The injectivity of the global function of a cellular automaton in the hyperbolic plane is undecidable

Maurice Margenstern *

June 27, 2008

Laboratoire d'Informatique Théorique et Appliquée, EA 3097,
Université Paul Verlaine − Metz,
Département d'Informatique,
Île du Saulcy,
57045 Metz Cedex, France,
e-mail: margens@univ-metz.fr

Abstract

In this paper, we look at the following question. We consider cellular automata in the hyperbolic plane, see [5, 21, 9, 13] and we consider the global function defined on all possible configurations. Is the injectivity of this function undecidable? The problem was answered positively in the case of the Euclidean plane by Jarkko Kari, in 1994, see [3]. In the present paper, we show that the answer is also positive for the hyperbolic plane: the problem is undecidable.

Keywords: cellular automata, global function, hyperbolic plane, tessellations, undecidability

1 Introduction

The global function of a cellular automaton A is defined in the set of all configurations. Note that when we implement an algorithm to solve a given problem, the initial configuration is usually finite. The study of the global function starts from another point of view.

In the case of the Euclidean plane, the definition of the set of configurations is very easy: it is $Q^{\mathbb{Z}^2}$, where Q is the set of states of the automaton.

In the hyperbolic plane, see [13, 11], we have the following situation: we consider that the grid is the pentagrid or the ternary heptagrid, see [13]. We fix a tile, which will be called the **central cell** and, around it, we dispatch α sectors, $\alpha \in \{5, 7\}$: $\alpha = 5$ in the case of the pentagrid, $\alpha = 7$ in the case of the ternary heptagrid. We assume that the sectors and the central cell cover the plane and the sectors do not overlap, neither the central cell, nor other sectors: call them the **basic sectors**. Denote by $\mathcal{F}_\alpha$ the set constituted by the central cell and α Fibonacci trees, see [5, 13, 21] , each one spanning a basic sector. Then, a configuration of a cellular automaton A in the hyperbolic plane can be represented as an element of $Q^{\mathcal{F}_\alpha}$, where Q is the set of states of A. If f_A denotes the **local** transition function of A, its **global** transition function G_A is defined by: $G_A(c)(x) = f(c(x))$, where c runs over $Q^{\mathcal{F}_\alpha}$ and $x \in \mathcal{F}_\alpha$.

The injectivity problem for a cellular automaton consists in asking whether there is an algorithm which, applied to a description of f_A would indicate whether G_A is injective or not.

In the present paper, we prove that there is no such algorithm and so, the corresponding problem is undecidable. The present paper relies on a previous work by the author, see [18]. In this paper, we give a construction which is described in [15, 12], which yields a plane-filling path, each time we can construct a valid tiling with an exception. However, in this exceptional case, a more careful analysis of the structure of the path shows that, changing a bit the way in which basic regions are traversed by the path, it is also possible to carry out the argument which is needed to prove the undecidability of the injectivity.

We shall not repeat the construction of the interwoven triangles on which the construction of the mauve triangles relies. In section 2, we remind the basic properties of the mauve triangles and we introduce new ones. In section 3, we more carefully describe the construction of the path based on the mauve triangles. In section 4, we show how to derive the proof of the main theorem:

Theorem 1 *There is no algorithm to decide whether the global transition function of a cellular automaton on the ternary heptagrid is injective or not.*

Note that it is enough to find a particular tiling whose cellular automata have the property that the injectivity of their global function is undecidable to prove that the same property for cellular automata in the hyperbolic plane in general is also undecidable. However, it seems impossible to transfer the construction of the path which we consider in this paper to the pentagrid. However, the construction of this paper can be generalized to any grid $\{p, 3\}$ of the hyperbolic plane with $p \geq 7$.

2 The mauve triangles

The mauve triangles are first constructed upon the interwoven triangles. The latter triangles are obtained by the construction which is illustrated by figure 1. We refer the reader to [14, 17, 15, 10] for a detailed account on the construction of the interwoven triangles and for their properties. We also refer him/her to the same papers for an account on the implementation of these triangles in the ternary heptagrid of the hyperbolic plane.

In [14, 17, 15, 10], we implement the interwoven triangles in the ternary heptagrid, using another tiling as a background. This tiling, called the **mantilla**, is a refinement of the ternary heptagrid by grouping its tile in a particular way. Now, it is possible to implement the interwoven triangles in a simpler context of the ternary heptagrid. However, the spacing imposed by the mantilla is a good point which allows to more easily solve a few details of the implementation of the path. This is why, in this paper, we assume the construction to be performed on the mantilla.

The construction of the interwoven triangles needs a lot of signals, which entails a huge number of tiles, around 18,000 of them, not taking into account the specific tiles devoted to the simulation of a Turing machine. The construction of this paper requires much more tiles, but we shall not try to count them.

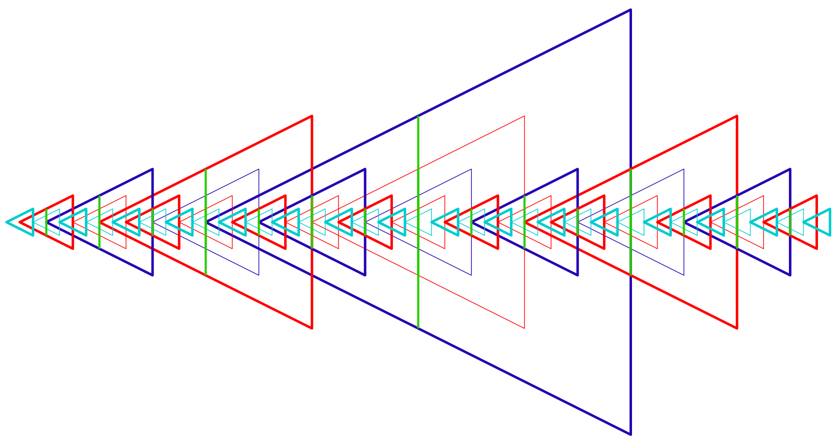

Figure 1 *Construction of the interwoven triangles in the Euclidean plane: look at the green signal.*

The mauve triangles are constructed from the red triangles of the interwoven ones. Any mauve triangle is triggered by a red triangle and conversely. The vertex of a mauve triangle T is that of a red triangle R. Its legs follow those of R. They go on on the same ray after the corner of R, until they meet the basis of the red phantoms which are generated by the basis of R. At this meeting, the legs meet the basis of the mauve triangle which coincide with the basis of the just mentioned red phantoms. In [15, 12], we thoroughly describe that this construction can be forced by a finitely generated tiling. We refer the reader to these papers.

2.1 Properties of the mauve triangles

As they stam from red triangles, we say that a mauve triangle of the generation n is constructed on a red triangle of the generation $2n+1$. Later, it will be useful to recognize the mauve triangles of generation 0. To this aim, we define a new colour, called **mauve-0** which is given to these triangles only which we call **mauve-0 triangles**.

From the doubling of the height with respect to the red triangles, the mauve triangles loose the nice property that the red triangles are either embedded or disjoint. This is no more the case for the mauve triangles. However, the overlappings and intersections of mauve triangles can precisely be described.

From [15, 12], we know that the intersection occurs by a leg of a mauve triangle cutting a basis of another mauve triangle. From the construction, see figure 2, any mauve triangle T of the generation n+1 contains three mauve triangles of the generation n with which they have no intersection. They also meet two mauve triangles of the previous generation. One of them is met at their basis: the legs of this triangle of the generation n cuts the basis of T. The other mauve triangle M of the generation n if any, meets T near its vertex. This time, the legs of T cut the basis of M. We give a number in [0..3] to the mauve triangles of the generation n whose vertex is contained in T, as four isoclines are involved by these vertices. Such a number is called the **rank** of the triangles. The rank is periodically repeated on

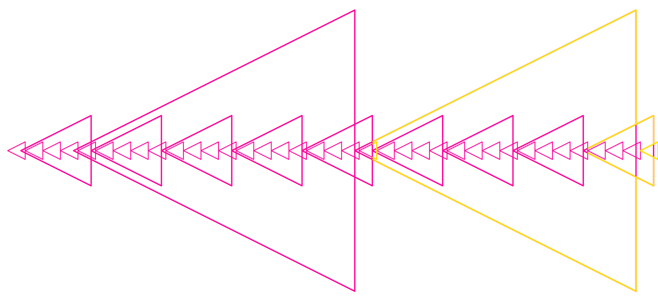

Figure 2 *An illustration of the mauve triangles.*

the mauve triangles of the generation n, to the top and to the bottom. A triangle of rank r is called an r**-triangle**. If a mauve triangle of the generation n contains the vertex of T, it is called the **hat** of T: it is a 3-triangle. The hat is unique when it exists. Note that if we can repeat the construction of the hat recursively until reaching a mauve-0 triangle, we obtain that the vertex of this mauve-0 triangle is at a distance at most $\dfrac{h}{4}$ of the vertex of T. We call this mauve-0 triangle the **remotest ancestor** of T, a notion already remarked for the interwoven triangles. Accordingly, if the vertex of T is on the basis of a mauve-triangle of the same generation, then its remotest ancestor exists.

The triangles of the generation n which cut the basis of T are also 3-triangles. We define the **low points** of a leg of a triangle, LP for short, as follows. Let R be the red triangle whose vertex is that of T and let P be a phantom generated by the basis of R. The LP's are the points of the legs of T which are on the isocline which is the mid-distance line of P. The LP's are at a distance $\dfrac{h}{4}$ from the basis of the triangle, where h is the length of the leg. The LP's play an important role: the line which joins the LP's of T cuts the 2-triangles also at their LP's. The intersection of the basis of T with its 3-triangles occur at their LP's.

In [15, 12], the consideration of the r-triangles has led to the extension of the notion of latitude used in the interwoven triangles to the case of the mauve triangles. First, we define the **primary latitude** of a mauve triangle as the set of isoclines which cross its legs, the basis being included but the top being excluded. This allows us to obtain a partition of the hyperbolic plane by the primary latitudes attached to a given generation. This defines a partition for each generation. But the primary latitudes overlap from one generation to the next one.

Now, we can precisely state the properties mentioned above about the intersections between mauve triangles.

Lemma 1 *Let T be a mauve triangle of the generation $n+1$. Then, the primary latitude of T intersects five primary latitudes of the generation n, denote them by L_{-1}, L_0, L_1, L_2 and L_3. There are four triangles of the generation n, T_0, T_1, T_2 and T_3 with the following properties:*

(i) T_i belongs to the primary latitude L_i for i in $\{0,..3\}$;
(ii) the vertex of T_{i+1} is on the basis of T_i for i in $\{0,..2\}$;
(iii) the LP's of T and T_2 are on the same isocline;
(iv) the legs of T_3 cut the basis of T at their LP's.

Assume that there is a 3-triangle T_{-1} belonging to the latitude L_{-1} which contains the vertex of T. In this case, the vertex of T is on the isocline which joins the LP's of T_{-1}.

We refer the reader to [19] for a proof of the lemma. From this proof we deduced a way to determine the mid-point and the LP's of a mauve triangle by means of signals defined by a finitely generated tiling. We give the mainlines of this construction in the next subsection.

2.2 Construction of the LP's of a mauve triangle

Let T be a mauve triangle of the generation n and let h be its height. Its mid-points, which lay at a distance $\frac{h}{2}$ are easy to determine: it is the corner of the red triangle R whose vertex is that of T.

The construction of the LP's proceeds as follows:

First, we look at the determination of the corners of T.

At the corners of R, the mauve signal defining the leg of T goes on along the extremal branch of the Fibonacci tree defining R. At the same time, each corner of R sends a signal towards the other one on the basis of R. Call this signal the **brown** signal. The brown signal has the laterality of the corner. When the brown signal meets the first vertex of the phantom P, it is a red phantom of the generation $2n+1$. The signal goes down along the leg of the phantom which has its laterality. It goes along this leg until it meets the corner of P. There, on the isocline ι of the basis of P, the brown signal leaves the leg to run on ι, to the side of its laterality, until it meets the mauve signal of the leg of T. Then, a mauve signal is sent to the other side, in order to meet the mauve signal sent by the other corner of T.

Now, the problem for the signal is to meet the correct leg, as it may encounter a lot of them along ι, the isocline of the basis of P, which belong to smaller generations. As the brown signal cannot count arbitrary numbers, it circumvents the triangles it meets on its way by climbing along their legs up to the vertex and then going down to the appropriate isocline. In order to recognize the right isocline, when the brown signal starts its circumventing path, it sends another brown signal, say a light one, with no laterality, which goes on running on ι, towards the appropriate corner. As ι is an isocline of the LP of the triangle which the brown signal circumvents, the brown signal cannot meet another light brown signal meeting the leg: as for red trilaterals, the isocline of a basis is specific to any mauve triangle. And so, when going down along the leg, the brown signal meets its light brown one, it knows that it has found ι on which it goes on its way, still to the side of its laterality. And now, the first mauve leg of its laterality met by the brown signal is the right one.

Now, The brown signal will help us to locate the LP's of T. Consider the time when the previous brown signal is going down along the appropriate leg of P. When the brown signal meets the mid-point of P, it knows that it is the isocline of the LP's of T. And so, the brown signal sends a purple signal of the same laterality as the brown signal towards the side of its laterality on the isocline ζ of the mid-distance line of P. This signal also circumvents the mauve triangles which it meets. Now, the signal is able to recognize ζ during the circumvention of phantoms thanks to the following. We know that the purple signal meets smaller mauve triangles at their LP's. By induction, we assume that a similar signal arrives to the LP's from inside the mauve triangle M, created at the time of the construction of M. Note that in any case, such a signal is stopped by the leg of M. Now, the arriving signal from the mid-point of the leg of P is deviated to the first part of the leg of M. When the signal goes down on the other leg, it identifies its LP by the arrival of a similar signal of the appropriate laterality which is stopped by the leg. This allows the signal to again find ζ and to go on its route on this isocline. Due to the laterality of the purple signal and to the fact that its laterality is unchanged and that it must match the mauve leg it meets from inside, such a signal cannot be present if it is not sent by a brown signal for detection purpose.

In [19], we thouroughly establish the correctness of this construction.

Now, the mauve signal which defines the basis of a mauve triangle is also emitted by the vertices of the mauve triangle of the same generation but whose primary laterality is just below the considered one. Now, in mauve triangles, a basis must be stopped by its corners. We know that in the construction of the interwoven triangles a triangle may be missing, which cancels all of those which could be constructed above its vertex. This also happens in the mauve triangles. But, as proved in [19], this can also be handled within the constraint of a finitely generated tiling. And so, we consider as granted that the corners of a mauve triangle stop its basis. This also means that if the legs do not exist to meet the basis, vertices of mauve triangles which lie on this part of the corresponding isocline do not emit the basis.

We can now state:

Lemma 2 *The mauve triangles together with the determination of their LP's and mid-points can be constructed from a finite set of prototiles.*

2.3 The β-clines and their construction

Now, we introduce the notion of β-cline and see how to construct it. This notion will play a key role in the construction of the path.

We start from the remark that the basis of a mauve triangle T of the generation $n+1$ cuts the legs of the 3-triangles which have their vertex inside T. Repeating this remark to the 3-triangle of the generation n, we can construct a sequence $\{T_i\}_{i\in[0..n+]}$ such that:

(i) $T_{n+1} = T$;
(ii) T_i is a 3-triangle of the generation i for i in $[0..n]$;
(iii) the basis of T_{i+1} cuts the legs of T_i, of course at their LP.

Any mauve triangle T of a generation $n+1$ generates such a sequence which we call the **shadow** of T. Of course, if $\{T_i\}_{i\in[0..n+1]}$ is the shadow of T_{n+1}, the sequence $\{T_j\}_{j\in[0..i+1]}$ is the shadow of T_{i+1} for i in $[0..n]$. We say that the shadow $\{T_j\}_{j\in[0..i+1]}$ is a **trace** of the shadow $\{T_i\}_{i\in[0..n+1]}$.

We say that a shadow $\{T_i\}_{i\in[0..n+1]}$ is a **finite tower** if it is not the trace of a shadow of a bigger generation. We shall see that there may be a sequence of mauve triangles $\{T_i\}_{i\in N}$ in which $\{T_i\}_{i\in[0..n+1]}$ is a trace of $\{T_i\}_{i\in[0..n+2]}$ for any n. In this case, we say that $\{T_i\}_{i\in N}$ is an **infinite tower**.

When $\{T_i\}_{i\in[0..n+1]}$ is a finite tower, we say that the isocline of the basis of T_0 is the **β-cline** of T_{n+1} and that its **type** is the rank of T_{n+1}.

From the β-clines, we define two new points on the legs of a triangle of a positive generation: the β- and **γ-points**.

By definition, the β-point of a mauve triangle T of the generation $n+1$ is the intersection of its leg with the β-cline of the 2-triangles whose vertex is inside T. It is not difficult to see that the β-point lies on the leg in between the LP and the corner. It is at a distance less than $\dfrac{h}{12}$ from the line joining the LP's of T, with h being the height of T, and as closer to this value as n tends to infinity.

2.3.1 Constructing the β-cline

To construct the β-cline, we define signals which start from the LP's of a mauve triangle of the considered generation and latitude. Call them the **β-signals**. The β-signals are lateral, with the laterality which is opposite to that of the leg on which they start. They travel along legs of mauve triangles and along isoclines of a basis. The β-signals go down along legs of a laterality opposite to their own one, from an LP to a corner. When they run on an isocline, they go in the direction of their laterality. When they meet a corner, they run on the basis, in the direction of the other corner. They can freely travel on this isocline, until they meet the leg of a triangle of a laterality which is opposite to their own one and at their LP. If the leg is of another laterality or if the meeting point is not in the closed interval with the LP of the leg and its corner as end points, the β-signals crosses the leg. It is plain that both β-signals starting from the opposite LP's of the same mauve triangle will meet, and they cannot do that along a leg or at a corner. When they meet, they use a join tile, see [14, 17], in which the right-hand, left-hand side β-signal is on the left-, right-hand side part of the tile. It is plain that both β-signals define a kind of convex hull of this part of the mauve triangle.

Note that for two consecutive mauve triangles of the same generation within the same isocline, the β-signal which starts from the low-point of one of them cannot travel on this isocline to the facing low-point of the other triangle. Indeed, on the right-hand side low-point, we have a left-hand side β-signal and on the left-hand side low-point with have a right-hand side β-signal. And so, this would require a join tile with a left-hand side β-signal on the left-hand side part of the tile: this is ruled out.

We may impose an additional constraint on the join tile for β-signals of opposite lateralities with the right-hand side signal on the left-hand side of the join tile: the join tile generates a horizontal unilateral yellow signal. This signal runs on an isocline only: it marks the β-cline. It is important to note here that the whole isocline constitutes the β-cline.

By construction, the yellow signal travels along an isocline of a basis of a mauve 0 triangle. Consequently, it travels on an isocline 5. Accordingly, it meets no basis of a mauve triangle of a positive generation and no LP, as LP's are always on an isocline 15. And so, the yellow signal will meet legs of triangles.

In our study of the shadow of a mauve triangle, we have already noticed that the same β-cline can be shared by several triangles of different generations.

For the purpose of the path, in the case of a β-cline of type 2, we consider that it defines a special signal on the isocline which is just below the β-cline. This means that there is a **pre-path signal** on the isocline 4 which is just below a β-cline of type 2. This signal plays an important role as can be seen further. Note that the pre-path signals of a given generation are different from those of the next generation: the signals corresponding to the generation $n+1$ occur on the isocline 4 of a β-cline of type 3 in terms of the generation n. We have a stronger result:

Lemma 3 *The isoclines of a pre-path signal of the generation n are different from those of the generation m for any n, m with $n \neq m$.*

We omit the easy proof which can be found in [19].

2.4 The β-points and their construction

For the next section , we need to make clear the connection between a mauve triangle T of the generation $n+1$ and its inner mauve triangles of the generation n.

To locate the triangles of the just previous generation, there is a way given by the local numbering of the triangles. We have already noticed that the intersection between mauve triangles occur between a leg and a basis and that with respect to the leg, the intersection happens at its low point. The consequence is that mauve triangles of the generation n which are inside T are cut by the basis of T if and only if they are 3-triangles. Now, the converse is true:

Lemma 4 *Let T be a mauve triangle of the generation $n+1$. Its basis cuts mauve triangles of the generations i for any i in $[0..n]$. When $i = n$, the mauve triangle is of type 3. When $i < n$, the mauve triangle is of type 2.*

Again, the proof is to be found in [19].

Now, consider the β-cline of type 2 which corresponds to the mauve triangles of the generation n which are inside T. It is important to recognize the intersection of this β-cline with the legs of T. We call them the **β-points** of T.

Note that there is no β-point on a mauve-0 triangle and that the β-points of a mauve triangle T_1 of generation 1 are easy to determine. Indeed, the line joining the LP of T_1 cuts inner mauve-0 triangles of type 2. Now, the basis of theses triangles are on the same isocline which is the β-cline passing through the β-point of T_1. And so the construction is simple:

---

a **silver** signal is sent from the LP of T_1 until it reaches the first mauve-0 triangle of type 2, T_2;

the silver signal goes down along the leg of T_2; when it reaches the corner of T_2, it also reaches the β-cline of T_2; it follows this β-cline outside T_2;

this intersection of the silver signal with the leg of T_1 defines the β-point of T_1.

The construction of the β-point in the general case relies on lemma 4. and is given by algorithm 1, below.

Algorithm 1 *The construction of the β-point of a triangle T of the generation $n+1$.*

---

the silver signal starts from the corner into two directions;

the first direction follows the basis until it meets the leg of a triangle of type 3;

- it goes along this leg up to the vertex and there, it follows the isocline of the vertex away from the leg of T, until it meets a corner which is a corner of a triangle of type 2, T_2;
- from the corner of T_2, the silver signal follows the β-signal coming from the LP of T_2 which is above the considered corner;
- then the silver signal eventually meets the β-cline defined by the β-signal of T_2; the silver signal goes back to the leg of T, following the just met β-cline;

the second direction follows the leg of T reaching the corner and goes up along this leg towards the LP of T;

both directions of the silver signal meet at the intersection of the leg of T with the expected β-cline of type 2 coming from an internal mauve triangle of the generation n: it is the expected β-point and the intersection stops both silver signals.

Note that this construction also holds when $n = 0$.

In [19], we prove the correction of this algorithm. We also get:

Corollary 1 *In any mauve triangle of a positive generation, there is a single β-point on each leg.*

It can be noticed that algorithm 1 to construct the β-points can be processed in the reverse order. This means that it can be constructed from a mauve triangle T of type 2 and of the generation n for looking at the β-point of the mauve triangle M of the generation $n+1$ which contains T if any. The algorithm may detect if M exists or not and, when it exits, how to find the β-point, see [19].

A last feature about the β-point is that it allows to differentiate the part of the β-cline of type 2 on which it lies which is contained in the triangle from the part which is outside.

Later, we shall see that this differentiation is very important. It can easily be realized, for instance as follows, according to the differentiation between open and covered basis in the interwoven triangles. Each β-point emits a horizontal signal on its isocline, outside the triangle to which it belongs. The signal is lateral and has the laterality of the leg. In between two consecutive mauve triangles on the same primary latitude and of the same generation, the signals emitted by the opposite β-points meet thanks to a join-tile which is similar to those used with the interwoven triangles. On the part where the horizontal signal is present, we shall say that the β-cline is **covered**. In the part where it is not present, we shall say that the β-cline is **open**. Clearly, the β-cline is open inside the mauve triangles of its generation and it is covered in-between two consecutive such triangles within the same latitude.

2.5 The latitude

From lemma 1, we know the intersections between mauve triangles of the generation $n+1$ and those of the generation n. We have to look at a more general situation.

From the construction of the interwoven triangles, we know that the bases and vertices of mauve triangles characterize the corresponding triangles. This is not the case for the isocline of their LP's: such an isocline is the mid-distance line of phantoms. Now, the same isocline can be the mid-distance line of phantoms which belong to different generations. Consequently, the same ambiguity is attached to the isoclines of the LP's as we can see from lemma 1. Recursively applying the lemma to inner triangles in a fixed mauve triangle, we obtain that LP's of a triangle of the generation n may be crossed by the basis of a triangle of the generation $n+k$, for any positive k. In general, it is not possible to predict if such a situation will occur. Now, if it occurs, we know that inside the mauve triangle of the generation n, the 2-triangles will also be cut by this basis, also at their LP's.

From lemma 1, we know that this situation does not occur for the 0- and 1-triangles which are contained in a mauve triangle. These triangles may be intersected by smaller triangles only, which cut their basis or their legs near their vertices.

Going back to 2-triangles, we can see that if a 2-triangle T is of a generation n with $n > 0$, we can find smaller triangles which are also 2-triangles inside T, their legs being cut by the basis of T, at their LP's too. And this can be repeated until we reach the generation 0.

We can say the same for 3-triangles. If such a triangle is not of the generation 0, its basis cuts triangles of the previous generation, and this property can be repeated recursively. Remember the notion of shadow of a triangle and the construction of the β-cline.

From this, we define the **border line** of a primary latitude of the generation n as a broken line as follows:

First, define the **bottom** of a mauve triangle as the broken line which consists of the legs of the triangle from the LP to the corner and the basis.

Then, we define the **border line** as the isocline of the LP's of the triangles of the generation n of this primary latitude in which each maximal segment which falls inside a mauve triangle M of a generation at most $n-1$ is replaced by the bottom of M, the same process of substitution being recursively applied to the basis of the triangle and of the substituted triangles. The term maximal indicates that we take the biggest triangle of a generation at most $n-1$ which is cut by the isocline.

From now on, the **latitude** of a triangle of the generation n is the set of tiles which is contained between the border line of its primary latitude and the border line of the same generation which is attached to the primary latitude which is just above. We include all the tiles of the lower border and we include none of the upper border.

Note that in a border line, when we apply the recursive process of substitution of bottoms of triangles starting from a triangle of the generation n, the bases which are the further from the isocline of the LP's are bases of the generation 0. They are all on the same β-cline.

Now that the notion of latitude is clearly defined, let us look at what happens between two consecutive triangles T_1 and T_2 of the same generation which belong to the same latitude.

A priori, we have three situations:

(*i*) for both T_1 and T_2, the vertex does not belong to a basis of a mauve triangle;

(*ii*) the vertex of T_1 does not belong to the basis of a mauve triangle but the vertex of T_2 does;

(*iii*) each vertex belongs to a basis of a mauve triangle.

In fact, we have:

Lemma 5 *Consider two mauve triangles T_1 and T_2 of the generation n and belonging to the same latitude. Assume that T_1 and T_2 are consecutive. Then, if the vertex of T_i belong to the basis of a triangle B_i for $i \in \{1, 2\}$, then $B_1 = B_2$.*

We refer the reader to [19] for the proof.

2.6 The γ-points and the high points

We conclude this section with the notion of **γ-point** and of **high point**, HP for short, which both play an important role in the next section.

Intuitively, the LP corresponds to the entry of the path into a triangle and the HP corresponds to its exit. The γ-point plays a similar role to that of the β-point.

2.6.1 The γ-point and its construction

The γ-point is defined by the intersection of the leg of mauve triangle with the β-cline defined by its hat, if any. The difficulty comes from the fact that the hat may not exist while the γ-point can always be defined for a mauve triangle of a positive generation.

As for the β-point, the γ-point is not defined for a mauve-0 triangle. For a mauve triangle T_1 of generation 1, consider the above definition when the hat exists. We remark that the β-cline is defined by the basis of the hat as it is a mauve-0 triangle. Now, the basis of the hat contains vertices of the 0-triangles of generation 0 contained in T_1. Now, as T_1 exists, its inner 0-triangles also exist. And so, it is possible to define the γ-points of T_1 by using its 0-triangles only.

First, we call **first points**, FP for short, the point of a leg of a mauve triangle T which is on the mid(point of the red triangle whose vertex is that of T. It is at a distance $\frac{h}{4}$ from the vertex of T. And so, the determination of the FP's is easy.

Then, we proceed as follows:

---

two γ-signals start from the FP of T_1: one to its vertex, along the leg, the other inside the triangle;
the inside signal goes on along the isocline until it meets the closest 0-triangle M_0 to this leg of T_1; there, it goes up along the leg until it reaches the vertex of M_0; the γ-signal goes back to the leg of T_1, following the isocline of the vertex of M_0;
the intersection of the γ-signal going back to the leg with the γ-signal going up along the leg defined the γ-point of T_1.

The general case is not much more difficult to establish by the following recursive algorithm.

Algorithm 2 *The construction of the γ-point of a triangle T of the generation n+1.*

---

two γ-signals start from the FP of T, one along the leg towards the vertex and the second inside the triangle along the isocline which joins the FP's;
the inside signal goes on until it meets the first 0-triangle M_0 inside T; there, meeting M_0 at an LP, it goes up along the leg of M_0 until it reaches the γ-point G_0 of M_0; there, it goes back to the leg of T, on the isocline which passes through G_0, circumventing the inner triangles which it encounters;
the intersection of the γ-signal going back from G_0 to the leg of T with the γ-signal climbing along this leg defines the γ-point of T; the γ-point stops both γ-signals.

The justification of the construction given by algorithm 2 is provided by the following lemma.

Lemma 6 *Let T be a mauve triangle of the generation $n+1$. The isocline which passes through its FP's encounters mauve triangles inside T of types 0 and 2 only. The meeting occurs at the LP's of the inner triangles. The encountered 0-triangles belong to the generation n. The encountered 2-triangles belong to a generation i with $i < n$.*

We refer the reader to [19] for the proof.

We have an additional interesting property:

Lemma 7 *The isocline of the γ-points of a mauve triangle meets other mauve triangles at their γ-points too.*

Proof: obvious.

We can formulate the same remark about algorithm 2 as the one which was formulated for algorithm 1. The construction can also be performed in the reverse order. Again a pre-signal detects the existence of a containing mauve triangle of the next generation. It is the same signal as previously, looking after a basis at the LP of a mauve triangle of type 3 reached from the considered mauve triangle of type 0. If the basis is found, the pre-signal goes back to its emitting point in order to trigger the signals of algorithm 2 in the reverse order. Again, this provides us with an iterative and bottom-up version of algorithm 2.

2.6.2 The HP

From the notion of γ-points, it is easy to define the HP's.

Indeed, the HP's of a mauve triangle T is defined by the following construction.

A signal starts from each FP of T and goes up along the leg, towards the vertex of T. If there is a γ-point, then if the β-cline which passes through the γ-point is a β-cline of type 2, the HP is the γ-point and the signal stops there. Otherwise, the signal goes on climbing along the vertex until it meets the first basis which cuts the legs of T if any. If such a basis is encountered, the meeting with the legs of T define the HP's. If not, the HP is the tile of the leg which is on the isocline which is just below the vertex. This is also the definition of the HP for a mauve-0 triangle.

3 An almost plane-filling path

Now, we turn to the construction of the path. The general strategy which we follow was presented in [15, 12], but we shall make it much more precise.

The path goes from an LP to a HP and then to an LP and so on. It can be seen as a bi-infinite word of the form ${}^{\infty}((LP)(HP))^{\infty}$ on the alphabet $\{LP, HP\}$.

Roughly speaking, we fill up a latitude until we meet legs which cross both the upper and the lower border of the latitude. Then, we go up or down, depending on the direction of the path and into which type of basic region we fall: the type of a bigger triangle or of a zone in between two bigger triangles.

In most cases, this strategy is enough to fill up the whole plane.

Later, we shall discuss about the exceptional cases.

3.1 The regions and the path

Our first task is to define the regions which we shall investigate and then, how the path is built on the basis of what will be called the **basic regions**.

We have two basic regions. The first one is the set of tiles defined by a mauve triangle: its borders and its inside. Remember that the basis of a mauve triangle contains more than the majority of tiles resulting from the just given definition. It is considered as a basic region as once the path enters a mauve triangle T, it fills up T almost completely before leaving T. In fact, there is a restriction and the path fills a bigger area. In fact, the path also fills up the space which is contained between the basis of T and the part of the border of the latitude of T which is delimited by the corners of T, the tiles on this border being included. The restriction comes from the definition of the latitude: we have to withdraw at least the tiles belonging to the border of the just upper latitude of the same generation. An additional restriction comes in the case when the HP is on an open β-cline of type 2, as we shall describe this later.

The other type of a basic region is defined by the area in between two consecutive mauve triangles of the same generation within the same latitude.

We already know that the just indicated regions can be split into four horizontal slices defined by the types of the triangles of the just previous generation which are contained in these regions. Now, if we go from one side to another in each slice, and if the directions alternate from one slice to the next one, this even number raises a problem: a priori, starting from one side, we go back to the same side. To solve this problem, we split one slice into two ones thanks to the β-cline of type 2: inside a mauve triangle, there is a unique open β-cline of type 2 which runs from one leg of the triangle to the other. It is the isocline of the β-points. This β-cline splits the region of type 3 into to sub-slices. We shall use the second one to go back to the original side. As there remain three slices, we go from the original one to the opposite one, as required.

This is the general principle for defining the path. Note that this principle holds both for triangles an the in between region. We shall now turn to the precise description.

We shall examine how we fill up the basic regions for generation 0 and we shall then proceed by induction from n to $n{+}1$. In fact, as we shall see, the induction step can be based on what is to do for the basic regions of generation 1.

<u>For generation 0</u>

For a triangle, the path enters the figure through one of its LP's, say A. Then, it runs along the leg of the triangle, downwards, until it reaches the corner. On this way, the path is in the inside part of the tile which supports the leg. At the corner, the path follows the basis, until it reaches the other corner. There, it goes up along the leg to the next isocline and there, it goes along the isocline to the leg of A. Just before reaching the leg, the path goes up to the next isocline and there, it runs along it until it reaches the leg, opposite to A. This back and forth motion, climbing up by one isocline each time a leg is reached goes on until the path reaches the top of the triangle. There, the path exits from the triangle through the isocline -1 below the vertex or the isocline -2, depending on the type of the triangle: if the triangle is of type 3, the path exists through the isocline -2, otherwise, it exits through the isocline -1. The exit B is placed on this isocline, on the leg of the triangle which is opposite to the leg on which A lies. The sub-figure (a) of Figure 3 illustrates this part of the path for a triangle when the topmost isocline is not occupied by another segment of the path.

For a part between two consecutive mauve-0 triangles within the same latitude, we have the three situations which result from lemma 5.

The easiest situation is when two consecutive mauve-0 triangles have their vertices on the basis of the same mauve-0 triangle. In this case, we have a similar zig-zag line as in a triangle. The situation is illustrated by the sub-figure (b) of Figure 3.

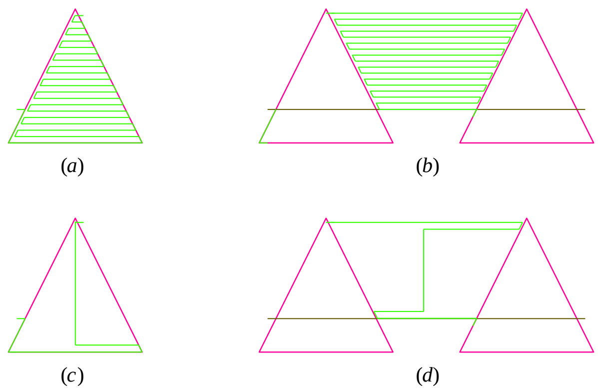

Figure 3 *A schematic representation of the path:*
On the left-hand side, inside a mauve-0 triangle. On the right-hand side, in between two consecutive mauve-0 triangles within the same latitude when the vertices belong to the same basis.

In order to describe what happens in the other situations, we define a schematic representation of the zig-zag path of the sub-figures (a) and (b) of Figure 3 by the sub-figures (c) and (d) of Figure 3 respectively. Now, as these situation will be involved starting from generations with a positive number, we postpone the representation of the other cases of basic regions of generation 0 to the situation concerning generation 1.

The representations of the sub-figures (a) and (b) of Figure 3 are also schematic. In fact, the actual trajectory of the path is a bit more complex. We cannot decide that on a leg of a mauve triangle of any generation the path strictly goes on the tiles crossed by the mauve signal and only them. If we do this, we cannot have a path which goes through any tile according to the indicated scenario. However, it is possible to slightly change the trajectory of the path in order to make things possible. Figure 4 illustrates a solution for this issue.

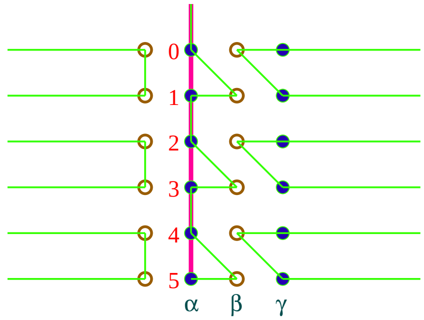

Figure 4 *The adaption of the path close to a leg of a mauve triangle.*

For generation 1

First, we look at what happens in between two consecutive mauve triangles of generation 1 within the same latitude. Denote them by T_1 and T_2, with T_1 on the left-hand side of T_2. Remind that we assume that the path enters a mauve triangle of generation 1 through an LP and that it exits the same triangle through its top, on the leg which is opposite to the entry point.

In figure 5, we consider the case when a mauve triangle T_1 of generation 1 is hatted by a mauve triangle H_1 of generation 0. The next mauve triangle of generation 1 to the right, say T_2 is not hatted as it can be easily concluded from the distance between the corners of two consecutive mauve-0 triangles.

There are necessarily 0-triangles of generation 0 in between T_1 and T_2. Figure 5 illustrates a schematic situation of the i-triangles of generation 0 which we may find in in between T_1 and T_2. Note that we have 0-, 1- and 2-triangles. The 3-triangles are not represented as they do not belong to the latitude of generation 1 defined by T_1 and T_2.

The figure illustrates the way of the path, assuming that it exits from T_1 through its right-hand side HP in order to enter T_2 through its left-hand side LP. We have to take into account the behaviour of the path in the primary latitude of H_1. It again appears in the figure by looking at the configuration of the 0- and 1-triangles of generation 0 which are in between T_1 and T_2.

First, the path follows the border of H_1 and then climbs along its right-hand leg until it reaches the isocline which is just below the LP of H_1. It goes on along this isocline until it reaches the left-hand side leg of T_2, just below the vertex of T_2. Next, it follows a zig-zag way until it goes back to the point M defined by the corner of H_1. This point M lies on the isocline ι which is just below the basis of H_1 and is on the way upwards taken by the path. During the zig-zag, the path meets the vertices of 0-triangles of generation 0. As the path inside a triangle never passes through its vertex, the path may cross them, as if it would do if a basis would contain these vertices. Coming back after leaving the closest vertex of such a 0-triangle to H_1 and traveling on the isocline ι+1, the path arrives to the tile which is before the tile of the path above M on $\iota+1$. There, the path goes down to ι and, on the tile which is adjacent to M, it goes on the isocline ι in the direction of T_2. Now, the path does not meet T_2 but a 0-triangle T_0 of generation 0, which it reaches just below the vertex. Accordingly, the path zig-zags downwards, oscillating between T_1 and T_0. By this oscillating motion, the path reaches the LP of T_0: it enters the triangle which it fills according to the motion defined by the sub-figure (a) of Figure 3. When the path exits from this triangle, it follows the way defined by the sub-figure (b) of Figure 3 until it reaches the next 0-triangle on its way to T_2. Accordingly, this sequence is repeated until the last 0-triangle of generation 0 before T_2. Now, when the path exits from the triangle, it is barred by the former passage of the path on β and so the path goes on ι until it reaches T_2. But, as the path exited from T_1 and as it is close to β, it knows that it cannot enter T_2. And so, it goes downwards in zig-zagging. Now, during this zig-zag, it will meet the LP of the last 0-triangle of generation 0: this LP is closed as the path filled up this triangle. We shall later see the mechanism which forces one LP to be open and the other to be closed. And so, going down, still zig-zagging, the path will meet the LP of the closest 1-triangle of generation 0 to T_2. Here, the LP is free, so that the path enters the triangle.

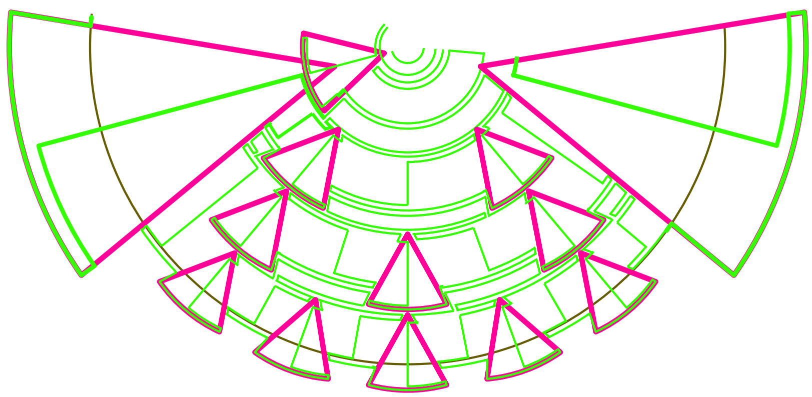

Figure 5 *The path in between two triangles of generation* 1.

Now, we turn to the route of the path inside a mauve triangle of generation 1.

In both pictures of figure 6, we can see an open β-cline of type 2 which cuts the strip delimited by the line of the LP's and the basis of the triangle into two parts.

This is a general feature. This cut allows to make the path going back near the LP through which it entered the triangle in order to cross the latitude of the 2-triangles in the direction from LP to HP, where LP refers to the side of the triangle through which the path entered and HP refers to the other side as the path will exit through the HP of this other side. The crossing of this latitude inside the triangle obeys the same principles as in between two triangles. When arriving almost to the closed LP, the path goes up to the isocline which is below the FP. From this point, it crosses the latitude of the 1-triangles, this time in the direction from HP to LP. When it arrives to the other side, the path goes up along the leg until it arrives by one isocline below the HP of this leg. From there, it crosses the latitude of the 0-triangles, in the direction from LP to HP. When the crossing completes, the path arrives at the FP from where it goes to the right HP by going up along the leg of the triangle.

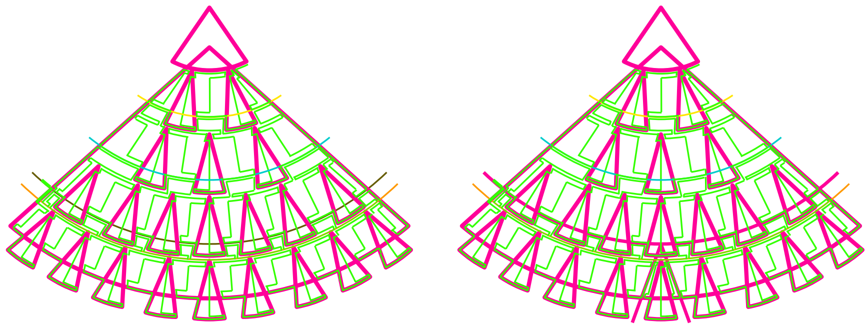

Figure 6 *The path inside a triangle of generation* 1.
On the left-hand side: a 0*- or a* 1*-triangle. On the right-hand side, a* 2*- or a* 3*-triangle which is cut by a mauve triangle of a bigger generation.*

In between two consecutive mauve triangles of generation 1 within the same latitude, the β-cline 2 which we noticed inside a triangle T of generation 1 plays a similar role but for another latitude: for the one which is below the latitude of T. Now, for a basic regions, there are a lot of β-clines of type 2 which cross the legs of the triangles defining these regions. The β- and γ-points tell us which one are important for the region: only those which pass through this points. The other intersections are not important.

In a basic region, we have four sub-latitudes, corresponding to the four types of mauve triangles of the previous generation. In order to go into the right direction, we need to split one such sub-latitude into two horizontal ones. The role of the β- and γ-points is to be the milestones on the path which indicate where it is possible to make this splitting. And so, when the path meets a β-cline of type 2 along a leg, if the point of intersection is neither a β-point nor a γ-one, it knows that it may cross this β-cline to go on the zig-zags. In the other case, depending on which type of point is met, the path knows that the β-cline must be followed in order to cross the leg of a triangle.

From the generation n to the generation $n{+}1$

Figures 5 and 6 allow us to prove the induction step which allow to establish the path in a basic region of the generation $n{+}1$ once the path is established in any basic region of the generation n.

However, a tuning is needed here, as the β-clines are no more in contact of the bases for the mauve triangles of the generation $n{+}1$. To see this point, consider that we also draw the mauve triangles of the generation $n{-}1$, now assuming that $n \geq 1$. Then, it is not difficult to see that the regions of the generation n split into regions of the generation $n{-}1$ in the same way as those of the generation $n{+}1$ split into regions of the generation n.

We have the following general property:

Lemma 8 *Let τ be a tile of the tiling. Then for any non-negative n, there is a mauve latitude Λ of the generation n such that $\tau \in \Lambda$. And then: either τ falls within a mauve triangle of generation n in this latitude or τ falls outside two consecutive mauve triangles of generation n and of the latitude Λ and in between them.*

3.2 Additional tuning

In order to ensure the guidance of the path, we provide an additional tool.

As indicated in the previous section, if one LP allows the path to enter a triangle, the other forbids such a possibility. We have the same property for a HP.

In fact, it is not difficult to devise signals based on the notion of laterality which allow to ensure this working. It may be one or the other LP, mandatory one of them and never both of them. This is performed by a signal which runs along the legs and which meet at the vertex. Each LP sends a signal to the other which runs along the leg to the vertex where they meet. If the LP admits the path, it sends a signal of its laterality and if not, it sends a signal of the other laterality. And so, it is enough to forbid the meeting of signals of opposite lateralities. In this way, only unilateral signals are allowed and they indicate the general motion of the path. Note that once the laterality is fixed, this allows to place signboards at appropriate places. First, the knowledge of which LP is admitting allows to know which HP allows the path to exit from the triangle. This is inside a triangle. Now, the same mechanism can be used to direct the path in between two consecutive triangles. This time, the information, still going from an LP to another goes through the corners and takes the route of the red basis of a phantom which runs on the considered isocline. On this isocline there can be corners of the appropriate generation only. Now, inside a basic region and within a sub-latitude, the direction of the path is the same. In fact it is the same all along the latitude, as can be easily noticed from the fact that there is a shift in the triangles with respect with the in between regions. Accordingly, the same direction occurs globally. The change of direction happens when the path meets the legs of a triangle. This occurs for the standard hairpins of the zig-zags. Now, the signal which goes from one LP to the other also allows to place signboards at the decisive positions: the mid-point and the FP's, when the path climbs along the leg to go from a sub-latitude to the next one. Now, the signal which goes in between two consecutive triangles has also to detect the possibility of a leg coming from a bigger generation: this event may change the direction of the further motion of the path. For this purpose, the signal circumvents the mauve triangles it meets on its way. The isocline of a corner continues a basis: accordingly it meets smaller mauve triangles at their LP's, the mauve triangles of their generation at corners again and bigger triangles at various places, except the LP's. Accordingly, when such a meeting occurs, the signal knows that it stops here.

A last tuning deals with the parity of the number of zig-zags in a basic region.

It is not difficult to notice that the path should arrive at particular isoclines in the right direction. As already seen, the various signboards which we have constructed allow to do this without problem. As an example, the corners of mauve triangles play an important role but there is not need to signalize them: they are recognizable by their very conformation which is unique. Now, in order that the zig-zag line leads a point on a leg to the opposite leg, we need an odd number of zig-zags. The height of a triangle, in terms of isoclines, the basis being included but the vertex being excluded, is an even number. But, it is not difficult to organize one piece of a zig-zag in a given direction on two isoclines. It is enough to go up to a node of the highest isocline from its leftmost son, then to go down to the next son from the son and then to go on until the leftmost son of the next node. The need of such a run can be signalized, as the parity of the number of isoclines can easily be computed. It is enough to put signboards of the required points three isoclines sooner in order the path know whether it goes on along a standard motion or it has to simultaneously cross two isoclines on the same motion.

4 About the injectivity of the global function of a cellular automaton in the hyperbolic plane

4.1 Almost filling up the plane

We can derive two corollaries from lemma 8, whose proofs can be found in [19].

Corollary 2 *The path contains no cycle.*

Corollary 3 *For any tile τ, the path on one side of τ fills up infinitely many mauve triangles of increasing sizes.*

The proofs of these corollaries allow to establish the following statement:

Corollary 4 *If there are only finite basic regions, the path goes through any tile of the plane.*

4.2 The exceptional situation

Corollary 4 indicates that if there are only finite triangles, then we have a plane-filling path.

Is it possible to have infinite triangles?

The answer is yes: this means that there are also infinite red triangles. We know that this happens with the butterfly model, see [14, 17]. In this case, no interwoven triangle crosses a given isocline 15. As a corollary, there is an infinite mauve basis which crosses infinitely many 2-triangles of any sizes. Now, this basis gives rise to infinitely many mauve triangles, by the very principle of synchronization.

And so, this situation is possible. Now, it is the unique one: an infinite triangle has an infinite basis and this assumption leads to what we have just described.

In this case, there cannot be a single path passing through all tiles of the plane once only.

Indeed, once the path enters an infinite triangle, it cannot leave it. The same for a region in between two infinite triangles with the vertex on the infinite basis. And so, in this case, there are infinitely many components for the path. However, corollary 3 is still valid for them.

4.3 Proof of the main theorem

We can now prove:

Theorem 2 *The injectivity of the global function of a cellular automaton on the ternary heptagrid from its local transition function is undecidable.*

The proof follows the argument of [3], with a slight modification.

In particular, we have to bring a new ingredient to the path as described in section 3: we define a direction for the path. This can be introduced by three hues in the colour used for the signal of the path. One colour calls the next one and the last one calls the first one. The periodic repetition of this pattern together with the order of the colours define the direction. This notion of direction allows to define the successor of a tile on the path. This can be formalized by a function δ from $\mathbb{Z}$ to the tiling such that $\delta(n{+}1)$ is the successor of $\delta(n)$ on the path.

Consider M a deterministic Turing machine with a single head and a single bi-infinite tape which is assumed to be initially empty. From [14, 17], we can define a finite set of tiles T_M such that T_M tiles the hyperbolic plane if and only if M does not halt. An automaton A_M is attached to M and its states are defined by $D \times \{0,1\} \times T_M$, where D is the set of tiles which defines the tiling which we have constructed in section 3. The $0,1$-component of a state is called its **bit**. We can still tile the plane as the tiles of T_M are ternary heptagons but the abutting conditions may be not observed: if it is observed with all the neighbours of the cell x, the corresponding configuration is said to be **correct** at x, otherwise it is said **incorrect**. When the considered configuration is correct at every tile for D or at every tile for T_M, it is called a **realization** of the corresponding tiling. Let δ denote the function defining the orientation of the path induced by a realization of D.

As in [3], the transition function does not change neither the D- nor the T_M-component of the state of a cell x: it only changes its bit. As in [3], we define $A_M(c(x)) = c(x)$ if the configuration in D or in T is incorrect at the considered tile. If both are correct, we define $A_M(c(x)) = \text{xor}(c(x), c(\delta(x)))$. It is plain that if M does not halt, T_M tiles the hyperbolic plane and there is a configuration of D and one of T_M which are realizations of the respective tilings. Then, the transition function computes the xor of the bit of a cell and its successor on the path. Hence, defining all cells with 0 and then all cells with 1 define two configurations which A_M transform to the same image: the configuration where all cells have the bit 0. Accordingly, A_M is not injective.

Conversely, if A_M is not injective, we have two different configurations c_0 and c_1 for which the image is the same. Hence, there is a cell x at which the configurations differ. Hence, the xor was applied, which means that D and T are both correct at this cell in these configurations and it is not difficult to see that the value for each configuration at the successor of x on the path must also be different. And so, following the path in one direction, we have a correct tiling for both D and T_M. Now, from corollary 3, as the path fills up infinitely many triangles of increasing sizes, this means that the tiling realized for T_M is correct in these triangles. In particular, the Turing machine M never halts. And so, we proved that A_M is not injective if and only if M does not halt. Accordingly, the injectivity of A_M is undecidable. ■

5 Conclusion

The question of the surjectivity of the global function of cellular automata in the hyperbolic plane is still open. In the Euclidean case, the undecidability of the surjectivity problem is derived from the undecidability of the injectivity as the surjectivity of the global function of a cellular automaton is equivalent to its injectivity on the set of finite configurations, see [22, 23]. Now, in the case of cellular automata in the hyperbolic plane, this is not at all the case. The surjectivity and the injectivity of the global function are independent: there are examples of surjective global functions which are not injective and of injective global functions which are not surjective, see [20].

Accordingly, this question is completely open in the hyperbolic plane, even if it is is likely to be undecidable.

References

[1] Berger R., The undecidability of the domino problem, *Memoirs of the American Mathematical Society*, **66**, (1966), 1-72.

[2] Goodman-Strauss, Ch., A strongly aperiodic set of tiles in the hyperbolic plane, *Inventiones Mathematicae*, **159**(1), (2005), 119-132.

[3] Kari J., Reversibility and Surjectivity Problems for Cellular Automata, *Journal of Computer and System Sciences*, **48**, (1994), 149-182.

[4] Kari J., The Tiling Problem Revisited, *Lecture Notes in Computer Science*, **4664**, (2007). 72-79.

[5] M. Margenstern, New Tools for Cellular Automata of the Hyperbolic Plane, *Journal of Universal Computer Science* **6**(12), (2000), 1226–1252.

[6] Margenstern M., About the domino problem in the hyperbolic plane from an algorithmic point of view, *Technical report,* 2006-101*, LITA, Université Paul Verlaine – Metz*, (2006), 100p., available at: http://www.lita.sciences.univ-metz.fr/~margens/hyp_dominoes.ps.gzip

[7] Margenstern M., About the domino problem in the hyperbolic plane, a new solution, *arXiv:cs.CG/*0701096, (2007), January, 60p.

[8] Margenstern M., About the domino problem in the hyperbolic plane, a new solution, *Technical report,* 2007-102*, LITA, Université Paul Verlaine – Metz*, (2007), 106p., available at: http://www.lita.sciences.univ-metz.fr/~margens/new_hyp_dominoes.ps.gzip

[9] Margenstern M., On a characterization of cellular automata in tilings of the hyperbolic plane, *arXiv:cs/*0702155, (2007), February, 17p.

[10] Margenstern M., The Domino Problem of the Hyperbolic Plane Is Undecidable, *arXiv*: 0706.4161, (2007), June, 18p.

[11] M. Margenstern, On a characterization of cellular automata in tilings of the hyperbolic plane, **ACMC'2007**, Aug. 2007, Budapest, (2007)

[12] Margenstern M., Constructing a uniform plane-filling path in the ternary heptagrid of the hyperbolic plane, *arXiv*: 0710.0232, (2007), October, 22p.

[13] Margenstern M., Cellular Automata in Hyperbolic Spaces, Volume 1, Theory, *OCP*, Philadelphia, (2007), 422p.

[14] Margenstern M., The Domino Problem of the Hyperbolic Plane is Undecidable, *Bulletin of the EATCS*, **93**, (2007), October, 220-237.

[15] Margenstern M., Constructing a uniform plane-filling path in the ternary heptagrid of the hyperbolic plane. *Computer Science Journal of Moldova*, **15**(3-45), (2007), to appear.

[16] Margenstern M., About the domino problem in the hyperbolic plane from an algorithmic point of view, *Theoretical Informatics and Application*, **42**(1), 2008, 21-36.

[17] Margenstern M., The domino problem of the hyperbolic plane is undecidable, *Theoretical Computer Science*, in press, **doi**: 10.1016/j.tcs.2008.04.038, 56p.

[18] Margenstern M., Is the injectivity of the global function of a cellular automaton in the hyperbolic plane undecidable? *arXiv:*0712.2577, (2007), December, 16p.

[19] Margenstern M., The injectivity of the global function of a cellular automaton in the hyperbolic plane is undecidable, *arXiv:*0806.1602, (2008), June, 29p.

[20] Margenstern M., Cellular Automata in hyperbolic spaces: new results, **AUTOMATA'2008**, June, 12-14, 2008, Bristol, UK, invited talk.

[21] M. Margenstern, K. Morita, NP problems are tractable in the space of cellular automata in the hyperbolic plane, *Theoretical Computer Science*, **259**, 99–128, (2001)

[22] Moore E.F., Machine Models of Self-reproduction, *Proceedings of the Symposium in Applied Mathematics*, **14**, (1962), 17-33.

[23] Myhill J., The Converse to Moore's Garden-of-Eden Theorem, *Proceedings of the American Mathematical Society*, **14**, (1963), 685-686.