HIGHER DIMENSIONAL EXTENSIONS OF SUBSTITUTIONS AND THEIR DUAL MAPS

YUKI SANO, PIERRE ARNOUX, AND SHUNJI ITO
TSUDA COLLEGE AND INSTITUT DE MATHÉMATIQUES DE LUMINY
SEPTEMBER 7, 2000; FINAL VERSION

Abstract. Given a substitution σ on d letters, we define its k-dimensional extension, $E_k(\sigma)$, for $0 \le k \le d$. The k-dimensional extension acts on the set of k-dimensional faces of unit cubes in $\mathbb{R}^d$ with integer vertices. The extensions of a substitution satisfy a commutation relation with the natural boundary operator: the boundary of the image is the image of the boundary. We say that a substitution is unimodular (resp. hyperbolic) if the matrix associated to the substitution by abelianization is unimodular (resp. hyperbolic). In the case where the substitution is unimodular, we also define dual substitutions which satisfy a similar coboundary condition. We use these constructions to build self-similar sets on the expanding and contracting space for an hyperbolic substitution.

0. Introduction and statement of results

Let σ be a substitution on the alphabet $\mathcal{W} = \{1, 2, \dots, d\}$. We denote by A_σ the linear map on $\mathbb{Z}^d$ obtained from σ by abelianization.

For any given point $\mathbf{x} \in \mathbb{Z}^d$, it is natural to associate to each word in $\mathcal{W}^*$ a broken path starting in $\mathbf{x}$ (associate to the letter i the unit segment from $\mathbf{x}$ to $\mathbf{x} + \mathbf{e}_i$, where $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ is the canonical basis of $\mathbb{Z}^d$, and extend by continuity). We can then define a map $E_1(\sigma)$ on the set of paths, replacing letter i by the word $\sigma(i)$. Care must be taken of the initial point $\mathbf{x}$, and it is easily computed that the correct definition is given by $(\mathbf{x}, i) \mapsto (A_\sigma(\mathbf{x}), \sigma(i))$ (see section 2 for the formal definition). The map $E_1(\sigma)$ acts in a natural way on the space of formal sums of weighted unit segments.

In this paper, we will define higher dimensional extensions $E_k(\sigma)$ of σ, acting on formal sums of weighted k-dimensional faces of unit cubes with vertices in $\mathbb{Z}^d$. In the case the substitution is *unimodular*, that is, A_σ is an invertible map of $\mathbb{Z}^d$, or has determinant $+1$ or -1, we will also define the dual maps $E_k^*(\sigma)$, and give explicit formulas. We will prove that these maps commute with the natural boundary morphisms, and establish some basic properties.

Before stating definitions and results, we wish to give some motivations, since the aim of this work is to set a framework allowing to understand more deeply and generalize previous results.

0.1. **The Rauzy substitution.** In the paper [Rauzy], we can find a curious compact domain X_σ, called the Rauzy fractal. The domain is constructed in the following manner.

Let σ be the Rauzy substitution on three letters defined by:

$$\sigma: \begin{aligned} 1 &\to 1\,2 \\ 2 &\to 1\,3 \\ 3 &\to 1. \end{aligned}$$

Let $w = (w_1, \ldots, w_n, \ldots)$ be the fixed point of this substitution, A_σ be the linear map associated with σ by abelianization, $\mathcal{P}$ be the contractive invariant plane of A_σ, and $\pi : \mathbb{R}^3 \to \mathcal{P}$ be the projection along the eigenvector corresponding to the maximum eigenvalue λ for A_σ.

The domain X_σ with fractal boundary is obtained as the closure of the set

$$\left\{ \pi \sum_{k=1}^{n} \mathbf{e}_{w_k} | n = 1, 2, \ldots \right\},$$

where $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the canonical basis of $\mathbb{R}^3$. This domain is not only interesting in the viewpoint of fractal geometry, but also in the sense of ergodic theory and number theory. In fact, two dynamical systems, a Markov endomorphism with the structure matrix A_σ and a quasi-periodic motion, act on the domain X_σ, and Rauzy proved that this second dynamical system is measurably conjugate by a continuous map to the dynamical system associated to the substitution σ, See [Rauzy],[Ito-Kimura],[Ito-Ohtsuki], and [Messaoudi].

To study the structure of the domain X_σ one of the authors, motivated by dynamical considerations, introduced a mapping $E_1^*(\sigma)$ on the set $\mathcal{G}_1^*$, the $\mathbb{Z}$-module denoted by

$$\mathcal{G}_1^* = \left\{ \sum_{\lambda \in \Lambda_1^*} n_\lambda \lambda \;\middle|\; n_\lambda \in \mathbb{Z}, \sharp\{\lambda \in \Lambda_1^* | n_\lambda \neq 0\} < +\infty \right\}$$

and $\Lambda_1^* := \mathbb{Z}^3 \times \{1^*, 2^*, 3^*\}$ (This set should be interpreted geometrically as the set of formal sums of weighted faces, i^* being the unit face orthogonal to the segment of direction $\mathbf{e}_i$ and $(\mathbf{x}, i^*)$ being the corresponding unit face with lower vertex at $\mathbf{x} + \mathbf{e}_i$, so that, for example, $(-\mathbf{e}_1, 1^*)$ is the unit face built on $\mathbf{e}_2$, $\mathbf{e}_3$ at the origin). The mapping $E_1^*(\sigma)$ is defined by:

$$\begin{aligned} E_1^*(\sigma) : (-\mathbf{e}_1, 1^*) &\mapsto (-\mathbf{e}_3, 1^*) + (-\mathbf{e}_3, 2^*) + (-\mathbf{e}_3, 3^*), \\ (-\mathbf{e}_2, 2^*) &\mapsto (-\mathbf{e}_1, 1^*), \\ (-\mathbf{e}_3, 3^*) &\mapsto (-\mathbf{e}_2, 2^*), \end{aligned}$$

and similar formulae for other $(\mathbf{x}, i^*)$.

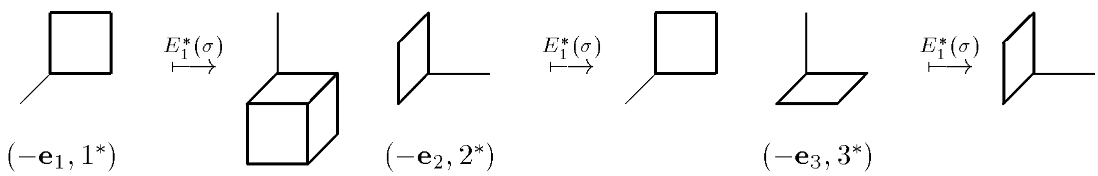

FIGURE 1. The map $E_1^*(\sigma)$ for Rauzy substitution

He proved that, if one considers the element $(\mathbf{x}, 1^*)$ as the unit square with sides $\mathbf{e}_2$, $\mathbf{e}_3$ with lower vertex in $\mathbf{x} + \mathbf{e}_1$, and similarly for the elements $(\mathbf{x}, 2^*)$, $(\mathbf{x}, 3^*)$, then the sequence of sets $A_\sigma^n \left(\pi \left(E_1^*(\sigma)^n \left((-\mathbf{e}_1, 1^*) + (-\mathbf{e}_2, 2^*) + (-\mathbf{e}_3, 3^*)\right)\right)\right)$ converges to the opposite $-X_\sigma$ of the set X_σ defined above.

Moreover, to study the boundary of this set, a mapping $E_2^*(\sigma)$ which satisfies the following commutative diagram was introduced in [Ito-Kimura] and [Ito-Ohtsuki] :

$$\begin{array}{ccc} \mathcal{G}_1^* & \xrightarrow{E_1^*(\sigma)} & \mathcal{G}_1^* \\ \delta\downarrow & & \downarrow\delta \\ \mathcal{G}_2^* & \xrightarrow[E_2^*(\sigma)]{} & \mathcal{G}_2^*, \end{array}$$

where $\mathcal{G}_2^*$ is the space of formal sums of weighted edges, and δ is the canonical boundary map. We call this mapping $E_2^*(\sigma)$ the *boundary endomorphism* of $E_1^*(\sigma)$. The boundary endomorphism $E_2^*(\sigma)$ corresponds to the endomorphism on the free group of rank 3 which produces the boundary of fractile discussed by Dekking in [Dek1], [Dek2]. Using this mapping $E_2^*(\sigma)$, the Hausdorff dimension of the boundary of X_σ is computed in [Ito-Kimura].

Recently, Arnoux and Ito showed that for any *Pisot* substitution σ on d letters, that is, σ is unimodular and A_σ has all eigenvalues, except one, of modulus strictly smaller than one, a compact domain X_σ with fractal boundary can be similarly constructed, using the following mapping $E_1^*(\sigma)$ on $\mathcal{G}_1^*$:

$$E_1^*(\sigma)(\mathbf{x}, i^*) = \sum_{n,j:W_n^{(j)}=i} \left(A_\sigma^{-1}\left(\mathbf{x} - f(P_n^{(j)})\right), j^*\right),$$

where $P_n^{(j)}$ is the prefix of length $n-1$ of $\sigma(j)$, see [Arn-Ito].

In this paper, we show how to generalize this construction in any dimension, defining the map $E_k^*(\sigma)$ under suitable hypotheses.

0.2. **Geometric models for substitutions.** We can associate to any substitution σ a map $E_1(\sigma)$ on 1 dimensional broken paths in $\mathbb{R}^3$, as explained above; it is immediately checked that the image by this map of a closed path is a closed path. One can then ask whether one could define a map on faces, which sends a face with boundary γ to a union of faces with boundary $\sigma(\gamma)$ (See figure 2.3 for the example of Rauzy substitution).

In this paper, we introduce the higher dimensional mapping $E_2(\sigma)$ called 2-*dimensional extension* of σ which solves this problem. It is defined on the set

$$\mathcal{G}_2 := \left\{ \sum_{\lambda\in\Lambda_2} n_\lambda \lambda \;\mid\; n_\lambda \in \mathbb{Z}, \sharp\{\lambda \in \Lambda_2 | n_\lambda \neq 0\} < +\infty \right\}$$

where $\Lambda_2 := \mathbb{Z}^d \times \{i \wedge j \mid 1 \le i < j \le d\}$ (one should think of $\mathcal{G}_2$ as the set of formal sums of weighted 2-dimensional faces). In fact, we will solve the problem in all dimensions, see theorem 2.1.

We will show that the mapping $E_1^*(\sigma)$ (resp. $E_2^*(\sigma)$) defined above is the dual, in the ordinary sense, of the linear mapping $E_1(\sigma)$ (resp.$E_2(\sigma)$), showing that the definition, which seems complicated, is in fact natural; these dual maps also satisfy commutation relation with a suitably defined coboundary operator, see theorem 3.1.

0.3. **The theorem of EI.** In the paper [Ei-Ito], it is proved that a substitution on two letters $1, 2$ is invertible, that is, it extends to an automorphisms of the free group on two elements, if and only the words $\sigma(12)$ and $\sigma(21)$, which have the same length, differ only in two consecutive indices, where one word contains 12 and the other 21. Graphically, this means that the paths associated as above to $\sigma(12)$ and $\sigma(21)$ differ only on the boundary of a unit square.

We can rephrase it as saying that the substitution σ is invertible if and only if the mapping $E_2(\sigma)$ defined below sends a square exactly to a square (with opposite orientation if the determinant of the substitution is -1); it is tempting to ask whether this result can be generalized to more letters, and we will provide a partial answer below.

0.4. **Structure of the paper.** In section 1, we fix notations, define the spaces $\mathcal{G}_k$ of formal sums of weighted k-dimensional faces and their dual spaces $\mathcal{G}_k^*$, and define the boundary and coboundary operators. We also define an isomorphism ϕ_k between $\mathcal{G}_k^*$ and $\mathcal{G}_{n-k}$, similar to Poincaré duality, which allows in suitable cases to give a geometric meaning to elements of $\mathcal{G}_k^*$.

In section 2, we define the main object of this paper, the k-dimensional extension $E_k(\sigma)$ of a substitution σ,and we show that the k-dimensional extensions satisfy a commutation relation with the boundary operator (theorem 2.1); this theorem solves the problem given in section 0.2. In section 3, we show that, under suitable hypothesis (namely, the matrix A_σ has determinant one), we can compute explicitly the dual map $E_k^*(\sigma)$ of $E_k(\sigma)$, and that these maps also satisfy a commutation relation with the coboundary operator.

In section 4, we show that, in the case of hyperbolic substitutions (when A_σ has no eigenvalue of modulus 1), we can use this construction to define, by iteration and renormalization, a self similar set in the stable space and in the unstable space; in the particular case of so-called Pisot substitutions, this can be used to build a geometric model for the dynamical system associated to the substitution, and a Markov partition for the toral automorphism A_σ. This allows us to generalize the construction of the Rauzy fractal. In section 5, we give a weak generalisation of the theorem of Ei: a substitution on d letters gives rise to an automorphism of the free group only if its top-dimensional extension associates a cube to a cube.

In the final two sections, we give a few explicit examples, and raise some open questions and possible generalizations.

A final remark: the definition of the upper-dimensional extensions is more or less natural from the statement of the problem it solves; the reason why we consider the dual maps is much less clear at first sight. We can give two main motivations:

First of all, the dynamical system associated to a substitution on the alphabet $\mathcal{W}$ has very specific property: it is a self-induced system; more specifically, it is a system $T : X \to X$, together with a generating partition $\{X_i | i \in \mathcal{W}\}$, and a subset A with partition $\{A_i | i \in \mathcal{W}\}$, such that the induced map T_A of T on A is conjugate to T, and, for any point $x \in A$, the symbolic dynamics of x under T with respect to the X_i is obtained from the symbolic dynamics of x under T_A with respect to the A_i by applying the substitution σ. Rauzy had first the idea to generate such a system by an"exduction" process, that is, starting with an arbitrary system, and building a larger systems from which it comes by induction, the symbolic dynamics with respect to suitable partitions being related by the given substitution. If one can, by a suitable renormalization, obtain convergence to a self-induced dynamical system, one obtains a geometric model for the substitution. Rauzy was able to build this for specific examples, and later the third author of this paper gave a general formula to build the "exduction"; it turns out that this formula is exactly the dual map of the one-dimensional extension.

Second, it is easy to prove that, for an hyperbolic substitution with unstable space of dimension k, the iterates of a k-dimensional face by $E_k(\sigma)$ stay within bounded distance of the unstable space. We can thus obtain a discrete approximation of the stable space. To approximate the stable space, it would be natural to study the inverse map of $E_k(\sigma)$. However, the k-dimensional extension is not invertible in general; but the dual map is always defined, and can replace the inverse map. One can check below that, in the case of unimodular substitutions, the map $E_0(\sigma)$ is unitary with repect to the natural quadratic form on $\mathcal{G}_0$, so that in that case the dual and inverse maps are isomorphic (this is also the case for A_σ); this does not hold in the higher dimensional case, but then the non-existant inverse map can be replaced by the dual map.

Acknowledgement. We thank the anonymous referee for his many useful comments.

1. **Framework and notations**

1.1. **The substitution.** We consider an alphabet $\mathcal{W} = \{1, 2, \dots, d\}$. The free monoid on $\mathcal{W}$, or the set of finite words on $\mathcal{W}$ is denoted by $\mathcal{W}^* = \bigcup_{n=0}^{\infty}\{1, 2, \dots, d\}^n$.

A substitution σ on $\mathcal{W}^*$ is a non-erasing morphism of the free monoid. It is completely defined by its value on the letters, that is, by a map from $\mathcal{W}$ to $\mathcal{W}^*$ which takes each letter to a non empty word, and extends in a natural way to an endomorphism on $\mathcal{W}^*$ by the rule $\sigma(U)\sigma(V) = \sigma(UV)$, for $U, V \in \mathcal{W}^*$. It also extends to a map on the set $\mathcal{W}^{\mathbb{N}}$ of infinite sequences with value in $\mathcal{W}$, and to an homomorphism on the free group on $\mathcal{W}$.

Notations 1.1. *In the sequel, we take the following notations*

$$\sigma(i) = W^{(i)} = W_1^{(i)} \cdots W_{l_i}^{(i)}$$

where l_i is called the length of $\sigma(i)$. We also write

$$\sigma(i) = W^{(i)} = P_n^{(i)} W_n^{(i)} S_n^{(i)}$$

where $P_n^{(i)} = W_1^{(i)} \cdots W_{n-1}^{(i)}$ is the prefix of length $n-1$ of the word $W^{(i)}$ (empty word for $n = 1$), and $S_n^{(i)} = W_{n+1}^{(i)} \cdots W_{l_i}^{(i)}$ is the suffix of length $l_i - n$ of $W^{(i)}$ (empty word for $n = l_i$).

Let $\{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ be the canonical basis of $\mathbb{Z}^d$. We will denote by f the natural homomorphism (*abelianization*) from $\mathcal{W}^*$ to $\mathbb{Z}^d$ given by $f(i) = \mathbf{e}_i$ for all symbols $i \in \mathcal{W}$. For any finite word $W \in \mathcal{W}^*$, $f(W) = {}^t(x_1, \dots, x_d)$, where x_i is the number of occurrences of the letter i in W.

There exists a unique linear transformation A_σ satisfying the following commutative diagram:

$$\begin{array}{ccc} \mathcal{W}^* & \xrightarrow{\sigma} & \mathcal{W}^* \\ f\downarrow & & \downarrow f \\ \mathbb{Z}^d & \xrightarrow[A_\sigma]{} & \mathbb{Z}^d. \end{array}$$

Remark that the matrix of A_σ, as a linear map on $\mathbb{Z}^d$, is by construction a positive matrix. This is the difference between substitutions and the general case of endomorphisms of free groups; most of what we do could in fact be extended to such an endomorphism.

1.2. **Faces of dimension k and boundary morphisms.** We define the symbolic sets A_k by

$$A_0 := \{\bullet\}$$

and

$$A_k := \{i_1 \wedge i_2 \wedge \cdots \wedge i_k \mid 1 \leq i_1 < i_2 \cdots < i_k \leq d\} \text{ for } 1 \leq k \leq d.$$

Let Λ_k $(0 \leq k \leq d)$ be the following product spaces:

$$\Lambda_k := \mathbb{Z}^d \times A_k.$$

One should think of an element $(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)$ of Λ_k as the face of dimension k, along unit vectors $\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_k}$, of the unit cube at the origin in $\mathbb{R}^d$ translated by $\mathbf{x} \in \mathbb{Z}^d$, and of an element $(\mathbf{x}, \bullet)$ of Λ_0 as the point $\mathbf{x} \in \mathbb{Z}^d$.

Definition 1.1. *We denote by $\mathcal{G}_k$ the free $\mathbb{Z}$-module with as generators the elements of Λ_k:*

$$\mathcal{G}_k := \left\{ \sum_{\lambda \in \Lambda_k} n_\lambda \lambda \;\middle|\; n_\lambda \in \mathbb{Z}, \sharp\{\lambda \in \Lambda_k | n_\lambda \neq 0\} < +\infty \right\}.$$

We think of $\mathcal{G}_k$ as the space of formal finite sums of weighted faces, with weight in $\mathbb{Z}$.

The element $(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)$ has been defined only in the case where $i_1 < i_2 < \cdots < i_k$. We will take advantage of the notation to define it in the general case, by $(\mathbf{x}, i_1 \wedge \cdots \wedge i_k) = 0$ if $i_n = i_m$ for some $n \neq m$, and $(\mathbf{x}, i_{\tau(1)} \wedge \cdots \wedge i_{\tau(k)}) = \epsilon(\tau)(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)$ otherwise, where $\epsilon(\tau)$ is the signature of the permutation τ of $\{1, \ldots, k\}$, so that we recover antisymmetry of the wedge product; for example, $(\mathbf{x}, i_1 \wedge i_2) = -(\mathbf{x}, i_2 \wedge i_1)$.

We will say that an element of $\mathcal{G}_k$ is *geometric* if all its non-zero coefficients are $+1$ or -1. The simplest geometric elements are the elements of Λ_k. To an element $(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)$, we can associate its geometric realization, that is, the set $\{\mathbf{x} + t_1\mathbf{e}_{i_1} + \cdots + t_k\mathbf{e}_{i_k} | 0 \leq t_n \leq 1, 1 \leq n \leq k\}$ together with the orientation given by the basis $\mathbf{e}_{i_1}, \ldots \mathbf{e}_{i_k}$. For an arbitrary geometric element $\sum_{\lambda \in \Lambda_k} n_\lambda \lambda$, we define its geometric realization as the union of the geometric realizations of the elements λ such that $n_\lambda \neq 0$. In particular, the geometric realization of 0 is the empty set.

Since it should not lead to confusion, we will denote a geometric element and its realization in the same way.

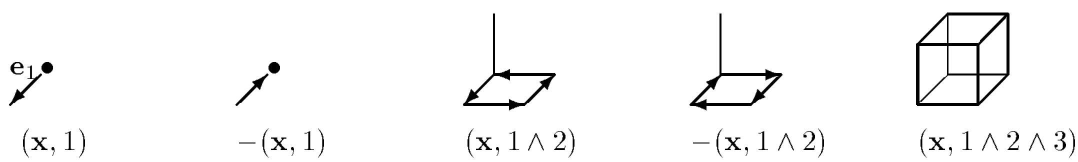

FIGURE 2. Examples of faces with 3 letters

We can now define boundary maps:

Definition 1.2. *The boundary maps $\delta_k : \mathcal{G}_k \to \mathcal{G}_{k-1}$ $(1 \le k \le d)$ are defined on basis vectors by:*

$$\delta_k(\mathbf{x}, i_1 \wedge \cdots \wedge i_k) := \sum_{n=1}^{k} (-1)^n \{(\mathbf{x}, i_1 \wedge \cdots \wedge \widehat{i_n} \wedge \cdots \wedge i_k) - (\mathbf{x} + \mathbf{e}_{i_n}, i_1 \wedge \cdots \wedge \widehat{i_n} \wedge \cdots \wedge i_k)\},$$

where as usual, $i_1 \wedge \cdots \wedge \widehat{i_n} \wedge \cdots \wedge i_k = i_1 \wedge \cdots \wedge i_{n-1} \wedge i_{n+1} \wedge \cdots \wedge i_k$ and in the case $k = 1$, we put $\hat{i} = \bullet$.

In particular, δ_1 is given by: $\delta_1(\mathbf{x}, i) = -(\mathbf{x}, \bullet) + (\mathbf{x} + \mathbf{e}_i, \bullet)$.

We extend δ_k to all of $\mathcal{G}_k$ by linearity; it is straightforward to check that the boundary maps satisfy the usual relations $\delta_{k-1} \circ \delta_k = 0$.

1.3. **Duality.** One can define in the obvious manner a dual space; since we are in an infinite-dimensional $\mathbb{Z}$-module, this defines a complicated space, and we will restrict ourself to the set of dual maps with finite support. We will denote this set by $\mathcal{G}_k^*$; it has a natural basis, and we will denote by $(\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*)$ the dual vector of $(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)$.

Defining the sets A_k^* by $A_k := \{i_1^* \wedge i_2^* \wedge \cdots \wedge i_k^* \mid 1 \le i_1 < i_2 \cdots < i_k \le d\}$ for $1 \le k \le d$ and $\Lambda_k^* := \mathbb{Z}^d \times A_k^*$, we see that $\mathcal{G}_k^*$ is the free $\mathbb{Z}$-module generated by the elements of Λ_k^*.

We will write $\langle\rangle$ the natural product between $\mathcal{G}_k$ and its dual, that is, for an element F of $\mathcal{G}_k$ and ϕ of the dual, we define $\langle F, \phi \rangle = \phi(F)$.

We can get the formula of dual boundary maps δ_k^* from $\mathcal{G}_k^*$ to $\mathcal{G}_{k+1}^*$:

Proposition 1.1. *The dual boundary maps $\delta_k^* : \mathcal{G}_{k-1}^* \to \mathcal{G}_k^*$ $(1 \le k \le d)$ are given as follows:*

$$\delta_k^*(\mathbf{x}, i_1^* \wedge \cdots \wedge i_{k-1}^*) = \sum_{n=1}^{d-k+1} (-1)^{j_n-n+1} \{(\mathbf{x}, i_1^* \wedge \cdots \wedge i_{j_n-n}^* \wedge j_n^* \wedge i_{j_n-n+1}^* \wedge \cdots \wedge i_{k-1}^*) - (\mathbf{x} - \mathbf{e}_{j_n}, i_1^* \wedge \cdots \wedge i_{j_n-n}^* \wedge j_n^* \wedge i_{j_n-n+1}^* \wedge \cdots \wedge i_{k-1}^*)\},$$

where $\{j_1, j_2, \ldots, j_{d-k+1}\} = \mathcal{W} \backslash \{i_1, i_2, \ldots, i_{k-1}\}$, with $i_1 < \cdots < i_{k-1}$, $j_1 < \cdots < j_{d-k+1}$, and $i_{j_n-n} < j_n < i_{j_n-n+1}$. In the case $k = 1$, we put $\{\bullet\} := \emptyset$, hence δ_1^* is given by

$$\delta_1^*(\mathbf{x}, \bullet^*) = \sum_{i=1}^{d} \{(\mathbf{x} - \mathbf{e}_i, i^*) - (\mathbf{x}, i^*)\}.$$

Proof. This is simple computation using the fact that, by definition, we have

$$< \delta_k^*(\mathbf{x}, i_1^* \wedge \cdots \wedge i_{k-1}^*), (\mathbf{y}, j_1 \wedge \cdots \wedge j_k) >=< (\mathbf{x}, i_1^* \wedge \cdots \wedge i_{k-1}^*), \delta_k(\mathbf{y}, j_1 \wedge \cdots \wedge j_k) > .$$

The only non-trivial part is the fact that $i_{j_n-n} < j_n < i_{j_n-n+1}$. Recall that the finite sequences $(i_1, \ldots, i_{k-1})$ and $(j_1, \ldots, j_{d-k+1})$ are increasing, and complementary of each other in $\{1, 2, \ldots, d\}$. Consider the set $\{1, 2, \ldots, j_n - 1\}$; by definition, it contains $j_n - 1$ elements, and among these, $n - 1$ elements belong to $(j_1, \ldots, j_{n-1})$. Hence, the remaining $j_n - n$ elements belong to $(i_1, \ldots, i_{k-1})$, which proves that $i_{j_n-n} < j_n < i_{j_n-n+1}$. □

It is easily computed that A_k^* and A_{d-k} have the same cardinal, and we can use this to define linear isomorphisms between $\mathcal{G}_k^*$ and $\mathcal{G}_{d-k}$:

Definition 1.3. *For $0 \le k \le d$, we define maps φ_k from $\mathcal{G}_k^*$ to $\mathcal{G}_{d-k}$ by:*

$$\varphi_k(\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*) := (-1)^{i_1+\cdots i_k}(\mathbf{x} + \mathbf{e}_{i_1} + \cdots + \mathbf{e}_{i_k}, j_1 \wedge \cdots \wedge j_{d-k}),$$

where $\{i_1, \ldots i_k\}$ and $\{j_1, \ldots j_{d-k}\}$ form a partition of $\{1, 2, \ldots, d\}$, with $i_1 < \cdots < i_k$, $j_1 < \cdots < j_{d-k}$. In the case $k = 0$, we put $(-1)^\bullet := 1$, $\mathbf{e}_\bullet := \mathbf{0}$, and $\{\bullet\} := \emptyset$, hence φ_0 is given by

$$\varphi_0(\mathbf{x}, \bullet^*) := (\mathbf{x}, 1 \wedge 2 \wedge \cdots \wedge d).$$

The map φ_k is one-to-one. It is possible to define as above geometric elements in the dual, and one can check that φ_k sends a geometric element of the dual to a geometric element. We can thus define the geometric realization of a geometric dual element; as above, we will not make a difference between an element and its geometric realization.

A straightforward computation, left to the reader, shows that the duality isomorphisms conjugate boundary and coboundary:

Proposition 1.2. *The dual boundary maps $\delta_k^* : \mathcal{G}_{k-1}^* \to \mathcal{G}_k^*$ $(1 \le k \le d)$ satisfy the equality:*

$$\varphi_k \circ \delta_k^* = \delta_{d-k+1} \circ \varphi_{k-1}.$$

2. Definition of higher dimensional extensions and commutation with the boundary operators

2.1. Definition of $E_0(\sigma)$ and $E_1(\sigma)$.

The unique linear map A_σ on $\mathbb{Z}^d$ that commutes with the abelianization extends in a natural way to a map $E_0(\sigma)$ on $\mathcal{G}_0$ by $E_0(\sigma)(\mathbf{x}, \bullet) := (A_\sigma(\mathbf{x}), \bullet)$.

We want to define a map on $\mathcal{G}_1$. It suffices to define it on basis elements $(\mathbf{x}, i)$, and to extend by linearity. Recall the informal motivation given at the beginning of the paper: to the basis element $(\mathbf{x}, i)$ we associate a unit segment along vector $\mathbf{e}_i$ starting at $\mathbf{x}$. We want to define $E_1(\sigma)$ in such a way that it sends $(\mathbf{x}, i)$, to some path associated to the word $\sigma(i)$. For the image of a continuous path associated to ij to be a continuous path, we need also to act on the origin of the segment, which leads, on the notation $\sigma(i) = P_n^{(i)} W_n^{(i)} S_n^{(i)}$, to the following definition:

Definition 2.1. *The map $E_1(\sigma)$ is the unique linear map on $\mathcal{G}_1$ defined by:*

$$E_1(\sigma)(\mathbf{x}, i) = \sum_{n=1}^{l_i} \left(A_\sigma(\mathbf{x}) + f(P_n^{(i)}), W_n^{(i)} \right).$$

It is easy to check that the image of a geometric element corresponding to a path from $\mathbf{x}$ to $\mathbf{y}$ is a path from $A_\sigma(\mathbf{x})$ to $A_\sigma(\mathbf{y})$. This is a geometric way to visualize the abelianization map. In particular, taking negative signs to take into account the orientation, we can consider closed paths, whose images will again be closed paths.

An other remark is that we can extend $E_1(\sigma)$ to elements with infinite support, in particular to infinite paths. If we consider an infinite sequence which is a fixed point of the substitution, the corresponding path from the origin is fixed by $E_1(\sigma)$. This can be a interesting way to represent this fixed sequence.

2.2. **Definition of $E_k(\sigma)$.** We want now to define linear maps $E_k(\sigma)$ on $\mathcal{G}_k$, for all k. By linearity, it is enough to define them on basis elements, and the following definition encompasses the previous one for $E_1(\sigma)$.

Definition 2.2. *We define the higher dimensional extension of dimension k of σ, denoted by $E_k(\sigma)$, for $1 \le k \le d$, by:*

$$\begin{aligned} & E_k(\sigma)(\mathbf{x}, i_1 \wedge \cdots \wedge i_k) \\ & := \sum_{n_1=1}^{l_{i_1}} \cdots \sum_{n_k=1}^{l_{i_k}} \left(A_\sigma(\mathbf{x}) + f(P_{n_1}^{(i_1)}) + \cdots + f(P_{n_k}^{(i_k)}), W_{n_1}^{(i_1)} \wedge \cdots \wedge W_{n_k}^{(i_k)} \right) \end{aligned}$$

where we use the antisymmetry of the wedge product: $(\mathbf{y}, j_1 \wedge \cdots \wedge j_k) = 0$ if there exists $l \neq m$ such that $j_l = j_m$ and otherwise, $(\mathbf{y}, j_{\tau(1)} \wedge \cdots \wedge j_{\tau(k)}) = \epsilon(\tau)(\mathbf{y}, j_1 \wedge \cdots \wedge j_k)$.

2.3. **Commutation with the boundary operator.** We will show in this section that the maps defined above solve one of the questions of the introduction. Namely, the image of the boundary of a k-dimensional face is the boundary of the image of this face, that is, the maps $E_k(\sigma)$ commute with the boundary operators.

Theorem 2.1. *The following commutative diagram holds:*

$$\begin{array}{ccccc} & \mathcal{G}_d & \xrightarrow{E_d(\sigma)} & \mathcal{G}_d & \\ \delta_d & \downarrow & & \downarrow & \delta_d \\ & \mathcal{G}_{d-1} & \xrightarrow{E_{d-1}(\sigma)} & \mathcal{G}_{d-1} & \\ \delta_{d-1} & \downarrow & & \downarrow & \delta_{d-1} \\ & \vdots & & \vdots & \\ & \mathcal{G}_1 & \xrightarrow{E_1(\sigma)} & \mathcal{G}_1 & \\ \delta_1 & \downarrow & & \downarrow & \delta_1 \\ & \mathcal{G}_0 & \xrightarrow{E_0(\sigma)} & \mathcal{G}_0. & \end{array}$$

Proof. We need to show the relation

$$\delta_k \circ E_k(\sigma)(\mathbf{x}, i_1 \wedge i_2 \wedge \cdots \wedge i_k) = E_{k-1}(\sigma) \circ \delta_k(\mathbf{x}, i_1 \wedge i_2 \wedge \cdots \wedge i_k) \tag{2.1}$$

For $k = 1$, it is easily checked; in that case, it just means that the image of a path from $\mathbf{x}$ to $\mathbf{y}$ is a path from $A_\sigma(\mathbf{x})$ to $A_\sigma(\mathbf{y})$, which was the initial purpose of the definition of $E_1(\sigma)$. We will prove the relation for $k = 2$, leaving the general case to the reader.

The right hand side of the relation is given by

$$\begin{aligned} & E_1(\sigma) \circ \delta_2(\mathbf{x}, i_1 \wedge i_2) \\ & = E_1(\sigma)(\mathbf{x}, i_1) - E_1(\sigma)(\mathbf{x} + \mathbf{e}_{i_2}, i_1) - E_1(\sigma)(\mathbf{x}, i_2) + E_1(\sigma)(\mathbf{x} + \mathbf{e}_{i_1}, i_2) \\ & = \sum_{k=1}^{l_{i_1}} \left(A_\sigma(\mathbf{x}) + f(P_k^{(i_1)}), W_k^{(i_1)} \right) - \left(A_\sigma(\mathbf{x} + \mathbf{e}_{i_2}) + f(P_k^{(i_1)}), W_k^{(i_1)} \right) \\ & \qquad - \sum_{k=1}^{l_{i_2}} \left(A_\sigma(\mathbf{x}) + f(P_k^{(i_2)}), W_k^{(i_2)} \right) - \left(A_\sigma(\mathbf{x} + \mathbf{e}_{i_1}) + f(P_k^{(i_2)}), W_k^{(i_2)} \right) \end{aligned}$$

The left hand side of the relation is given by

$$\begin{aligned}
&\delta_2 \circ E_2(\sigma)\,(\mathbf{x}, i_1 \wedge i_2)\\
&= \delta_2 \left[\sum_{j=1}^{l_{i_1}} \sum_{k=1}^{l_{i_2}} \left(A_\sigma(\mathbf{x}) + f(P_j^{(i_1)}) + f(P_k^{(i_2)}), W_j^{(i_1)} \wedge W_k^{(i_2)} \right) \right]\\
&= \sum_{j=1}^{l_{i_1}} \sum_{k=1}^{l_{i_2}} \Big[\left(A_\sigma(\mathbf{x}) + f(P_j^{(i_1)}) + f(P_k^{(i_2)}), W_j^{(i_1)} \right)\\
&\qquad\qquad - \left(A_\sigma(\mathbf{x}) + f(P_j^{(i_1)}) + f(P_k^{(i_2)}) + f(W_k^{(i_2)}), W_j^{(i_1)} \right) \Big]\\
&- \sum_{j=1}^{l_{i_1}} \sum_{k=1}^{l_{i_2}} \Big[\left(A_\sigma(\mathbf{x}) + f(P_j^{(i_1)}) + f(P_k^{(i_2)}), W_k^{(i_2)} \right)\\
&\qquad\qquad - \left(A_\sigma(\mathbf{x}) + f(P_j^{(i_1)}) + f(P_k^{(i_2)}) + f(W_j^{(i_1)}), W_k^{(i_2)} \right) \Big]
\end{aligned}$$

Since we have, for $1 \leq k < l_{i_2}$, $f(P_k^{(i_2)}) + f(W_k^{(i_2)}) = f(P_{k+1}^{(i_2)})$ by definition, we see that, in the first sum, for fixed j, all terms except the first and last cancel in pairs; since we have $f(P_{l_{i_2}}^{(i_2)}) + f(W_{l_{i_2}}^{(i_2)}) = f(\sigma(i_2)) = A_\sigma(\mathbf{e}_{i_2})$, the first sum reduces to

$$\sum_{j=1}^{l_{i_1}} \left(A_\sigma(\mathbf{x}) + f(P_j^{(i_1)}), W_j^{(i_1)} \right) - \left(A_\sigma(\mathbf{x} + \mathbf{e}_{i_2}) + f(P_j^{(i_1)}), W_j^{(i_1)} \right).$$

A similar argument for the second sum, fixing k and varying j, proves both sides of the relation 2.1 are equal. □

The first non-trivial example is given by the Rauzy substitution, see figure 2.3.

3. Dual maps for unimodular substitutions, and coboundary

3.1. **The dual map.** Since $E_k(\sigma)$ are linear maps, we can define their dual maps on the dual spaces of $\mathcal{G}_k$. For an element F of $\mathcal{G}_k$ and ϕ of the dual, the dual map is defined by the relation $\langle F, E_k^*(\sigma)\phi \rangle = \langle E_k(\sigma)F, \phi \rangle$.

Note that, if A_σ is not one-to-one, there can be infinitely many elements of Λ_k whose image contain a given element, since the equation $A_\sigma(\mathbf{y}) + f(P_n^{(j)}) = \mathbf{x}$ can have infinitely many solutions. Hence, in general, the space of dual elements with finite support is not invariant by $E_k^*(\sigma)$.

If however the map A_σ is one-to-one, this argument proves that the space $\mathcal{G}_k^*$ of elements with finite support is invariant by $E_k^*(\sigma)$, and it makes sense to compute explicitly this map for the canonical basis.

3.2. **Explicit computation of the dual map: the unimodularity condition.** A substitution is said to be *unimodular* if its abelianization A_σ has determinant $+1$ or -1.

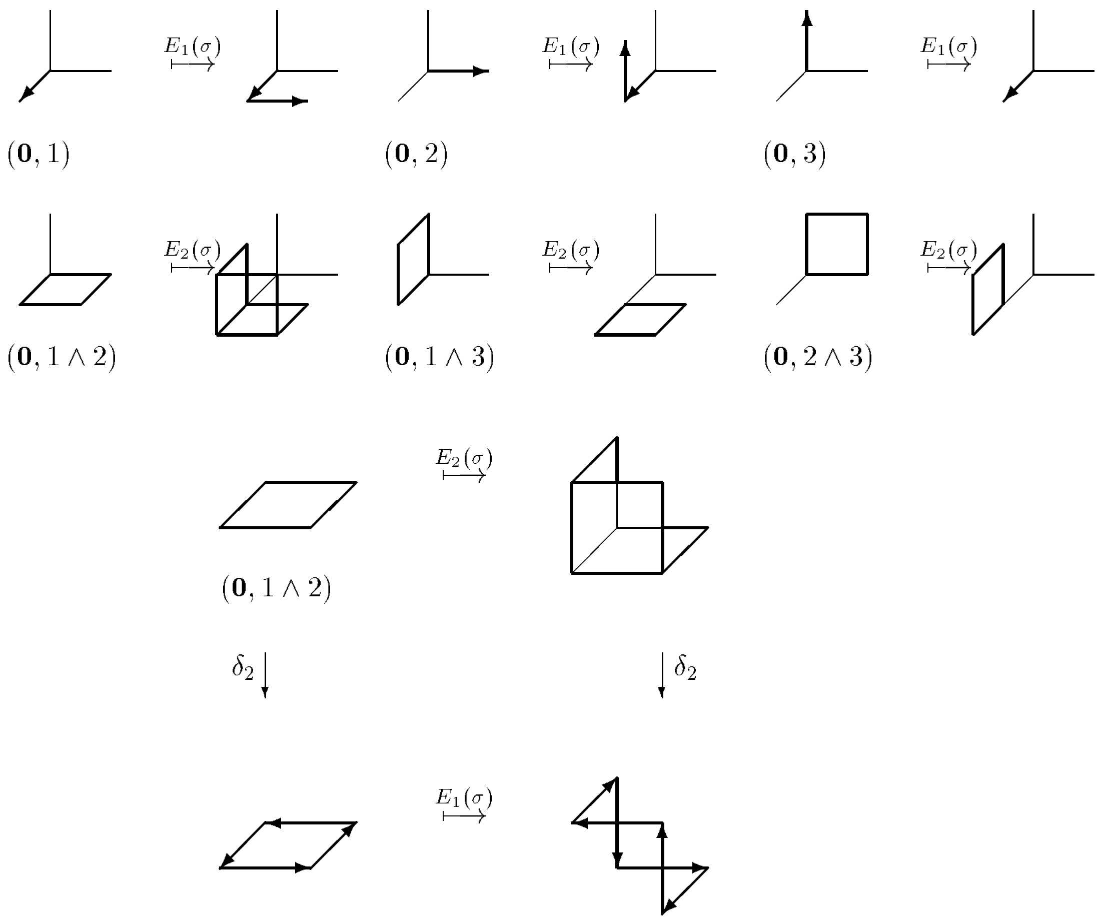

FIGURE 3. The figure of $E_k(\sigma)$, $k = 1, 2$ and the commutation with boundary for Rauzy substitution

To compute the image of a basis element of the dual, we must calculate its value on a basis element of $\mathcal{G}_k$, which is given as follows:

$$\begin{aligned} &< (\mathbf{y}, j_1 \wedge \cdots \wedge j_k), E_k^*(\sigma)(\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*) > \\ &= < E_k(\sigma)(\mathbf{y}, j_1 \wedge \cdots \wedge j_k), (\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*) > \\ &= \sum_{p=1}^{k} \sum_{n_p=1}^{l_{j_p}} < (A_\sigma(\mathbf{y}) + \sum_{p=1}^{k} f(P_{n_p}^{(j_p)}), W_{n_1}^{(j_1)} \wedge \cdots \wedge W_{n_k}^{(j_k)}), (\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*) > . \end{aligned}$$

The product takes value 0 except if one of the faces in the left-hand side corresponds to the dual element on the right hand side, that is, if we can find indices $n_p, p = 1, \ldots, k$ such that:

$$A_\sigma(\mathbf{y}) + \sum_{p=1}^{k} f(P_{n_p}^{(j_p)}) = \mathbf{x}$$

and the face $W_{n_1}^{(j_1)} \wedge \cdots \wedge W_{n_k}^{(j_k)}$ is equal, up to orientation, to the face $i_1 \wedge \cdots \wedge i_k$; that is, there exists a permutation τ of $\{1, \ldots, k\}$ such that,for all $1 \leq l \leq k$

$$W_{n_l}^{(j_l)} = i_{\tau(l)}.$$

In that case, we have

$$< (A_\sigma(\mathbf{y}) + \sum_{p=1}^{k} f(P_{n_p}^{(j_p)}), W_{n_1}^{(j_1)} \wedge \cdots \wedge W_{n_k}^{(j_k)}), (\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*) >= \epsilon(\tau)$$

In general, this can lead to a complicated case study, since A_σ needs not be invertible on $\mathbb{Z}^d$. If however the substitution is unimodular, A_σ can be inverted, and the computation above gives an explicit formula for the dual.

Proposition 3.1. *If the substitution σ is unimodular, the mappings $E_k^*(\sigma)$ $(1 \le k \le d)$ are given on $\mathcal{G}_k^*$ by:*

$$E_k^*(\sigma)(\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*)$$
$$= \sum_{\tau \in S_k} \sum_{1 \le l \le k} \sum_{W_{n_l}^{(j_l)} = i_{\tau(l)}} \epsilon(\tau) \left(A_\sigma^{-1} \left(\mathbf{x} - f(P_{n_1}^{(j_1)}) - \cdots - f(P_{n_k}^{(j_k)}) \right), j_1^* \wedge \cdots \wedge j_k^* \right).$$

In the special case $k = 0$, the formula also makes sense, and it is worth to remark that $E_0^*(\sigma)$ is also the inverse of $E_0(\sigma)$, as was already noted at the end of the introduction. For $k = 1$, we obtain the simpler formula already given in the introduction:

Corollary 3.1. *If the substitution σ is unimodular, the mapping $E_1^*(\sigma)$ is given by:*

$$E_1^*(\sigma)(\mathbf{x}, i) = \sum_{n,j : W_n^{(j)} = i} \left(A_\sigma^{-1} \left(\mathbf{x} - f(P_n^{(j)}) \right), j^* \right).$$

3.3. **Commutation with coboundary.** We get the following relation among dual mappings, using the commutation relations proved in the previous section, and the property of a dual mappping:

Theorem 3.1. *The following commutative diagram holds:*

$$\begin{array}{ccccc}
 & \mathcal{G}_0^* & \xrightarrow{E_0^*(\sigma)} & \mathcal{G}_0^* & \\
\delta_1^* & \downarrow & & \downarrow & \delta_1^* \\
 & \mathcal{G}_1^* & \xrightarrow{E_1^*(\sigma)} & \mathcal{G}_1^* & \\
\delta_2^* & \downarrow & & \downarrow & \delta_2^* \\
 & \vdots & & \vdots & \\
\delta_{d-1}^* & \downarrow & & \downarrow & \delta_{d-1}^* \\
 & \mathcal{G}_{d-1}^* & \xrightarrow{E_{d-1}^*(\sigma)} & \mathcal{G}_{d-1}^* & \\
\delta_d^* & \downarrow & & \downarrow & \delta_d^* \\
 & \mathcal{G}_d^* & \xrightarrow{E_d^*(\sigma)} & \mathcal{G}_d^*. &
\end{array}$$

Proof. From Theorem 3.1, $\delta_k \circ E_k(\sigma) = E_{k-1}(\sigma) \circ \delta_k$. Using the property of composition of dual mappings, $(\phi \circ \psi)^* = \psi^* \circ \phi^*$, we have $E_k^*(\sigma) \circ \delta_k^* = \delta_k^* \circ E_{k-1}^*(\sigma)$. □

3.4. **Geometric interpretation of dual mappings.** Recall that we can identify dual spaces $\mathcal{G}_k^*$ to $\mathcal{G}_{d-k}$, using isomorphisms $\varphi_k : \mathcal{G}_k^* \to \mathcal{G}_{d-k}$ given by definition 1.3. If we conjugate dual mappings $E_k^*(\sigma)$ on $\mathcal{G}_k^*$ by these isomorphisms, we obtain mappings on $\mathcal{G}_{d-k}$. A straightforward computation, using proposition 1.2, proves that these mappings commute with the usual boundary morphism.

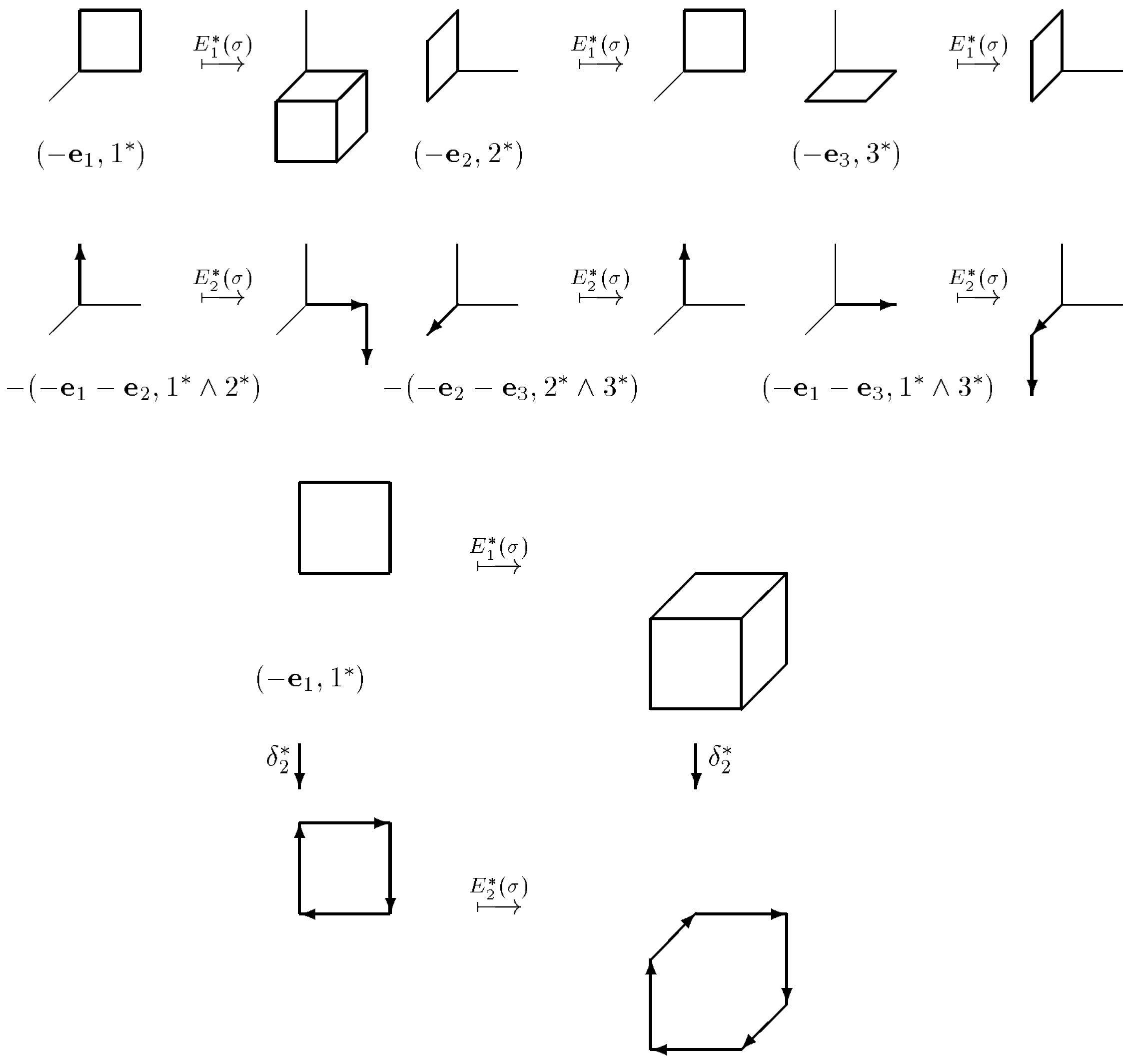

FIGURE 4. The figures of $E_k^*(\sigma)$, $k = 1, 2$ and the commutation with coboundary for Rauzy substitution

An application of this property is the following: in some cases, we are interested in the geometric set associated to $E_1^*(\sigma)^n(\mathbf{x}, i^*)$; in particular, we want to study its boundary, and this is just the geometric set associated to $E_2^*(\sigma)^n(\mathbf{x}, i^*)$; this can give an easy way to study this boundary, and in particular its dimension.

The figures 2.3 and 3.4 show, in the case of Rauzy substitution, the commutative relations with the boundary map δ_2 and the coboundary map δ_2^*; this answers the questions of sections 0.1 and 0.2.

4. Hyperbolic substitutions: Hausdorff convergence of renormalized iteration

Definition 4.1. *A substitution is said to be hyperbolic if it is unimodular and its abelianization A_σ has no eigenvalue of modulus 1.*

In that case, the space $\mathbb{R}^n$ splits into a pair of invariant spaces: the stable space E_s, where the restriction of the map A_σ is strictly contracting for an appropriate norm, and the unstable space E_u, where A_σ is strictly expanding; we will denote by π_s (resp. π_u) the projection on E_s along E_u (resp. on E_u along E_s).

Let us consider an hyperbolic substitution, with an unstable space of dimension k. Remark that, since A_σ is strictly expanding on E_u, it is one-to-one from E_u to itself. Hence, without any hypothesis, A_σ^{-1}, as a linear map on the real vector space E_u, is well defined (even if A_σ is not invertible as a map on $\mathbb{R}^d$).

Theorem 4.1. *Let σ be an hyperbolic substitution, with an unstable space of dimension k. For any k-dimensional face $(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)$, the sequence of compact subsets of E_u:*

$$X_n = A_\sigma^{-n}\left(\pi_u\left(E_k(\sigma)^n(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)\right)\right)$$

converges in the sense of Hausdorff.

Proof. The main ingredient in the proof is the following lemma, whose proof is left to the reader:

Lemma 4.1. *Let A, B, C, D be 4 compact sets in $\mathbb{R}^d$. The Hausdorff metric satisfies $d_h(A \cup B, C \cup D) \leq \max(d_h(A, C), d_h(B, D))$.*

A simple computation shows that the Hausdorff distance between $\pi_u(\mathbf{x}, i_1 \wedge \ldots \wedge i_k)$ and $A_\sigma^{-1}(\pi_u(E_k(\sigma)(\mathbf{x}, i_1 \wedge \cdots \wedge i_k)))$ is bounded by a constant K independent of $\mathbf{x}$, since the renormalization by A_σ^{-1} cancels the A_σ that occurs in the definition of $E_k(\sigma)$.

But we know, by hypothesis, that A_σ^{-1} is strictly contracting with ratio $\lambda < 1$ for a suitable norm. Using the lemma, we immediately obtain $d_H(X_n, X_{n+1}) \leq K\lambda^n$. The distance decreases exponentially fast, hence the sequence $(X_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in the Hausdorff topology. But it is well known that the space of compact subsets of $\mathbb{R}^d$ is complete for the Hausdorff topology. Hence the sequence $(X_n)_{n\in\mathbb{N}}$ converges. □

A similar theorem can be proved for the dual maps in the unimodular case:

Theorem 4.2. *Let σ be an unimodular hyperbolic substitution, with a stable space of dimension $d-k$. For any k-dimensional dual face $(\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*)$, the sequence of compact subsets of E_s*

$$X_n = A_\sigma^{n}\left(\pi_s\left(E_k^*(\sigma)^n(\mathbf{x}, i_1^* \wedge \cdots \wedge i_k^*)\right)\right)$$

converges in the sense of Hausdorff.

The proof is exactly similar to the preceding one, the only difference being that it is now A_σ, and not A_σ^{-1}, that is contracting.

We can get deeper results in a special case: we say that a substitution is *Pisot* if it is unimodular and all its eigenvalues, except one, are of modulus strictly smaller than one.

In that case, the stable space is of codimension 1; the results above prove that the sets $X^{(i)} = \lim_{n\to\infty} A_\sigma^n(\pi_s(E_1^*(\sigma)^n(\mathbf{0}, i^*))$ are well defined. However, although, for fixed n, the sets $\pi_s(E_1^*(\sigma)^n(\mathbf{0}, i^*)$ are disjoint, up to sets of measure 0, it is unclear that this property holds in the limit. It works however under a technical condition:

Definition 4.2. *We say that the substitution σ has immediate coincidence for all letters if, for all pairs of letters i, j, there is an index k such that: $W_k^{(i)} = W_k^{(j)}$ and $f(P_k^{(i)}) = f(P_k^{(j)})$*

The geometric meaning of this condition is that, if one represents each word $\sigma(i)$ as a broken path starting from 0, every two paths share at least one edge.

In the paper [Arn-Ito], it is proved that, under this hypothesis, the domains $X^{(i)}$ are pairwise disjoint sets up to sets of Lebesgue measure 0, and the following theorem is proved:

Theorem 4.3. *Let σ be a unimodular Pisot substitution satisfying the coincidence condition. The dynamical system associated to the substitution is measurably conjugate by a continuous map to a domain exchange defined on the sets $X^{(i)}$.*

5. Invertible substitution and top-dimensional extensions

In the maximal dimension, $k = d$, there is only one type of face: the unit cube of dimension d. The corresponding map $E_d(\sigma)$ associates to each unit cube a finite sum of weighted unit cubes, and, abbreviating $C = 1 \wedge 2 \wedge \cdots \wedge d$, can be expressed as $E_d(\sigma)(\mathbf{x}, C) = \sum_{\mathbf{y} \in \mathbb{Z}^d} n_{\mathbf{y}}(\mathbf{y}, C)$.

A straightforward computation shows the following:

Proposition 5.1. *For any substitution, the sum $\sum_{\mathbf{y} \in \mathbb{Z}^d} n_{\mathbf{y}}$ is equal to the determinant of the map A_σ.*

In particular, for a unimodular substitution, this sum is $+1$ or -1.

Definition 5.1. *A substitution is said to be invertible if it extends to an automorphism of the free group.*

Invertible substitution on 2 letters are completely characterized, see [Wen-Wen] and [Mig-See]; in particular, they are *sturmian*, that is, they preserve sturmian words, or words of minimal complexity.

It has been shown in [Ei-Ito] that a necessary and sufficient condition for σ, on two letters, to be invertible, is that the words $\sigma(ij)$ and $\sigma(ji)$ differ only in two consecutive places. If we consider the boundary of the unit square at the origin, and take into account the boundary relations, this exactly means that the substitution σ is invertible if and only if the image by $E_2(\sigma)$ of a unit square is a unit square (with weight $+1$ or -1).

This result can be partially generalized in any dimension:

Proposition 5.2. *Let σ be an invertible substitution on d letters; if we denote $E_d(\sigma)(\mathbf{x}, C)$ by $\sum_{\mathbf{y} \in \mathbb{Z}^d} n_{\mathbf{y}}(\mathbf{y}, C)$, we have $\sum_{\mathbf{y} \in \mathbb{Z}^d} |n_{\mathbf{y}}| = 1$.*

Proof. We just remark that, from a straightforward computation, $E_d(\sigma\tau) = E_d(\sigma)E_d(\tau)$. The generators of the group of automorphisms of the free group are known (see [MaKaSo]), and it is an easy exercice to check that the property is true for all the generators, hence for all the automorphisms of the free group. □

Another way to say it is that, if σ is invertible, the image of a unit cube by $E_d(\sigma)$ is exactly one cube (with positive or negative orientation).

For the end of this proof, we need to check that the definitions we gave for the substitutions are still valid for homomorphisms of free groups, see the last section. A direct proof restricted to substitutions would be much more difficult, since the structure of the monoid of invertible substitutions seems to be quite complicated, except for two letters where it is completely known.

It would be interesting to know if this necessary condition is also sufficient.

6. Examples

Example 1. Let σ be the substitution

$$\sigma: \begin{array}{ll} 1 & \to 121 \\ 2 & \to 12. \end{array}$$

This is called Fibonacci substitution. The matrix A_σ and the inverse matrix ${A_\sigma}^{-1}$ are

$$A_\sigma = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \qquad and \qquad A_\sigma^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

Then the linear mapping $E_1(\sigma)$ and the dual mapping $E_1^*(\sigma)$ are given by

$$\begin{array}{lll} E_1(\sigma): & (\mathbf{0},1) & \mapsto (\mathbf{0},1) + (\mathbf{e}_1,2) + (\mathbf{e}_1+\mathbf{e}_2,1), \\ & (\mathbf{0},2) & \mapsto (\mathbf{0},1) + (\mathbf{e}_1,2), \\ E_1^*(\sigma): & -(-\mathbf{e}_1,1^*) & \mapsto -(-\mathbf{e}_1,1^*) - (-\mathbf{e}_1+\mathbf{e}_2,1^*) - (-\mathbf{e}_1+\mathbf{e}_2,2^*), \\ & (-\mathbf{e}_2,2^*) & \mapsto (-\mathbf{e}_2,2^*) + (-\mathbf{e}_2,1^*). \end{array}$$

These are displayed in figure 6. It is interesting to remark that the image, by the dual substitution $E_1^*(\sigma)$, of each unit tip is a connected path. This is in fact characteristic, among unitary substitutions on 2 letters, of the so called *invertible* substitutions (see [Wen-Wen]), also called *sturmian* substitutions since their fixed point is a sturmian sequence. The next example shows what happens in the non-invertible case; it would be interesting to find analogous properties for 3 or more letters.

Note also that the connectivity property for the dual map arises only if we consider the geometric representation linked to the duality morphisms φ_k; if we do not shift the point $\mathbf{x}$ by $\mathbf{e}_{i_1} + \cdots + \mathbf{e}_{i_k}$, many convenient properties, in this and other cases, are lost.

Example 2. Let σ be the substitution

$$\sigma: \begin{array}{ll} 1 & \to 112 \\ 2 & \to 21. \end{array}$$

The matrix A_σ is the same as the one for Fibonacci substitution given in Example 1. But this is not invertible substitution. It is easy to check this fact by using EI's Theorem. Then the linear

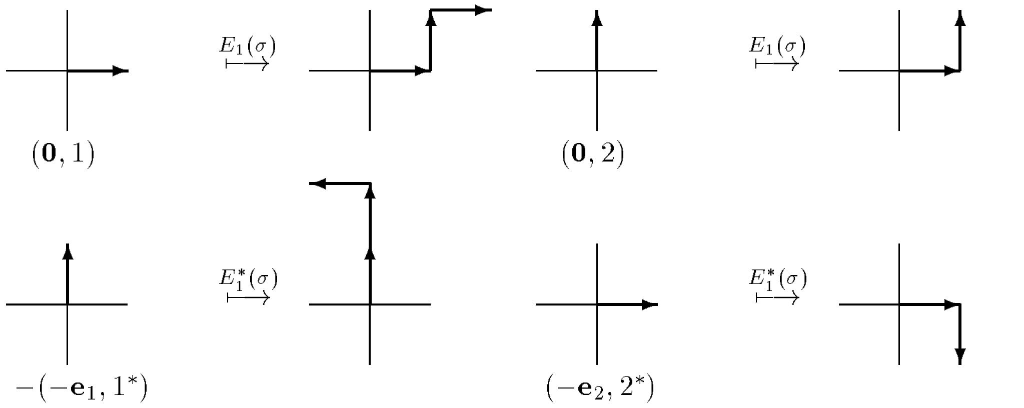

FIGURE 5. The figure of $E_1(\sigma)$ and $E_1^*(\sigma)$ in Example 1

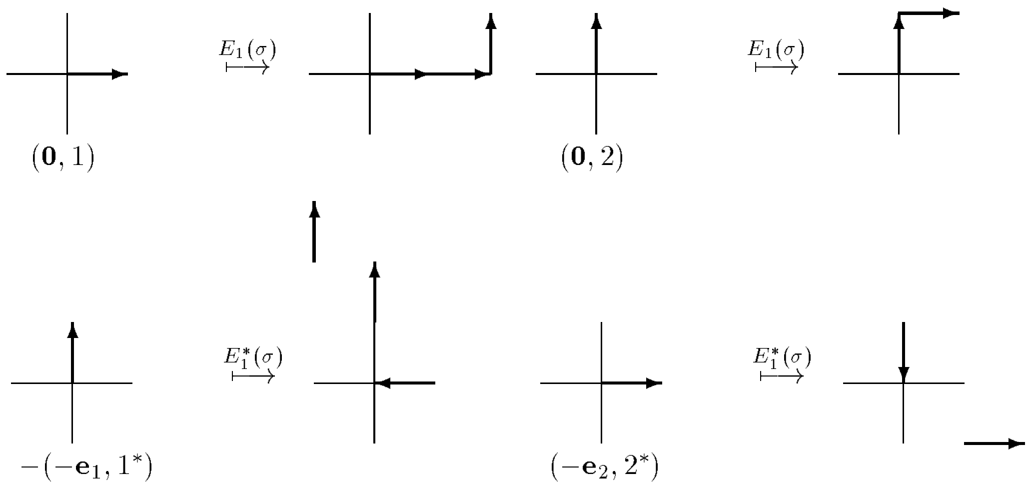

FIGURE 6. The figure of $E_1(\sigma)$ and $E_1^*(\sigma)$ in Example 2

mapping $E_1(\sigma)$ and the dual mapping $E_1^*(\sigma)$ are given by

$$\begin{aligned}
E_1(\sigma): \quad & (\mathbf{0}, 1) && \mapsto (\mathbf{0}, 1) + (\mathbf{e}_1, 1) + (2\mathbf{e}_1, 2), \\
& (\mathbf{0}, 2) && \mapsto (\mathbf{0}, 2) + (\mathbf{e}_2, 1), \\
E_1^*(\sigma): \quad & -(-\mathbf{e}_1, 1^*) && \mapsto -(-\mathbf{e}_1 + \mathbf{e}_2, 1^*) - (-2\mathbf{e}_1 + 2\mathbf{e}_2, 1^*) - (-\mathbf{e}_2, 2^*), \\
& (-\mathbf{e}_2, 2^*) && \mapsto (-\mathbf{e}_1, 1^*) + (\mathbf{e}_1 - 2\mathbf{e}_2, 2^*).
\end{aligned}$$

These are displayed in figure 6.

Example 3. Let σ be the substitution

$$\begin{aligned}
\sigma: \quad 1 & \to 23 \\
2 & \to 123 \\
3 & \to 1122233.
\end{aligned}$$

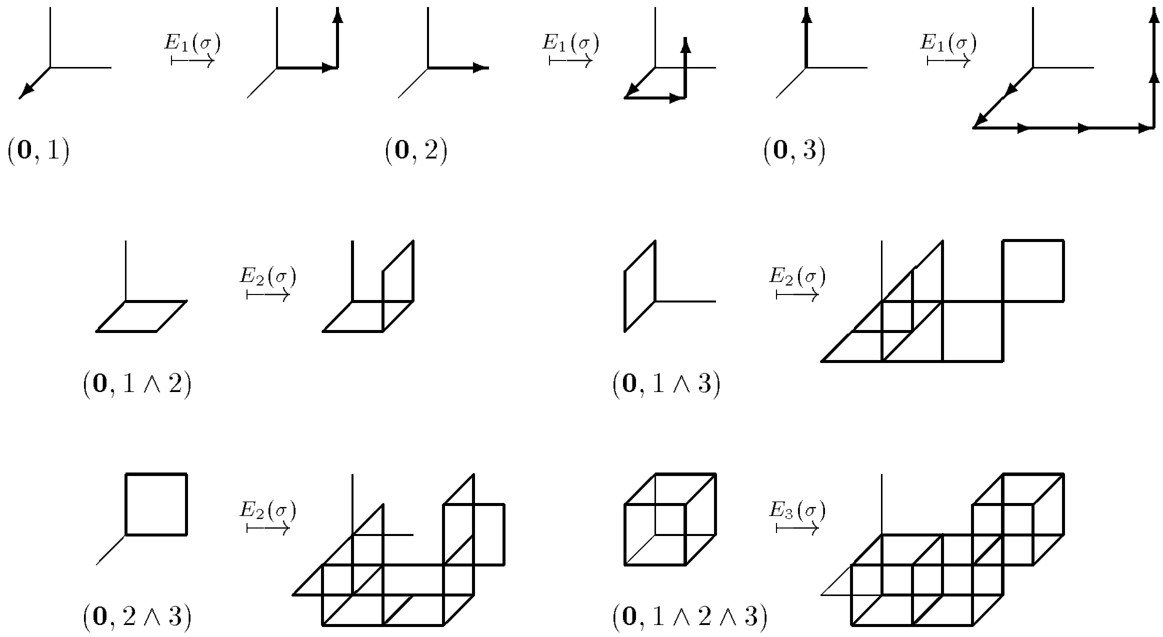

FIGURE 7. The figure of $E_k(\sigma)$, $k = 1, 2, 3$ in Example 3

This substitution is an example which is unimodular but not an invertible endomorphism on the free group of rank 3. The matrix A_σ and the inverse matrix A_σ^{-1} are

$$A_\sigma = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \\ 1 & 1 & 2 \end{bmatrix} \qquad \textit{and} \qquad A_\sigma^{-1} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix}.$$

Then the linear mappings $E_k(\sigma)$, $k = 1, 2, 3$, are given by

$$\begin{aligned}
E_1(\sigma):\quad & (\mathbf{0}, 1) && \mapsto (\mathbf{0}, 2) + (\mathbf{e}_2, 3), \\
& (\mathbf{0}, 2) && \mapsto (\mathbf{0}, 1) + (\mathbf{e}_1, 2) + (\mathbf{e}_1 + \mathbf{e}_2, 3), \\
& (\mathbf{0}, 3) && \mapsto (\mathbf{0}, 1) + (\mathbf{e}_1, 1) + (2\mathbf{e}_1, 2) + (2\mathbf{e}_1 + \mathbf{e}_2, 2) \\
& && \qquad + (2\mathbf{e}_1 + 2\mathbf{e}_2, 2) + (2\mathbf{e}_1 + 3\mathbf{e}_2, 3) + (2\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, 3), \\
E_2(\sigma):\quad & (\mathbf{0}, 1 \wedge 2) && \mapsto -(\mathbf{0}, 1 \wedge 2) - (\mathbf{e}_2, 1 \wedge 3), \\
& (\mathbf{0}, 1 \wedge 3) && \mapsto -(\mathbf{0}, 1 \wedge 2) - (\mathbf{e}_2, 1 \wedge 3) - (\mathbf{e}_1, 1 \wedge 2) - (\mathbf{e}_1 + \mathbf{e}_2, 1 \wedge 3) \\
& && \quad -(2\mathbf{e}_1 + \mathbf{e}_2, 2 \wedge 3) - (2\mathbf{e}_1 + 2\mathbf{e}_2, 2 \wedge 3) + (2\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, 2 \wedge 3), \\
& (\mathbf{0}, 2 \wedge 3) && \mapsto -(\mathbf{e}_1, 1 \wedge 2) - (\mathbf{e}_1 + \mathbf{e}_2, 1 \wedge 3) + (2\mathbf{e}_1 + \mathbf{e}_2, 1 \wedge 2) - (2\mathbf{e}_1 + \mathbf{e}_2, 1 \wedge 3) \\
& && \quad +(2\mathbf{e}_1 + 2\mathbf{e}_2, 1 \wedge 2) + (2\mathbf{e}_1 + 3\mathbf{e}_2, 1 \wedge 3) + (2\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, 1 \wedge 3) \\
& && \quad -(3\mathbf{e}_1 + \mathbf{e}_2, 2 \wedge 3) - (3\mathbf{e}_1 + 2\mathbf{e}_2, 2 \wedge 3) + (3\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, 2 \wedge 3), \\
E_3(\sigma):\quad & (\mathbf{0}, 1 \wedge 2 \wedge 3) && \mapsto (2\mathbf{e}_1 + \mathbf{e}_2, 1 \wedge 2 \wedge 3) + (2\mathbf{e}_1 + 2\mathbf{e}_2, 1 \wedge 2 \wedge 3) \\
& && \qquad -(2\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, 1 \wedge 2 \wedge 3).
\end{aligned}$$

These are displayed in figure 6. Observing this figure, where we can recognize that the image of the unit cube at the origin consists in 3 cubes (with different orientations), we can make the conjecture that the necessary condition of Proposition 5.2 is also sufficient.

Example 4. Let σ be the substitution

$$\begin{aligned}\sigma:\ 1 &\to 12\\ 2 &\to 13\\ 3 &\to 14\\ 4 &\to 1.\end{aligned}$$

This substitution can be considered as the natural generalization, on 4 letters, of the Rauzy substitution. One can easily compute the extensions and dual extensions. This is a case where results of the last part of section 4 apply, since it is a unimodular Pisot substitution satisfying the coincidence condition.

The pictures are not easy to draw, since we are in dimension 4; however, it is possible to show the limit set on the stable space, which looks a bit like a potato. The exact figure can be found as the domain of the Potato exchange transformation in [Ito-Miz].

7. Additional remarks

A first remark is that a large part can be immediately generalized from substitutions to endomorphisms of free groups. The main interest of considering substitutions is that, in this case, there is no cancellation; this makes it easy, in particular, to study $E_1(\sigma)$. However, this reason disappears for higher dimensional extension.

To generalize this framework to the free group, one must be able to define the element $(\mathbf{x}, i^{-1})$. Since we want the path associated to the word $w = ii^{-1}$ to be empty, the natural solution is to define $(\mathbf{x}, i^{-1})$ as $-(\mathbf{x}-\mathbf{e}_i, i)$. The rest follows easily from this definition.

A second remark is that there is certainly an underlying homological theory, which would make all these contructions natural; the maps φ_k defined in section 1 seem to be a kind of Poincaré duality.

It is also possible to make the union of the $\mathcal{G}_k$ into a graded module, by defining an exterior product. It is given on $\mathcal{G}_1$ by $(\mathbf{x}, i) \wedge (\mathbf{y}, j) = (\mathbf{x}+\mathbf{y}, i \wedge j)$, and the generalization to the other sets is straightforward. It is interesting to note that this leads to a simple definition of $E_2(\sigma)$, as $E_1(\sigma) \wedge E_1(\sigma)$.

A last remark is that, instead of considering one substitution and its powers, we can consider the products of a sequence of substitutions; this is a non-commutative generalization of the product of a sequence of matrices. In this setting, we can build the extensions of these products; this can allow, for example, to build explicitly discrete approximations of a plane, using generalized continued fractions, as is done in the paper [Arn-Ber-Ito].

References

[Arn-Ber-Ito] P.Arnoux, V. Berthé, Sh.Ito, *Discrete planes, $\mathbb{Z}^2$-actions, Jacobi-Perron algorithm and substitutions*, preprint.

[Arn-Ito] P.Arnoux, Sh.Ito, *Pisot Substitutions and Rauzy Fractals*, Prétirage IML **98-18**, preprint submitted.

[Dek1] F.M.Dekking, *recurrent sets*, Advances in Math. **44** (1982), 78-104.

[Dek2] F.M.Dekking, *replicating superfigures and endomorphisms of free groups*, J. Combin. Theory Ser. A **32** (1982), 315–320.

[Ei-Ito] H.Ei, Sh.Ito, *Decomposition Theorem on Invertible Substitutions*, Osaka J.Math.**35** (1998),821-834.

[Ito-Ohtsuki] Sh.Ito, M.Ohtsuki, *Modified Jacobi-Perron Algorithm and Generating Markov Partitions for Special Hyperbolic Toral Automorphisms*, Tokyo J. Math. **16** (1993), 441-472.

[Ito-Kimura] Sh.Ito, M.Kimura, *On Rauzy fractal*, Japan J. Indust. Appl. Math. **8** (1991), 461-486.

[Ito-Miz] Sh.Ito, M.Mizutani, *Potato Exchange Transformations with Fractal Domains*, preprint.

[MaKaSo] W.Magnus, A.Karrass, D.Solitar, *Combinatorial group theory*, Wiley Interscience, New York 1966.

[Messaoudi] A.Messaoudi, *Autour du Fractal de Rauzy*, Thèse, Université d'Aix-Marseille **2** (1996).

[Rauzy] G.Rauzy, *Nombres algébriques et substitutions*, Bull. Soc. Math. France **110** (1982), 147-178.

[Wen-Wen] Z.-X. Wen, Z.-Y. Wen, *Local Isomorphisms of Invertible Substitutions*, C. R. Acad. Sci. Paris, **318** Série I (1994), 299-304.

(Yuki Sano) Department of Mathematics, Tsuda College, Tsuda-Machi, Kodaira, Tokyo 187, Japan

E-mail address: sano@tsuda.ac.jp

(Pierre Arnoux) Institut de Mathématiques de Luminy (UPR 9016), 163 Avenue de Luminy, Case 930, 13288 Marseille Cedex 9, France

E-mail address: arnoux@iml.univ-mrs.fr

(Shunji Ito) Department of Mathematics, Tsuda College, Tsuda-Machi, Kodaira, Tokyo 187, Japan

E-mail address: Ito@tsuda.ac.jp