Parallelogram Morphisms and Circular Codes

Alexandre Blondin Massé[1], Mélodie Lapointe[1], and Hugo Tremblay*,[1]

Laboratoire de combinatoire et d'informatique mathématique
Université du Québec à Montréal
CP 8888, Succ. Centre-ville, Montréal (Québec) H3C 3P8
blondin_masse.alexandre@uqam.ca, lapointe.melodie@courrier.uqam.ca
hugo.tremblay@lacim.ca

Abstract. In 2014, it was conjectured that any polyomino can be factorized uniquely as a product of prime polyominoes [8]. In this paper, we present simple tools from words combinatorics and graph topology that seem very useful in solving the conjecture. The main one is called parallelogram network, which is a particular subgraph of $G(\mathbb{Z}^2)$ induced by a parallelogram morphism, i.e. a morphism describing the contour of a polyomino tiling the plane as a parallelogram would. In particular, we show that parallelogram networks are homeomorphic to $G(\mathbb{Z}^2)$. This leads us to show that the image of the letters of parallelogram morphisms is a circular code provided each element is primitive, therefore solving positively a 2013 conjecture [7].

Keywords: Codes, Combinatorics on words, Graphs, Digital geometry, Topological graph theory, Morphisms

1 Introduction

The interaction between combinatorics on words and digital geometry has been extensively studied in the last decades [1,6,9,10]. The most famous example is without doubt the family of Sturmian words, which can be seen as the discrete counterpart of lines having irrational slope [15]. Another remarkable example is about digital convexity: It was recently established that it can be decided very efficiently if some discrete figure is convex by factorizing its boundary in Lyndon and Christoffel words [10]. In the same spirit, one can decide in linear time and space whether some discrete path is self-intersecting, by using combinatorial arguments together with an enriched radix quadtree [9]. Finally, generalizations of discrete lines in 3D have also been proposed, such as in [6].

In parallel, the theory of codes has been developed for more than 50 years. Here, we focus on circular codes, i.e. sets of words that allow unique encoding of words written on a circle. Circular codes were first introduced and studied by Golomb and Gordon [13] and have received a lot of attention from researchers

* Corresponding author. This research is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

since then. From an algebraic perspective, Schützenberger has contributed significantly to a better understanding of their structure [17]. His results have been generalized by Bassino who described the generating functions of weighted circular codes [3]. Circular codes have also been extensively studied in bioinformatics. For instance, a remarkable circular code for the protein coding genes of mitochondria has been brought to light by Arques and Michel [2].

More recently, researchers (including both authors) have been interested in the shape of parallelogram tiles (also called *square tiles* in [16]) using words combinatorics formalism [7,8,11,16]. In particular, in 2008, Provençal defined the product (or composition) of a polyomino and a parallelogram polyomino, which consists in substituting each unit square of the first polyomino with a copy of the parallelogram polyomino (see Figure 1). This leads to the natural definition of prime and composed polyominoes: A polyomino is called *prime* if it cannot be obtained by the composition of two smaller nontrivial polyominoes [16]. Provençal's definition was further studied in [8], where it was proved that every polyomino can be factorized as a product of prime polyominoes, a result in the same spirit than the Fundamental Theorem of Arithmetic. However, the authors were not able to prove that such a factorization is unique and left it as a conjecture:

Conjecture 1. Let U be the unit square polyomino and $P \neq U$ be a polyomino. Then P can be factorized uniquely as a product of a prime polyomino Q and primes parallelogram polyominoes P_1, P_2, ..., P_n, i.e. $P = Q \circ P_1 \circ P_2 \circ \cdots \circ P_n$.

In this paper, we neither prove nor disprove Conjecture 1, but we provide tools that we believe are essential in showing the unicity of the prime factorization. It relies on basic words combinatorics as well as graph topology. In particular, it introduces *parallelogram networks*, i.e. undirected subgraphs of the grid graph $\mathbb{Z}^2$ induced by special morphisms called parallelogram [8]. They turn out to be expressive and easy to manipulate: As a byproduct, we obtain a simple proof that the image of parallelogram morphisms is a circular code under very mild conditions (Theorem 13), thus solving another conjecture stated in [7].

The content is divided as follows. In Section 2, we introduce the basic definitions about words and codes. In Section 3, we recall basic definitions about graphs and their interaction with words. Section 4 is devoted to the study of the

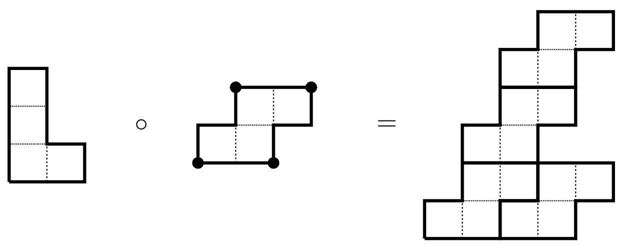

Fig. 1: The composition of a polyomino (left) with a parallelogram polyomino (middle) is a composed polyomino (right).

properties of parallelogram networks, culminating with Theorem 13 in Section 5. We briefly conclude with an open problem.

2 Words and Codes

We recall the basic definitions and notation for words and codes (see [15] for more details). An *alphabet* is a finite set Σ whose elements are called *letters*. A *word* on Σ is a finite sequence $w = w_1 w_2 \cdots w_n$ of letters of Σ. The i-th letter of w is denoted by w_i. The *length* of w, written $|w|$ is the number of elements in the sequence w. The unique word of length 0 is called the *empty word* and is written ε. Whenever $|w| > 0$, we write $\text{FST}(w)$ and $\text{LST}(w)$ for the first and last letter of w. Moreover, for any letter $a \in \Sigma$, $|w|_a$ is the number of occurrences of the letter a in w.

Given two words $u = u_1 u_2 \cdots u_m$ and $v = v_1 v_2 \cdots v_n$, the *concatenation* of u and v, denoted by uv or $u \cdot v$, is the word $u_1 u_2 \cdots u_m v_1 v_2 \cdots v_n$. If u is a word and n is an integer, then $u^n = u \cdot u \cdots u$ (n times). A word w is called *primitive* if there does not exist any word u and integer $n \geq 2$ such that $w = u^n$. A well-known fact is the following:

Proposition 2 ([15]). *Let w be a word such that there exist words u and v with $w = uv = vu$. Then w is not primitive.*

The set of all words on Σ having length n is denoted by Σ^n. The *free monoid* is defined by $\Sigma^* = \bigcup_{n \geq 0} \Sigma_n$. Its name comes from the fact that it has a monoid structure when combined with the concatenation operation, and with neutral element ε. A *submonoid* of Σ^* is a subset $M \subseteq \Sigma^*$ which is stable under the concatenation and which includes ε. The submonoid M is *pure* if for all $x \in \Sigma^*$ and $n \geq 1$, $x^n \in M$ implies $x \in M$. Moreover, we say that M is *very pure* if for all $u, v \in \Sigma^*$, the relations $uv \in M$ and $vu \in M$ imply $u, v \in M$. It is straightforward to show that any very pure submonoid is also pure. However, the converse is false: The submonoid of $\{a, b\}^*$ generated by $\{ab, ba\}$ is pure but not very pure.

Let w be some word. Then we say that u is a *factor* of w if there exist words x and y such that $w = xuy$. Moreover, if $x = \varepsilon$ (resp. $y = \varepsilon$), u is called *prefix* (resp. *suffix*) of w. The set of prefixes (resp. suffixes) of a word w is denoted by $\text{Pref}(w)$ (resp. $\text{Suff}(w)$). Also, the unique prefix (resp. suffix) of length ℓ of w is denoted by $\text{Pref}_\ell(w)$ (resp. $\text{Suff}_\ell(w)$), where $0 \leq \ell \leq |w|$.

Given two alphabets A and B, an application $\varphi : A^* \to B^*$ is called *morphism* (resp. *antimorphism*) if $\varphi(uv) = \varphi(u)\varphi(v)$ (resp. $\varphi(uv) = \varphi(v)\varphi(u)$) for all $u, v \in A^*$. Given $w = w_1 w_2 \cdots w_n$, the *reversal of* w, denoted by $\widetilde{w}$, is defined by $\widetilde{w} = w_n w_{n-1} \cdots w_2 w_1$. The operator $\widetilde{\cdot}$ is an antimorphism. It is easy to see that morphisms and antimorphisms are completely defined by their action on single letters.

Let Σ be an alphabet and $X \subseteq \Sigma^*$. Then X is a *code over* Σ if for all $m, n \geq 1$ and $x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n \in X$, the condition $x_1 x_2 \cdots x_m = y_1 y_2 \cdots y_n$ implies $m = n$ and $x_i = y_i$ for $i = 1, 2, \ldots, n$. Roughly speaking, X is a code if any word in X^* can be written uniquely as a product of words in X. Similarly, we say that X is a *circular code* if for all $m, n \geq 1$ and $x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n \in X$, $p \in \Sigma^*$ and $s \in \Sigma^+$, the relations $sx_2x_3 \cdots x_m p = y_1 y_2 \cdots y_n$ and $x_1 = ps$ imply $m = n$, $p = \varepsilon$ and $x_i = y_i$ for $i = 1, 2, \ldots, n$. In other words, X is a circular code if any circular permutation of a word in X^* can be written uniquely as a product of words in X. It is not hard to prove that any circular code is a code. The reader is referred to [5] for more details about code theory, but one important result for our purpose is the following characterization of circular codes:

Theorem 3 (Proposition 1.1 of [5]). *A submonoid M of A^* is very pure if and only if its minimal set of generators is a circular code.*

3 Discrete Paths and Graphs

An alphabet of particular interest for our purposes is the *Freeman chain code* $\mathcal{F} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}\}$, which encodes the four elementary steps on the square grid $\mathbb{Z}^2$ with respect to the bijection

$$\mathbf{0} \mapsto \rightarrow, \quad \mathbf{1} \mapsto \uparrow, \quad \mathbf{2} \mapsto \leftarrow, \quad \mathbf{3} \mapsto \downarrow .$$

Two basic operations on Freeman words have useful geometrical interpretations. The application $\overline{\cdot}$ is the morphism defined by

$$\overline{\mathbf{0}} = \mathbf{2}, \quad \overline{\mathbf{1}} = \mathbf{3}, \quad \overline{\mathbf{2}} = \mathbf{0}, \quad \overline{\mathbf{3}} = \mathbf{1},$$

which corresponds geometrically to the application of a rotation of angle π. Also, the antimorphism $\widehat{\cdot} = \overline{\cdot} \circ \widetilde{\cdot}$ corresponds to traveling the sequence of elementary steps in the opposite order.

Given $w \in \mathcal{F}^*$, we write $\vec{w} = (|w|_\mathbf{0} - |w|_\mathbf{2}, |w|_\mathbf{1} - |w|_\mathbf{3})$. Any word $w \in \mathcal{F}^*$ is called *closed* if $\vec{w}$ is the null vector. Moreover, w is called *simple* if none of its proper factor is closed, and is a *contour word* if it is nonempty, closed and simple.

A *discrete path* is a sequence of connected unit segments whose endpoints are on $\mathbb{Z}^2$. Discrete paths can naturally be represented by an ordered pair $\gamma = (p, w)$, where $p \in \mathbb{Z}^2$ and $w \in \mathcal{F}^*$. Thus, the set of points of $\mathbb{Z}^2$ visited by γ is $\text{Points}(\gamma) = \{p + \vec{u} \mid u \in \text{Pref}(w)\}$. A discrete path is called *closed* (resp. *simple*) if w is closed (resp. *simple*). Given a closed discrete path γ, the *region of* γ, denoted by $R(\gamma)$, is defined as the closed subset of $\mathbb{R}^2$ whose boundary is exactly described by γ.

Every discrete path yields a unique undirected graph $G(\gamma) = (V, E)$, where $V = \text{Points}(\gamma)$ and $(q, q') \in E$ if and only if there exist two consecutive prefixes u, u' of w such that $q = p + \vec{u}$ and $q' = p + \vec{u}'$. Also, the *(graph) distance* between two vertices p and p' is the length of a shortest discrete path γ between p and p'.

The *grid graph* $G(\mathbb{Z}^2)$ is the infinite graph whose set of vertices is $\mathbb{Z}^2$ and whose set of edges E is defined as follows: $\{p, p'\} \in E$ if and only if $\text{dist}(p, p') = 1$, where dist is the usual Euclidean distance. The set of all discrete paths of $G(\mathbb{Z}^2)$ is denoted by $\Gamma(\mathbb{Z}^2)$. Clearly, for any $\gamma \in \Gamma(\mathbb{Z}^2)$, the undirected version of the graph $G(\gamma)$ is a subgraph of $G(\mathbb{Z}^2)$.

We now recall topological graph theoretic definitions. We use the same terminology as in [12]. Let $G = (V, E)$ be a undirected graph. A *subdivision of* G is any graph obtained from G by replacing some edges in E with new paths between their ends such that those paths have no inner vertex in V or in another path. The original vertices of G are then called *branch vertices* and the new vertices are called *inner vertices*. It is clear that inner vertices have degree 2 while branch vertices retain their respective degree from G.

Given two graphs $G = (V, E)$ and $G' = (V', E')$, G and G' are called *isomorphic*, and we write $G \simeq G'$, if there exists a bijection $f : V \to V'$ such that $\{u, v\} \in E$ if and only if $\{f(u), f(v)\} \in E'$. From this, one defines the notion of *graph homeomorphism*: Two graphs G and G' are *homeomorphic* (i.e. topologically isomorphic) if there exist two isomorphic subdivisions T and T' of G and G' respectively. It is easy to show that G and any of its subdivision T are homeomorphic. Also, the notions of graph homeomorphism and standard topological homeomorphism are equivalent when considering the topological representations of graphs (i.e. the topological space obtained by representing vertices as distinct points and edges by homeomorphic images of the closed unit interval $[0, 1]$) [14].

4 Parallelogram Networks

Some morphisms are of particular interest from a geometrical perspective. We recall some definitions from [8].

Definition 4 ([8]). *Let $\varphi : \mathcal{F}^* \to \mathcal{F}^*$ be a morphism. Then φ is called*

(i) homologous *if $\varphi(a) = \widehat{\varphi(\overline{a})}$;*
(ii) parallelogram *if it is homologous, $\varphi(\mathbf{0123})$ is a contour word and* $\textsc{Fst}(\varphi(a)) = a$ *for all $a \in \mathcal{F}$.*

Let $\varphi : \mathcal{F}^* \to \mathcal{F}^*$ be a parallelogram morphism. For simplicity of writing, we extend the application φ as follows. For any $p = (x, y) \in \mathbb{Z}^2$, let $\varphi(p) = \varphi(x, y) = (0, 0) + x\overrightarrow{\varphi(0)} + y\overrightarrow{\varphi(1)} \in \mathbb{Z}^2$. Moreover, if $\gamma = (p, w)$ is a discrete path, then $\varphi(\gamma)$ is the discrete path $\varphi(\gamma) = (\varphi(p), \varphi(w))$.

The *graph of* φ is defined by

$$G(\varphi) = \bigcup_{\gamma \in \Gamma(\mathbb{Z}^2)} G(\varphi(\gamma)) = \bigcup_{p \in \mathbb{Z}^2} G(\varphi(p, \mathbf{0123})). \tag{1}$$

Any such graph is called *parallelogram network*. The second equality of Equation (1) is easy to check: The inclusion $\supseteq$ follows directly from the fact that $(p, \mathbf{0123})$ is a path in $G(\mathbb{Z}^2)$ while the inclusion $\subseteq$ follows from the fact that any path γ in $G(\mathbb{Z}^2)$ can be divided into discrete paths of length 1, each belonging to at least one discrete path of the form $(p, \mathbf{0123})$, for some $p \in \mathbb{Z}^2$.

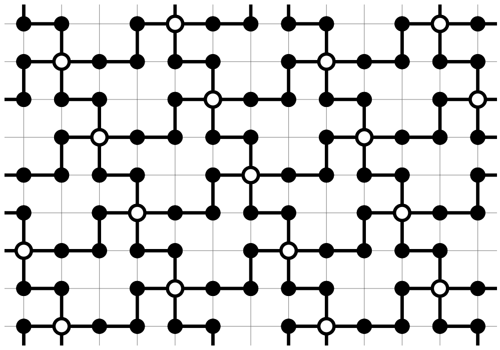

Fig. 2: The parallelogram network $G(\varphi)$ induced by the parallelogram morphism such that $\varphi(\mathbf{0}) = \mathbf{0010}$ and $\varphi(\mathbf{1}) = \mathbf{121}$. The white dots correspond to branch vertices.

Example 5. The graph $G(\varphi)$ is represented in Figure 2, where φ is the parallelogram morphism such that $\varphi(\mathbf{0}) = \mathbf{0010}$ and $\varphi(\mathbf{1}) = \mathbf{121}$.

Clearly, if φ is a parallelogram morphism, then the morphism φ_i defined by $\varphi_i(a) = \varphi(a + i)$ is also a parallelogram morphism for $i = \mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}$ and $a + i$ is the addition modulo 4. Therefore, for unicity purposes, we assume that $\textsc{Fst}(\varphi(a)) = a$ for all $a \in \mathcal{F}$, and that any discrete path whose associated word is $\varphi(\mathbf{0123})$ is traveled counterclockwise.

The following basic properties of homologous morphisms are useful.

Proposition 6. *Let φ be an homologous morphism and $w \in \mathcal{F}^*$.*

(i) *For any $a \in \mathcal{F}$, $\overrightarrow{\varphi(a)} + \overrightarrow{\varphi(\overline{a})} = \vec{0}$.*
(ii) *If $\vec{w} = (x, y)$, then $\overrightarrow{\varphi(w)} = x\overrightarrow{\varphi(\mathbf{0})} + y\overrightarrow{\varphi(\mathbf{1})}$.*

Proof. (i) Since φ is homologous, for any $a \in \mathcal{F}$, we have $\overrightarrow{\varphi(a)} = \overrightarrow{\widehat{\varphi(\overline{a})}} = -\overrightarrow{\varphi(\overline{a})}$.

(ii) Write $w = w_1 w_2 \cdots w_n$. Then

$$\begin{aligned}
\overrightarrow{\varphi(w)} &= \sum_{i=1}^{n} \overrightarrow{\varphi(w_i)} \\
&= \sum_{a \in \mathcal{F}} |w|_a \overrightarrow{\varphi(a)} \\
&= \sum_{a \in \{\mathbf{0},\mathbf{1}\}} (|w|_a \overrightarrow{\varphi(a)} + |w|_{\overline{a}} \overrightarrow{\varphi(\overline{a})}) \\
&= \sum_{a \in \{\mathbf{0},\mathbf{1}\}} (|w|_a - |w|_{\overline{a}}) \overrightarrow{\varphi(a)} \\
&= x\overrightarrow{\varphi(\mathbf{0})} + y\overrightarrow{\varphi(\mathbf{1})},
\end{aligned}$$

as claimed. □

It is worth noticing that for any parallelogram morphism φ, the graph $G(\varphi)$ is a regular parallelogram tiling of the plane $\mathbb{R}^2$. In other words, it is possible to completely cover the plane by non-overlapping translated copies of $\varphi(\mathbf{0123})$ along the direction of the two vectors $\overrightarrow{\varphi(\mathbf{0})}$ and $\overrightarrow{\varphi(\mathbf{1})}$, i.e.

$$\mathbb{R}^2 = \bigcup_{(a,b)\in\mathbb{Z}^2} \left\{ R((0,0);\varphi(\mathbf{0123})) + a\overrightarrow{\varphi(\mathbf{0})} + b\overrightarrow{\varphi(\mathbf{1})} \right\},$$

Indeed, as shown in [4], a tile admitting a contour word $w \in \mathcal{F}^*$ tiles the plane by translation along the direction of exactly two vectors if and only if w can be factorized as $w = XY\hat{X}\hat{Y}$, where $X, Y \in \mathcal{F}$. Moreover, the authors characterize such regular tiling by describing the surrounding of parallelogram tiles (see Figure 3). From this, Proposition 7 follows.

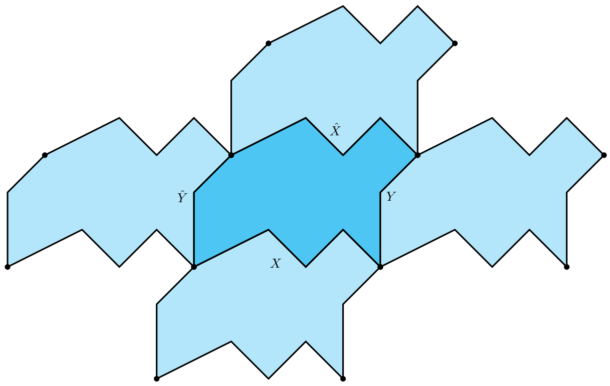

Fig. 3: The surrounding of a tile t coded by $w = XY\hat{X}\hat{Y}$ obtained by taking the four translated copies $t \pm \overrightarrow{\varphi(\mathbf{0})}$ or $\overrightarrow{\varphi(\mathbf{1})}$ and matching the corresponding homologous factors. It induces a regular parallelogram tiling of the plane $\mathbb{R}^2$.

Proposition 7. *Let φ be a parallelogram morphism. Then $\left\{\overrightarrow{\varphi(\mathbf{0})}, \overrightarrow{\varphi(\mathbf{1})}\right\}$ is a basis of the vector space $\mathbb{R}^2$.*

Proof. Let $\vec{u} = \overrightarrow{\varphi(\mathbf{0})}$ and $\vec{v} = \overrightarrow{\varphi(\mathbf{1})}$. It suffices to prove that $\vec{u}$ and $\vec{v}$ are linearly independent since $\mathbb{R}^2$ is a vector space of dimension 2. Arguing by contradiction, assume that this is not the case and let

$$T = \bigcup_{(a,b)\in\mathbb{Z}^2} \left\{ R((0,0),\varphi(\mathbf{0123})) + a\overrightarrow{\varphi(\mathbf{0})} + b\overrightarrow{\varphi(\mathbf{1})} \right\}.$$

Now, since the region $R((0,0),\varphi(\mathbf{0123}))$ is bounded, there exist points $p_1, p_2 \in \mathbb{R}^2$ such that $R((0,0),\varphi(\mathbf{0123}))$ lies completely in the region B between the two parallel lines $l_1 = p_1 + t_1\vec{u}$ and $l_2 = p_2 + t_2\vec{u}$, where $t_1, t_2 \in \mathbb{R}$. Further, the linear dependance of $\vec{u}$ and $\vec{v}$ implies that any point of T lies entirely in B, so that T is a subset of B. But then T is a proper subset of $\mathbb{R}^2$, contradicting $T = \mathbb{R}^2$.

A remarkable property of parallelogram morphisms is that they preserve closed and simple paths. The former is an immediate consequence of Proposition 7 while the latter is more complicated to show and we need additional results. First, we recall a result of [18] about tessellation that translates directly to our context:

Theorem 8. *Let φ be a parallelogram morphism, $\vec{a} = \overrightarrow{\varphi(\mathbf{0})}$, $\vec{b} = \overrightarrow{\varphi(\mathbf{1})}$, $p, q \in \mathbb{Z}^2$ and P, Q be the regions enclosed inside the discrete paths $(p, \varphi(\mathbf{0123}))$ and $(q, \varphi(\mathbf{0123}))$ respectively. Then exactly one of the following conditions holds:*

(i) *$P = Q$ and then $p = q$;*
(ii) *P and Q share a single point and then $\overrightarrow{q - p} = \pm\vec{a} \pm \vec{b}$;*
(iii) *P and Q share a chain in $\varphi(\mathcal{F})$ and then $\overrightarrow{q - p} \in \{\pm\vec{a}, \pm\vec{b}\}$;*
(iv) *P and Q are disjoint.*

Proof. By definition, the regions enclosed inside the discrete path $\varphi(\mathbf{0123})$ is a polyomino tiling the plane by translation in a parallelogram manner. It follows from Theorem 4.13 of [18] that P and Q verify one and only one of Conditions (i)–(iv). □

We observe from Figure 2 that each vertex $x\overrightarrow{\varphi(\mathbf{0})} + y\overrightarrow{\varphi(\mathbf{1})}$, where $x, y \in \mathbb{Z}$ of $G(\varphi)$ has degree 4. We call such vertices *branch vertices*. A non branch vertex p is called *inner vertex of type a* if there exists some discrete path $(p', \varphi(a))$ visiting p. Note that if p is an inner vertex of type a, then it is also an inner vertex of type $\overline{a}$. An immediate consequence of Theorem 8 is a simple description of parallelogram networks.

Corollary 9. *Let φ be some parallelogram morphism and $p \in \mathbb{Z}^2$. Then*

$$\deg(p) = \begin{cases} 4, & \textit{if } p \textit{ is a branch vertex;} \\ 2, & \textit{otherwise.} \end{cases}$$

The remainder of this section is devoted to proving that $G(\mathbb{Z}^2)$ and $G(\varphi)$ are homeomorphic. First, observe that any parallelogram morphism φ induces a subdivision T_φ of $\mathbb{Z}^2$: Subdivide horizontal edges $\{u, v\}$ of $G(\mathbb{Z}^2)$ by adding $|\varphi(\mathbf{0})| - 1$ inner vertices between the two branch vertices u and v. Similarly, vertical edges are subdivided using $|\varphi(\mathbf{1})| - 1$ new inner vertices. Therefore, the new horizontal (*resp.* vertical) chains obtained between two branch vertices adjacent in the original graph have length $|\varphi(\mathbf{0})|$ (*resp.* $|\varphi(\mathbf{1})|$), since $|\varphi(\mathbf{0})| = |\varphi(\mathbf{2})|$ and $|\varphi(\mathbf{1})| = |\varphi(\mathbf{3})|$.

Our first main result follows:

Theorem 10. *Let φ be a parallelogram morphism. Then, $T_\varphi \simeq G(\varphi)$.*

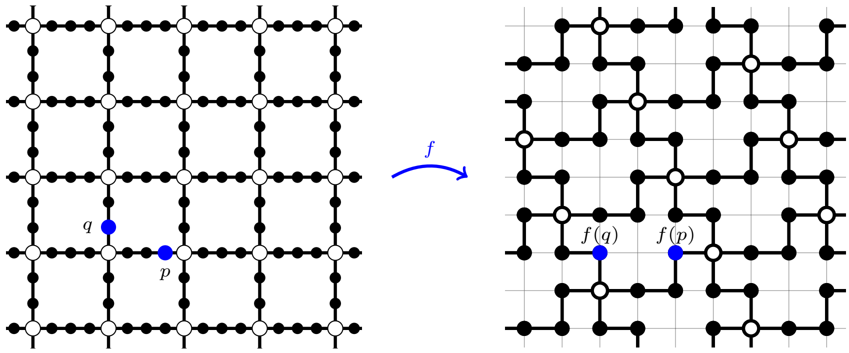

Fig. 4: The effect of f on two vertices of T_φ

Proof. Let $V(T_\varphi)$ and $V(G(\varphi))$ be the set of vertices of T_φ and $G(\varphi)$ respectively. Also, let (x, y) be a vertex of T_φ. By construction, we have

$$(x,y) \in \left\{ (\lfloor x \rfloor, \lfloor y \rfloor) + \left(\frac{k_1}{|\varphi(\mathbf{0})|}, \frac{k_2}{|\varphi(\mathbf{1})|} \right) \right\}$$

with $0 \leq k_1 < |\varphi(\mathbf{0})|$, $0 \leq k_2 < |\varphi(\mathbf{1})|$ and $k_1 k_2 = 0$. Now consider the function $f : V(T_\varphi) \to V(G(\varphi))$ defined by

$$f(x,y) = \varphi(\lfloor x \rfloor, \lfloor y \rfloor) + \begin{cases} \overrightarrow{\mathrm{Pref}_{k_1}(\varphi(\mathbf{0}))}, & \text{if } k_2 = 0; \\ \overrightarrow{\mathrm{Pref}_{k_2}(\varphi(\mathbf{1}))}, & \text{if } k_1 = 0. \end{cases}$$

Intuitively, the transformation f finds the closest bottom or left branch vertex $(\lfloor x \rfloor, \lfloor y \rfloor)$ of any vertex (x, y), and then consider the k-th vertex in the path $\varphi((x, y), a)$ in $G(\varphi)$, where $k \in \{k_1, k_2\}$ and $a \in \{\mathbf{0}, \mathbf{1}\}$ (see Figure 4). It is straightforward to check that f is a bijection. It remains to show that $p, q \in \mathbb{Z}^2$ are adjacent in T_φ if and only if $f(p)$ and $f(q)$ are adjacent in $G(\varphi)$.

First, for any $p \in \mathbb{Z}^2$ and $a \in \mathcal{F}$, let $C(p, a)$ be the sequence whose i-th element is $p + i\vec{a}$, for $i = 0, 1, \ldots, |\varphi(a)|$ and consider the sequence $C'(p, a)$ whose i-th element is $f(p + i\vec{a})$, for $i = 0, 1, \ldots, |\varphi(a)|$. Then

$$f(p + i\vec{a}) = \varphi(p) + \overrightarrow{\mathrm{Pref}_i(\varphi(a))}.$$

Consequently, $C(p, a)$ is a chain of T_φ if and only if $C'(p, a)$ is a chain of $G(\varphi)$, since (p, a) is a discrete path of T_φ if and only if $(p, \varphi(a))$ is a discrete path of $G(\varphi)$.

In other words, paths between vertices having integer coordinates in T_φ are isomorphic to path between branch vertices in $G(\varphi)$. By Corollary 9, the degrees of vertices match, so that we have considered all possible neighbors. □

From Theorem 10, we deduce that $G(\mathbb{Z}^2)$ and $G(\varphi)$ have essentially the same structure.

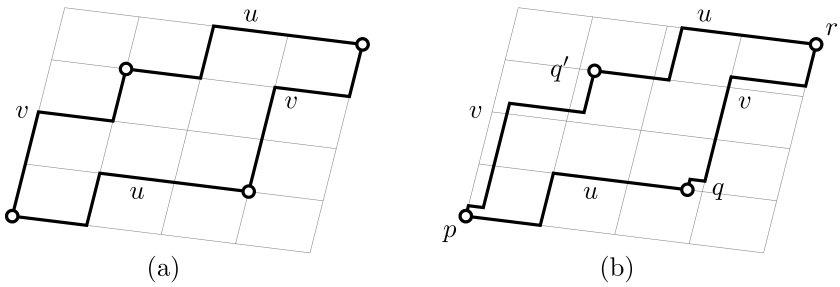

Fig. 5: Geometric representation of the paths u and v. (a) $u, v \in M$. (b) $u, v \notin M$.

Corollary 11. *Let φ be a parallelogram morphism. Then, $G(\mathbb{Z}^2)$ and $G(\varphi)$ are homeomorphic.*

Finally, from Corollary 11, one deduces that φ preserves both closed and simple paths. In other words, $G(\varphi)$ is a deformed image of $G(\mathbb{Z}^2)$.

5 Main Result

Before proving Theorem 13, we describe the graph distance between particular pairs of vertices in parallelogram networks.

Lemma 12. *Let φ be any parallelogram morphism and p be a vertex of $G(\varphi)$. Moreover, let $q = p + k\overrightarrow{\varphi(a)}$ for some $a \in \mathcal{F}$ and some positive integer k.*

(i) *If p and q are branch vertices, then $(p, \varphi(a)^k)$ is the unique shortest path going from p to q and $\mathrm{dist}_{G(\varphi)}(p, q) = k|\varphi(a)|$.*
(ii) *If p and q are inner vertices of type b, where $b \in \mathcal{F}$ and $b \neq a, \overline{a}$, then $\mathrm{dist}_{G(\varphi)}(p, q) > k|\varphi(a)|$.*

Proof. (i) By definition of parallelogram network, there is a path from p to q in $G(\varphi)$ described by $\varphi(a)^k$. This path is also the shortest: Any other path from p to q must be composed of at least k non-overlapping subpaths of the form $(p_i, \varphi(a))$, where $b = (a + \mathbf{1}) \bmod 4$, $p_i = p + i\overrightarrow{\varphi(a)} + j_i\overrightarrow{\varphi(b)}$ and $j_i \in \mathbb{Z}$ for $i = 0, 1, \ldots, k$.

(ii) A shortest path from p to q is obtained by going to the nearest branch vertex, then traveling along $\varphi(b)^k$ and then going to q. Since p and q are inner vertices, the number of edges in this shortest path is more than $k|\varphi(a)|$. □

We are now ready to solve Conjecture 36 of [7].

Theorem 13. *Let φ be any parallelogram morphism. Then $\varphi(\mathcal{F})$ is a circular code if and only if $\varphi(\mathbf{0})$ and $\varphi(\mathbf{1})$ are both primitive words.*

Proof. ($\Rightarrow$) If $\varphi(\mathcal{F})$ is a circular code, then each of its element must be primitive, in particular $\varphi(\mathbf{0})$ and $\varphi(\mathbf{1})$.

($\Leftarrow$) Let $M = \varphi(\mathcal{F})^*$. We show that M is very pure. Arguing by contradiction, assume the contrary, i.e. there exist $u, v \in \mathcal{F}^*$ such that $uv, vu \in M$ but $u, v \notin M$.

Clearly, $\overrightarrow{uv} = \overrightarrow{vu}$, which implies that the discrete paths (p, uv) and (p, vu) of $G(\varphi)$ end at the same point, for any $p \in \mathbb{Z}^2$. Moreover, there exist branch vertices $p, r \in \mathbb{Z}^2$ of $G(\varphi)$ and inner vertices q, q' of type a, a' of $G(\varphi)$ such that the discrete paths (p, u) and $(r, \widehat{v})$ both end at q and the discrete paths (p, v) and $(r, \widehat{u})$ both end at q' (the situation is depicted in Figure 5). There are two cases to consider.

First, suppose that $uv = \varphi(b)^k$ for some $b \in \mathcal{F}$. Since $|uv| = |vu|$ and since $(p, \varphi(b)^k)$ is the unique shortest path from p to r (Lemma 12(i)), we deduce that $uv = \varphi(b)^k = vu$. Write $u = u'u''$ and $v = v = v'v''$, where $u', v'' \in M$ and $u''v' = \varphi(b)$ (such a decomposition exists and is unique since $uv \in M$ but $u, v \notin M$). Then $\varphi(b)^k = uv = vu = v'v''u'u''$, which implies that v' is a prefix of $\varphi(b)$ and u'' is a suffix of $\varphi(b)$. Hence, $u''v' = \varphi(b) = v'u''$, so that, by Proposition 2, $\varphi(b)$ is not primitive, contradicting the theorem assumption.

Otherwise, let u' and v' be the maximal words of $\mathcal{F}^*$ such that $\varphi(u')$ is a prefix of u and $\varphi(v')$ is a suffix of v. Let

$$Q = \{q' + \overrightarrow{\varphi(u'')} \mid u'' \in \text{Pref}(u')\} \cup \{q' + \overrightarrow{\varphi(\widehat{v''})} \mid v'' \in \text{Suff}(v')\}.$$

Since $u, v \notin M$, all elements of Q are inner vertices. Moreover, they all are of type a' (the same type as q'). However, there must exist at least two distinct $s, s' \in Q$ such that $s' = s + \overrightarrow{\varphi(b)}$, where $b \neq a', \overline{a'}$: Otherwise, we would have $uv = \varphi(a') = vu$ which was considered in the previous paragraph. But then Lemma 12 applies to s and s', so that $\text{dist}_{G(\varphi)}(s, s') > |\varphi(b)|$, contradicting the fact that s' can be reached from s through the path $(s, \varphi(b))$. □

6 Concluding Remarks

Theorem 13 might be seen as a first important step in solving Conjecture 1. Indeed, as mentioned in Section 4, parallelogram networks are not uniquely represented by a parallelogram morphism φ, since its circular permutations also yield the same parallelogram network. Moreover, there exist examples of parallelogram morphisms having a circular permutation which induces a distinct parallelogram network. In fact, there are infinitely many of them, and their structure has been described in [7].

For instance, it is easy to verify that for any $p \in \mathbb{Z}^2$, $(p, \varphi(\mathbf{0123}))$ is a discrete path of both $G(\varphi)$ and $G(\varphi')$ defined by

$$\varphi(\mathbf{0}) = \mathbf{01010}, \quad \varphi(\mathbf{1}) = \mathbf{121}, \quad \varphi'(\mathbf{0}) = \mathbf{030}, \quad \varphi'(\mathbf{1}) = \mathbf{10101}.$$

However, it seems that no other closed discrete path can exist in both parallelogram networks.

References

1. Anselmo, M., Giammarresi, D., Madonia, M.: Two dimensional prefix codes of pictures. In: Béal, M.P., Carton, O. (eds.) Developments in Language Theory, Lecture Notes in Computer Science, vol. 7907, pp. 46–57. Springer Berlin Heidelberg (2013)
2. Arquès, D.G., Michel, C.J.: A circular code in the protein coding genes of mitochondria. Journal of Theoretical Biology 189(3), 273 – 290 (1997), http://www.sciencedirect.com/science/article/pii/S0022519397905130
3. Bassino, F.: Generating functions of circular codes. Advances in Applied Mathematics 22(1), 1 – 24 (1999)
4. Beauquier, D., Nivat, M.: On translating one polyomino to tile the plane. Discrete Comput. Geom. 6, 575–592 (1991)
5. Berstel, J., Perrin, D., Reutenauer, C.: Codes and Automata, Encyclopedia of Mathematics and its Applications, vol. 129. Cambridge University Press (2009)
6. Berthé, V., Labbé, S.: An arithmetic and combinatorial approach to three-dimensional discrete lines. In: Debled-Rennesson, I., Domenjoud, E., Kerautret, B., Even, P. (eds.) DGCI. Lecture Notes in Computer Science, vol. 6607, pp. 47–58. Springer (2011)
7. Blondin Massé, A., Garon, A., Labbé, S.: Combinatorial properties of double square tiles. Theor. Comput. Sci. 502, 98–117 (2013)
8. Blondin Massé, A., Tall, A., Tremblay, H.: On the arithmetics of discrete figures. In: Dediu, A.H., Martín-Vide, C., Sierra-Rodríguez, J.L., Truthe, B. (eds.) Language and Automata Theory and Applications, Lecture Notes in Computer Science, vol. 8370, pp. 198–209. Springer International Publishing (2014)
9. Brlek, S., Koskas, M., Provençal, X.: A linear time and space algorithm for detecting path intersection in $\mathbb{Z}^d$. Theor. Comput. Sci. 412(36), 4841–4850 (2011)
10. Brlek, S., Lachaud, J., Provençal, X., Reutenauer, C.: Lyndon + Christoffel = digitally convex. Pattern Recognition 42(10), 2239–2246 (2009)
11. Cousineau, G.: Characterization of some periodic tiles by contour words. Oligomerization of Chemical and Biological Compounds (2014)
12. Diestel, R.: Graph Theory. Springer, fourth edn. (2010)
13. Golomb, S.W., Gordon, B.: Codes with bounded synchronization delay. Information and Control 8(4), 355 – 372 (1965)
14. Gross, J.L., Tucker, T.W.: Topological graph theory. Wiley (1987)
15. Lothaire, M.: Combinatorics on words. Cambridge University Press (1997)
16. Provençal, X.: Combinatoire des mots, géométrie discrète et pavages. Ph.D. thesis, D1715, Université du Québec à Montréal (2008)
17. Schützenberger, M.: On a factorization of free monoids. Proc. Amer. Math. Soc. 16, 21–24 (1965)
18. Wijshoff, H., van Leeuwen, J.: Arbitrary versus periodic storage schemes and tessellations of the plane using one type of polyomino. Information and Control 62(1), 1 – 25 (1984)