \Rightarrow i)$ Let $x, y \in \mathbb{R}^d$ and $n \in \mathbb{Z}$ be such that $p_L(y) = R^n_{\alpha,L}(p_L(x))$. Then there exists $l \in L \subset G$ such that $y = x + n\alpha_a + l$, hence $y = x + g$, with $g = n\alpha_a + l \in G$. The equivalence between (i) and (iii) is given by Lemma 20 and by the minimality of (β, M).

□

Lemma 59. *Let $x, y \in A'$ that are equal modulo G. Denote by $(l_1, ..., l_{d+1})$ and $(m_1, ..., m_{d+1})$ the coordinates of the element $y - x$ with respect to the bases $(\alpha_1, .., .\alpha_{d+1})$ and $(\beta_1, ...\beta_{d+1})$ respectively. Then the integers $l_1, ..., l_{d+1}, m_1, ..., m_{d+1}$ are simultaneously nonnegative or nonpositive. Furthermore, if we set $l = \sum_{i \in \mathcal{I}} l_i$ and $m = \sum_{i \in \mathcal{I}} m_i$, we have $|l| \geq |m|$ and $l_a = m$.*

Corollary 60 (immediate)**.** *If $x, y \in A'$ are equal modulo G, then there exists two unique integers, l and m, such that $p_L(y) = R^l_{\alpha,L}(p_L(x))$ and $p_M(y) = R^m_{\beta,M}(p_M(x))$. Furthermore, l and m are simultaneously positive, negative or equal to zero.*

Proof of Lemma 59. Let $x, y \in A'$ that are equal modulo G. By Lemma 58, the exists $m \in \mathbb{Z}$ such that $p_M(y) = R^m_{\beta,M}(p_M(x))$. First, assume that $m \geq 0$. Since $x, y \in A'$, by definition of natural coding, there exist $m_1, ..., m_{d+1} \in \mathbb{N}$ such that:

$$\begin{aligned} y &= x + (\beta_1, ..., \beta_{d+1})(m_1, ..., m_{d+1})^t \\ &= x + (\alpha_1, ..., \alpha_{d+1})\mathcal{M}(m_1, ..., m_{d+1})^t \\ &= x + (\alpha_1, ..., \alpha_{d+1})(l_1, ..., l_{d+1})^t. \end{aligned}$$

Since $\mathcal{M} \in GL(\mathbb{Z})$ and has nonnegative entries, we have $l_i \geq 0$ for all $i \in \mathcal{I}$, and $l_1, ...l_{d+1}$ are simultaneously equal to zero if and only if $m_1, ..., m_{d+1}$ are simultaneously equal to zero if and only if $m = 0$; we also obtain $l \geq m$. We lead a symmetric argument for $m \leq 0$ and conclude that in both cases, $p_L(y) = R^l_{\alpha,L}(p_L(x))$, with $l_1, ..., l_{d+1}, m_1, ..., m_{d+1}$ simultaneously nonnegative or nonpositive, and $|l| \geq |m|$. At last, since the words u_i, for $i \in \mathcal{I}$, contain a unique occurrence of the letter a, the $a - th$ line of the matrix $\mathcal{M}$ only contains the entry 1, and $m = \sum_{i \in \mathcal{I}} m_i = l_a$. □

Proposition 61. *The following assertions are true.*

1. *The sets $\Omega_1, ..., \Omega_{d+1}$ are nonempty and open.*
2. *For all $i \in \mathcal{I}$, the set Ω_i is included and dense in Ω'_i.*
3. *The sets $\Omega'_1, ..., \Omega'_{d+1}$ are pairwise disjoint.*

Proof. Let $i \in \mathcal{I}$, and $k \in K_i$. The sets $A_{i,k}$ and $A'_{i,k}$ are the translated sets, by the vector $v_{i,k}$, of A_i and A'_i respectively, from which they inherit of the following properties: $A_{i,k}$ is nonempty, open, included and dense in $A'_{i,k}$. Now, let $j \in \mathcal{I}$. As the finite and nonempty union, for some $i \in \mathcal{I}$ and some $k \in K_i$, of the sets $A_{i,k}$, the set Ω_j is nonempty, open and furthermore included and dense in and Ω'_j, which is the union, for the same indices, of the sets $A'_{i,k}$.

We now prove that the sets $A'_{i,k}$ are pairwise disjoint. Let $y \in A'_{i_1,k_1} \cap A'_{i_2,k_2}$, with $i_1, i_2 \in \mathcal{I}$, $0 \leq k_1 < |u_{i_1}|$ and $0 \leq k_2 < |u_{i_2}|$. Denote $x_j = y - v_{i_j,k_j}$, which belongs to A'_{i_j}, for $j = 1, 2$. Since $x_1, x_2 \in A'$, and since $x_2 - x_1$ is an integer linear combination of $\alpha_1, ..., \alpha_{d+1}$, it comes that $x_2 - x_1 \in G$ and by Lemma 59 and Corollary 60, that the coordinates $(l_1, ..., l_{d+1})$ of $x_2 - x_1$ with respect to the basis $(\alpha_1, ..., \alpha_{d+1})$ are simultaneously nonnegative or nonpositive (w.l.o.g. say nonnegative). Therefore, if $k_1 = 0$, we have successively $k_2 = 0$, $x_1 = x_2$ and, since the sets $A'_1, ..., A'_{d+1}$ form a partition of A', $i_1 = i_2$. Now, assume that k_1 is positive. If $k_2 = 0$, then $y = x_2$, $l_a = 1$ and thereby $p_M(y) = R_{\beta,M}(p_M(x_1))$. Since $x_1 \in A'_{i_1}$, by definition of natural coding we have $y = x_1 + \beta_{i_1}$, which is conflicting with the hypothesis $k_1 < |u_1|$. So, if k_1 is positive, then k_2 is positive too. In this case, we have $l_a = |\text{pref}_{k_1}(u_{i_1})|_a - |\text{pref}_{k_2}(u_{i_2})|_a = 0$ since the word u_j, for $j \in \{1, 2\}$, contains exactly one occurrence of the letter a, at the first position. By Lemma 59 and Corollary 60 again, we have $p_M(x_2) = p_M(x_1)$; by M-simplicity of A', we obtain that $x_1 = x_2$ and $i_1 = i_2$. Then we have $0 = \sum_{l=k_1}^{k_2-1} \alpha_{u_{i_1}[l]}$, which implies $k_1 = k_2$. So the sets $A'_{i,k}$ are pairwise disjoint. We conclude, by observing that each $A'_{i,k}$, for $i \in \mathcal{I}$ and $k \in K_i$, belongs to exactly one set Ω'_j, that the sets $\Omega'_1, ..., \Omega'_{d+1}$ are pairwise disjoint. □

Proposition 62. *The set Ω' is a fundamental domain of the torus $\mathbb{T}_L$.*

Proof. We first show that the set A' is L-simple, from which we deduce that Ω' is L-simple; we conclude by proving that the projection map $p_L : \Omega' \mapsto \mathbb{T}_L$ is onto.

Let $x, y \in A'$ be such that $p_L(x) = p_L(y)$. Since $L \subset G$, the points x and y are equal modulo G and, by Corollary 60, there exist two integers l and m, that are unique, such that $p_L(y) = R^l_{\alpha,L}(p_L(x))$ and $p_M(y) = R^m_{\beta,M}(p_M(x))$. Here, we already have $l = 0$, and since $|m| \leq |l|$, we obtain $m = 0$; hence $p_M(y) = p_M(x)$ and by M-simplicity of A', $x = y$. This proves the L-simplicity of A'.

Now, let z_1 and $z_2 \in \Omega'$ be such that $p_L(z_1) = p_L(z_2)$. By construction, there exists two 3-tuples (x_1, i_1, k_1) and (x_2, i_2, k_2) with $i_1, i_2 \in \mathcal{I}$, such that for $j = 1, 2$, $x_j \in A'_{i_j} \subset A'$, $k_j \in K_{i_j}$ and $y_j = x_j + v_{i_j,k_j}$. We are going to show that $x_1 = x_2$. On one hand, the points x_1 and x_2 are in the same equivalent class modulo G (which is the class of z_1 and z_2), hence:

$$\begin{cases} \quad\; p_M(x_2) = R^{l+1}_{\beta,M}(p_M(x_1)) & \text{with } l \geq 0, \\ \text{or} \quad p_M(x_2) = R^{l}_{\beta,M}(p_M(x_1)) & \text{with } l \leq 0. \end{cases}$$

This implies, since x_1 and x_2 belongs to A', and $x_1 \in A'_{i_1}$:

$$\begin{cases} \quad\; p_L(x_2) = R^{m+|u_{i_1}|}_{\alpha,L}(p_L(x_1)) & \text{with } m \geq 0, \\ \text{or} \quad p_L(x_2) = R^{m}_{\alpha,L}(p_L(x_1)) & \text{with } m \leq 0. \end{cases}$$

On the other hand, without loss of generality, assume that $k_2 \leq k_1$. Then, if we set $y_1 = x_1 + v_{i_1,k_1-k_2}$, we have $p_L(x_2) = p_L(y_1) = R^{k_1-k_2}_{\alpha,L}(p_L(x_1))$ with $0 \leq k_1 - k_2 \leq |u_{i_1}| - 1$. The only possibility is thus $k_1 = k_2$, from which we deduce successively $p_L(x_1) = p_L(x_2)$, the equality $x_1 = x_2$ by L-simplicity of A', the equality $i_1 = i_2$ by pairwise disjointedness of the sets A'_j for $j \in \mathcal{I}$, and in the end, $z_1 = z_2$. Therefore, the set Ω' is L-simple.

At last, we show that $p_L : \Omega' \mapsto \mathbb{T}_L$ is onto. Let $\tilde{y} \in \mathbb{T}_L$ and denote $k = \min\{n \in \mathbb{N} | R^{-n}_{\alpha,L}(\tilde{y}) \in p_L(A')\}$, which is finite since (α, L) is minimal (Corollary 57) and $p_L(A')$ has nonempty interior. Denote also $\tilde{x} = R^{-k}_{\alpha,L}(\tilde{y})$ and $x = r_{A',L}(\tilde{x})$ its covering into A' (which is L-simple by the first part of the proof), and $i \in \mathcal{I}$ such that $x \in A'_i$. Then, since $R^{-k+|u_i|}_{\alpha,L}(\tilde{y}) \in p_L(A')$, by minimality of k we must have $k < |u_i|$. So if we set $y = x + v_{i,k}$, which belongs to Ω' by construction, it comes that $p_L(y) = p_L(x) + k\alpha = R^k_{\alpha,L}(\tilde{x}) = \tilde{y}$. We conclude that Ω' is a fundamental domain of the torus $\mathbb{T}_L$. □

Hereafter, we denote by $T_{sus} = r_{\Omega',L} \circ R_{\alpha,L} \circ p_L$ the covered rotation in the fundamental domain Ω'.

Proposition 63. *The following assertions are true.*

1. *For all $j \in \mathcal{I}$, for all $y \in \Omega'_j$, we have $T_{sus}(y) = y + \alpha_j$.*

2. *We have $\Omega_a = A$ and $\Omega'_a = A'$. Furthermore, the induced map of T_{sus} on the set with nonempty interior Ω'_a is the map T.*

3. *For all nonnegative integer n, $T^n_{sus}(x_0) \in \Omega_{\sigma(w)[n]}$.*

4. *For all $y \in \Omega'$ and for all $q \in \mathbb{N}$, there exists an extraction τ such that: (i) $T^{\tau(m)}_{sus}(x_0) \to_{m\to\infty} y$; (ii) for all $n \in \{0, ..., q\}$ and for all nonnegative integer m, $T^{\tau(m)+n}_{sus}(x_0) \in \Omega_{\iota_n}$, where ι_n is defined by $T^m_{sus}(y) \in \Omega'_{\iota_n}$.*

Proof. (1) Let $j \in \mathcal{I}$ and $y \in \Omega'_j$. Since $R_{\alpha,L}(p_L(y)) = p_L(y+\alpha_j)$, to prove the assertion, we need to show that $y+\alpha_j \in \Omega'$. Let (i,k) be the unique pair, with $i \in \mathcal{I}$ and $k \in K_i$, such that $y \in A'_{i,k}$. By definition of Ω'_j, the indices i and k satisfy $u_i[k] = j$. Thus, if $k < |u_i|-1$, then $y+\alpha_j \in A'_{i,k+1} \subset \Omega'$. Otherwise, let $x = y - v_{i,k}$. Then we have $y + \alpha_j = x + v_{i,|u_i|} = x + \beta_i \in A' \subset \Omega'$, which ends the proof.

(2) Since each u_i, for $i \in \mathcal{I}$, admits exactly one occurrence of the letter a, at the first position, we have by construction $\Omega_a = A$ and $\Omega'_a = A'$. Moreover, for all $i \in \mathcal{I}$ and for all $x \in A'_i$, we have $\min\{n \in \mathbb{N}^* | T^n_{sus}(x) \in \Omega'_a\} = |u_i|$, and by (1), $T^{|u_i|}_{sus}(x) = x + v_{i,|u_i|} = x + \beta_i = T(x)$.

(3) Let $(n_k)_{k\in\mathbb{N}}$ be the sequence of indices such that for all $k \in \mathbb{N}$, $T^{n_k}_{sus}(x_0) = T^k(x_0) \in A'$ (it actually belongs to A), which is well-defined by (2). Let $n \in \mathbb{N}$. Denote by k the unique nonnegative integer such that $n_k \leq n < n_{k+1}$ and by i the unique index in $\mathcal{I}$ such that $T^k(x_0) \in A_i$ (consequence of Proposition 61). Then, on one hand we have $\sigma(w)[n] = u_i[n-n_k]$, and on the other hand, $T^n_{sus}(x_0) = T^{n_k}_{sus}(x_0) + v_{i,n-n_k} \in A_{i,n-n_k} \subset \Omega_{u_i[n-n_k]}$. Finally, we have $T^n_{sus}(x_0) \in \Omega_{\sigma(w)[n]}$.

(4) Let $y \in \Omega'$ and $q \in \mathbb{N}$. By construction, there exist $i \in \mathcal{I}$ and $k \in K_i$ such that $x := y - v_{i,k} \in A'$. Denote by φ the extraction given by the definition of borders assignments associated with the natural coding w_0 for the point $x \in A'$ and the integer $k+q$. Then, we have that $T^{\varphi(m)}(x_0) \to_{m\to\infty} x$ and for any nonnegative integer m, the first $k+q+1$ letters of $f'(x)$ and $f(T^{\varphi(m)}(x_0))$ coincide, which immediately implies, since no image of letters by the substitution σ is the empty word, that for any m, the $k+q+1$ first letters of the words $\sigma(f'(x))$ and $\sigma(f(T^{\varphi(m)}(x_0))$ coincide as well. Since the sequence $(T^m(x_0))_{m\in\mathbb{N}}$ is a subsequence of $(T^m_{sus}(x_0))_{m\in\mathbb{N}}$, we can define an extraction ψ such that for all m, $T^{\psi(m)}_{sus}(x_0) = T^{\varphi(m)}(x_0)$. We finally set $\tau(m) = \psi(m) + k$. Thus, on one hand we have that $T^{\tau(m)}_{sus}(x_0) \to_{m\to\infty} y$; on the other hand, for all $n \in \{0, ..., q\}$ and for all nonnegative integer m, $T^{\tau(m)+n}_{sus}(x_0) \in \Omega_{\iota_n}$, where ι_n is defined by $T^n_{sus}(y) \in \Omega'_{\iota_n}$.

□

Proof of Theorem G. Proof of Theorem G results of Corollary 57 and Propositions 61, 62 and 63. □

6.3 Consequences for Arnoux-Rauzy and primitive C-adic words

Theorem G applies in particular to Arnoux-Rauzy and primitive C-words.

Proposition 64. *Let w be an Arnoux-Rauzy word and $\sigma \in AR^*$ (i.e., a finite product of substitutions in AR). Assume that w is a natural coding of a minimal rotation of the 2-torus. Then $\sigma(w)$ is also a natural coding of a minimal rotation of the 2-torus, whose elements and borders assignment can be explicitly described from the elements and the choice made for the borders assignment of the natural coding w. In particular, the piecewise translation vectors $\alpha_1, \alpha_2, \alpha_3$ of $\sigma(w)$ satisfy:*

$$(\alpha_1, \alpha_2, \alpha_3) = (\beta_1, \beta_2, \beta_3) M_\sigma^{-1},$$

where $M_\sigma = (|\sigma(j)|_i)_{i,j}$ is the incidence matrix of the substitution σ, and β_1, β_2 and β_3 are the piecewise translation vectors of w.

Proof. It is sufficient to prove the proposition for $\sigma \in AR$. For $\sigma \in AR$, we immediately have that the words $u_i = \sigma(i)$, for $i \in \{1, 2, 3\}$ satisfy the two assumptions of Theorem G. □

Proposition 65. *Let w be a primitive C-adic word and $\sigma \in C^*$ (i.e., a finite product of substitutions in C). Assume that w is a natural coding of a minimal rotation of the 2-torus. Then, there exists $k \in \mathbb{N}$ such that $S^k(\sigma(w))$ is also a natural coding of a minimal rotation of the 2-torus, whose elements and borders assignment can be explicitly described from the elements and the choice made* *for the borders assignment of the natural coding w. In particular, the piecewise translation vectors $\alpha_1, \alpha_2, \alpha_3$ of $S^k(\sigma(w))$ satisfy:*

$$(\alpha_1, \alpha_2, \alpha_3) = (\beta_1, \beta_2, \beta_3) M_\sigma^{-1},$$

where $M_\sigma = (|\sigma(j)|_i)_{i,j}$ is the incidence matrix of the substitution σ, and β_1, β_2 and β_3 are the piecewise translation vectors of w.

Proof. Again, it is sufficient to prove the proposition for the substitutions c_1 and c_2. Let w be a primitive C-adic word, and assume that w is a natural coding of a minimal rotation of the 2-torus. Denote by $(d_n)_{n\in\mathbb{N}}$ its directive sequence. By Theorem E, there exists a unique accepted path $e = (l_e, r_e)$ in the automaton of partial quotients for C-adic words (see Figure 5.3) that starts from the initial state 2 and such that l_e is a prefix of $(d_n)_{n\in\mathbb{N}}$. Denote by $w' = D_2(w)$ the derivated word of w relatively to the letter 2. By Theorem A, the word w' is a natural coding of a minimal rotation of the 2-torus. Besides, the path $e_1 = (c_1, s \circ c_1) \cdot e$ (where the symbol $\cdot$ denotes the concatenation operation), starting from the initial state 3, is accepted by the automaton. Therefore, the words $u_i = r_{e_1}(i)$, for $i \in \{1, 2, 3\}$, are the three return words to the letter 3 in the primitive C-adic word $c_1(w)$ and, thereby, satisfy the assumptions of Theorem G. Thus, the word $\sigma(w)$, where the substitution σ is given by $\sigma(i) = u_i$ for $i \in \{1, 2, 3\}$, is equal to the word $S^k(c_1(w))$ for a certain $k \in \mathbb{N}$, and is a natural coding of a minimal rotation of the 2-torus, whose elements and borders assignment are explicitly given by those of w', which themselves are explicitly given by the elements and the choice made for borders assignment of the natural coding w. In particular, if $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$ and $(\gamma_1, \gamma_2, \gamma_3)$ respectively denote the piecewise translation vectors of the natural codings $S^k(c_1(w))$, w and w', then we have, on one hand $(\gamma_1, \gamma_2, \gamma_2) = (\beta_1, \beta_2, \beta_3) M_{d_0} ... M_{d_{n_0-1}}$, where n_0 is the length of the path e, and on the other hand $(\gamma_1, \gamma_2, \gamma_2) = (\alpha_1, \alpha_2, \alpha_3) M_{c_1} M_{d_0} ... M_{d_{n_0-1}}$. Each matrix being invertible, we conclude that $(\alpha_1, \alpha_2, \alpha_3) = (\beta_1, \beta_2, \beta_3) M_{c_1}^{-1}$. A symmetric argument applies to $c_2(w)$. □

Theorem H. *For Arnoux-Rauzy and primitive C-adic subshifts, the property of being a natural coding of a minimal rotation of the 2-torus does not depend on any prefix of the directive sequence $(d_n)_{n\in\mathbb{N}}$.*

Proof. Let w be an Arnoux-Rauzy (resp. primitive C-adic) word, and denote by $(d_n)_{n\in\mathbb{N}}$ its directive sequence. Assume that w is a natural coding of a minimal rotation of the 2-torus. On the first hand, we showed in Proposition 64 (resp. Proposition 65) that for all $\sigma \in AR^*$ (resp. $\sigma \in C^*$), the subshift generated by $\sigma(w)$ is a natural coding of a minimal rotation of the 2-torus. On the other hand, we claim that for all $n_0 \in \mathbb{N}$, the Arnoux-Rauzy (resp. primitive C-adic) subshift with directive sequence $(d_n)_{n\geq n_0}$ is again a natural coding of a minimal rotation of the 2-torus. Indeed, by inducing on letters as many times as needed, we can find $n_1 \geq n_0$ such that the word with directive sequence $(d_n)_{n\geq n_1}$ is a natural coding of a minimal rotation of the 2-torus (Proposition 43, Theorems D and E). But then, by applying Proposition 64 (resp. 65) with $\sigma = d_{n_0} \circ ... \circ d_{n_1-1}$, we obtain that the Arnoux-Rauzy (resp. primitive C-adic) subshift with directive sequence $(d_n)_{n\geq n_0}$ is also a natural coding of a minimal rotation of the 2-torus. □

References

[Ada03] Boris Adamczewski. Balances for fixed points of primitive substitutions. *Theoretical Computer Science*, 307(1):47–75, 2003.

[AHS13] Artur Avila, Pascal Hubert, and Alexandra Skripchenko. On the Hausdorff dimension of the Rauzy gasket. *Bulletin de la Société mathématique de France*, 144:539–568, 2013.

[AI01] Pierre Arnoux and Shunji Ito. Pisot substitutions and Rauzy fractals. *Bulletin of the Belgian Mathematical Society - Simon Stevin*, 8(2):181–207, 2001.

[And18] Mélodie Andrieu. Autour du déséquilibre des mots C-adiques (text in french). *Proceedings of Mons Days of Theoretical Computer Science*, pages 51–54, 2018.

[And21] Mélodie Andrieu. A Rauzy fractal unbounded in all directions of the plane. *Comptes Rendus de l'Académie des sciences*, 359(4):399–407, 2021.

[AR91] Pierre Arnoux and Gérard Rauzy. Représentation géométrique de suites de complexité 2n+1. *Bulletin de la Société Mathématique de France*, 119:199–215, 1991.

[Arn20] Pierre Arnoux. Private communication. 2020.

[AS13] Pierre Arnoux and Štěpán Starosta. The Rauzy Gasket. In *Further Developments in Fractals and Related Fields*, pages 1–23. Springer, 2013.

[BB12] Nicolas Bédaride and Jean-François Bertazzon. Minoration of the complexity function associated to a translation on the torus. *Monatshefte für Mathematik*, 171, 2012.

[BD14] Valérie Berthé and Vincent Delecroix. Beyond substitutive dynamical systems: S-adic expansions. In *Lecture note 'Kokyuroku Bessatu'*, pages 81–123, 2014.

[Ber11] Valérie Berthé. Multidimensional Euclidean algorithms, numeration and substitutions. *Integers [electronic only]*, 2011.

[BFD$^+$15a] Valérie Berthé, Clelia De Felice, Francesco Dolce, Julien Leroy, Dominique Perrin, Christophe Reutenauer, and Giuseppina Rindone. Acyclic, connected and tree sets. *Monatshefte für Mathematik*, 176:521–550, 2015.

[BFD$^+$15b] Valérie Berthé, Clelia De Felice, Francesco Dolce, Julien Leroy, Dominique Perrin, Christophe Reutenauer, and Giuseppina Rindone. Maximal bifix decoding. *Discrete Mathematics*, 338:725–742, 2015.

[BJS12] Valérie Berthé, Timo Jolivet, and Anne Siegel. Substitutive Arnoux-Rauzy sequences have pure discrete spectrum. *Uniform Distribution Theory*, 7(1):173–197, 2012.

[Bos84] Michael Boshernitzan. A unique ergodicity of minimal symbolic flows with linear block growth. *J. Analyse Math.*, 44:77–96, 1984.

[BPS08] L'Ubomíra Balková, Edita Pelantová, and Wolfgang Steiner. Sequences with constant number of return words. *Monatshefte für Mathematik*, 155, 2008.

[BST19] Valérie Berthé, Wolfgang Steiner, and Jörg M. Thuswaldner. Geometry, dynamics, and arithmetic of S-adic shifts. *Annales de l'Institut Fourier*, 69:1347–1409, 2019.

[BST20] Valérie Berthé, Wolfgang Steiner, and Jörg M. Thuswaldner. Multidimensional continued fractions and symbolic codings of toral translations. *preprint arXiv:2005.13038v1*, 2020.

[BŠW13] Marcy Barge, Sonja Štimac, and R. F. Williams. Pure discrete spectrum in substitution tiling spaces. *Discrete and Continuous Dynamical Systems*, 33(2):579–597, 2013.

[CFM08] Julien Cassaigne, Sébastien Ferenczi, and Ali Messaoudi. Weak mixing and eigenvalues for Arnoux-Rauzy sequences. *Annales de l'Institut Fourier*, 58(6), 2008.

[CFZ00] Julien Cassaigne, Sébastien Ferenczi, and Luca Q. Zamboni. Imbalances in Arnoux-Rauzy sequences. *Annales de l'Institut Fourier*, 50:1265–1276, 2000.

[CH73] Ethan M. Coven and Gustav A. Hedlund. Sequences with minimal block growth. *Mathematical systems theory*, 7:138–153, 1973.

[CLL17] Julien Cassaigne, Sébastien Labbé, and Julien Leroy. A set of sequences of complexity 2n+1. In *WORDS 2017 Proceedings*, pages 144–156. Springer, 2017.

[Dur98] Fabien Durand. A characterization of substitutive sequences using return words. *Discrete Mathematics*, 179(1):89 – 101, 1998.

[Fer92] Sébastien Ferenczi. Bounded remainder sets. *Acta Arithmetica*, 61, 1992.

[Fog02] N. Pytheas Fogg. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[Fog20] N. Pytheas Fogg. Symbolic coding of linear complexity for generic translations of the torus, using continued fractions. *preprint arXiv:2005.12229v1*, 2020.

[JV00] Jacques Justin and Laurent Vuillon. Return words in Sturmian and episturmian words. *Theoretical Informatics and Applications*, 34:343–356, 2000.

[LM95] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding.* Cambridge University Press, 1995.

[MH38] Marston Morse and Gustav A. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60(4):815–866, 1938.

[MH40] Marston Morse and Gustav A. Hedlund. Symbolic dynamics ii. Sturmian trajectories. *American Journal of Mathematics*, 62(1):1–42, 1940.

[Rau82] Gérard Rauzy. Nombres algébriques et substitutions. *Bulletin de la Société Mathématique de France*, 110:147–178, 1982.

[Rau84] Gérard Rauzy. Ensembles à restes bornés. *Séminaire de Théorie des Nombres de Bordeaux*, pages 1–12, 1984.

[RSZ09] Gwénaël Richomme, Kalle Saari, and Luca Zamboni. Balance and abelian complexity of the Tribonacci word. *Advances in Applied Mathematics*, 45(2):212–231, 2009.

[Sch00] Fritz Schweiger. *Multidimensional Continued Fractions.* Oxford Science Publications. Oxford University Press, 2000.

[Thu19] Jörg M. Thuswaldner. S-adic sequences: a bridge between dynamics, arithmetic, and geometry. *arXiv:1908.05954*, 2019.

[Vui01] Laurent Vuillon. A characterization of Sturmian words by return words. *European Journal of Combinatorics*, 22:363–375, 2001.