Lattices and multi-dimensional words

Valérie Berthé [a],

[a] *LIRMM, 161 rue Ada, F-34392 Montpellier Cedex 05, France*

Robert Tijdeman [b],

[b] *Mathematical Institute, Leiden University, Postbus 9512, 2300 RA Leiden, The Netherlands*

Abstract

In the present paper we develop a formalism to generate multi-dimensional words using lattices which generalizes the construction of real numbers (one-dimensional words) from a sequence of partial quotients using continued fractions. The construction was introduced in a special case by Simpson and Tijdeman in order to derive a multi-dimensional generalisation of the theorem of Fine and Wilf. We show that the produced multi-dimensional words are intrinsically connected with k-dimensional Sturmian words.

Key words: Multi-dimensional combinatorics on words, tilings, lattices, discrete planes, multi-dimensional Sturmian sequences.

1 Introduction

Let $k+1$ vectors $\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k$ in $\mathbb{Z}^k$ be given such that all the subsets of k vectors are linearly independent. Consider the sublattice $\Lambda := \mathbb{Z}(\vec{v}_1 - \vec{v}_0) + \cdots + \mathbb{Z}(\vec{v}_k - \vec{v}_0)$ of the lattice $L := \mathbb{Z}\vec{v}_0 + \mathbb{Z}\vec{v}_1 + \cdots + \mathbb{Z}\vec{v}_k$. For convenience we assume $L = \mathbb{Z}^k$. Let D denote the cardinality of $\mathbb{Z}^k/\Lambda$. We show that the multiples of $\vec{v}_0$ represent the D cosets of Λ. This induces a numbering $0, 1, \cdots, D-1$ of the cosets such that $\sum_{i=0}^{k} x_i \vec{v}_i$ belongs to coset $\sum_{i=0}^{k} x_i \pmod D$ which we shall indicate by $g(\sum_{i=0}^{k} x_i \vec{v}_i)$. The colouring map $\chi : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ is a projection of g, thus constant on cosets of Λ. By considering a fundamental domain A of $\mathbb{Z}^k/\Lambda$ the function g induces a roundwalk $\mathbf{w}$ given by

Email addresses: berthe@lirmm.fr (Valérie Berthé), tijdeman@math.leidenuniv.nl (Robert Tijdeman).

$\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_{D-1}, \vec{w}_D = \vec{w}_0$ through A. In our construction $\mathbf{w}$ has the special property that $\vec{w}_{i+1} - \vec{w}_i \in \{\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k\}$ for $i = 0, 1, \cdots, D-1$. The function χ induces a colouring of A. This is worked out in Section 2 and illustrated by an example.

In Section 3 we introduce extension rules which extend the roundwalk $\mathbf{w}$ through a lattice domain A of $\mathbb{Z}^k/\Lambda$ to a roundwalk $\mathbf{w}^*$ through a larger lattice domain A^* of $\mathbb{Z}^k/\Lambda^*$ where Λ^* is a lattice generated from Λ by a substitution rule $\sigma : \{0, 1, \cdots, k\} \to \{0, 1, \cdots, k\}^\star$ applied to $\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k$. Here $\{0, 1, \cdots, k\}^\star$ denotes the set of nonempty finite words with letters from $\{0, 1, \cdots, k\}$. The word W^* defines a function g^* numbering the cosets of Λ^*. The induced colouring map $\chi^* : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ will have the property that it is constant on cosets of Λ^* and coincides with χ on A. We prove, under suitable conditions, that the functions g^* and χ^* can be described in terms of g, χ and a matrix M, the so-called substitution matrix, which represents the numbers of each letter in the words $\sigma(0), \sigma(1), \cdots, \sigma(k)$. So the precise order of letters in these words is irrelevant for the definition of g^* and χ^*. In Sec. 3.1 we consider basic extension rules and in Sec. 3.3 rules generated by substitutions of Arnoux-Rauzy type.

Section 4 studies the iteration of extension rules producing a tower of lattices with corresponding matrices $M^{(n)}$. By extending the colouring of the domains of the roundwalk we thus generate a sequence of finite k-dimensional words $(W^{(n)})_{n\geq 1}$ through larger and larger domains $A^{(n)}$ $(n = 1, 2, \cdots)$ with letters from $\{0, 1, \cdots, k\}$. By induction we get infinite sequences of functions $g^{(n)} : \mathbb{Z}^k \to \{0, 1, \cdots, D^{(n)} - 1\}$ and $\chi^{(n)} : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$. We give explicit formulas for $(\chi^{(n)})_{n\geq 1}$ in terms of $g^{(n)}$ and $M^{(n)}$. Furthermore we show that in case of Rauzy extension steps a recurrence relation for the frequencies of the letters can be obtained in a simple way.

In Section 5 we consider the infinite limit words. Obviously the sequence $(W^{(n)})_{n\geq 1}$ has a limit word, since the words become larger and larger and as soon as a place has got a χ-value, it remains constant. However, the word need not be defined on $\mathbb{Z}^k$. Therefore we investigate whether $\chi := \lim_{n\to\infty} \chi^{(n)} : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ exists in which case every word $W^{(n)}$ is the restriction of χ to $A^{(n)}$. In Theorem 5.3 we show that under mild conditions χ exists indeed and that it represents a multi-dimensional Sturmian word. Furthermore we discuss which multi-dimensional Sturmian words can be limit words of towers of Rauzy extensions and which of basic extensions. In Part II of the paper we shall turn to more general questions and investigate when the roundwalks are space filling.

2 Roundwalks and lattices

The aim of this section is to introduce the algebraic framework we will use throughout the paper. This formalism was introduced in a special case in [13].

2.1 Lattices

Let $k+1$ vectors $\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k$ in $\mathbb{Z}^k$ be given such that $f_0 := \det(\vec{v}_1, \cdots, \vec{v}_k) \neq 0, f_i := \det(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{i-1}, -\vec{v}_0, \vec{v}_{i+1}, \cdots, \vec{v}_k) \neq 0$ for $i = 1, \cdots, k$ and f_0,f_1, $\cdots, f_k$ all have the same sign. This implies that $\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_k$ are not on the same side of some hyperplane. Note that one deduces from Cramer's rule that

$$f_0\vec{v}_0 + f_1\vec{v}_1 + \cdots + f_k\vec{v}_k = \vec{0}.$$

Let L be the lattice $\mathbb{Z}\vec{v}_0 + \mathbb{Z}\vec{v}_1 + \cdots + \mathbb{Z}\vec{v}_k$. Then on the one hand, $\det(L) \mid f_i$ for $i = 0, 1, \cdots, k$, whence $\det(L) \mid d := \gcd(f_0, f_1, \cdots, f_k)$. On the other hand, if $\vec{w}_1, \cdots, \vec{w}_k$ is some basis of L, then $\vec{w}_j = \rho_{0,j}\vec{v}_0 + \rho_{1,j}\vec{v}_1 + \cdots + \rho_{k,j}\vec{v}_k$ with $\rho_{0,j}, \rho_{1,j}, \cdots, \rho_{k,j} \in \mathbb{Z}$ for $j = 1, \cdots, k$. Hence $\det(L) = \mid \det(\vec{w}_1, \cdots, \vec{w}_k) \mid$ is a linear combination of $f_0, f_1, \cdots, f_k$ and therefore divisible by d. We conclude that $\det(L) = d$ and that L has exactly d cosets in $\mathbb{Z}^k$. If $d = 1$, then $L = \mathbb{Z}^k$.

Let Λ be the lattice $\mathbb{Z}(\vec{v}_1 - \vec{v}_0) + \cdots + \mathbb{Z}(\vec{v}_k - \vec{v}_0)$. Then

$$\det(\Lambda) = \mid \det(\vec{v}_1 - \vec{v}_0, \cdots, \vec{v}_k - \vec{v}_0) \mid = \mid f_0 + f_1 + \cdots + f_k \mid.$$

Note that Λ is a sublattice of L of index

$$D := \det(\Lambda)/\det(L) = d_0 + d_1 + \cdots + d_k,$$

where $d_i := |f_i|/d$ for $i = 0, 1, \cdots, k$. Hence $\gcd(d_0, d_1, \cdots, d_k) = 1$ and

$$d_0\vec{v}_0 + d_1\vec{v}_1 + \cdots + d_k\vec{v}_k = \vec{0}.$$

Moreover, if $a_0\vec{v}_0 + a_1\vec{v}_1 + \cdots + a_k\vec{v}_k = \vec{0}$ for some $a_0, a_1, \cdots, a_k \in \mathbb{Z}$, then $a_i = td_i$ for some integer t and $i = 0, 1, \cdots, k$.

2.2 Roundwalks and their codings

Let us now introduce the notion of a roundwalk.

We claim that $\vec{v}_0$ generates L/Λ. Suppose, on the contrary, that for some i, j with $0 \leq i < j < D$, $iv_0 \equiv jv_0 \pmod{\Lambda}$. Then $(j-i)\vec{v}_0 = \lambda_1(\vec{v}_1 - \vec{v}_0) + \cdots +$

$\lambda_k(\vec{v}_k - \vec{v}_0)$ for some integers $\lambda_1, \cdots, \lambda_k$. Hence

$$(j - i + \lambda_1 + \cdots + \lambda_k)\vec{v}_0 - \lambda_1\vec{v}_1 \cdots - \lambda_k\vec{v}_k = \vec{0}.$$

Therefore $j - i + \lambda_1 + \cdots + \lambda_k = td_0, -\lambda_1 = td_1, \cdots, -\lambda_k = td_k$ for some $t \in \mathbb{Z}$. It follows that $j - i = t(d_0 + d_1 + \cdots + d_k) = tD$ for some $t \in \mathbb{Z}$ which yields a contradiction. Thus the multiples $i\vec{v}_0$ $(0 \leq i < D)$ are elements in L which are distinct modulo Λ. Hence $\{\vec{0}, \vec{v}_0, \cdots, (D-1)\vec{v}_0\}$ is a complete set of representatives of L/Λ. Observe that $D\vec{v}_0 \equiv \vec{0} \pmod{\Lambda}$.

By a *roundwalk* $\mathbf{w}$ through L/Λ we mean a sequence $\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_{D-1}$, $\vec{w}_D = \vec{w}_0$ of $D+1$ elements from $\mathbb{Z}^k$ such that $\vec{w}_m - m\vec{v}_0 \in \Lambda$ for $m = 0, 1, \cdots, D$ and $\vec{w}_{i+1} - \vec{w}_i \in \{\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k\}$ for $i = 0, 1, \cdots, D-1$. Since modulo Λ it does not make any difference whether one adds $\vec{v}_0, \vec{v}_1, \cdots$, or $\vec{v}_k$, the vectors $\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_{D-1}$ are in D different cosets of L/Λ and therefore distinct. Since $\vec{w}_D = \vec{w}_0$, there are exactly d_j numbers $i \in \{0, 1, \cdots, D-1\}$ such that $\vec{w}_{i+1} - \vec{w}_i = \vec{v}_j$ for $j = 0, 1, \cdots, k$. Conversely, suppose we are given D vectors $\vec{w}_0 = \vec{w}_D, \vec{w}_1, \ldots, \vec{w}_{D-1}$ such that there are exactly d_j numbers $i \in \{0, 1, \cdots, D-1\}$ with $\vec{w}_{i+1} - \vec{w}_i = \vec{v}_j$ for $j = 0, 1, \cdots, k$. Then $\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_{D-1}, \vec{w}_D (= \vec{w}_0)$ form a roundwalk through L/Λ.

We introduce some further notation. We define the *domain* $A = A(\mathbf{w})$ of the roundwalk $\mathbf{w}$ as the subset $\{\vec{w}_0, \ldots, \vec{w}_{D-1}\}$ of $\mathbb{Z}^k$. So $A(\mathbf{w})$ represents a complete set of representatives of L/Λ. We define the *coding* w of a roundwalk $\vec{w}_0, \vec{w}_1, \ldots, \vec{w}_{D-1}, \vec{w}_D = \vec{w}_0$, with $\vec{w}_{i+1} - \vec{w}_i \in \{\vec{v}_0, \cdots, \vec{v}_k\}$ for $i = 0, 1, \cdots, D-1$ as the finite word $w = w_0 \ldots w_{D-1}$ over the alphabet $\{0, 1, \ldots, k\}$ defined by $w_i = j$ if $\vec{w}_{i+1} - \vec{w}_i = \vec{v}_j$ for $0 \leq i \leq D-1$. Observe that given the vectors $\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_k$, and the vector $\vec{w}_0$, the roundwalk $\mathbf{w}$ is perfectly determined by its coding w.

2.3 An example

We illustrate the definitions by an example.

Example 1 We take $k = 2, \vec{v}_0 = (1, 4), \vec{v}_1 = (3, 1), \vec{v}_2 = (-2, -3)$. Hence $f_0 = -7, f_1 = -5, f_2 = -11, d = 1, L = \mathbb{Z}^2, d_0 = 7, d_1 = 5, d_2 = 11, D = 23, \Lambda = \mathbb{Z}(2, -3) + \mathbb{Z}(-3, -7)$. We make a roundwalk $\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_{23} = \vec{w}_0$ through L/Λ where

$$\vec{w}_{i+1} - \vec{w}_i = \begin{cases} \vec{v}_0 \text{ for } i = 0, 2, 6, 9, 12, 15, 19, \\ \vec{v}_1 \text{ for } i = 4, 8, 13, 17, 21, \\ \vec{v}_2 \text{ for } i = 1, 3, 5, 7, 10, 11, 14, 16, 18, 20, 22. \end{cases}$$

We indicate the vectors in Figure 1 with h_j where h indicates the index of $\vec{w}_h$ and j the index of $\vec{v}_j$ when $\vec{v}_j = \vec{w}_{h+1} - \vec{w}_h$. The coding w of the roundwalk $\mathbf{w}$ is the sequence of the subscripts

$$0\ 2\ 0\ 2\ 1\ 2\ 0\ 2\ 1\ 0\ 2\ 2\ 0\ 1\ 2\ 0\ 2\ 1\ 2\ 0\ 2\ 1\ 2,$$

where j means that the next jump is $\vec{v}_j$. Note that there are $d_0 = 7$ numbers $0, d_1 = 5$ numbers $1, d_2 = 11$ numbers 2 in the coding word indeed.

			16_2	10_2
		3_2	20_2	14_2
	13_1	7_2	1_2	18_2
	17_1	11_2	5_2	22_2
4_1	21_1	15_0	9_0	
8_1	2_0	19_0		
12_0	6_0	0_0		

Figure 1. Roundwalk with jumps from $\{\vec{v}_0, \vec{v}_1, \vec{v}_2\}$

In fact, the points of the roundwalk through L/Λ generate a periodic tiling of $L = \mathbb{Z}^k$, since they are a complete set of representatives of L/Λ, as illustrated in Figure 2, where one finds a finite part of the periodic tiling of $\mathbb{Z}^2$ by the pattern in Figure 1, the period vectors being $\vec{v}_1 - \vec{v}_0 = (2, -3)$ and $\vec{v}_2 - \vec{v}_0 = (-3, -7)$. We can consider the (upper) numbers in Figures 1 and 2 as a function g from the cosets L/Λ to $\{0, 1, \cdots, 22\}$, numbering the cosets.

$\cdots$	19_0	13_1	7_2	1_2	18_2	12_0	6_0	0_0	17_1	$\cdots$
$\cdots$	0_0	17_1	11_2	5_2	22_2	16_2	10_2	4_1	21_1	$\cdots$
$\cdots$	4_1	21_1	15_0	9_0	3_2	20_2	14_2	8_1	2_0	$\cdots$
$\cdots$	8_1	2_0	19_0	13_1	7_2	1_2	18_2	12_0	6_0	$\cdots$
$\cdots$	12_0	6_0	0_0	17_1	11_2	5_2	22_2	16_2	10_2	$\cdots$
$\cdots$	16_2	10_2	4_1	21_1	15_0	9_0	3_2	20_2	14_2	$\cdots$
$\cdots$	20_2	14_2	8_1	2_0	19_0	13_1	7_2	1_2	18_2	$\cdots$
$\cdots$	1_2	18_2	12_0	6_0	0_0	17_1	11_2	5_2	22_2	$\cdots$

Figure 2. The periodic tiling generated by the roundwalk

2.4 *Colouring of the roundwalks*

We return to the general case. For the sake of simplicity we assume in the sequel that $L = \mathbb{Z}^k$. Hence $d = 1$. We define the function g from $\mathbb{Z}^k$ to $\{0, 1, \cdots, D-1\}$ by $g(\vec{x}) = m$ whenever $\vec{x} - m\vec{v}_0 \in \Lambda$. It follows firstly that g is constant on cosets of L/Λ, and secondly that g is a linear function. Thus $g(\vec{0}) = 0$ and if $\vec{x} = (x_1, \cdots, x_k) \in \mathbb{Z}^k$, then

$$g(\vec{x}) = \sum_{i=1}^{k} x_i g(\vec{e}_i) \pmod{D} \tag{1}$$

where $\vec{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ is the unit vector with 1 at place i for $i = 1, 2, \cdots, k$, and where $a \pmod{b}$ is defined as the number c with $c \equiv a \pmod{b}$ and $0 \leq c < b$. Since $\vec{v}_i - \vec{v}_0 \in \Lambda$ we have $g(\vec{v}_i) = 1$ for $0 \leq i \leq k$, hence

$$g(\sum_{i=0}^{k} x_i \vec{v}_i) = \sum_{i=0}^{k} x_i \pmod{D}.$$

For instance in Example 1, when considering the point 0_0 as the origin in $\mathbb{Z}^2$, one has (Figure 2) $g(\vec{e}_1) = 17 = -6 \pmod{23}$, $g(\vec{e}_2) = 19 = -4 \pmod{23}$, and thus the function $g : \mathbb{Z}^2 \to \{0, \ldots, 22\}$ is given by

$$g(m, n) = -6m - 4n \pmod{23}.$$

Note that the restriction of g to $A(\mathbf{w})$ in Figure 1 precisely indicates the route of the walk.

To the roundwalk $\mathbf{w}$ given by $\vec{w}_0, \cdots, \vec{w}_D (= \vec{w}_0)$ we have associated a one-dimensional word $w = w_0 \cdots w_{D-1}$ over the alphabet $\{0, 1, \ldots, k\}$ which we called the *coding* of $\mathbf{w}$. Now we shall associate to $\mathbf{w}$ a k-dimensional word W which we call the *colouring* of $\mathbf{w}$. To that purpose, we introduce a colouring function $\chi : \mathbb{Z}^k \to \{0, 1, \ldots, k\}$ which will "colour" the roundwalk. More precisely, $\{0, 1, \ldots, D-1\}$ is split into $k+1$ sets $I_0, \ldots, I_k$ (in practice always consecutive blocks of integers modulo D) and

$$\chi(\vec{x}) = l \iff g(\vec{x}) \in I_l \quad \text{for } l \in \{0, 1, \cdots k\}, \vec{x} \in A(\mathbf{w}). \tag{2}$$

Observe that one could colour the roundwalk $\mathbf{w}$ with any number of colours. Our motivation to consider exactly $k+1$ colours is to make the connection with k-dimensional Sturmian words.

Example 1 (continued). We split the roundwalk $\mathbf{w}$ in Figure 1 into three nonempty parts, e.g. we replace numbers i for $5 \leq i < 14$ with 0, i for

$14 \le i < 22$ with 1 and numbers i for $i = 22$ or $0 \le i < 5$ with 2. We omit the subscripts and indicate the place of 0_0 by underlining. This yields Figure 3.

			1	0
		2	1	1
	0	0	2	1
	1	0	0	2
2	1	1	0	
0	2	1		
0	0	$\underline{2}$		

Figure 3. The colouring W of $\mathbf{w}$

Thus χ is given by

$$\chi(m,n) = \begin{cases} 0 \text{ if} & 5 \le -6m - 4n \pmod{23} < 14 \\ 1 \text{ if} & 14 \le -6m - 4n \pmod{23} < 22 \\ 2 \text{ otherwise,} & \end{cases}$$

and $W = \chi|_{A(\mathbf{w})}$ is presented in Figure 3.

3 Extension rules

We construct an extended k-dimensional word by extending the coding of a roundwalk.

3.1 The basic extension step

Starting from the situation described in the previous section we introduce the extension step $S_{i,j}$ for $i, j \in \{0, 1, \cdots, k\}$ with $i \ne j$. Put $\vec{v}_i^* = \vec{v}_i - \vec{v}_j, \vec{v}_h^* = \vec{v}_h$ for $h = 0, 1, \cdots, k; h \ne i$. Note that the lattice $L^* := \mathbb{Z}\vec{v}_0^* + \cdots + \mathbb{Z}\vec{v}_k^*$ equals $L := \mathbb{Z}\vec{v}_0 + \cdots + \mathbb{Z}\vec{v}_k$ which we have assumed to be $\mathbb{Z}^k$. Put $d_0^* :=$

$|\det(\vec{v}_1^*, \cdots, \vec{v}_k^*)|, d_h^* := |\det(\vec{v}_1^*, \cdots, \vec{v}_{h-1}^*, -\vec{v}_0^*, \vec{v}_{h+1}^*, \cdots, \vec{v}_k^*)|$ for $h = 1, \cdots, k$. Then

$$d_h^* = d_h \text{ for } h \neq j \text{ and } d_j^* = d_i + d_j \quad (h, i, j \in \{0, 1, \cdots, k\}, i \neq j).$$

Let Λ^* be the lattice $\mathbb{Z}(\vec{v}_1^* - \vec{v}_0^*) + \cdots + \mathbb{Z}(\vec{v}_k^* - \vec{v}_0^*)$. Then $D^* := \det(\Lambda^*) = d_0^* + d_1^* + \cdots + d_k^* = D + d_i$.

Suppose we have a roundwalk $\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_{D-1}, \vec{w}_D (= \vec{w}_0) \in \mathbb{Z}^k$ through $\mathbb{Z}^k/\Lambda$ with $\vec{w}_{h+1} - \vec{w}_h \in \{\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k\}$ for $h = 0, 1, \cdots, D-1$. Then there are exactly d_j numbers $h \in \{0, 1, \cdots, D-1\}$ such that $\vec{w}_{h+1} - \vec{w}_h = \vec{v}_j$ for $j = 0, 1, \cdots, k$. We may construct a roundwalk $\vec{w}_0^* = \vec{w}_0, \vec{w}_1^*, \cdots, \vec{w}_{D^*-1}^*, \vec{w}_{D^*}^* = \vec{w}_0^* \in \mathbb{Z}^k$ with $\vec{w}_{h+1}^* - \vec{w}_h^* \in \{\vec{v}_0^*, \vec{v}_1^*, \cdots, \vec{v}_k^*\}$ for $h = 0, 1, \cdots, D^* - 1$ through $\mathbb{Z}^k/\Lambda^*$ by inserting for every h with $\vec{w}_{h+1} - \vec{w}_h = \vec{v}_i$ either $\vec{w}_h + \vec{v}_h^*$ or $\vec{w}_h + \vec{v}_j^*$ in between $\vec{w}_h$ and $\vec{w}_{h+1}$. Then the new jumps in the roundwalk are of the form $\vec{v}_i^*, \vec{v}_j^*, \vec{v}_i - \vec{v}_i^* = \vec{v}_j^*$ or $\vec{v}_i - \vec{v}_j^* = \vec{v}_i^*$, and therefore all jumps in the new roundwalk belong to $\{\vec{v}_0^*, \vec{v}_1^*, \cdots, \vec{v}_k^*\}$. By arguments given in the previous section, the vectors $\vec{w}_0^*, \vec{w}_1^*, \cdots, \vec{w}_{D^*-1}^*$ represent D^* different cosets of $\mathbb{Z}^k/\Lambda^*$ and are therefore distinct, whereas the new roundwalk contains exactly d_h^* jumps $\vec{v}_h^*$ for $h = 0, 1, \cdots, k$. The cycle is not uniquely determined as we have d_i times a choice out of two for the vector to be inserted. We say that the insertion is done in the *canonical* way if each time $\vec{w}_h + \vec{v}_i^*$ is inserted and in the *anti-canonical* way if each time $\vec{w}_h + \vec{v}_j^*$ is inserted.

The extension step $S_{i,j}$ has the following effect on the coding sequence: every i is replaced by ij in the canonical case, by ji in the anti-canonical case, whereas we have free choice between ij and ji for every i in the general case. This action can be described by means of the formalism of substitutions. Let us recall that a *substitution* σ is an application from an alphabet $\mathcal{A}$ into the set $\mathcal{A}^\star - \{\varepsilon\}$ of nonempty finite words on $\mathcal{A}$; it extends to a morphism of $\mathcal{A}^\star$ by concatenation, that is, $\sigma(WW') = \sigma(W)\sigma(W')$ and $\sigma(\varepsilon) = \varepsilon$. It also extends in a natural way to a map defined over $\mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$.

Let $\mathbf{w}$ and $\mathbf{w}^*$ denote respectiveley the roundwalks $\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_{D-1}, \vec{w}_D (= \vec{w}_0)$ and $\vec{w}_0^*, \vec{w}_1^*, \cdots, \vec{w}_{D^*-1}^*, \vec{w}_{D^*}^* (= \vec{w}_0^*)$. Let w and w^* denote the codings of the roundwalks $\mathbf{w}$ and $\mathbf{w}^*$, respectively. The action of the basic extension step $S_{i,j}$ on the coding of the roundwalk $\mathbf{w}$ is described in the canonical case by the substitution $\sigma_{i,j} : i \mapsto ij$, and $h \mapsto h$ for $h \neq i$, and in the anti-canonical case by $\tilde{\sigma}_{i,j} : i \mapsto ji$, and $h \mapsto h$ for $h \neq i$.

The lattice Λ^* induces a linear function $g^* : \mathbb{Z}^k \to \{0, 1, \cdots, D^* - 1\}$ by $g^*(\vec{x}) = m$ whenever $\vec{x} - m\vec{v}_0^* \in \Lambda^*$. This function is constant on cosets of $\mathbb{Z}^k/\Lambda^*$ and satisfies $g^*(0) = 0$. We define the projection $\chi^* : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ of g^* by $\chi^*(\vec{x}) = \chi(\vec{y})$ where $\vec{y}$ is the lastly visited place of $A(\mathbf{w})$ when reaching $\vec{x}$ along $\mathbf{w}^*$. In particular $\chi^*(\vec{x}) = \chi(\vec{x})$ if $\vec{x} \in A$.

Example 2 Let us apply the basic extension step $S_{1,0}$ to the roundwalk of Example 1. Recall that $k = 2, \vec{v}_0 = (1,4), \vec{v}_1 = (3,1), \vec{v}_2 = (-2,-3)$, hence $d_0 = 7, d_1 = 5, d_2 = 11, D = 23, \Lambda = \mathbb{Z}(2,-3) + \mathbb{Z}(-3,-7)$. The coding w of the roundwalk $\mathbf{w}$ is the word

$$0\ 2\ 0\ 2\ 1\ 2\ 0\ 2\ 1\ 0\ 2\ 2\ 0\ 1\ 2\ 0\ 2\ 1\ 2\ 0\ 2\ 1\ 2.$$

We now apply step $S_{1,0}$. Hence $\vec{v}_0^* = (1,4), \vec{v}_1^* = (2,-3), \vec{v}_2^* = (-2,-3)$, $d_0^* = 12, d_1^* = 5, d_2^* = 11, D^* = 28, \Lambda^* = \mathbb{Z}(1,-7) + \mathbb{Z}(-3,-7)$. Five vectors are inserted, since $d_1 = 5$. If we apply the canonical insertion, see Figure 4, then there are five new vectors $\vec{w}_5^*, \vec{w}_{10}^*, \vec{w}_{16}^*, \vec{w}_{21}^*, \vec{w}_{26}^*$ which are the translates of $\vec{w}_4^* = \vec{w}_4$, $\vec{w}_9^* = \vec{w}_8$, $\vec{w}_{15}^* = \vec{w}_{13}$, $\vec{w}_{20}^* = \vec{w}_{17}, \vec{w}_{25}^* = \vec{w}_{21}$ by $\vec{v}_1^* = (2,-3)$. The translated vectors are just the ones in Figure 1 with index 1. The old vectors keep their subscripts, whereas the five newly introduced vectors get subscript 0. The new coding $w^* = \sigma_{1,0}(w)$ of the roundwalk $\mathbf{w}^*$ is

$$0\ 2\ 0\ 2\ 10\ 2\ 0\ 2\ 10\ 0\ 2\ 2\ 0\ 10\ 2\ 0\ 2\ 10\ 2\ 0\ 2\ 10\ 2.$$

If we make the anti-canonical insertions $\vec{w}_h + \vec{v}_j^*$ instead, there are again five new vectors $\vec{w}_5^*$,$\vec{w}_{10}^*$,$\vec{w}_{16}^*$,$\vec{w}_{21}^*$,$\vec{w}_{26}^*$ which are the translates of $\vec{w}_4^* = \vec{w}_4$, $\vec{w}_9^* = \vec{w}_8$, $\vec{w}_{15}^* = \vec{w}_{13}$, $\vec{w}_{20}^* = \vec{w}_{17}, \vec{w}_{25}^* = \vec{w}_{21}$, but now by $\vec{v}_0^* = (1,4)$. The new vectors get index 1 and the original five vectors get index 0 instead of 1 (see Figure 5). The coding w^* of the roundwalk $\mathbf{w}^*$ is in this case

$$w^* = \sigma_{0,1}(w) = 0\ 2\ 0\ 2\ 01\ 2\ 0\ 2\ 01\ 0\ 2\ 2\ 0\ 01\ 2\ 0\ 2\ 01\ 2\ 0\ 2\ 01\ 2.$$

Observe that the patterns are connected. Actually the (upper) numbers form the restriction W^* to some set $A(\mathbf{w}^*)$ of the function $g^* : \mathbb{Z}^2 \to \{0, 1, \cdots, D^* - 1\}$ (as defined in (1)). Since g^* is linear and $g^*(0) = 0$, it is determined by $g^*(1,0)$ and $g^*(0,1)$. We obtain

$$g^*(m,n) \equiv -7m - 5n \pmod{28}.$$

Hence the numbers coincide in Figures 4 and 5 at corresponding places. The subscripts read in the order $0, 1, \cdots, 27$ indicate the jumps of the roundwalk and therefore reflect the coding words. The colouring induced by χ (as presented in Figure 3) is given by

$$\chi(m,n) = \begin{cases} 0 \text{ if} & 6 \le -7m - 5n \pmod{28} < 17 \\ 1 \text{ if} & 17 \le -7m - 5n \pmod{28} < 27 \\ 2 \text{ otherwise}, & \end{cases}$$

			19_2	12_2
		3_2	24_2	17_2
	15_1	8_2	1_2	22_2
	20_1	13_2	6_2	27_2
4_1	25_1	18_0	11_0	
9_1	2_0	23_0	16_0	
14_0	7_0	0_0	21_0	
		5_0	26_0	
		10_0		

Figure 4.

Insertion in the canonical way

		16_1		
		21_1		
	5_1	26_1	19_2	12_2
	10_1	3_2	24_2	17_2
	15_0	8_2	1_2	22_2
	20_0	13_2	6_2	27_2
4_0	25_0	18_0	11_0	
9_0	2_0	23_0		
14_0	7_0	0_0		

Figure 5.

Insertion in the anti-canonical way

			1	0
		2	1	1
	0	0	2	1
	1	0	0	2
2	1	1	0	
0	2	1	0	
0	0	$\underline{2}$	1	
		2	1	
		0		

Figure 6.

W^* in the canonical case

		0		
		1		
	2	1	1	0
	0	2	1	1
	0	0	2	1
	1	0	0	2
2	1	1	0	
0	2	1		
0	0	$\underline{2}$		

Figure 7.

W^* in the anti-canonical case

Note that $6, 17, 27$ in Figures 4 and 5 correspond with the numbers $5, 14, 22$ in Figure 1. The words $W^* = \chi^*|_{A^*}$ corresponding to Figures 4 and 5 are given in Figures 6 and 7, respectively. We omit the subscripts, but underline the entry at the origin, as we did in Figure 3.

3.2 Substitution matrices

We introduce vectors and matrices to create a general framework to describe substitutions and their effects on roundwalks. Let σ be a substitution defined over the alphabet $\mathcal{A} = \{0, \ldots, k\}$ of cardinality $k+1$. The *substitution matrix* of the substitution σ is, by definition, the $(k+1) \times (k+1)$ matrix M_σ the entry of index (i, j) of which is $|\sigma(a_j)|_{a_i}$, that is, the number of occurrences of a_i in $\sigma(a_j)$.

Let $k+1$ vectors $(\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_k)$ in $\mathbb{Z}^k$ be given such that $\mathbb{Z}\vec{v}_0 + \mathbb{Z}\vec{v}_1 + \cdots + \mathbb{Z}\vec{v}_k = \mathbb{Z}^k$ and that

$$\begin{cases} d_0 := \det(\vec{v}_1, \cdots, \vec{v}_k), \\ d_h := \det(\vec{v}_1, \cdots, \vec{v}_{h-1}, -\vec{v}_0, \vec{v}_{h+1}, \cdots, \vec{v}_k), \text{ for } h \neq 0 \end{cases}$$

are coprime positive integers. Then, by Cramer's rule, $d_0\vec{v}_0 + d_1\vec{v}_1 + \cdots + d_k\vec{v}_k = \vec{0}$. Let σ be a substitution with substitution matrix M having determinant 1. We define the column vectors $\vec{v}_0^*, \vec{v}_1^*, \ldots, \vec{v}_k^* \in \mathbb{Z}^k$ by

$$(\vec{v}_0^*, \vec{v}_1^*, \ldots, \vec{v}_k^*) = (\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_k)M^{-1}.$$

Then $\mathbb{Z}\vec{v}_0^* + \mathbb{Z}\vec{v}_1^* + \cdots + \mathbb{Z}\vec{v}_k^* = \mathbb{Z}^k$. Put

$$\begin{cases} d_0^* := \det(\vec{v}_1^*, \cdots, \vec{v}_k^*), \\ d_h^* := \det(\vec{v}_1^*, \cdots, \vec{v}_{h-1}^*, -\vec{v}_0^*, \vec{v}_{h+1}^*, \cdots, \vec{v}_k^*) \text{ for } h \neq 0. \end{cases}$$

It follows that $d_0^*, d_1^*, \cdots, d_k^*$ are coprime positive integers such that $d_0^*\vec{v}_0^* + d_1^*\vec{v}_1^* + \cdots + d_k^*\vec{v}_k^* = \vec{0}$. Denote by V and V^* the k by $k+1$ matrices with column vectors $\vec{v}_0, \vec{v}_1, \ldots, \vec{v}_k$ and $\vec{v}_0^*, \vec{v}_1^*, \ldots, \vec{v}_k^*$, respectively, and by $\vec{d}$ and $\vec{d}^*$ the row vectors $(d_0, d_1, \cdots, d_k)$ and $(d_0^*, d_1^*, \cdots, d_k^*)$. Then

$$V\ {}^t\vec{d} = \vec{0}, V^* = VM^{-1}, V^*\ {}^t\vec{d}^* = \vec{0}, \tag{3}$$

where a left superscript t indicates transposition of vectors or matrices. By the uniqueness of the vector $\vec{d}$ with coprime positive coefficients such that $V\ {}^t\vec{d} = \vec{0}$, we deduce from $VM^{-1}\ {}^t\vec{d}^* = \vec{0}$ that

$${}^t\vec{d} = M^{-1}\ {}^t d^*. \tag{4}$$

Put $D = d_0 + d_1 + \cdots + d_k$ and $D^* = d_0^* + d_1^* + \cdots + d_k^*$.

Let $\vec{x} = (x_0, x_1, \cdots, x_k) \in \mathbb{Z}^{k+1}$ and ${}^t\vec{x}^* = M\ {}^t\vec{x}$. Then

$$\sum_{i=0}^{k} x_i^* \vec{v}_i^* = V^*\ {}^t\vec{x}^* = VM^{-1}\ {}^t\vec{x}^* = V\ {}^t\vec{x} = \sum_{i=0}^{k} x_i \vec{v}_i.$$

Let $g : \mathbb{Z}^k \to \{0, 1, \cdots, D-1\}$ and $g^* : \mathbb{Z}^k \to \{0, 1, \cdots, D^*-1\}$ be linear functions such that $g(0) = g^*(0) = 0$ and $g(v_i) = g^*(v_i^*) = 1$ for $i = 0, 1, \cdots, k$. Then

$$g^*(\sum_{i=0}^{k} x_i \vec{v}_i) = g^*(\sum_{i=0}^{k} x_i^* \vec{v}_i^*) = \sum_{i=0}^{k} x_i^* \pmod{D^*}. \tag{5}$$

Observe that $V^*, \vec{d}^*, D^*$, and g^* depend only on V and M and are independent of the way of insertion prescribed by σ.

In Section 3.1 the substitution matrix $M_{i,j}$ of the substitution $\sigma_{i,j}$ (which is also that of $\tilde{\sigma}_{i,j}$), satisfies $M_{i,j} = Id + E_{j,i}$, where Id is the identity matrix and $E_{j,i}$ the matrix of which all entries are 0 except for the entry of index (j, i) which equals 1. Note that the matrix $M_{i,j}$ has determinant 1.

Given a roundwalk any substitution rule σ with substitution matrix M having determinant 1 or -1 induces an extended roundwalk: the roundwalk $\mathbf{w}$ is determined by the starting place, the vectors $\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k$ and the coding w. By keeping the starting place fixed, by computing the vectors $\vec{v}_0^*, \vec{v}_1^*, \cdots, \vec{v}_k^*$ according to the above formula, and by applying the substitution σ to w to obtain w^* we get the roundwalk $\mathbf{w}^*$. There is a simple way to describe $\mathbf{w}^*$ and to find $A(\mathbf{w}^*)$. Let $0 \le i \le k$. Consider the places of the roundwalk with subscripts i. The jump $\vec{v}_i$ in the old roundwalk will be replaced by successive jumps $\vec{v}_{j_1}^*, \vec{v}_{j_2}^*, \cdots, \vec{v}_{j_n}^*$ where $\sigma(i) = j_1 j_2 \cdots j_n$. The new places will be adjoined to $A(\mathbf{w})$. Consider a place $\vec{a}$ which belongs to the part $A_i(\mathbf{w})$ with subscript i in $A(\mathbf{w})$. This point has been reached after reading a certain part of the coding w. Consider the corresponding part of the coding w^*. Suppose the reached point is $\vec{a}$ too. Then this point appears both in $A(\mathbf{w})$ and in $A(\mathbf{w}^*)$. Subsequently $A(\mathbf{w}^*)$ is augmented with the places $\vec{a}$ translated by the vectors $\vec{v}_{j_1}^*, \vec{v}_{j_1}^* + \vec{v}_{j_2}^*, \cdots, \vec{v}_{j_1}^* + \vec{v}_{j_2}^* + \cdots \vec{v}_{j_{n-1}}^*$. The next point which is reached is the point $\vec{a} + \vec{v}_{j_1}^* + \vec{v}_{j_2}^* + \cdots \vec{v}_{j_n}^* = \vec{a} + \vec{v}_i$ which is the point following $\vec{a}$ in $\mathbf{w}$. Hence, by induction, $A(\mathbf{w}) \subset A(\mathbf{w}^*)$ and the set $A_i(\mathbf{w})$ is replaced with

$$A_i(\mathbf{w}) \cup A_i(\mathbf{w}) + \vec{v}_{j_1}^* \cup A_i(\mathbf{w}) + \vec{v}_{j_1}^* + \vec{v}_{j_2}^* \cup \cdots \cup A_i(\mathbf{w}) + \vec{v}_{j_1}^* + \vec{v}_{j_2}^* + \cdots + \vec{v}_{j_{n-1}}^*.$$

The situation is illustrated in Section 3.3 below where $A_1(\mathbf{w})$ and $A_2(\mathbf{w})$ are each copied once (Figure 10).

Assume we are given a colouring map $\chi : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ which is a projection of g. Then χ induces a colouring map $\chi^* : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ as follows. If for $a \in A$ we have $\chi(\vec{a}) = i$, then in $\mathbf{w}^*$ we put

$$\chi^*(\vec{a}) = \chi^*(\vec{a} + \vec{v}^*_{j_1}) = \cdots = \chi^*(\vec{a} + \vec{v}^*_{j_1} + \vec{v}^*_{j_2} + \cdots + \vec{v}^*_{j_{n-1}}) = i. \tag{6}$$

Note that n equals the i-th column sum of the matrix M. Hence the definition of χ^* depends only on χ and M and is independent of the way of insertion. Only the number of occurrences of each letter in the substitution rule matters. Note that χ^* is constant on cosets of Λ^* and is therefore a projection of g^*.

3.3 The Rauzy extension step

We start from the situation as described in the previous sections. For given $j \in \{0, 1, \cdots, k\}$ we introduce the Rauzy extension step R_j for $j \in \{0, 1, \cdots, k\}$ which is actually the composition of all $S_{i,j}$ with $i \in \{0, 1, \cdots, k\} \setminus \{j\}$. Put $\vec{v}^R_i = \vec{v}_i - \vec{v}_j$ for $i = 0, 1, \cdots, k$ with $i \neq j$ and $\vec{v}^R_j := \vec{v}_j$. Hence the corresponding substitution matrix M_j has entries 1 at the diagonal and at the j-th row and further entries 0. Then $d^R_i = d_i$ for $i \neq j$, and $d^R_j = d_0 + \cdots + d_k = D$. All the superscripts R will refer to the situation after a Rauzy step. Note that the lattice $L^R := \mathbb{Z}\vec{v}^R_1 + \cdots + \mathbb{Z}\vec{v}^R_k$ equals $L = \mathbb{Z}^k$ with lattice determinant $d^R := d = 1$. Put $D^R := d^R_0 + d^R_1 + \cdots + d^R_k = 2D - d_j$. Let Λ^R be the lattice $\mathbb{Z}(\vec{v}^R_1 - \vec{v}^R_0) + \cdots + \mathbb{Z}(\vec{v}^R_k - \vec{v}^R_0)$. Then

$$\det(\Lambda^R) = d^R_0 + d^R_1 + \cdots + d^R_k = 2D - d_j = D^R.$$

Suppose we have a roundwalk $\vec{w}_0, \vec{w}_1, \cdots, \vec{w}_D = \vec{w}_0$ in $\mathbb{Z}^k$ through some fundamental domain A of Λ with $\vec{w}_{i+1} - \vec{w}_i \in \{\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_k\}$ for $i = 0, 1, \cdots, D-1$. We extend it to a roundwalk $\vec{w}^R_0, \vec{w}^R_1, \cdots, \vec{w}^R_{D^R} = \vec{w}^R_0$ in $\mathbb{Z}^k$ through a fundamental domain A^R of Λ^R by inserting for every pair h, i with $\vec{w}_{h+1} - \vec{w}_h = \vec{v}_i$ and $i \neq j$ either $\vec{w}_h + \vec{v}^R_i$ or $\vec{w}_h + \vec{v}^R_j$ in between $\vec{w}_h$ and $\vec{w}_{h+1}$. Doing so we obtain a roundwalk where $\vec{w}^R_0, \vec{w}^R_1, \cdots, \vec{w}^R_{D^R-1}$ represent the D^R cosets of Λ^R. The new roundwalk contains d^R_i jumps $\vec{v}^R_i$ for $i = 0, 1, \cdots, k$. This time we have $D^R - d^R_j$ times a choice out of two to make the insertion. If always $\vec{w}_h + \vec{v}^R_i$ is inserted (and never $\vec{w}_h + \vec{v}^R_j$), then we say that the insertion is done in the *canonical way*. If always $\vec{w}_h + \vec{v}^R_j$ is inserted, then we do it in the *anti-canonical way*.

The action of the Rauzy extension step R_j on the coding w of the roundwalk $\mathbf{w}$ is described in the canonical case by the substitution $\sigma^R_j : i \mapsto ij$, for $i \neq j$ and $j \mapsto j$, and in the anti-canonical case by $\tilde{\sigma}^R_j : i \mapsto ji$, for $i \neq j$ and $j \mapsto j$. Note that σ^R_j equals the composition of the substitutions $\sigma_{i,j}$ for $i \neq j$. These substitutions are called *generalized Rauzy substitutions* following [2]. They are introduced in [3] where it is proved that each Arnoux-Rauzy sequence is in the shift orbit closure of a unique sequence of the form

$$\lim_{n\to\infty} \sigma_{i_1}^R \circ \ldots \circ \sigma_{i_n}^R(0),$$

where the sequence (i_n) takes infinitely many times the value i for every $i = 0, \cdots, k$. Note that $\tilde{\sigma}_0^R\tilde{\sigma}_1^R\tilde{\sigma}_2^R = \sigma^3$, where σ denotes the usual Rauzy substitution $\sigma(0) = 01$, $\sigma(1) = 02$, $\sigma(2) = 0$.

We illustrate this by starting from the same situation as in Example 1 and applying R_0 in the canonical way to obtain Figure 8. Therefore in the coding sequence (cf. subscripts) we replace every $i > 0$ by $i0$.

Example 3 We still pursue Example 1. We start from $k = 2, \vec{v}_0 = (1, 4), \vec{v}_1 = (3, 1), \vec{v}_2 = (-2, -3)$ and roundwalk $\mathbf{w}$ with coding

$$w = 0\ 2\ 0\ 2\ 1\ 2\ 0\ 2\ 1\ 0\ 2\ 2\ 0\ 1\ 2\ 0\ 2\ 1\ 2\ 0\ 2\ 1\ 2.$$

On applying R_0 we find $\vec{v}_0^R = (1, 4), \vec{v}_1^R = (2, -3), \vec{v}_2^R = (-3, -7), d_0^R = 23, d_1^R = 11, d_2^R = 5, D^R = 39$. There are $d_1 + d_2 = 16$ new points. If we replace each 1 by 10 and each 2 by 20 we get the roundwalk $\mathbf{w}^R$ with coding

$$w^R = 0\ 20\ 0\ 20\ 10\ 20\ 0\ 20\ 10\ 0\ 20\ 20\ 0\ 10\ 20\ 0\ 20\ 10\ 20\ 0\ 20\ 10\ 20$$

in Figure 8. In Figure 9 we have applied the coding

$$\tilde{\sigma}_0^R(w) = 0\ 02\ 0\ 02\ 10\ 02\ 0\ 02\ 01\ 0\ 20\ 02\ 0\ 10\ 20\ 0\ 20\ 01\ 20\ 0\ 20\ 01\ 20.$$

as it is used in Simpson and Tijdeman [13]. Here there are no fixed substitutions for the replacement of the letters 1 and 2 and the result is a convex set A^* which ressembles a hexagon.

Figure 1 can be divided into three zones which correspond to the points which have subscript 0,1,2, respectively (see the left Figure 10 below). In the right Figure 10 the part in Figure 1 with index 1 is translated over $\vec{v}_1^R = (2, -3)$, the part with index 2 over $\vec{v}_2^R = (-3, -7)$, whereas all the parts remain at the same place too. Because of the choice for canonical insertion, the subscripts at new places become 0. The right Figure 10 explains Figure 8. In Figure 9 the copied parts $A_1(\mathbf{w})$ and $A_2(\mathbf{w})$ are each split into two parts because of the mixed substitutions 10 and 01 for 1, and 20 and 02 for 2.

				26_2	16_2
			4_2	33_2	23_2
		21_1	11_2	1_2	30_2
		28_1	18_2	8_2	37_2
	6_1	35_1	25_0	15_0	
	13_1	3_0	32_0	22_0	
	20_0	10_0	0_0	29_0	
	27_0	17_0	7_0	36_0	
5_0	34_0	24_0	14_0		
12_0	2_0	31_0			
19_0	9_0	38_0			

Figure 8. Canonical extension

			5_2	
			12_2	2_2
		29_1	19_2	9_2
		36_1	26_2	16_2
	14_1	4_0	33_2	23_2
	21_1	11_0	1_0	30_2
	28_0	18_0	8_0	37_2
6_1	35_0	25_0	15_0	
13_0	3_0	32_0	22_0	
20_0	10_0	0_0		
27_0	17_0	7_0		
34_0	24_0			
	31_0			
	38_0			

Figure 9. ST-extension

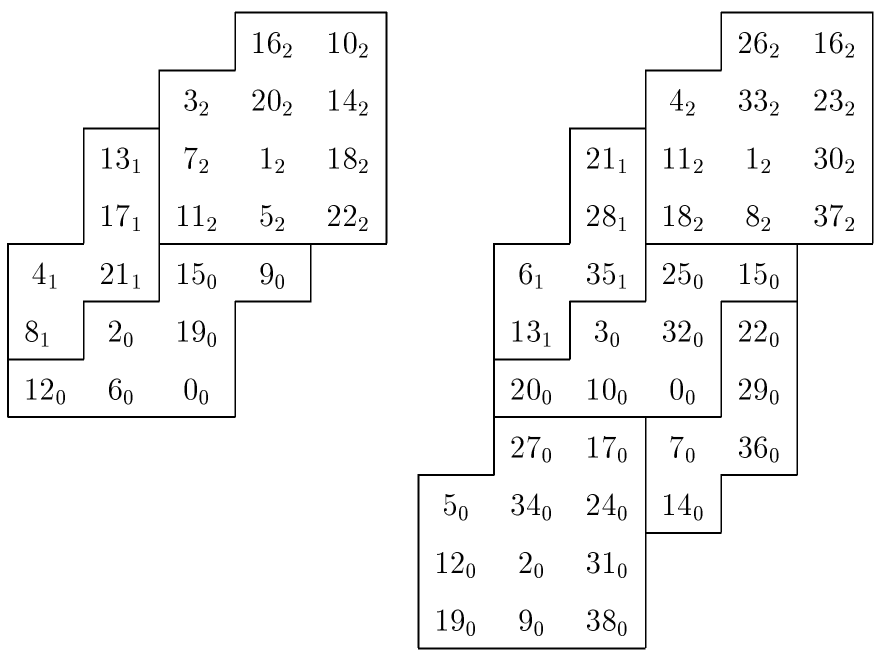

Figure 10. The roundwalk $\mathbf{w}$ *from Figure 1 and its extension* $\mathbf{w}^R$ *in Figure 8*

Points in Figure 8 and Figure 9 with the same (upper) number are representatives of the same coset of Λ^R. Hence their difference vector is in $\Lambda^R = \mathbb{Z}(1,-7)+\mathbb{Z}(-4,-11)$. Figures 8 and 9 are the restrictions of a linear function $g^R : \mathbb{Z}^2 \to \{0,1,2,\cdots,38\}$. If we consider the point 0_0 as the origin, then the function g^R is given by $g(m,n) = -10m - 7n \pmod{39}$.

Let us write for short $A = A(\mathbf{w})$ and $A^R = A(\mathbf{w}^R)$. We extend the function $\chi : A \to \{0,1,2\}$ to a function $\chi^R : A^R \to \{0,1,2\}$. Recall that $\chi|_A$ is given by

$$\chi(m,n) = \begin{cases} 0 \text{ if} & 5 \le -6m - 4n \pmod{23} < 14 \\ 1 \text{ if} & 14 \le -6m - 4n \pmod{23} < 22 \\ 2 \text{ otherwise.} & \end{cases}$$

Consider now the extension to A_R given in Figure 11. We write a 0, 1 or 2 according to the value it has in Figure 3 for places in $A \cap A_R$ and the value of the preceding place in the roundwalk if the place is in $A_R \setminus A$. This yields Figure 11 where we have underlined the number at the origin. It follows that the induced word $W^R = f^R|_{A^R}$ satisfies (cf. the upper values in the right Figure 10)

$$\chi^R(m,n) = \begin{cases} 0 \text{ if} & 8 \le -10m - 7n \pmod{39} < 23 \\ 1 \text{ if} & 23 \le -10m - 7n \pmod{39} < 37 \\ 2 \text{ otherwise.} & \end{cases}$$

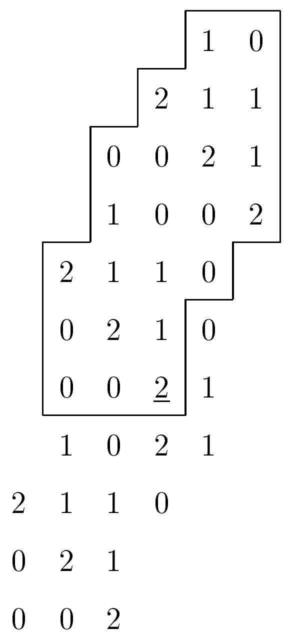

Figure 11. W from Figure 3 and its extension W^* in the canonical case

4 Towers of lattices

The aim of this section is to construct towers of lattices generating larger and larger roundwalks, and thus colourings with larger and larger shapes.

4.1 Basic towers

Let $\vec{v}_1^{(0)}$, $\vec{v}_2^{(0)}, \cdots, \vec{v}_k^{(0)}$ denote vectors in $\mathbb{Z}^k$ with

$$d_0^{(0)} := \det(\vec{v}_1^{(0)}, \cdots, \vec{v}_k^{(0)}) = 1.$$

Put $\vec{v}_0^{(0)} = -\vec{v}_1^{(0)} - \ldots - \vec{v}_k^{(0)}$. Then, for $1 \le i \le k$,

$$d_i^{(0)} := \det(\vec{v}_1^{(0)}, \cdots, \vec{v}_{i-1}^{(0)}, -\vec{v}_0^{(0)}, \vec{v}_{i+1}^{(0)}, \cdots, \vec{v}_k^{(0)}) = 1.$$

Put

$$\begin{cases} L^{(0)} = \mathbb{Z}\vec{v}_0^{(0)} + \mathbb{Z}\vec{v}_1^{(0)} + \cdots + \mathbb{Z}\vec{v}_k^{(0)}, \\ \Lambda^{(0)} = \mathbb{Z}(\vec{v}_1^{(0)} - \vec{v}_0^{(0)}) + \cdots + \mathbb{Z}(\vec{v}_k^{(0)} - \vec{v}_0^{(0)}). \end{cases}$$

Then $L^{(0)} = \mathbb{Z}^k, D^{(0)} := d_0^{(0)} + d_1^{(0)} + \cdots + d_k^{(0)} = k+1$. We start with the roundwalk $\mathbf{w}^{(0)}$ given by $\vec{0}, \vec{v}_0^{(0)}, \vec{v}_0^{(0)} + \vec{v}_1^{(0)}, \cdots, \sum_{i=0}^{k} \vec{v}_i^{(0)} = \vec{0}$ with coding $w^{(0)} = 012\cdots k$. For convenience we assume $\sum_{i=0}^{m-1} \vec{v}_i^{(0)} = \vec{e}_m$ for $m = 1, 2, \cdots, k$ in the sequel. This can be achieved by a transformation of coordinates. Hence $A^{(0)} := A(\mathbf{w}^{(0)}) = \{\vec{0}, \vec{e}_1, \cdots, \vec{e}_k\}$. One has $\vec{e}_m - m\vec{e}_1 \in \Lambda^{(0)}$, for $m = 1, \cdots, k$. We thus define $g^{(0)} = \chi^{(0)} : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ by $\chi^{(0)}(\vec{0}) = 0, \chi^{(0)}(\vec{e}_m) = m$ for $m = 1, \cdots, k$ and $\chi^{(0)}$ is constant on cosets of $\mathbb{Z}^k/\Lambda^{(0)}$. Thus if $\vec{x} = (x_1, \cdots, x_k) \in \mathbb{Z}^k$, then

$$g^{(0)}(\vec{x}) = \chi^{(0)}(\vec{x}) = \sum_{i=1}^{k} i x_i \pmod{k+1}.$$

We iteratively apply basic extension steps $S_{i,j}$ in the canonical way. Let $S^{(n)} = S_{i_n, j_n}$ denote the n-th extension rule that is applied. We will use the notation $\vec{v}_i^{(n)}$, $d_i^{(n)}$, for $0 \le i \le k$, $L^{(n)}, \Lambda^{(n)}, g^{(n)}$, $\chi^{(n)}$, $V^{(n)}$, $\vec{d}^{(n)}$, $A^{(n)}$, $\mathbf{w}^{(n)}$, $w^{(n)}$, $D^{(n)}$, $W^{(n)}$ for the values of the previously defined symbols at level n. Let us recall that we use the convention expressed by (6) for the definition of $\chi^{(n)}$.

By (3) and (4) we have, in terms of matrices,

$$V^{(n)} = V^{(0)} M_{i_1,j_1}^{-1} \cdots M_{i_n,j_n}^{-1} \quad \text{and} \quad {}^t\vec{d}^{(n)} = M_{i_n,j_n} \cdots M_{i_1,i_n}({}^t\vec{d}^{(0)}).$$

Since $\vec{d}^{(0)} = (1, 1, \cdots, 1)$, the vector $\vec{d}^{(n)}$ is given by the row sums of the product matrix $M_{i_n,j_n} \cdots M_{i_1,j_1}$.

Put

$$M^{(n)} := {}^tM_{i_1,j_1} \cdots {}^tM_{i_n,j_n},$$

($M^{(0)} = Id$). Define $c_j^{(n)}$ as the j-th column sum of $M^{(n)}$ (the numbering starting with 0), and $r_j^{(n)}$ as the j-th row sum of $M^{(n)}$. When we apply extension step S_{i_n,j_n} to $M^{(n)}$, then $M^{(n+1)}$ is obtained by adding the i_n-th column vector of $M^{(n)}$ to its j_n-th column vector.

Observe that the i-th row sum $r_i^{(n)}$ denotes the total number of entries i in the word $W^{(n)}$ for $0 \le i \le k$, whereas the j-th column sum $c_j^{(n)}$ denotes the number of letters j in the coding word $w^{(n)}$, that is, the number of places in $A^{(n)}$ with subscript j, for $0 \le j \le k$. Of course, $\sum_{i=0}^k c_i^{(n)} = \sum_{i=0}^k r_i^{(n)} = D^{(n)}$, the total number of points in $A^{(n)}$, that is the cardinality of $\mathbb{Z}^k/\Lambda^{(n)}$.

We use the linearity of the function $g^{(n)}$ and the results of Sec. 3.2 to give explicit expressions for $g^{(n)}$ and $\chi^{(n)}$. Let $\vec{x} \in \mathbb{Z}^k$. Put $\vec{x} = \sum_{m=1}^k x_m \vec{e}_m$. Then

$$g^{(n)}(\vec{x}) = \sum_{m=1}^{k} x_m g^{(n)}(\vec{e}_m) \pmod{D^{(n)}}.$$

By our special choice of the $\vec{v}_i$ and $\chi^{(0)}$, the roundwalk $\mathbf{w}^{(n)}$ starts from the origin and jumps along $r_0^{(n)}$ places with subscript 0 until it reaches $\vec{e}_1$, then passes $r_1^{(n)}$ places with subscript 1 until it reaches $\vec{e}_2$, and so on. Hence the number $g^{(n)}(\vec{e}_j)$ equals the total number of letters $0, 1, \cdots, j-1$ in $W^{(n)}$. Thus $g^{(n)}(\vec{e}_j) = \sum_{i=0}^{j-1} r_i^{(n)}$ for $j = 1, \cdots, k$ and

$$g^{(n)}(\vec{x}) = \sum_{j=1}^{k} x_j \sum_{i=0}^{j-1} r_i^{(n)} \pmod{D^{(n)}}.$$

It follows from the definition of χ that

$$\chi^{(n)}(\vec{x}) = m \Longleftrightarrow \sum_{i=0}^{m-1} r_i^{(n)} \le g^{(n)}(\vec{x}) < \sum_{i=0}^{m} r_i^{(n)}.$$

So we have derived the following proposition.

Proposition 1 *Under the assumptions made in this section the function* $\chi^{(n)} : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ *satisfies, for* $m = 1, \cdots, k$ *and* $\vec{x} = (x_1, \cdots, x_k) \in \mathbb{Z}^k$,

$$\chi^{(n)}(\vec{x}) = m \Longleftrightarrow \frac{\sum_{i=0}^{m-1} r_i^{(n)}}{D^{(n)}} \le \left\{ \frac{\sum_{j=1}^{k} x_j \sum_{i=0}^{j-1} r_i^{(n)}}{D^{(n)}} \right\} < \frac{\sum_{i=0}^{m} r_i^{(n)}}{D^{(n)}},$$

where $D^{(n)} = \sum_{i=0}^k r_i^{(n)}$ *and* $\{y\}$ *denotes the fractional part of* y.

Example 4 Let $k = 2, \vec{v}_0^{(0)} = (1,0), \vec{v}_1^{(0)} = (-1,1), \vec{v}_2^{(0)} = (0,-1)$. We apply periodically extension steps $S_{0,1}, S_{1,2}, S_{2,0}$. Hence we obtain the sequence of matrices $M^{(n)}$:

$$\begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}\begin{pmatrix}2&1&1\\1&1&1\\1&0&1\end{pmatrix}\begin{pmatrix}2&3&1\\1&2&1\\1&1&1\end{pmatrix}$$
$$n=1 \qquad n=2 \qquad n=3 \qquad n=4$$

$$\begin{pmatrix}2&3&4\\1&2&3\\1&1&2\end{pmatrix}\begin{pmatrix}6&3&4\\4&2&3\\3&1&2\end{pmatrix}\begin{pmatrix}6&9&4\\4&6&3\\3&4&2\end{pmatrix}\begin{pmatrix}6&9&13\\4&6&9\\3&4&6\end{pmatrix}$$
$$n=5 \qquad n=6 \qquad n=7 \qquad n=8$$

We compute $r_0^{(8)} = 28, r_1^{(8)} = 19, r_2^{(8)} = 13, D^{(8)} = 60, \vec{v}_0^{(8)} = (2,-5), \vec{v}_1^{(8)} = (6,-1), \vec{v}_2^{(8)} = (-5,3)$. Thus the induced function $\chi^{(8)} : \mathbb{Z}^2 \to \{0,1,2\}$ is given by

$$\chi^{(8)}(x,y) = \begin{cases} 0 \text{ if } 0 \le \{\frac{28}{60}x + \frac{47}{60}y\} < \frac{28}{60} \\ 1 \text{ if } \frac{28}{60} \le \{\frac{28}{60}x + \frac{47}{60}y\} < \frac{47}{60} \\ 2 \text{ if } \frac{47}{60} \le \{\frac{28}{60}x + \frac{47}{60}y\} < 1. \end{cases}$$

This is independent of the made insertions.

4.2 Rauzy towers

We start from the same situation as in the previous section, but now we apply Rauzy extensions steps R_j. Since a Rauzy extension step is a combination of k basic extension steps, the formulas for g and χ in Sec. 4.1 remain valid with superscripts (n) in place of (kn). However, in the Rauzy case the numbers $r_i^{(n)}$ satisfy a useful recurrence relation which facilitates the computation of the function $\chi^{(n)}$. Write $\vec{r}^{(n)} = (r_0^{(n)}, r_1^{(n)}, \cdots, r_k^{(n)})$. Let $0 \le j \le k$. If the Rauzy step R_j is applied at level n, then

$$\vec{r}^{(n)} = M^{(n)}\ {}^t(1,1,\cdots,1) = M^{(n-1)}\ {}^tM_j\ {}^t(1,1,\cdots,1).$$

(Recall that ${}^tM_j$ is the matrix with entries 1 at the diagonal and at the j-th column and entries 0 elsewhere).

Since

$${}^tM_j\ {}^t(1,1,\cdots,1) = 2\ {}^t(1,1,\cdots,1) - \vec{e}_j,$$

we have $\vec{r}^{(n)} = 2\vec{r}^{(n-1)} - M^{(n-1)}\vec{e}_j = 2\vec{r}^{(n-1)} - \vec{c}_j^{(n-1)}$, where $\vec{c}_j^{(n-1)}$ is the j-th column vector of $M^{(n-1)}$. In view of $M^{(l+1)} = M^{(l)}\ {}^tM_i$ when R_i is applied at level l, we obtain $\vec{c}_j^{(n-1)} = \vec{c}_j^{(n-2)} = \cdots = \vec{c}_j^{(q)} \neq \vec{c}_j^{(q-1)}$ if the previous time that a Rauzy extension step R_j was applied was at level q. Furthermore, $\vec{c}_j^{(q)} = \vec{r}^{(q-1)}$. Thus

$$\vec{r}^{(n)} = 2\vec{r}^{(n-1)} - \vec{r}^{(q-1)} \tag{7}$$

if at level n a Rauzy step R_j is applied and the previous time that R_j was applied was at level q. The above argument is of course independent of the chosen way of insertion. The corresponding function $\chi^{(n)} : \mathbb{Z}^k \to \{0, 1, 2\}$ is given in Proposition 1.

Example 5 We consider $k = 2$ and apply periodically Rauzy steps R_2, R_1, R_0. This yields a sequence of matrices $(M^{(n)})_{n\geq 0}$ starting with

$$\underset{n=0}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \quad \underset{n=1}{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}} \quad \underset{n=2}{\begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}} \quad \underset{n=3}{\begin{pmatrix} 4 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}}$$

$$\underset{n=4}{\begin{pmatrix} 4 & 2 & 7 \\ 3 & 2 & 6 \\ 2 & 1 & 4 \end{pmatrix}} \quad \underset{n=5}{\begin{pmatrix} 4 & 13 & 7 \\ 3 & 11 & 6 \\ 2 & 7 & 4 \end{pmatrix}} \quad \underset{n=6}{\begin{pmatrix} 24 & 13 & 7 \\ 20 & 11 & 6 \\ 13 & 7 & 4 \end{pmatrix}}$$

The row sums satisfy a recurrence relation $\vec{r}^{(n)} = 2\vec{r}^{(n-1)} - \vec{r}^{(n-4)}$ for every $n \geq 4$. The characteristic polynomial reads $x^4 - 2x^3 + 1 = (x-1)(x^3 - x^2 - x - 1) = (x-1)(x-\alpha)(x-\beta)(x-\overline{\beta})$ where α is the Tribonacci number (in reference to the Fibonacci number $\frac{1+\sqrt{5}}{2}$) ; one has $\alpha \approx 1.84$ and $|\beta| < 1$. We find that there are constant coefficients $c_{j,i}$ such that $r_i^{(n)} = c_{0,i}\alpha^n + c_{1,i} + c_{2,i}\beta^n + c_{3,i}\overline{\beta}^n$ for $i = 0, 1, 2$ and all n. The coefficients can be computed from $\vec{r}^{(0)}, \vec{r}^{(1)}, \vec{r}^{(2)}, \vec{r}^{(3)}$. The corresponding functions $\chi^{(n)}$ are given by Proposition 1. It will be clear that the functions $\chi^{(n)}$ converge to a limit word on $\mathbb{Z}^2$, but this will be the subject of the next section.

5 Multi-dimensional Sturmian words

In this section we study the limit words which can be obtained by using towers of lattices as described in the previous section.

5.1 Some definitions and non-existence results

It is obvious from the theory of extended roundwalks that $W := \lim_{n\to\infty} W^{(n)}$ exists and is defined on $A := \cup_{n=0}^{\infty} A^{(n)} = \lim_{n\to\infty} A^{(n)}$. However, the word W depends on the way the insertions are being made at each step. It is easy to show that it can happen that $A \neq \mathbb{Z}^k$. Take a fixed place $\vec{x} \in \mathbb{Z}^k, |\vec{x}| > 1$. At every level we have free choice of making insertions. At most one of both insertions involves $\vec{x}$, since in the notation of Sec. 3.1 at least one among $\vec{w_h} + \vec{v_i^*}$ and $\vec{w_h} + \vec{v_j^*}$ is different from $\vec{x}$. So we can secure by making the "right" insertions that $\vec{x} \notin A^{(n)}$ for every n. Thus $\vec{x} \notin A$.

In view of $W^{(n)} = \chi^{(n)}|_{A^{(n)}}$ for every n, it would be nice if $\lim_{n\to\infty} \chi^{(n)}$ exists, that is, $\chi^{(n)}(\vec{x})$ is constant for every $\vec{x} \in \mathbb{Z}^k$ and $n \geq n_0(\vec{x})$, since then we have a limit word $\chi : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ which is independent of the chosen insertions and $W = \chi|_A$. We shall show that $\lim_{n\to\infty} \chi^{(n)}$ does not exist in general.

We use the notation of Sec. 4.1. Put $\rho_m^{(n)} = \sum_{i=0}^{m-1} \frac{r_i^{(n)}}{D^{(n)}}$ for $m = 1, \cdots, k + 1$. Then we deduce from Proposition 1 that, for $m = 0, 1, \cdots, k$ and $\vec{x} = (x_1, \cdots, x_k) \in \mathbb{Z}^k$,

$$\chi^{(n)}(\vec{x}) = m \Longleftrightarrow \rho_m^{(n)} \leq \{\sum_{j=1}^{k} x_j \rho_j^{(n)}\} < \rho_{m+1}^{(n)}. \tag{8}$$

Obviously $\lim_{n\to\infty} \chi^{(n)}$ exists if $\rho_m := \lim_{n\to\infty} \rho_m^{(n)}$ exists for $m = 0, 1, \cdots, k$ and, in case ρ_m is rational, $\rho_m^{(n)} \leq \rho_m$ for all large n. If so, the limit word χ is given by

$$\chi(\vec{x}) = m \Longleftrightarrow \rho_m \leq \{\sum_{j=1}^{k} x_j \rho_j\} < \rho_{m+1} \tag{9}$$

and the density of the letter m equals $\rho_{m+1} - \rho_m$ for $m = 0, 1, \cdots, k$. Note that $0 \leq \rho_0 \leq \rho_1 \leq \cdots \leq \rho_k \leq \rho_{k+1} := 1$ and that the sum of the densities of the letters equals 1. Before going into the study of the structure of the limit words, we consider a situation where the limit does not exist.

Let $k = 4$ and $0 < \varepsilon < \frac{1}{2}$. We apply basic extension steps $S_{0,1}$ and $S_{1,0}$ until the sum of the densities of 0 and 1 exceeds $1 - \varepsilon$, subsequently we apply basic extension steps $S_{2,3}$ and $S_{3,2}$ until the sum of the densities of 2 and 3 exceeds $1 - \varepsilon$, subsequently we apply basic extension steps $S_{0,1}$ and $S_{1,0}$ until the sum of the densities of 0 and 1 exceeds $1 - \varepsilon$, and so on. Obviously, $\liminf_{n\to\infty} \rho_2^{(n)} < \varepsilon$, $\limsup_{n\to\infty} \rho_2^{(n)} > 1-\varepsilon$, and $\lim_{n\to\infty} \rho_2^{(n)}$ does not exist, so that $\lim_{n\to\infty} \chi^{(n)}$ does not exist either.

We give the definitions of k-dimensional regular word and k-dimensional Sturmian word, respectively.

Definition 1 *An infinite k-dimensional regular word is an infinite word $U : \mathbb{Z}^k \to \{0, 1, \ldots, k\}$ which satisfies either*

$$\forall (x_1, \ldots, x_k) \in \mathbb{Z}^k,\ (U(x_1, \ldots, x_k) = m \Longleftrightarrow$$
$$x_1\alpha_1 + \ldots + x_k\alpha_k + \rho \in [\alpha_m, \alpha_{m+1}[\ (\mathrm{mod}\ 1)),$$

$$\textit{or}\ \ \forall (x_1, \ldots, x_k) \in \mathbb{Z}^k,\ (U(x_1, \ldots, x_k) = m \Longleftrightarrow$$
$$x_1\alpha_1 + \ldots + x_k\alpha_k + \rho \in]\alpha_m, \alpha_{m+1}]\ (\mathrm{mod}\ 1)),$$

for some real numbers $\alpha_0 = 0 < \alpha_1 < \ldots < \alpha_k < \alpha_{k+1} = 1$ and ρ. If, moreover, $1, \alpha_1, \ldots, \alpha_k$ are independent over $\mathbb{Q}$ we call it a k-dimensional Sturmian word.

The multidimensional Sturmian words have been studied mainly for $k = 2$ in [5–8,4] and have many interesting combinatorial properties which allow us to consider them as a higher-dimensional generalisation of Sturmian words. In particular, they are nonperiodic (i.e., there is no nonzero vector of periodicity with integer coefficients) and uniformly recurrent (i.e., for every positive integer n, there exists an integer N such that every square factor of size $(N, \ldots, N)$ contains every factor of size $(n, \ldots, n)$). Furthermore they have $m_1 \ldots m_k + \sum_{i=1}^{k} \prod_{j\neq i} m_j$ factors of length $(m_1, \ldots, m_k)$. Recall that (classic) Sturmian words code the approximation of a line by a discrete line made of horizontal and vertical segments with integer vertices (for more details, see for instance [12,11]). These multidimensional sequences code discrete hyperplane approximations. In the sequel we will use the following observation: the densities of letters $0, 1, \ldots, k$ in a k-dimensional Sturmian word exist and are equal to $\alpha_1, \alpha_2 - \alpha_1, \cdots, 1 - \alpha_k$, respectively, (in the notation of Def. 1). The first theorem is an assertion of the type that every finite balanced word is a factor of a Sturmian word.

Theorem 1 *We use the notation of Sec. 2 and Def.1. Let* $\mathbf{w}$ *be a roundwalk in the domain $A(\mathbf{w})$. Define $\chi : \mathbb{Z}^k \to \{0, 1, \ldots, k\}$ in $\vec{x} = (x_1, \cdots, x_k) \in \mathbb{Z}^k$ by (9). Then the k-dimensional finite word $\chi|_{A(\mathbf{w})}$ is a factor of a k-dimensional Sturmian word with $\rho = 0$.*

Proof We have

$$\chi(\sum_{i=1}^{k} x_i \vec{e}_i) = m \Longleftrightarrow \rho_m \leq \{\sum_{j=1}^{k} x_j \rho_j\} < \rho_{m+1} \text{ for } m \in \{0, 1, \cdots k\}, \vec{x} \in A(\mathbf{w}).$$

Let $(\alpha_1, \cdots, \alpha_k) \in \mathbb{R}^k$ with $1, \alpha_1, \cdots, \alpha_k$ linearly independent over $\mathbb{Q}$ be such that $(D\alpha_1, \cdots, D\alpha_k)$ approximates the vector $\rho_1, \cdots, \rho_k$ with rational entries so well that

$$\{\sum_{i=1}^{k} x_i \rho_i\} \in \ [\rho_m, \rho_{m+1}[\Longleftrightarrow \{\sum_{i=1}^{k} x_i \alpha_i\} \in \ [\alpha_m, \alpha_{m+1}[,$$

for $m \in \{0, 1, \cdots k\}, \vec{x} \in A(\mathbf{w})$. For $m \in \{0, 1, \cdots, k\}$ we then have

$$\chi|_{A(\mathbf{w})}(\vec{x}) = m \Longleftrightarrow \alpha_m \leq \{\sum_{i=1}^{k} x_i \alpha_i\} < \alpha_{m+1}.$$

It follows from (9) that $\chi|_{A(\mathbf{w})}$ is a factor of the k-dimensional Sturmian word with parameters $\alpha_1, \cdots, \alpha_k$ associated to the partition by right-open and left-closed intervals and $\rho = 0$. ■

The limit words we study have the additional property that the constant ρ in the definition is 0. We call such words *homogeneous.*

Corollary 1 *The words $W^{(n)}$ corresponding to $\chi^{(n)}|_{A^{(n)}(\mathbf{w}^{(n)})}$ occurring in Sec. 4.1 are factors of homogeneous multidimensional Sturmian words.*

5.2 Some sufficient conditions for convergence

We will apply the following result.

Theorem 2 [9] *Let $(M_j)_{j \in \mathbb{N}}$ be a sequence of square matrices of size $k+1$ with coefficients in $\mathbb{N}$ with values from a finite set for which there exists a positive matrix P such that M_j attains this value P for infinitely many values j. Let $\mathcal{C}_+^{k+1}$ denote the nonnegative cone in $\mathbb{R}^{k+1}$ of vectors with nonnegative entries. Then there exists a positive vector $\vec{l} = {}^t(l_0, l_1, \ldots, l_k)$ with $\sum_{j=0}^{k} l_j = 1$ such that*

$$\bigcap_{n \in \mathbb{N}} M_1 \cdots M_n(\mathcal{C}_+^{k+1}) = \{\lambda \vec{l};\ \lambda \in \mathbb{R}_+\}.$$

In other words, for every nonzero vector $\vec{x}$ in $\mathbb{R}^{k+1}$ with nonnegative entries, $M_1 M_2 \ldots M_n \vec{x}$ converges towards the vector $\vec{l}$ in $\mathbb{R}^{k+1}$. Observe that such a convergence property needs not hold without the assumption of Theorem 2 as illustrated by Keane's example of a minimal and nonuniquely ergodic exchange of 4 intervals [10].

We apply Theorem 2 to the starting situation described in Sec. 4.1. We assume that we have an infinite tower $(M_{i_n,j_n})_{n\geq 1}$ of basic extension matrices. Since there are only finitely many choices for $M_{i,j}$, for every positive integer h there exist matrices P such that $P = {}^tM_{i_n,j_n}\ {}^tM_{i_{n+1},j_{n+1}} \cdots {}^tM_{i_{n+h-1},j_{n+h-1}}$ for infinitely many n. Suppose there exists an h for which such a P exists with all entries positive. Then there exists a constant c for which there exist infinitely many such n of the form $mh + c$. We define

$$P_1 = {}^tM_{i_1,j_1}\ {}^tM_{i_2,j_2} \cdots {}^tM_{i_{c-1},j_{c-1}}$$

and

$$P_m = {}^tM_{i_{(m-1)h+c},j_{(m-1)h+c}}\ {}^tM_{i_{(m-1)h+c+1},j_{(m-1)h+c+1}} \cdots {}^tM_{i_{mh+c-1},j_{mh+c-1}}.$$

According to Theorem 2 there exist positive numbers $l_0, l_1, \cdots, l_k \in \mathbb{R}^{k+1}$ with $\sum_{j=0}^{k} l_j = 1$ such that

$$\bigcap_{n\in\mathbb{N}} P_1 \cdots P_n(\mathcal{C}_+^{k+1}) = \mathbb{R}_+\ {}^t(l_0, l_1, \ldots, l_k).$$

It follows (with the notation of Section 5.1) that

$$\bigcap_{n\in\mathbb{N}} {}^tM_{i_1,j_1} \cdots {}^tM_{i_n,j_n}(\mathcal{C}_+^{k+1}) = \mathbb{R}_+\ {}^t(l_0, l_1, \ldots, l_k).$$

In particular,

$$\lim_{n\to\infty}(r_0^{(n)}, \cdots, r_k^{(n)}) = \lim_{n\to\infty} {}^tM_{i_1,j_1} \cdots {}^tM_{i_n,j_n}\ {}^t(1, \cdots, 1) \in \mathbb{R}_+\ {}^t(l_0, l_1, \ldots, l_k).$$

Since $r_0^{(n)} + r_1^{(n)} + \cdots + r_k^{(n)} = D^{(n)}$, we obtain

$$\lim_{n\to\infty} \frac{r_m^{(n)}}{D^{(n)}} = l_m \quad \text{for } m = 0, 1, \cdots k.$$

It follows for these values of m that

$$\lim_{n\to\infty} \rho_m^{(n)} = l_0 + l_1 + \cdots + l_{m-1} =: \rho_m.$$

We define $\chi : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ by

$$\chi(\vec{x}) = m \Longleftrightarrow \rho_m \leq \{\sum_{j=1}^{k} x_j \rho_j\} < \rho_{m+1}, \tag{10}$$

for $\vec{x} = (x_1, \cdots, x_k) \in \mathbb{Z}^k$ and $m = 0, 1, \cdots, k$ where $\rho_0 = 0, \rho_{k+1} = 1$. Thus χ represents a multi-dimensional regular word and the density of the letter m equals $\rho_{m+1} - \rho_m = l_m$ for $m = 0, 1, \cdots, k$.

It remains to prove that $\lim_{n\to\infty} \chi^{(n)} = \chi$. We have from (8)

$$\chi^{(n)}(\vec{x}) = m \Longleftrightarrow \rho_m^{(n)} \leq \{\sum_{j=1}^{k} x_j \rho_j^{(n)}\} < \rho_{m+1}^{(n)}.$$

Fix $\vec{x} = (x_0, x_1, \cdots, x_k) \in \mathbb{Z}^k$. If $\rho_m < \{\sum_{j=1}^k x_j \rho_j\} < \rho_{m+1}$, then $\rho_m^{(n)} < \{\sum_{j=1}^k x_j \rho_j^{(n)}\} < \rho_{m+1}^{(n)}$ for $n \geq n_0(\vec{x})$, so that $\chi^{(n)}(\vec{x}) = m$ for $n \geq n_0(\vec{x})$. Hence $\chi(\vec{x}) = \lim_{n\to\infty} \chi^{(n)}(\vec{x}) = m$.

We now consider the boundary case. Suppose $\{\sum_{j=1}^k x_j \rho_j\} = \rho_m$. Then $\rho_{m-1} < \{\sum_{j=1}^k x_j \rho_j^{(n)}\} < \rho_{m+1}$ for $n \geq n_0(\vec{x})$ so that $\chi^{(n)}(\vec{x})$ is $m-1$ or m (with obvious cyclic adjustments if $m = 0$ or $m = k$). If

$$\rho_m^{(n)} \leq \{\sum_{j=1}^{k} x_j \rho_j\} < \rho_{m+1}^{(n)}$$

for all $n \geq n_1(\vec{x})$, then we are certain that $\lim_{n\to\infty}\chi^{(n)}(\vec{x}) = m$. In case there are infinitely many n such that the reverse inequality holds, then it is not true that the limit exists and equals χ. Thus we have proved the following result.

Theorem 3 *Apply from the starting situation as described in Sec. 4.1 an infinite sequence* (M_{i_n,j_n}) *of basic extension matrices. Suppose there exists a positive integer* h *and a matrix* P *with only positive entries such that* $P = M_{i_n,j_n} M_{i_n+1,j_n+1} \cdots M_{i_n+h-1,j_n+h-1}$ *for infinitely many values of* n. *Let* $\rho_m = \lim_{n\to\infty} \rho_m^{(n)}$ *for* $m = 1, 2, \cdots, k$ *and define* χ *by (9). Then* $\lim_{n\to\infty} \chi^{(n)} = \chi : \mathbb{Z}^k \to \{0, 1, \cdots, k\}$ *exists if and only if*

$$\rho_m^{(n)} \leq \{\sum_{j=1}^{k} x_j \rho_j^{(n)}\} < \rho_{m+1}^{(n)}$$

for $n \geq n_1(\vec{x})$ *for every* $\vec{x} \in \mathbb{Z}^k$ *for which* $\{\sum_{j=1}^k x_j \rho_j\} = \rho_m$ *for some* $m \in \{0, 1, \cdots, k\}$.

Note that the latter condition is fulfilled if $1, \rho_1, \cdots, \rho_k$ are linearly independent over the rationals, that is, if the limit word is Sturmian, since then there cannot be a point $\vec{x} \neq \vec{e}_m$ with $\{\sum_{j=1}^k x_j \rho_j\} = \rho_m$, whereas $\chi^{(n)}(\vec{e}_m) = \chi(\vec{e}_m) = m$ by definition.

Corollary *Suppose that to the starting situation as described in Sec. 4.1 we apply an infinite periodic sequence* $(R_{j_n})_{n\geq 1}$ *of Rauzy steps. If each of* $R_0, R_1, \cdots, R_k$ *occurs in the period and the limit values* $\rho_0, \rho_1, \cdots, \rho_k$ *are linearly independent over the rationals, then the limit word* $\lim_{n\to\infty} \chi^{(n)} = \chi$

exists and it is the Sturmian regular word given for $m = 0, 1, \cdots, k$ *by*

$$\chi(\vec{x}) = m \Longleftrightarrow \rho_m \le \{\sum_{j=1}^{k} x_j \rho_j\} < \rho_{m+1}.$$

Proof Suppose each of $R_0, R_1, \cdots, R_k$ occurs in the period. Note that when applying R_i we multiply by the substitution matrix M_i with nonnegative entries and with entries 1 at the i-th column. Hence the product matrix corresponding to a period has only positive entries at the i-th column for $i = 0, 1, \cdots, k$. Thus the product matrix corresponding to one period has only positive entries. Because of the linear independence condition it follows from Theorem 3 that $\lim_{n\to\infty} \chi^{(n)} = \chi$. ∎

Example 5 (continued). We consider $k = 2$ and apply the Rauzy steps R_2, R_1, R_0 periodically. Then the substitution matrices M_n have row sums $\vec{r}^{(n)}$ satisfying $r_i^{(n)} = c_{0,i}\alpha^n + c_{1,i} + c_{2,i}\beta^n + c_{3,i}\overline{\beta}^n$ for $i = 0, 1, 2$ and all n. Recall that $\alpha > 1, |\beta| < 1$. Since $r_i^{(n)} \to \infty$, we have $\lim_{n\to\infty} \frac{r_i^{(n)}}{D^{(n)}} = c_{0,i}$ for $i = 0, 1, 2$. A direct calculation gives

$$c_{0,0} = \frac{\alpha^2}{(\alpha+1)^2}, \quad c_{0,1} = \frac{1}{\alpha+1}, \quad c_{0,2} = \frac{\alpha}{(\alpha+1)^2}.$$

Since α is a root of the irreducible polynomial $x^3 - x^2 - x - 1$, the numbers $\rho_0 = c_{0,0}, \rho_1 = c_{0,1}, \rho_2 = c_{0,2}$ are linearly independent over the rationals. Thus the limit word $\chi : \mathbb{Z}^2 \to \{0, 1, 2\}$ is given by

$$\chi(x,y) = \begin{cases} 0 \text{ if } 0 \le \{\frac{\alpha}{(\alpha+1)^2}x + \frac{2\alpha+1}{(\alpha+1)^2}y\} < \frac{\alpha^2}{(\alpha+1)^2} \\ 1 \text{ if } \frac{\alpha^2}{(\alpha+1)^2} \le \{\frac{\alpha}{(\alpha+1)^2}x + \frac{2\alpha+1}{(\alpha+1)^2}y\} < \frac{\alpha^2+\alpha+1}{(\alpha+1)^2} \\ 2 \text{ if } \frac{\alpha^2+\alpha+1}{(\alpha+1)^2} \le \{\frac{\alpha}{(\alpha+1)^2}x + \frac{2\alpha+1}{(\alpha+1)^2}y\} < 1. \end{cases}$$

Note that in the above corollary the requirement that each among the rules R_0, $R_1, \cdots, R_k$ occurs in a period is necessary for the conclusion. If R_m does not appear in the period, then $r_m^{(n)} = 1$ for all n, that is, in every $\Lambda^{(n)}$ there is only one coset with $\chi^{(n)}$-value m. Hence the density of the letter m becomes 0 and the limit word cannot be regular.

A similar corollary can of course be given for a periodic sequence of basic extension steps (S_{i_n,j_n}) where each matrix $M_{i,j}$ appears in a period. However, here the condition is far too strong. It suffices, but it is not necessary, that the product matrix taken over a period has positive entries. We shall deal with such situations in part II of the paper.

In this subsection we address the question whether and how some given regular word can be obtained as the limit by applying a tower of extension steps. Not every regular word χ can be the limit word when applying an infinite sequence of Rauzy extensions steps. Indeed, for every n the vector $\vec{r} = (r_0, r_1, \cdots, r_k)$, where $r_0, r_1, \cdots, r_k$ are the densities of the letters $0, 1, \cdots, k$ of the word χ, should be in the convex hull of the column vectors $(D^{(n)})^{-1}(\vec{c}_0^{(n)}, \vec{c}_1^{(n)}, \cdots, \vec{c}_k^{(n)})$ of $M^{(n)}$, because $\vec{c}_0^{(n)} + \vec{c}_1^{(n)} + \cdots + \vec{c}_k^{(n)} = \vec{r}^{(n)}$ and $\lim_{n\to\infty} \frac{\vec{r}^{(n)}}{D^{(n)}} = \vec{r}$. Here we restrict ourselves to the linear manifold $x_0 + x_1 + \cdots + x_k = 1$, since the sum of the densities equals 1. In particular, in case $k = 2$, the vector $\vec{r}^{(n)}$ should be in the convex hull of $\vec{c}_0^{(n)}, \vec{c}_1^{(n)}, \vec{c}_2^{(n)}$. When we start with R_0 we have $\vec{c}_0^{(1)} = {}^t(1, 0, 0)$, $\vec{c}_1^{(1)} = {}^t(1, 1, 0)$, $\vec{c}_2^{(1)} = {}^t(1, 0, 1)$ and for R_1 and R_2 symmetric situations occur. Thus no density vector (r_0, r_1, r_2) inside the triangle with vertices $\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{2}}(1, 0, 1), \frac{1}{\sqrt{2}}(0, 1, 1)$ can be obtained as a limit. In particular, regular words where the letters have about equal densities cannot be the limit words of Rauzy extensions.

A further elaboration of the convexity argument would lead to the conclusion that the limit values $\vec{r}$ of towers of Rauzy extensions have the shape of a Sierpinski triangle fractal.

If we consider basic extension steps, then the situation is entirely different. Suppose that we apply an infinite sequence of basic extension steps $\{S_{i_n, j_n}\}_{n\geq 1}$ to the usual starting position. It is clear that any possible limit density vector $\vec{r} = (r_0, r_1, \cdots, r_k)$ of a regular word satisfies $r_0 + r_1 + \cdots + r_k = 1$ and has nonnegative coefficients. It is therefore in the convex hull of $\vec{e}_0 = \vec{c}_0^{(0)}, \cdots, \vec{e}_k = \vec{c}_k^{(0)}$. Suppose that after n extension steps $(D(n))^{-1}$ the vector $\vec{r}^{(n)}$ is in the convex hull of the column vectors $(D^{(n)})^{-1}\vec{c}_0^{(n)}$, $(D^{(n)})^{-1}\vec{c}_1^{(n)}, \cdots, (D^{(n)})^{-1}\vec{c}_k^{(n)}$. Then applying $M_{i,j}(= M_{i_n, j_n})$, the column vector $\vec{c}_i^{(n)}$ is replaced with the vector $\vec{c}_i^{(n)} + \vec{c}_j^{(n)}$ and the other column vectors are unchanged. If $\vec{r}$ is not in the convex hull of the vectors

$$\frac{\vec{c}_0^{(n)}}{D^{(n)}}, \cdots, \frac{\vec{c}_{i-1}^{(n)}}{D^{(n)}}, \frac{\vec{c}_i^{(n)} + \vec{c}_j^{(n)}}{D^{(n)}}, \frac{\vec{c}_{i+1}^{(n)}}{D^{(n)}}, \cdots, \frac{\vec{c}_k^{(n)}}{D^{(n)}},$$

then it is in the convex hull of the vectors

$$\frac{\vec{c}_0^{(n)}}{D^{(n)}}, \cdots, \frac{\vec{c}_{j-1}^{(n)}}{D^{(n)}}, \frac{\vec{c}_i^{(n)} + \vec{c}_j^{(n)}}{D^{(n)}}, \frac{\vec{c}_{j+1}^{(n)}}{D^{(n)}}, \cdots, \frac{\vec{c}_k^{(n)}}{D^{(n)}}.$$

Thus by replacing $S_{i_{n+1}, j_{n+1}}$ with $S_{j_{n+1}, i_{n+1}}$ if necessary, we keep $\vec{r}$ in the convex hull of the column vectors. Doing so inductively we can guarantee that starting with any sequence of extensions steps $(S_{i_n, j_n})_{n\geq 1}$ and making appropriate interchanges of i's and j's, the vector $\vec{r}$ is in the intersection of the convex hulls of $(D^{(n)})^{-1}\vec{c}_0^{(n)}, (D^{(n)})^{-1}\vec{c}_1^{(n)}, \cdots, (D^{(n)})^{-1}\vec{c}_k^{(n)}$ taken over all n. In particular, if the intersection consists of one point $(r_0, r_1, \cdots, r_k)$ and $r_0, r_1, \cdots, r_k$ are linearly independent over the rationals, then the limit word exists and is a homogeneous Sturmian word.

Acknowledgements

We thank Pierre Arnoux, Anne Siegel and Laurent Vuillon for many fruitful discussions on the subject.

References

[1] E. ALTMAN, B. GAUJAL, A. HORDIJK *Balanced sequences and optimal routing*, J. ACM. **47** (2000), 752-775.

[2] P. ARNOUX, S. ITO *Pisot substitutions and Rauzy fractals*, Bull. Belg. Math. Soc. Simon Stevin **8** (2001), 181–207.

[3] P. ARNOUX, G. RAUZY *Représentation géométrique de suites de complexité* $2n+1$, Bull. Soc. Math. France **119** (1991), 199–215.

[4] P. ARNOUX, V. BERTHÉ, S. ITO *Discrete planes, $\mathbb{Z}^2$-actions, Jacobi-Perron algorithm and substitutions*, Ann. Inst. Fourier (Grenoble) **52** (2002), 1001–1045.

[5] V. BERTHÉ, R. TIJDEMAN *Balance properties of multi-dimensional words*, Theoret. Comput. Sci. **273** (2002), 197–224.

[6] V. BERTHÉ, L. VUILLON *Tilings and rotations on the torus: a two-dimensional generalization of Sturmian sequences*, Discrete Math. **223** (2000), 27–53.

[7] V. BERTHÉ, L. VUILLON *Suites doubles de basse complexité*, Journal de Théorie des Nombres de Bordeaux **12** (2000), 179–208.

[8] V. BERTHÉ, L. VUILLON *Palindromes and two-dimensional Sturmian sequences*, J. Autom. Lang. Comb. **6** (2001), 121–138.

[9] H. FURSTENBERG *Stationary processes and prediction theory*, Annals of Mathematics Studies, Princeton University Press, 1960.

[10] M. KEANE, *Non-ergodic interval exchange transformations*, Israel J. Math. **26** (1977), 188–196.

[11] M. LOTHAIRE *Algebraic combinatorics on words*, Encyclopedia of Mathematics and its Applications **90**, Cambridge University Press, 2002.

[12] N. PYTHEAS FOGG, *Substitutions in Dynamics, Arithmetics and Combinatorics,* edited by V. Berthé, S. Ferenczi, C. Mauduit, A. Siegel, Lecture Notes Mathematical Series **1794**, Springer-Verlag, 2002.

[13] R.J. SIMPSON, R.TIJDEMAN *Multi-dimensional versions of a theorem of Fine and Wilf and a formula of Sylvester*, Proc. Amer. Math. Soc. **131** (2003), 1661-1667.