Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Theoretical Computer Science 349 (2005) 318–336

Theoretical Computer Science

www.elsevier.com/locate/tcs

Some results about the chaotic behavior of cellular automata ☆

F. Blanchard[a], J. Cervelle[b], E. Formenti[c,*]

[a] *Institut de Mathématique de Luminy, CNRS, Campus de Luminy, Case 907-13288 Marseille Cedex 9, France*
[b] *Institut Gaspard Monge, Université de Marne-la-Vallée, France*
[c] *Université de Nice-Sophia Antipolis, Laroratoire I3S, 2000, Route des lucioles, Sophia Antipolis, France*

Received 6 May 2004; received in revised form 17 June 2005; accepted 17 June 2005

Communicated by B. Durand

Abstract

We study the behavior of cellular automata (CA for short) in the Cantor, Besicovitch and Weyl topologies. We solve an open problem about the existence of transitive CA in the Besicovitch topology. The proof of this result has some interest of its own since it is obtained by using Kolmogorov complexity. To our knowledge it is the first result about discrete dynamical systems obtained using Kolmogorov complexity. We also prove that in the Besicovitch topology every CA has either a unique periodic point (thus a fixed point) or an uncountable set of periodic points. This result underlines the fact that CA have a great degree of stability; it may be considered a further step towards the understanding of CA periodic behavior.

Moreover, we prove that in the Besicovitch topology there is a special set of configurations, the set of Toeplitz configurations, that plays a role similar to that of spatially periodic configurations in the Cantor topology, that is, it is dense and has a central role in the study of surjectivity and injectivity. Finally, it is shown that the set of spatially quasi-periodic configurations is not dense in the Weyl topology.

Keywords: Cellular automata; Deterministic chaos; Unconventional topologies; Transitivity; Kolmogorov complexity

1. Introduction

In the last 20 years cellular automata (CA for short) received a growing attention as formal models for complex systems with applications in almost every scientific domain. They consist in an infinite lattice of identical finite automata. Each automaton updates its state according to a *local rule* on the basis of its present state and those of a finite set of neighboring automata. The states of all automata are updated synchronously. A *configuration* is a snapshot of all the states of all automata in the lattice.

The simplicity of the definition of this model is in contrast with the (perhaps only apparent) variety of dynamical behaviors, most of which are not completely understood yet.

The dynamical behavior of CA has been studied mainly in the context of discrete dynamical systems by endowing the set of configurations with the classical Cantor topology (i.e. the product topology when the set of states of the

☆ A short version of this paper was presented at MFCS'2002 conference, see [2].

* Corresponding author.

E-mail addresses: francois.blanchard@iml.univ-mrs.fr (F. Blanchard), julien.cervelle@univ-mlv.fr (J. Cervelle), enrico.formenti@unice.fr (E. Formenti).

0304-3975/$ - see front matter
doi:10.1016/j.tcs.2005.06.038

automaton is equipped with the discrete topology). Deterministic chaos is one of the most appealing features of dynamical behaviors. It is far from being completely understood. Among CA, one can find various interesting examples of this kind of behavior.

The problem is that in the Cantor topology the *shift* map is chaotic according to Devaney's popular definition of chaos [14]: a dynamical system is chaotic in this sense if it is transitive and regular (i.e. there exists a dense set of periodic points).

The shift map is a very simple CA; it just shifts the content of configurations one step to the left. It is a model of chaoticity in Information Theory and Ergodic Theory but, viewed as a CA, its chaoticity is somewhat counter-intuitive (see [9,3] for a discussion on this topic). In fact, the chaoticity of the shift is more an artefact of the Cantor topology, or of the Devaney definition of chaos, than a consequence of the intrinsic complexity of the automaton [15,10]. In [9], in order to avoid what they felt to be biasses of the Cantor topology (in the context of chaotic behavior), the authors proposed to replace it by the Besicovitch topology. In [18,3], the authors proved that this topology better links the classical notion of sensitivity to initials conditions with the intuitive notion of chaotic behavior. The Besicovitch topology deserves to be studied for at least one other reason: in statistical mechanics, the intuitive idea that two spatial configurations are close to each other translates easily into the assumption that their Besicovitch distance is small.

In the Cantor topology we study strong transitivity, a property that is strictly stronger than transitivity. We prove that strongly transitive CA are not injective. Recalling that, for CA, injectivity is equal to reversibility, this result underlines the fact that in strongly transitive systems there is a real global information loss. This means moreover that when substituting strong transitivity to transitivity in Devaney's definition of chaos the shift ceases to be chaotic (but in this case the shift on one-sided sequences is still chaotic). We also prove that all positively expansive CA are strongly transitive, one more step in the hierarchy of chaos [13].

In the case of the Besicovitch topology, we address three topics: stability, surjectivity and transitivity. Concerning stability, we prove that all CA have either a unique fixed point or an uncountable set of periodic points. This result points out that CA have a great degree of local stability. This fact is further stressed by the extension to the Besicovitch topology of a result in [5]: surjective CA having a blocking word are regular. Here we give a proof which does not use measure theory. This strengthens the conjecture that all surjective CA are regular [5,12,6].

Spatially periodic configurations, or rather their equivalence classes, are not dense in the Besicovitch topology. We study the relation between surjectivity and the wider set of Toeplitz configurations. We prove that in the Besicovitch topology it is dense, and with respect to surjectivity plays a role similar to that of the set of periodic configurations in the Cantor topology. Moreover, we prove that, exactly like in the Cantor case, injectivity implies surjectivity both for the global map and for its restriction to Toeplitz configurations.

Afterwards, we negatively answer the question of existence of transitive CA in the Besicovitch topology; this was called a challenging open problem in [23]. This result has deep implications for CA dynamics. First, it states that CA cannot change arbitrarily the density of differences between two configurations during their evolutions. In its turn this fact implies that the information contained in configurations cannot spread too much during evolutions. Second, it opens the quest for new, more appropriate, properties for describing the "complex" behavior of CA dynamics. Some very interesting proposals along this line of thought may be found in [24]. Third, the proof technique is of some interest of its own. We make use of the theory of Kolmogorov complexity and the famous incompressibility method [22] to prove a result of pure topological dynamics.

The paper ends with some results about similar problems in the Weyl topology. We first prove that the set of quasi-periodic configurations (containing Toeplitz sequences) is not dense in this topology. Then we address the question of finding an interesting dense set which could replace Toeplitz configurations in the Weyl setting, and prove that the set of non-generic configurations with respect to the uniform Bernoulli measure is dense.

2. Cellular automata

Formally, a CA is a quadruple $\langle d, S, N, \lambda\rangle$. The integer d is the *dimension* of the CA and controls how the cells of the lattice are indexed. Indeed, indexes of cells take values in $\mathbb{Z}^d$. The symbol S ($|S| < \infty$) is the finite set of states of cells and $\lambda\colon S^N \to S$ is the *local rule* which updates the state of a cell on the basis of a (finite) *neighborhood* $N \subset \mathbb{Z}^d$.

A *configuration* c is a function from $\mathbb{Z}^d$ to S and may be viewed as a snapshot of the content of each cell in the lattice. Denote by X the set $S^{\mathbb{Z}^d}$ of all configurations.

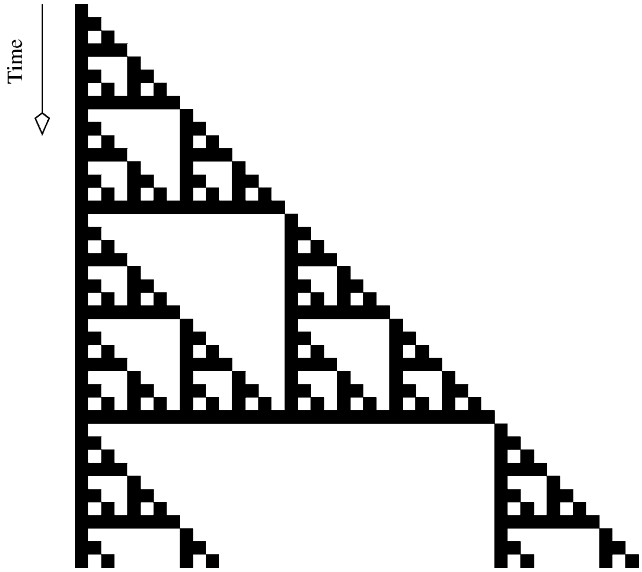

Fig. 1. A space–time diagram of a CA whose local rule computes the sum modulo 2 of the state of the current cell and that of its nearest neighbor. The initial configuration consists in a cell in state 1 (a black square) and all the other in state 0 (a white square).

The local rule induces naturally a *global rule* on the space of configurations $f : X \to X$ as follows

$$\forall c \in X \forall i \in \mathbb{Z}^d, \quad f(c)(i) = \lambda(c(i+n_1), \ldots, c(i+n_t)),$$

where $N = \{n_1, \ldots, n_t\}$ and $+$ is addition in $\mathbb{Z}^d$. When $d = 1$, one usually takes $N = \{-r, \ldots, 0, \ldots, r\}$ ($r \in \mathbb{N}$). In this case, r is called the *radius* of the CA.

In the sequel, when no misunderstanding is possible, we often make no distinction between a CA and its local rule.

Some sets of configurations such as finite and spatially periodic configurations play a special role in the study of CA dynamics. For $a \in S$, a configuration $c \in X$ is *a-finite* if it has a finite number of cells which are in a state different from a. In this paper, for the sake of simplicity we distinguish a special symbol $0 \in S$. Denote F the set of 0-finite configurations.

A configuration $c \in X$ is *spatially periodic* if there exists $p \in \mathbb{N}$ such that $\forall i \in \mathbb{Z},\ c_i = c_{p+i}$. The least p with the above property is the (*spatial*) *period* of c. Denote $\mathfrak{P}$ the set of spatially periodic configurations. A point with spatial period 1 is said to be *homogeneous*. A configuration in which all cells are in the same state $a \in S$ is both homogeneous and a-finite and it is denoted $\underline{a}$.

Another kind of periodicity—under the action of f— is considered in this article. A point $x \in X$ is *ultimately periodic* for f if there exist $p, t \in \mathbb{N}$ such that $\forall h \in \mathbb{N},\ f^{ph+t}(x) = f^t(x)$. When $t = 0$, x is called a *periodic point*. The least integer p with such a property is called the (*temporal*) *period* of x and t is its *preperiod*.

A CA is *injective* (resp. surjective) if its global function is injective (resp. surjective). Let $f|_A$ denote the restriction of the function f to the set A. A CA is *pre-injective* if $f|_F$ is injective.

In this paper, we restrict ourselves to one-dimensional CA ($d = 1$); some of the results are true in higher dimensions too.

For any configuration $c \in X$, $c_{a:b}$ is the word $c_a c_{a+1} \ldots c_b$ if $b \geqslant a$, ε (the empty word) otherwise. The pattern of size $2n+1$ centered at index 0 of a configuration $c \in X$ is denoted $\mathcal{M}_c(n) = c_{-n:n}$. For any word $w \in S^*$, $|w|$ denotes its length.

For any $c \in X$, the set $\mathcal{O}_f(c) = \{f^i(c),\ i \in \mathbb{N}\}$ is the *orbit* of initial condition c for f (we assume $f^0(c) = c$). The *space-time diagram* of initial configuration $c \in X$ is a graphical representation of $\mathcal{O}_f(c)$ and it is obtained by superposing the configurations $c, f(c), \ldots, f^n(c), \ldots$ (see Fig. 1). This representation is very useful for the visualization of CA simulations on a computer.

Our main interest is to study CA in the context of discrete dynamical systems i.e. structures $\langle U, F \rangle$, where U is a metric space and F a continuous function from U to itself. The natural equivalence in this category is called

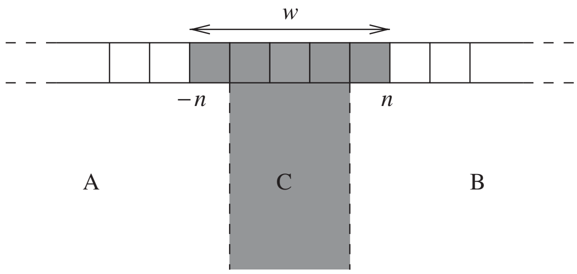

Fig. 2. Space–time diagram of a CA on an initial configuration centered on a blocking word w: the region C is determined by w, and the region A (resp. B) depends only on the coordinates to the left (resp. right) side of w in the initial configuration.

topological conjugacy: two dynamical systems $\langle U, F\rangle$ and $\langle V, G\rangle$ are said to be topologically conjugate if there exists a bi-continuous bijection $\Phi\colon U \to V$ such that $\Phi \circ F = G \circ \Phi$.

When S is endowed with the discrete topology and $X = S^{\mathbb{Z}^d}$ with the induced product topology, any global rule f of a CA can be considered as a discrete dynamical system $\langle X, f\rangle$. The product topology on X is usually called the *Cantor* topology since in this case X is a compact, totally disconnected and perfect space. For $d = 1$, one can easily verify that the following metric on configurations

$$\forall x, y \in X, \quad d(x, y) = 2^{-min\{|i|, x_i \neq y_i, i \in \mathbb{Z}\}}$$

induces exactly the Cantor topology on X. The definition for higher dimensions is similar.

For any word $w \in S^\star$ and $i \in \mathbb{Z}$, the cylinder $[w]_i$ starting at i is defined as $\left\{c \in X, c_{i:i+|w|-1} = w\right\}$; when the index i is not specified it means that $i = -\lfloor |w| /2\rfloor$. Cylinder sets form a basis of open (and closed) sets in the Cantor topology.

2.1. Cellular automata classifications

When examining space–time diagrams of CA, some seem more "complex" or "unordered", while others seem rather simple or predictable. Wolfram pioneered the quest for a classification of such dynamical behaviors, devising four classes [30]:

(i) all configurations evolve towards a homogeneous configuration;
(ii) configurations evolve to a set of periodic (w.r.t. time) configurations;
(iii) complex chaotic aperiodic evolutions;
(iv) complex evolutions with long-lived structures.

The main drawback of Wolfram's classification is that it is not precisely formalized. Many successive works in the field tried to find well-formalized interesting classifications, for example, see [11,7]. Clearly, "interesting" has a subjective meaning which strongly depends on one's own domain of study. In [20], with the help of classical notions of the theory of topological dynamics, Kůrka proposed a very interesting classification which discriminates the degree of equicontinuity of the system in the Cantor topology. In this section, after having introduced the property of equicontinuity and related notions, we try to refine Kůrka's classes.

For any point $x \in X$, $B_\delta(x)$ is the open ball of radius $\delta > 0$ centered at x. A system $\langle X, F\rangle$ is *equicontinuous at* $x \in X$ if for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all points $y \in B_\delta(x)$ and for all $n \in \mathbb{N}$, it holds that $d(F^n(x), F^n(y)) < \varepsilon$. We also say that x is an equicontinuity point for F.

A system $\langle X, F\rangle$ is *equicontinuous* if $\forall\varepsilon > 0 \forall x \in X \exists\delta > 0 \forall y \in B_\delta(x) \forall n \in \mathbb{N}, d(F^n(x), F^n(y)) < \varepsilon$.

In the case of CA, the existence of equicontinuity points is linked to the existence of *blocking words*. A word w of size $2k + 1$ is *blocking* for f if there exists an infinite sequence of words v_n of length $|v_n| = 2i + 1 \geqslant r$ (r is the radius of the automaton) and such that for any configuration c if $c_{-k:k} = w$ then $f^i(c)_{-i:i} = v_n$, for all $n \in \mathbb{N}$ (see Fig. 2).

Roughly speaking, a blocking word partitions the set of coordinates into two independent regions. In some sense, the system can be considered as the composition of two independent sub-systems.

The following well-known results describe the connections between blocking words and equicontinuity. All of them are essentially or explicitly in [20].

Proposition 1. *If x is an equicontinuity point then it contains an occurrence of a blocking word.*

Proposition 2. *If a CA has a blocking word then it has a dense set of equicontinuity points.*

At this point, we are ready to formally recall Kůrka's classification.

Definition 1 (*Kůrka classes*). A CA f belongs to

(i) *class* I if it is equicontinuous;
(ii) *class* II if it has equicontinuity points but is not equicontinuous;
(iii) *class* III or it is *sensitive to initial conditions* (or simply *sensitive*) if:

$$\exists \varepsilon > 0\, \forall x \in X\, \forall \delta > 0\, \exists y \in B_\delta(x)\, \exists n \in \mathbb{N}, \quad d(f^n(x), f^n(y)) > \varepsilon;$$

(iv) *class* IV or it is *positively expansive* if:

$$\exists \varepsilon > 0\, \forall x, y \in X\, \exists n,\, d(f^n(x), f^n(y)) > \varepsilon.$$

In compact dynamical systems there is a tight relationship between equicontinuity and equicontinuity points. Thus

Proposition 3. *A CA is equicontinuous if and only if it is equicontinuous at all points.*

There is a huge literature about CA in class IV, for instance, see [27,20,4] and therein references.

2.1.1. Refining class III

When looking at computer simulations one is immediately convinced that classes III and IV contain all (topologically) chaotic CA. This is essentially true; when restricted to certain subshifts (i.e. closed shift-invariant subsets of X), CA of class II may act chaotically but their action on random configurations never looks chaotic. The point is that under the common label of "chaotic dynamics" there are a variety of very different behaviors. For this reason and since class IV exhibits homogeneous topological properties, the next natural step is to try to refine class III. Let us introduce two more topological properties.

A CA f is *transitive* if and only if

$$\forall \varepsilon > 0\, \forall x, y \in X\, \exists x' \in B_\varepsilon(x)\, \exists y' \in B_\varepsilon(y)\, \exists n \in \mathbb{N}, \quad f^n(x') = y'.$$

Transitivity can be viewed as a property of indecomposability of the system, since from any open set one can reach every other open set in the phase space after a finite number of iterations. This notion can be strengthened by requiring the system to be able to "reach" not merely any open set but every possible point, starting from any open neighborhood. More formally, f is *strongly transitive* if and only if

$$\forall \varepsilon > 0\, \forall x, y \in X\, \exists x' \in B_\varepsilon(x)\, \exists n \in \mathbb{N}, \quad f^n(x') = y,$$

or equivalently if for any open set $U \neq \emptyset$ one has $\bigcup_{n \geqslant 0} f^n(U) = X$.

Both transitivity and strong transitivity are preserved under topological conjugacy.

The shift CA is transitive. The following result generalizes the almost obvious fact that it is not strongly transitive. Actually there is a more general, less elementary, result in topological dynamics: Auslander [1] remarked that if X is compact, a homeomorphism is strongly transitive if and only if it is minimal (i.e. every orbit is dense); and it is very easy to prove that a CA is never minimal (remark that, for $a \in S$, $\mathcal{O}(\underline{a})$ contains only homogeneous configurations).

Proposition 4. *A strongly transitive CA is not injective.*

Proof. Let A be a strongly transitive CA of global rule f. By the definition of cellular automata the finite closed set $\mathcal{H}$ of homogeneous configurations $\{\underline{a}, \underline{b}, \ldots\}$ is stable under $f : f(\mathcal{H}) \subset \mathcal{H}$. Let U be a non-empty open set

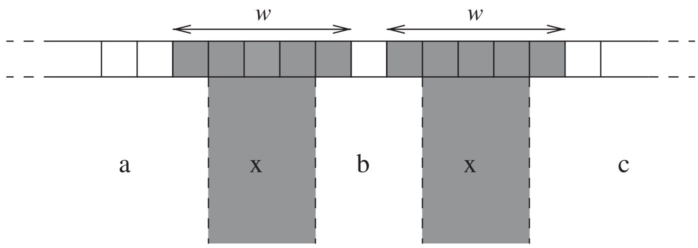

Fig. 3. Grayed areas labeled *X* are identical.

containing no homogeneous configurations, for instance $\mathcal{H}^c$. Applied to U, the strong transitivity assumption implies that $\mathcal{H} \subset \bigcup_{n \geqslant 0} f^n(U)$. As $U \cap \mathcal{H} = \emptyset$, a consequence is that there are an integer $n \in \mathbb{N}$, at least one point $\underline{d} \in \mathcal{H}$ and one point $x \in \mathcal{H}^c$ with $f^n(x) = \underline{d}$.

Now suppose that f is injective. Then $f_{|\mathcal{H}}$ must be injective too; this, together with stability, implies that $f(\mathcal{H}) = \mathcal{H}$ and hence $f_{|\mathcal{H}}$ is bijective. Therefore $\underline{d}$ has at least two points in its pre-image, $x \notin \mathcal{H}$ and one homogeneous configuration. This is a contradiction. $\square$

When the space of configurations is perfect (i.e. without isolated points), it is easy to see that positively expansive CA are sensitive to initial conditions. It is also clear that a strongly transitive CA is transitive. The next result is due to Kůrka [20]. Here we give a more “visual” proof.

Theorem 1. *A transitive CA is sensitive to initial conditions.*

Proof. For simplicity sake, assume that $\{0, 1\} \in S$. By contradiction, suppose that a CA f is transitive but not sensitive to initial conditions. As it is not sensitive, it has an equicontinuity point and therefore a blocking word w. Consider the cylinder $[w0w]$, where 0 is a state of the CA. As w is blocking, all the words of size $2|w| + 1$, located in the center of the configurations of $f^n([w0w])$, $n \in \mathbb{N}$ is of the type *axbxc*, where a, b, c and x are words over S. Remark that $0 < |x| \leqslant |w|$. Moreover, the size pf x depends on n but not on the initial configuration choosen in $[w0w]$ (Fig. 3). Thus, starting from any configuration $c \in [w0w]$ it is impossible to reach the cylinder $[1^{|w|+1}00^{|w|} + 1]$ after any number of iterations. Since cylinders are open sets, f is not transitive. $\square$

Theorem 2. *A positively expansive CA is strongly transitive.*

Proof. Let f be an expansive CA with state set S. In [27], Nasu proved that f is topologically conjugate to a certain one-sided subshift $\langle X = Q^{\mathbb{N}}, \sigma\rangle$. In [20], Kurka proved that expansive CA are open. Now, in [28], Parry proved that open subshifts are SFT. In [4], Blanchard and Maass proved that $\langle X = Q^{\mathbb{N}}, \sigma\rangle$ is topologically mixing. Finally, in [28], Parry proved that open transitive SFT are necessarily strong transitive. $\square$

3. The Besicovitch topology

The topology with which the space of a dynamical system is endowed plays a fundamental role in the study of its asymptotic behavior. It can filter out some intrinsic behaviors and change the meaning of chaoticity. The Besicovitch topology was introduced in the context of symbolic dynamics in [9] in order to refine the study of sensitivity to initial conditions. In contrast with the Cantor topology, it neglects errors near the cell of index zero by giving the same weight to all cells.

Let $d_{\mathcal{B}}$ be the following function:

$$d_{\mathcal{B}}(x, y) = \limsup_{n\to\infty} \frac{\Delta(x_{-n:n}, y_{-n:n})}{2n+1},$$

where $\Delta(x_{-n:n}, y_{-n:n})$ is the number of indexes which the words $x_{-n:n}$ and $y_{-n:n}$ differ at.

The function $d_{\mathcal{B}}$ is a *pseudo-distance* [9]; it is called the *Besicovitch pseudo-distance*. The Besicovitch topology is not defined on $X = S^{\mathbb{Z}}$, but on the quotient space $\dot{X}$ with respect to the equivalence relation "being at zero $d_{\mathcal{B}}$-pseudo distance". Denoting this relation by $\doteq$, one has $\dot{X} = X/\doteq$. Call $\dot{x}$ the equivalence class of $x \in X$: the formula $d_{\mathcal{B}}(\dot{x}, \dot{y}) = d_{\mathcal{B}}(x, y)$ defines a true metric on equivalence classes. In the sequel, when no misunderstanding is possible, we will denote $d_{\mathcal{B}}$ simply by d.

This topology is suitable for the study of CA: it is perfect, complete, path-wise connected and infinite dimensional; unfortunately, it is neither locally compact nor separable (see [9,3]). The following proposition allows us to induce, from a CA global map from X to itself, a self-map on $\dot{X}$.

Proposition 5 (*Cattaneo et al. [9]*). *Any CA f on X is compatible with the equivalence relation $\doteq$, that is, $x \doteq y \implies f(x) \doteq f(y)$.*

For any CA f, denote $\dot{f}$ the map which transforms a class c of $\dot{X}$ to the class of all the images under f of the configurations of c. The map $\dot{f}$ is called a *CA on (Besicovitch) classes*.

3.1. Fixed and periodic points

Some recent articles pointed out that CA are not complex from the point of view of algorithmic complexity [15,8,10]. This fact seems to originate from an intrinsic stability of such systems. It is well-known that in the Cantor topology, if a CA has a non-homogeneous fixed point then it has at least a countable set of fixed points. The Besicovitch topology allows to go even further. In this section, we prove that either a CA has one unique fixed point or it has uncountably many periodic points. Clearly these "new" periodic points are due to the special structure of the Besicovitch space but if we analyze in more detail how they are built one can see that they are made up of larger and larger areas in which the configuration remains locally periodic for a long time.

Before introducing the main result of this section (Theorem 3) we need some technical lemmas and notation. If p is an integer and u a word of size at least $2p$, then ${}_p|u|_p$ is the word $u_{p:|u|-p-1}$, that is, the word u in which the first and last p letters are deleted. For any pair of integers l, h ($l < h$), let $[[l, h]]$ be the set of integers between l and h.

Lemma 1. *Consider a CA of global rule f and radius r. Let $(a_i)_{i\in[[1,h]]}$ be a finite sequence of h words of length greater than $2r$. Let $x = a_1 \ldots a_h$ be the concatenation of these words. Then*

$$\Delta({}_r|x|_r, f(x)) \leqslant \sum_{i=1}^{h} \Delta({}_r|a_i|_r, f(a_i)) + 2r(h-1).$$

Proof. First let a and b be two words of length bigger than $2r$. Then

$$\Delta({}_r|ab|_r, f(ab)) \leqslant \Delta({}_r|b|_r, f(b)) + \Delta({}_r|a|_r, f(a)) + 2r,$$

because the image of the concatenation of two words is the concatenation of the images of these words separated by $2r$ cells of perturbation. The inequality in the statement is obtained by an induction on the number of concatenated words. □

We will also make use of the well-known Cesàro lemma on series.

Lemma 2 (*The Cesàro lemma*). *Let $(a_n)_{n\in\mathbb{N}}$ and $(u_n)_{n\in\mathbb{N}}$ be two sequences of reals such that*

$$\lim_{n\to\infty} \frac{u_n}{a_n} = l.$$

Moreover, assume that (a_n) is divergent and positive. Then

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n} u_i}{\sum_{i=0}^{n} a_i} = l.$$

Proposition 6. *In the Besicovitch space, any CA with two distinct periodic points of periods* p_1 *and* p_2 *has uncountably many periodic points of period* $\mathrm{lcm}(p_1, p_2)$.

Proof. Let f be a CA with radius r on Besicovitch classes. Let $\dot{x}$ and $\dot{y}$ be two distinct periodic points for f, with periods p_1 and p_2, respectively. Let $p = \mathrm{lcm}(p_1, p_2)$. Let x be a member of the class $\dot{x}$ and y a member of the class $\dot{y}$. Since $\dot{x}$ and $\dot{y}$ are distinct, $d(x, y) = \delta > 0$. Hence, there exists a sequence of integers $(u_n)_{n\in\mathbb{N}}$ such that

$$\forall n > 0, \quad \Delta(x_{-u_n:u_n}, y_{-u_n:u_n}) \geqslant \delta u_n. \tag{1}$$

Let σ be an increasing function on integers such that $u_{\sigma(0)} > 4r$ and $u_{\sigma(n+1)} > 2u_{\sigma(n)}$. Let $v_n = u_{\sigma(n)}$. By a simple recurrence on n one has

$$\sum_{i=0}^{n} v_i \leqslant 2v_n. \tag{2}$$

We construct an injection g from $\{0, 1\}^{\mathbb{Z}}$ into X (not $\dot{X}$) such that the classes of the image set of g are all periodic points of f. Let α be a sequence of $\{0, 1\}^{\mathbb{N}}$. For all non-negative integer i, define the sequences k^α and k'^α as follows:

$$k_i^\alpha = \begin{cases} x_{1:v_i} & \text{if } \alpha_i = 0, \\ y_{1:v_i} & \text{if } \alpha_i = 1, \end{cases}$$

and

$$k_i'^\alpha = \begin{cases} x_{-v_i:-1} & \text{if } \alpha_i = 0, \\ y_{-v_i:-1} & \text{if } \alpha_i = 1. \end{cases}$$

Define the configuration $g(\alpha)$ as follows:

$$g(\alpha) = \ldots k_n'^\alpha k_{n-1}'^\alpha \ldots k_2'^\alpha k_1'^\alpha k_0'^\alpha 0^{2rp+1} k_0^\alpha k_1^\alpha k_2^\alpha \ldots k_{n-1}^\alpha k_n^\alpha \ldots .$$

Let us prove that the class containing $g(\alpha)$ is a periodic point for f with period p, i.e. $d(g(\alpha), f^p(g(\alpha))) = 0$. One has to prove that

$$\limsup_{n\to\infty} \frac{\Delta(g(\alpha)_{-n:n}, f^p(g(\alpha))_{-n:n})}{2n+1} = 0.$$

As x and y are two periodic points with periods p_1 and p_2, which are both divisors of p, one has $d(x, f^p(x)) = d(y, f^p(y)) = 0$. Thus

$$\lim_{n\to\infty} \frac{\Delta(x_{-n:n}, f^p(x)_{-n:n})}{2n+1} = \lim_{n\to\infty} \frac{\Delta(y_{-n:n}, f^p(y)_{-n:n})}{2n+1} = 0.$$

This implies that

$$\lim_{n\to\infty} \frac{\Delta(x_{1:n}, f^p(x)_{1:n})}{n} = \lim_{n\to\infty} \frac{\Delta(y_{1:n}, f^p(y)_{1:n})}{n} = 0. \tag{3}$$

Fix an integer $n > r$. Decompose $g(\alpha)_{-n:n}$ into its k_i^α and $k_i'^\alpha$ factors

$$g(\alpha)_{-n:n} = \overrightarrow{k_h'^\alpha} k_{h-1}'^\alpha \ldots k_2'^\alpha k_1'^\alpha k_0'^\alpha 0^{2rp+1} k_0^\alpha k_1^\alpha k_2^\alpha \ldots k_{h-1}^\alpha \overleftarrow{k_h^\alpha}$$

where h depends on n, and $\overrightarrow{k_h'^\alpha}$ [resp. $\overleftarrow{k_h^\alpha}$] is the suitable suffix [resp. prefix] of $k_h'^\alpha$ [resp. k_h^α] of $g(\alpha)_{-n:n}$ (Fig. 4).

Since f^p is a CA of radius rp, Lemma 1 implies that

$$\begin{aligned} \Delta(g(\alpha)_{-n:n}, f^p(g(\alpha))_{-n:n}) \;\leqslant\; & \Delta(0, f^p(0^{2rp+1})) + \sum_{i=0}^{h-1} \Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha)) \\ & + \sum_{i=0}^{h-1} \Delta({}_{rp}|k_i'^\alpha|_{rp}, f^p(k_i'^\alpha)) + \Delta({}_{rp}|\overrightarrow{k_h'^\alpha}|_{rp}, f^p(\overrightarrow{k_h'^\alpha})) \\ & + \Delta({}_{rp}|\overleftarrow{k_h^\alpha}|_{rp}, f^p(\overleftarrow{k_h^\alpha})) + 2rp(2h+2). \end{aligned}$$

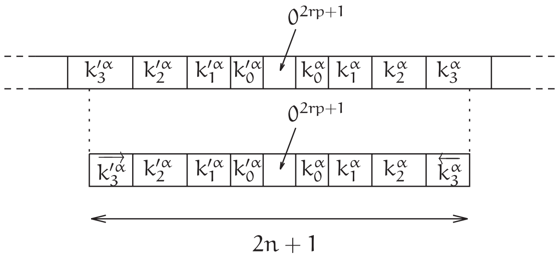

Fig. 4. The decomposition of $g(\alpha)$; here $h = 3$.

By the definition of k_i^α, one has that $\overleftarrow{k_h^\alpha}$ is a prefix of $x'_{1:\infty}$ or of $y'_{1:\infty}$. Hence

$$\Delta({}_{rp}|\overleftarrow{k_h^\alpha}|_{rp}, f^p(\overleftarrow{k_h^\alpha})) \leqslant \Delta({}_{rp}|x_{1:n}|_{rp}, f^p(x_{1:n})) + \Delta({}_{rp}|y_{1:n}|_{rp}, f^p(y_{1:n})).$$

Similarly one finds

$$\Delta({}_{rp}|\overrightarrow{k_h'^\alpha}|_{rp}, f^p(\overrightarrow{k_h'^\alpha})) \leqslant \Delta({}_{rp}|x_{-n:-1}|_{rp}, f^p(x_{-n:-1})) + \Delta({}_{rp}|y_{-n:-1}|_{rp}, f^p(y_{-n:-1})).$$

Then one has

$$\begin{aligned}\Delta(g(\alpha)_{-n:n}, f^p(g(\alpha))_{-n:n}) \leqslant\ & 1 + \sum_{i=0}^{h-1} \Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha)) + \sum_{i=0}^{h-1} \Delta({}_{rp}|k_i'^\alpha|_{rp}, f^p(k_i'^\alpha)) \\ & + \Delta({}_{rp}|x_{-n:n}|_{rp}, f^p(x_{-n:n})) + \Delta({}_{rp}|y_{-n:n}|_{rp}, f^p(y_{-n:n})) \\ & + 4rp(h+1). \end{aligned} \tag{4}$$

With the help of Eq. (3) one deduces from (4) that

$$\begin{aligned}\lim_{\substack{i\to\infty\\ \alpha_i=0}} \frac{\Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha))}{v_i} &\leqslant \lim_{i\to\infty} \frac{\Delta({}_{rp}|x_{-v_i:v_i}|_{rp}, f^p(x_{-v_i:v_i}))}{v_i} \\ &= \lim_{n\to\infty} \frac{\Delta(x_{-n:n}, f^p(x)_{-n:n})}{n} = 0.\end{aligned}$$

And similarly

$$\lim_{\substack{i\to\infty\\ \alpha_i=1}} \frac{\Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha))}{v_i} = 0.$$

Summing up the two previous equations one obtains

$$\lim_{i\to\infty} \frac{\Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha))}{v_i} = 0.$$

By the Cesàro lemma (Lemma 2) one has

$$\lim_{h\to\infty} \frac{\sum_{i=0}^{h-1} \Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha))}{\sum_{i=0}^{h-1} v_i} = 0$$

since the sequence $(v_i)_{i\in\mathbb{N}}$ is positive and divergent. The same argument applies to k', and therefore one has

$$\lim_{h\to\infty} \frac{\sum_{i=0}^{h-1} \Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha))}{\sum_{i=0}^{h-1} v_i} + \frac{\sum_{i=0}^{h-1} \Delta({}_{rp}|k_i'^\alpha|_{rp}, f^p(k_i'^\alpha))}{\sum_{i=0}^{h-1} v_i} = 0.$$

Since $\sum_{i=0}^{h-1} v_i \leqslant 2n$, one gets

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n} \Delta({}_{rp}|k_i^\alpha|_{rp}, f^p(k_i^\alpha))}{2n+1} + \frac{\sum_{i=0}^{n} \Delta({}_{rp}|k_i'^\alpha|_{rp}, f^p(k_i'^\alpha))}{2n+1} = 0. \tag{5}$$

Using Eq. (3) again, we obtain that

$$\lim_{n\to\infty} \frac{\Delta({}_{rp}|x_{-n:n}|_{rp}, f^p(x_{-n:n}))}{2n+1} = \lim_{n\to\infty} \frac{\Delta({}_{rp}|y_{-n:n}|_{rp}, f^p(y_{-n:n}))}{2n+1} = 0. \tag{6}$$

Finally

$$\lim_{n\to\infty} \frac{4rp(h+1)}{2n+1} = 0 \tag{7}$$

since $h \leqslant \ln_2(n)$.

Using Eqs. (5)–(7) inside Eq. (4), one finds that

$$\lim_{n\to\infty} \frac{\Delta(g(\alpha)_{-n:n}, f^p(g(\alpha))_{-n:n})}{2n+1} = 0$$

which implies that $g(\alpha)$ is a periodic point of f of period p.

Let $\sim \subseteq \{0,1\}^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}}$ be the equivalence relation such that $x \sim y$ if and only if y and x differ only at a finite number of positions, and let $\nsim$ be the opposite relation. Let α and β be two sequences of $\{0,1\}^{\mathbb{N}}$ such that $\alpha \nsim \beta$. Let $(a_i)_{i\in\mathbb{N}}$ be the increasing sequence of positions where they differ. Then

$$\begin{aligned} d(g(\alpha), g(\beta)) &\geqslant \limsup_{n\in\mathbb{N}} \frac{\Delta(g(\alpha)_{-n:n}, g(\beta)_{-n:n})}{2n+1} \\ &\geqslant \limsup_{n\in\mathbb{N}} \frac{\Delta(k^{\alpha}_{a_n}, k^{\beta}_{a_n}) + \Delta(k'^{\alpha}_{a_n}, k'^{\beta}_{a_n})}{2\sum_{i=0}^{a_n} v_i + 2rp + 1} \\ &\geqslant \limsup_{n\in\mathbb{N}} \frac{\Delta(x_{-v_{a_n}:v_{a_n}}, y_{-v_{a_n}:v_{a_n}})}{2\sum_{i=0}^{a_n} v_i + 2rp + 1}. \end{aligned}$$

Applying Eqs. (1) and (2), we obtain

$$\begin{aligned} d(g(\alpha), g(\beta)) &\geqslant \limsup_{n\in\mathbb{N}} \frac{\delta v_{a_n}}{2v_{a_n} + 2rp + 1} \\ &\geqslant \frac{\delta}{2}. \end{aligned}$$

This shows that $g(\alpha)$ and $g(\beta)$ are in different classes.

Finally, we have proved that all configurations in $g(\{0,1\}^{\mathbb{N}})$ are periodic points of period p, and that $\alpha \nsim \beta \Longrightarrow g(\alpha) \not\doteq g(\beta)$.

Let E be a set containing a member of each equivalence class of $\sim$. Since $\{0,1\}^{\mathbb{N}}/\sim$ is not countable, neither is E. By the previous implication $g|_E$ is injective so that $g(E)$ is a non-countable set of periodic points of f of period p. □

This result has two easy corollaries. The first one is obtained by simply recalling that a fixed point is a periodic point of period 1. The second one comes from the fact that if p is a periodic point of period greater than or equal to 2, then there are at least two distinct periodic points.

Corollary 1. *In the Besicovitch topology, if a CA f has two fixed points then it has uncountably many fixed points.*

Corollary 2. *In the Besicovitch topology, if a CA has a periodic point of period* $p > 1$ *then its set of periodic points is not countable.*

Putting together the results of the two previous corollaries we obtain the main result of this section.

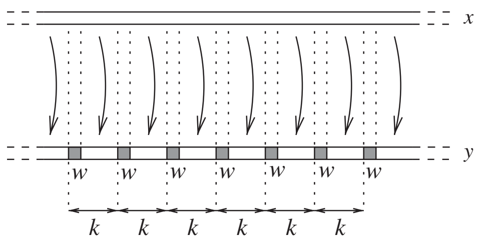

Fig. 5. Construction of y from x and $k = \lfloor 2\varepsilon^{-1}|w| \rfloor$.

Theorem 3. *In the Besicovitch topology, any CA has exactly one of these two properties:*
(i) *a unique fixed point and no other periodic point;*
(ii) *uncountably many periodic points.*

Proof. Let f be a CA. There are several cases. First, f can have one fixed point as its unique periodic point. Second, if f has two fixed points, then by Corollary 1 it has an uncountable set of periodic points (in this case fixed points). Finally assume that f has no fixed points; since for the finite set $\mathcal{H}$ of homogeneous points one has $f(\mathcal{H}) \subset \mathcal{H}$, it must contain at least one periodic point of period at least 2; then by Corollary 2 there are uncountably many periodic points for f. $\square$

Proposition 7. *If a surjective CA has a blocking word (or, equivalently, an equicontinuity point for the Cantor topology), then its set of periodic points is dense in the Besicovitch topology.*

Remark that the same argument can be used to prove a similar result for the Cantor topology, without making use of measure theory as in the first proof [5].

Proof. Let f be a CA and w a blocking word for f. One has to prove that for any configuration x and any real number $\varepsilon > 0$ there exists a periodic configuration at Besicovitch distance less than ε from x. Let y be the following configuration

$$\forall n, l \in \mathbb{N}, \quad y_{nk+l} = \begin{cases} w_{l \bmod k} & \text{if } l \bmod k < |w|, \\ x_{nk+l} & \text{otherwise,} \end{cases}$$

where $k = \lfloor 2\varepsilon^{-1}|w| \rfloor$. The configuration y is everywhere equal to x except that periodically words of equal length are replaced by w (Fig. 5).

The number of differences between $x_{-n:n}$ and $y_{-n:n}$ is bounded by $|w|$ times the number of occurrences of w within $y_{-n:n}$. Thus

$$\begin{aligned} d(x, y) &\leqslant \lim_{n\to\infty} \frac{\Delta(x_{-n:n}, y_{-n:n})}{2n+1} \\ &\leqslant \lim_{n\to\infty} \frac{\left\lceil \frac{2n+1}{k} \right\rceil |w|}{2n+1} \\ &\leqslant \lim_{n\to\infty} \frac{n\varepsilon + |w|}{2n+1} \\ &< \varepsilon. \end{aligned}$$

Now we prove that y is a periodic point of f. Since w is a blocking word, $\forall i, n \in \mathbb{N}$, the pattern $f^{i+1}(y)_{nk+\lceil |w|/2 \rceil : n(k+1)+\lfloor |w|/2 \rfloor}$ depends only on the corresponding word of the same size in the pre-image $f^{i}(y)_{nk+\lceil |w|/2 \rceil : n(k+1)+\lfloor |w|/2 \rfloor}$. This means that values of cells in a grayed zone in Fig. 6 do not dependent on the value of any cell in any other grayed zone. For all integers i and n, define

$$u_n^{(i)} = f^{i}(y)_{nk+\lceil |w|/2 \rceil : n(k+1)+\lfloor |w|/2 \rfloor}.$$

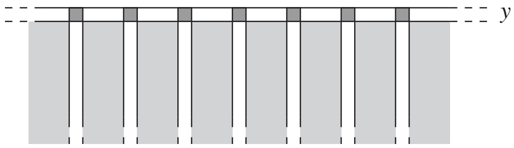

Fig. 6. Construction of periodic equicontinuous points (see Proposition 7). Values of cells in one grayed zone do not depend on valued on other grayed zones.

For any fixed n, the sequence $(u_n^{(i)})_{i\in\mathbb{N}}$ is of finite values and each term depends only on its predecessor. Hence, it is ultimately periodic. By the hypothesis the CA is surjective, this implies that y is periodic (for if the sequence is ultimately periodic of period p with preperiod k, then $u_n^{(k)}$ has two pre-images: $u_n^{(k-1)}$ and $u_n^{(k+p-1)}$, which is impossible since a surjective CA is pre-injective [25,26]). Since the number of possible values for this sequence is 2^k, the configuration is periodic and its period is at most $2^{k(k+1)/2}$.

Finally, remark that if the CA is not quiescent (i.e. if there is not a state s such that $\lambda(s, s, \ldots, s) = s$) then one should consider $f^{|S|}$ in place of f in the above proof. Of course, when f is surjective then $f^{|S|}$ is surjective and if $f^{|S|}$ has a dense set of periodic points then f has a dense set of periodic points too. □

The next result needs some more notions from topological dynamics. Let $\langle X, F\rangle$ be a compact dynamical system. Consider a F-invariant Borel measure μ on X. A configuration $x \in X$ is said to be *generic* for $\langle X, F\rangle$ endowed with the measure μ if for all continuous functions $\varphi : X \to \mathbb{R}$ the convergence

$$\frac{\varphi(x) + \varphi(F(x)) + \cdots + \varphi(F^n(x))}{n+1} \underset{n\to\infty}{\longrightarrow} \int \varphi \,\mathrm{d}\mu$$

holds. If F is invertible then one can consider also the negative orbit and the previous formula becomes

$$\frac{\sum_{i=-n}^{n} \varphi(F^i(x))}{2n+1} \underset{n\to\infty}{\longrightarrow} \int \varphi \,\mathrm{d}\mu.$$

In the Besicovitch topology, the shift map is an isometry. Hence, it is not sensitive to initial conditions. The following proposition proves that it is not even regular.

Proposition 8. *The shift map is not regular in the Besicovitch topology.*

Proof. By contradiction, assume that the shift map is regular in the Besicovitch topology. Then, for all configurations x and for all $\varepsilon > 0$, there exists $k \in \mathbb{N}$ such that $d_B(x, \sigma^k(x)) < \varepsilon$ (it is enough to take a periodic point of σ y of period k such that $d_B(x, y) < \varepsilon/2$ and apply the fact that σ is an isometry in the Besicovitch topology).

Now, consider a configuration z that is generic for the uniform measure on S. For any integer n, let $\varphi_n : S^{\mathbb{Z}} \to \mathbb{R}$ be such that $\forall c \in S^{\mathbb{Z}},\ \varphi_n(c) = 1$, if $c_0 \neq c_n$ and 0 otherwise. It is clear that φ_n is continous (in the Cantor topology); in fact, if two configurations are equal in on some central big enough segment then their images are equal. By applying the genericity property to z and φ_n one obtains that $d_B(z, \sigma^n(z)) = 1/|S|$. Hence σ is not regular in the Besicovitch topology. □

3.2. Toeplitz sequences

In the Cantor topology, spatially periodic configurations play a special role. They form a dense positively invariant set. In some cases this property permits to prove results about CA global behavior by simply studying the dynamics on $\mathfrak{P}$ (see for instance [19,6,17]). This is often a very convenient situation since a configuration in $\mathfrak{P}$ has a finite description.

In the Besicovitch topology $\mathfrak{P}$ is not dense: here the problem is to find a set which plays a role similar to that of $\mathfrak{P}$ in the Cantor topology. Clearly this set should be dense, positively invariant and its configurations should have an easy representation scheme. In this section we prove that the set of Toeplitz configurations, which contains $\mathfrak{P}$, is a good candidate.

A configuration x is *Toeplitz* if for all $i \in \mathbb{Z}$, there exists a (spatial) period p_i such that $\forall k \in \mathbb{Z}$, $x_{p_i k+i} = x_i$. Denote by $\mathscr{T}$ the set of all Toeplitz configurations. Note that any finite pattern in a Toeplitz configuration repeats periodically; its period is given by the least common multiple of the periods of cells that compose the pattern. Obviously $\mathfrak{P} \subset \mathscr{T}$, a fact that is used repeatedly in this subsection.

The following lemma states that there is at most one Toeplitz configuration per class. Some classes do not contain any.

Lemma 3 (*Blanchard et al. [3]*). *The restriction to Toeplitz configurations of the canonical surjection Π of X to $\dot{X}$ is injective.*

Lemma 4. *The set $\mathscr{T}$ is f-positively invariant, i.e. $f(\mathscr{T}) \subseteq \mathscr{T}$.*

Proof. Let $x \in \mathscr{T}$ and f be a CA with radius r and local rule λ. For all i there exists p_i such that forall $k \in \mathbb{Z}$ it holds that $x_{p_i k+i} = x_i$. Let $p'_i = \mathrm{lcm}_{-r \leqslant l \leqslant r}\{p_{i+l}\}$. Then

$$f(x)_i = \lambda(x_{i-r}, \ldots, x_{i+r}) = \lambda(x_{p'_i k+i-r}, \ldots, x_{p'_i k+i+r}) = f(x)_{p'_i k+i}.$$

Therefore $f(x)$ is a Toeplitz configuration. $\square$

Proposition 9. *The set $\mathscr{T}$ is dense in the Besicovitch topology.*

Proof. Let x be a configuration and $\varepsilon > 0$. We build a Toeplitz configuration y such that $d(x, y) < \varepsilon$.

Let $(n_k)_{k\in\mathbb{N}}$ be a sequence of integers satisfying the following constraints:

(i) for all couples of integers p and q such that $p < q$, n_p is a divisor of n_q;

(ii) $\sum_{i=0}^{\infty} n_i^{-1}$ converges, and the limit is less than ε.

The construction of y is recursive. The initialization step sets, for all $i \in \mathbb{Z}$, $y_{in_0} = x_0$. Then, for all integers k increasing from 1 to infinity, we apply the two following steps:

Step $k - \frac{1}{2}$: if y_k has not already been set, define, for all $i \in \mathbb{Z}$, $y_k = y_{k+in_{2k-1}} = x_k$;

Step k: if y_{-k} has not already been set, set for all $i \in \mathbb{Z}$, $y_{-k} = y_{-k+in_{2k}} = x_{-k}$.

For all $k \in \mathbb{N}$, let $u_k = -k$ and $u_{k-\frac{1}{2}} = k$. The position u_k is the target of the test of step k.

First, one has to prove that the construction is correct, that is to say, no cell of y is assigned a value twice (otherwise y will not be a Toeplitz sequence). By contradiction, suppose that the cell at position m is assigned a value twice, first at step a, and then at step b, $a < b$. This means that there are two integers k_1 and k_2 such that $m = u_a + k_1 n_{2a} = u_b + k_2 n_{2b}$. Condition (i) of the hypothesis on $(n_k)_{k\in\mathbb{N}}$ tells that there exists an integer p such that $n_{2b} = pn_{2a}$, and so $u_b = u_a + (k_1 - pk_2)n_{2a}$. Thus, the position u_b was given a value at step a, and hence was already fixed at step b, contradicting the assumption that an assignment was made at this step. By construction, all positions are assigned a value at some step. The configuration y is in $\mathscr{T}$; in fact each position has the period n_{2a}, if a is the step at which it was defined.

There remains to prove that $d(x, y) < \varepsilon$. To this purpose, we need to find an upper bound for $N(k)$, the number of positions at which $y_{-k:k}$ and $x_{-k:k}$ differ. This number is less than the number of integers b, $-k < u_b < k$, such that the position u_b was already assigned a value at step b: otherwise $x_{u_b} = y_{u_b}$. There are at most $\lfloor (2k+1)/n_{2b} \rfloor + 1$ positions between $-k$ and k already fixed at step b. Since, at this step, we set $x_{u_b} = y_{u_b}$, the number of cells of y which are set to a value different from that of the corresponding cell in x is at most $\lfloor (2k+1)/n_{2b} \rfloor$. Thus by hypothesis (ii)

$$N(k) \leqslant \sum_{i\in\mathbb{N}} \left\lfloor \frac{2k+1}{n_i} \right\rfloor \leqslant \sum_{i\in\mathbb{N}} \frac{2k+1}{n_i} \leqslant (2k+1) \sum_{i\in\mathbb{N}} n_i^{-1} \leqslant (2k+1)\varepsilon.$$

Dividing this inequality by $2k+1$ and taking the lim sup when k goes to infinity, one concludes that $d(x, y) \leqslant \varepsilon$. $\square$

Proposition 10 (*Blanchard et al. [3]*). *Let f be the global rule of a CA. Then f is surjective if and only if $\dot{f}$ is surjective.*

An easy consequence is

Proposition 11. *Let f be the global rule of a CA. Then* $\dot{f}$ *is surjective if* $\dot{f}|_T$ *is surjective.*

Proof. Consider a spatial periodic configuration $\dot{x}$. Since $\dot{f}_{|T}$ is surjective, there exists a non-empty set Y of Toeplitz configurations such that $\forall \dot{y} \in Y$, $\dot{f}(\dot{y}) = \dot{x}$. We claim that Y contains at least a spatial periodic configuration. Let x be the spatial periodic representative in $\dot{x}$ and p be its period. Let U be the set of blocks that are possible preimages of $x_{0:p-1}$ that are contained in configurations in Y. Since all the blocks $x_{i:ip-1}$ are equal and U is finite, one can build a spatial periodic configuration y such that $f(y) = x$. As a consequence, the restriction of f to spatial periodic configurations is surjective. It is well-known that this last property is equivalent to surjectivity (at least for 1-dimensional CA) [16]. An application of Proposition 10 finishes the proof. □

Proposition 12. *Let f be the global rule of a CA. If* $\dot{f}$ *is injective then it is surjective too.*

Proof. If $\dot{f}$ is injective then the restriction of f to spatially periodic configurations is injective too. In fact, if x and y are two different spatially periodic configurations then, by Lemma 3, $\dot{x}$ and $\dot{y}$ are distinct. This implies that $f|_F$ is injective, so that f is surjective (for a proof see [16]). Finally, using Proposition 10, we have the thesis. □

Proposition 13. *Let f be the global rule of a CA. If* $\dot{f}|_T$ *is injective then* $\dot{f}|_T$ *is surjective.*

Proof. If $\dot{f}|_T$ is injective $f|_{\mathfrak{P}}$ is injective too since there is only one spatial periodic configuration per class; then $f|_F$ is injective. By the Moore–Myhill theorem [25,26], if $f|_F$ is injective then f is surjective and, by Proposition 11, $\dot{f}|_T$ is surjective. □

Exactly as in the Cantor case, the converse of Proposition 12 is false; in fact elementary rule 90 is surjective but not injective.

As we have already pointed out one can prevent the shift from being chaotic in two ways: 1. In the Cantor topology, by including an assumption of strong transitivity in the definition of chaos; 2. By adopting the Besicovitch topology. These two methods are so different that they cannot be expected to filter out chaoticity for the same set of CA. For instance by Proposition 4 we know that, along with σ, all invertible CA are removed from chaos by the first method. This is not clearly the case with the second method: the next example shows that there exist invertible CA that are sensitive in the Besicovitch topology.

Example 1. Let $X = \{0, 1\}^{\mathbb{Z}}$ and d_C (resp. d_B) be the Cantor (resp. Besicovitch) distance.

Consider the CA on $X \times X$ whose global rule f is defined as follows:

$$\forall x, y \in X, \quad f((x, y)) = (\sigma(x), \sigma(x) \oplus y),$$

where $\oplus$ is the bit-wise addition on $\mathbb{Z}_2$. One can easily verify that f has the following inverse:

$$\forall x, y \in X, \quad f^{-1}((x, y)) = (\sigma^{-1}(x), x \oplus y).$$

By induction on $k \in \mathbb{N}$, it is easy to prove that

$$\forall x, y \in X, \quad \forall k \in \mathbb{N}, \quad f^k((x, y)) = (\sigma^k(x), \sigma^k(x) \oplus \sigma^{k-1}(x) \oplus \cdots \oplus \sigma(x) \oplus y). \tag{8}$$

We claim that f is sensitive to initial conditions in Besicovitch topology with sensitivity constant equal to $\frac{1}{2}$. Let $x, y \in X$, for any $\delta > 0$, let k be the smallest integer such that $1/k \leqslant \delta$. Then build a configuration x' as follows: $\forall i \in \mathbb{Z}$, $x'(i) = 1 \oplus x_i$ if $i \equiv 0 \bmod k$, x_i otherwise. It is clear that $d_B((x, y), (x', y)) \leqslant \delta$. Using Eq. (8), it is not difficult to see that for any $z \in X$, $d_B(f^k((x, z)), f^k((x', z))) = 1/2$.

Remark that, by Eq. (8), for any $x, y \in X$, (x, y) is periodic for f if and only if x is periodic for σ. By Theorem 3, f has uncountably many periodic points in Besicovitch topology, but by Proposition 8, it is easy to see that f is not regular. □

3.3. Transitive cellular automata

As already recalled the Besicovitch topology was introduced in order to shed a different, more intuitive, light on the chaotic behavior of CA, in particular on sensitivity to initial conditions. In [3], the authors asked about the existence of transitive CA in this topology. Here we prove that the answer is negative.

There were former attempts to prove or disprove the existence of transitive CA either by looking for examples or by complicated combinatorial techniques. Here we drastically diminish the difficulty of the problem by making use of Kolmogorov complexity and the classical approach of the "incompressibility method" (see [22] for more on the last subject). It should be possible to modify our proof into a purely combinatorial one by doing some "reverse-engineering".

For any two words u, w on $\{0, 1\}^\star$, denote $K(u)$ the Kolmogorov complexity of u and $K(u|w)$ the Kolmogorov complexity of u conditional to w. We make reference to [22] for the precise definitions of these quantities and all the well-known related inequalities.

Theorem 4. *In the Besicovitch topological space there is no transitive CA.*

Proof. By contradiction, suppose that there exists a transitive CA f of radius r with $C = |S|$ states. Let x and y be two configurations such that

$$\forall n \in \mathbb{N}, \quad K(x_{-n:n}|y_{-n:n}) \geqslant \frac{n}{2}.$$

One proves that such configurations x and y exist by a simple counting argument. Since f is transitive, there are two configurations x' and y' such that for all $n \in \mathbb{N}$,

$$\begin{aligned} &\Delta(x_{-n:n}, x'_{-n:n}) \leqslant 4\varepsilon n, \\ &\Delta(y_{-n:n}, y'_{-n:n}) \leqslant 4\delta n \end{aligned} \tag{9}$$

and an integer u (which only depends on ε and δ) such that

$$f^u(y') = x', \tag{10}$$

where $\varepsilon = \delta = (4e^{10 \log_2 C})^{-1}$.

In what follows only n varies, while $C, u, x, y, x', y', \delta$ and ε are fixed and independent of n.

By Eq. (10), one may compute the word $x'_{-n:n}$ from the following items:

- $y'_{-n:n}$, f, u et n ;
- the two words of y' of length ur which surround $y'_{-n:n}$ and which are missing to compute $x'_{-n:n}$ with Eq. (10).

We obtain that

$$K(x'_{-n:n}|y'_{-n:n}) \leqslant 2ur + K(u) + K(n) + K(f) + \mathrm{O}(1) \leqslant \mathrm{o}(n) \tag{11}$$

(the notations O and o are defined with respect to n). Remark that n is fixed and hence $K(n)$ is a constant bounded by $\log_2 n$. Similarly, r and S are fixed and hence $K(f)$ is constant and bounded by $C^{2r+1} \log_2 C + \mathrm{O}(1)$.

Let us evaluate $K(y'_{-n:n}|y_{-n:n})$. Let $a_1, a_2, a_3, \ldots, a_k$ be the positive positions which $y_{-n:n}$ and $y'_{-n:n}$ differ at, sorted increasingly. Let $b_1 = a_1$ and $b_i = a_i - a_{i-1}$, for $2 \leqslant i \leqslant k$. By Eq. (9) we know that $k \leqslant 4\delta n$. Remark that $\sum_{i=1}^k b_i = a_k \leqslant n$.

Symmetrically let $a'_1, a'_2, a'_3, \ldots, a'_{k'}$ be the absolute values of the strictly negative positions which $y_{-n:n}$ and $y'_{-n:n}$ differ at, sorted increasingly. Let $b'_1 = a'_1$ and $b'_i = a'_i - a'_{i-1}$, where $2 \leqslant i \leqslant k'$. Eq. (9) states that $k' \leqslant 4\delta n$.

Since the logarithm is a concave function, one has

$$\sum \frac{\ln b_i}{k} \leqslant \ln \frac{\sum b_i}{k} \leqslant \ln \frac{n}{k}$$

and hence

$$\sum \ln b_i \leqslant k \ \ln \frac{n}{k} \tag{12}$$

which also holds for b'_i and k'.

The knowledge of the b_i, the b_i', and of the $k+k'$ states of the cells of $y'_{-n:n}$ where $y_{-n:n}$ differs from $y'_{-n:n}$ is enough to compute $y'_{-n:n}$ from $y_{-n:n}$. Hence,

$$K(y'_{-n:n}|y_{-n:n}) \leqslant \sum \ln(b_i) + \sum \ln(b_i') + (k+k')\log_2 C + \mathrm{O}(1).$$

Eq. (12) states that

$$K(y'_{-n:n}|y_{-n:n}) \leqslant k \ln\frac{n}{k} + k' \ln\frac{n}{k'} + (k+k')\log_2 C + \mathrm{O}(1).$$

The function $k \mapsto k\ln(n/k)$ is increasing on $[0, n/e]$. As $k \leqslant 4\delta n \leqslant \frac{n}{\mathrm{e}^{10\log_2 C}}$, we have that

$$k \ln\frac{n}{k} \leqslant 4\delta n \ln\frac{n}{4\delta n} \leqslant \frac{n}{\mathrm{e}^{10\log_2 C}} \ln \mathrm{e}^{10\log_2 C} \leqslant \frac{10n}{\mathrm{e}^{10\log_2 C}}$$

and that

$$(k+k')\log_2 C \leqslant \frac{2n\log_2 C}{\mathrm{e}^{10\log_2 C}}.$$

Replacing a, b and k by a', b' and k', the same sequence of inequalities leads to a similar result. One may deduce that

$$K(y'_{-n:n}|y_{-n:n}) \leqslant \frac{(2\log_2 C + 20)n}{\mathrm{e}^{10\log_2 C}} + \mathrm{O}(1). \tag{13}$$

Similarly, Eq. (13) is also true with $K(x_{-n:n}|x'_{-n:n})$.

The triangular inequality for the Kolmogorov complexity gives

$$K(x_{-n:n}|y_{-n:n}) \leqslant K(x_{-n:n}|x'_{-n:n}) + K(x'_{-n:n}|y'_{-n:n}) + K(y'_{-n:n}|y_{-n:n}) + \mathrm{O}(1).$$

Eqs. (13) and (11) allow one to conclude that

$$K(x_{-n:n}|y_{-n:n}) \leqslant \frac{(2\log_2 C + 20)n}{\mathrm{e}^{10\log_2 C}} + \mathrm{o}(n).$$

The hypothesis on x and y was $K(x_{-n:n}|y_{-n:n}) \geqslant n/2$. This implies that

$$\frac{n}{2} \leqslant \frac{(2C+20)n}{\mathrm{e}^{10\log_2 C}} + \mathrm{o}(n).$$

The last inequality is false for big enough n. $\square$

4. The Weyl topology

The Besicovitch topology still gives some importance to sites near the cell of index zero. The Weyl pseudo-distance and topology remove this bias. The pseudo-distance is defined via a density function like $d_{\mathcal{B}}$ but it is computed with a sup on all possible starting cells.

For all configurations x and y, the *Weyl pseudo-distance* $d_{\mathcal{W}}$ is defined as follows

$$d_{\mathcal{W}}(x, y) = \limsup_{n\to\infty} \sup_{l\in\mathbb{Z}} \frac{\Delta(x_{l:l+n}, y_{l:l+n})}{n+l+1}.$$

As in the Besicovitch setting, the quotient of $S^{\mathbb{Z}}$ with respect to the equivalence relation *being at zero $d_{\mathcal{W}}$-distance* is a metric space.

In [3], the authors proved that this quotient space has enough interesting topological properties to make worth the study of dynamical systems on it. For example it is perfect, path-wise connected, and infinite-dimensional although it is neither locally compact nor complete.

On the other hand, requiring the density function to be independent of the central position imposes very strong constraints; the dynamics acts on really huge equivalence classes. For example the set of Toeplitz configurations, which is Besicovitch dense, is not so in the Weyl space; actually we prove this result for a much larger set, that of quasi-periodic configurations.

A spatially quasi-periodic configuration $x \in S^{\mathbb{Z}}$ is one such that any finite pattern of x occurs with bounded gaps in x: more precisely, for any $n \in \mathbb{N}$ there exists N such that whatever the indexes $i, j \in \mathbb{Z}$, there is at least one occurrence of $x_{i:i+n}$ inside the word $x_{j:j+N}$.

Toeplitz sequences are quasi-periodic, but they are a very small subset of all quasi-periodic sequences. In contrast with Lemma 3, two quasi-periodic configurations may differ on a finite set of coordinates only, and thus be in the same equivalence class for $d_{\mathcal{W}}$ and $d_{\mathcal{B}}$ as well. This does not matter here, since the statement we prove about them is negative.

Proposition 14. *The set of quasi-periodic configurations is not dense in the Weyl space.*

Proof. Let a and b be two states in S. Consider the configuration c defined as follows

$$\forall i \in \mathbb{Z}, \quad c_i = \begin{cases} a & \text{if } i < 0, \\ b & \text{otherwise.} \end{cases}$$

Let x be any quasi-periodic configuration, then $d_{\mathcal{W}}(x, c) \geqslant \frac{1}{2}$. For any integer i, let $\mathcal{M}_n$ be the pattern $x_{0:n}$. By quasi-periodicity for any $n \in \mathbb{N}$ there is an integer k_n with $k_n + n < 0$ such that $x_{k_n:k_n+n} = \mathcal{M}_n$. Then

$$\begin{aligned} \sup_{l \in \mathbb{Z}} \Delta(x_{l:l+n}, c_{l:l+n}) &\geqslant \max(\Delta(x_{0:n}, c_{0:n}), \Delta(x_{k_n:k_n+n}, c_{k_n:k_n+n})) \\ &= \max(\#_{\neq b}(\mathcal{M}_n), \#_{\neq a}(\mathcal{M}_n)) \\ &\geqslant \frac{n}{2} \end{aligned}$$

where $\#_{\neq j}(m)$ is the number of sites in pattern m where the state is different from j. Dividing by $n+1$ and taking the superior limit one finds that $d_{\mathcal{W}}(x, c) \geqslant \frac{1}{2}$. $\square$

Consider the set D of non-generic points for $\langle S^{\mathbb{Z}}, \sigma \rangle$ endowed with the Cantor topology and the uniform Bernoulli measure μ. Arguments from topological dynamics show that any quasi-periodic point is contained in D.

Proposition 15. *The set D is dense in the Weyl space.*

Proof. Let $a, b \in S$. Consider a generic point $x \in X$. For any $n \in \mathbb{N}$, define a non-generic configuration $x^{(n)}$ as follows: in x, replace any string of exactly n a's (occurring between two non-b symbols) by the word ba^{n-1}. Remark that $d_{\mathcal{W}}(x, x^{(n)}) = 1/\#(S)^n.(\#(S) - 1)^2$ since for a generic point this is the frequency of the modified strings. It is easy to see that $x^{(n)}$ is not generic, because the densities of words of the form $w.a^n.z$ have been tampered with. $\square$

The value of this result is mostly indicative. Practically it cannot be of much use. The definition of non-generic points for the uniform measure relies on the Cantor topology, and involves a Besicovitch-type convergence! Moreover, the set D is not invariant under CA maps, and a Weyl equivalence class may contain a huge set of non-generic points (or none at all). What happens is that we are not aware of any *smaller* interesting set. [1]

We conclude this section with a remark about the dynamics of CA in this topology.

Corollary 3. *In the Weyl topology there is no transitive CA.*

Proof. One easily sees that $d_B(x, y) \leqslant d_W(x, y)$. This implies that if a CA is transitive in the Weyl topology, it is transitive in the Besicovitch topology too. Therefore, by Theorem 4, there exist no transitive CA for the Weyl topology. $\square$

[1] Here "huge" and "small" are not meant in the sense of classical set theory but rather in the sense of descriptive set theory (see Silver [29] and followers).

5. Conclusions

This paper further investigates the periodic and chaotic behaviors of cellular automata in different topological settings, namely the Cantor, Besicovitch and Weyl spaces and topologies, keeping in mind the idea that describing the shift CA as chaotic is not satisfactory.

In the Cantor topology the shift is not strongly transitive. It is no longer chaotic when the Devaney definition of chaoticity is strengthened: strong transitivity plus a dense set of periodic points. The significance of strong transitivity in the Cantor topology is far from being fully understood. Its relations with the Besicovitch or Weyl topologies are even more obscure. Open questions arise in this setting when considering several results of the present article. In particular, does strong transitivity with respect to the Cantor topology imply sensitivity to initial conditions in the Besicovitch topology? The claim is true for additive CA [18] but nothing is known for the general case.

We answer in the negative the question whether there exist transitive CA in the Besicovitch topology (Theorem 4). This result has deep implications for the dynamics of CA. First, its proof shows that iterations of a CA map cannot arbitrarily change the density of differences between two configurations. In its turn this fact implies that the information contained in configurations cannot spread much during evolutions. Also the proof technique is interesting in itself since we use the Kolmogorov complexity to prove a purely topological property about discrete dynamical systems.

The low degree of complexity from the point of view of chaoticity is underlined by the second main result of the paper: in the Besicovitch space a CA has either a unique fixed point or an uncountable set of periodic points.

In the Besicovitch topology CA are never very chaotic. Moreover, these two results suggest that the variety of behaviors observed in computer simulations is not so wide as it looks. They also open the quest for new, more appropriate, properties for describing the "complex" behavior of CA dynamics. Some interesting suggestions in this direction may be found, for instance, in [24,21]; they concern in particular the evolution of probability measures under the action of a CA. The authors are currently investigating this subject.

Finally in the Weyl topology there does not seem to be any finitely defined "test set", like finite or spatially periodic configurations in the Cantor space or Toeplitz sequences in the Besicovitch space. The set of non-generic sequences is dense but does not look very convenient.

References

[1] J. Auslander, Personal Communication, 2003.
[2] F. Blanchard, J. Cervelle, E. Formenti, Periodicity and transitivity for cellular automata in Besicovitch topologies, in: Mathematical Foundations of Computer Science (MFCS '03), Lecture Notes in Computer Science, Springer, Berlin, 2003.
[3] F. Blanchard, E. Formenti, P. Kůrka, Cellular automata in the Cantor, Besicovitch and Weyl topological spaces, Complex Systems 11 (1999) 107–123.
[4] F. Blanchard, A. Maass, Dynamical properties of expansive one-sided cellular automata, Israel J. Math. 99 (1997) 149–174.
[5] F. Blanchard, P. Tisseur, Some properties of cellular automata with equicontinuity points, Ann. Inst. H. Poincaré 36 (5) (2000) 562–582.
[6] M. Boyle, B. Kitchens, Periodic points for onto cellular automata, Indag. Math. 10 (4) (1999) 483–493.
[7] G. Braga, G. Cattaneo, P. Flocchini, C. Quaranta Vogliotti, Pattern growth in elementary cellular automata, Theoret. Comput. Sci. 145 (1995) 1–26.
[8] C. Calude, P. Hertling, H. Jürgensen, K. Weihrauch, Randomness on full shift spaces, Chaos, Solitons & Fractals 1 (2000) 1–13.
[9] G. Cattaneo, E. Formenti, L. Margara, J. Mazoyer, A shift-invariant metric on $S^{\mathbb{Z}}$ inducing a non-trivial topology, in: I. Privara, P. Rusika (Eds.), Mathematical Foundations of Computer Science (MFCS'97), Lecture Notes in Computer Science, Vol. 1295, Bratislava, 1997, Springer, Berlin, pp. 179–188.
[10] J. Cervelle, B. Durand, E. Formenti, Algorithmic information theory and cellular automata dynamics, in: Mathematical Foundations of Computer Science (MFCS'01), Lecture Notes in Computer Science, Vol. 2136, Springer, Berlin, 2001, pp. 248–259.
[11] K. Culik, S. Yu, Undecidability of cellular automata classification schemes, Complex Systems 2 (1988) 177–190.
[12] M. Delorme, E. Formenti, J. Mazoyer, Open problems on cellular automata, Electronic version available, No. RR2000-25, dated June 2000.
[13] M. Delorme, J. Mazoyer (Eds.), Cellular Automata, Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, 1999, pp. 213–259 (Chapter 4).
[14] R.L. Devaney, An Introduction to Chaotic Dynamical Systems, Addison-Wesley, Reading, MA, 1989.
[15] J-C. Dubacq, B. Durand, E. Formenti, Kolmogorov complexity and cellular automata classification, Theoret. Comput. Sci. 259 (1–2) (2001) 271–285.
[16] B. Durand, Global properties of cellular automata, in: E. Goles, S. Martinez (Eds.), Cellular Automata and Complex Systems, Kluwer, Dordrecht, 1998.
[17] B. Durand, E. Formenti, G. Varouchas, On undecidability of equicontinuity classification for cellular automata, in: Michel Morvan, Eric Rémila (Eds.), DMCS, Discrete Mathematics and Theoretical Computer Science Proceedings, Vol. AB, DMTCS, 2003, pp. 117–128.

[18] E. Formenti, On the sensitivity of cellular automata in Besicovitch spaces, Theoret. Comput. Sci. 301 (1–3) (2003) 341–354.
[19] C. Knudsen, Chaos without nonperiodicity, Amer. Math. Monthly 101 (1994) 563–565.
[20] P. Kůrka, Languages, equicontinuity and attractors in cellular automata, Ergodic Theory Dynamical Systems 217 (1997) 417–433.
[21] P. Kůrka, A. Maass, Limit set of cellular automata associated to probability measures, J. Statist. Phys. 100 (2000) 1031–1047.
[22] M. Li, P. Vitányi, An Introduction to Kolmogorov Complexity and its Applications, second ed., Springer, Berlin, 1997.
[23] G. Manzini, Characterization of sensitive linear cellular automata with respect to the counting distance, in: Mathematical Foundations of Computer Science (MFCS'98), Lecture Notes in Computer Science, Vol. 1450, Springer, Berlin, 1998, pp. 825–833.
[24] B. Martin, Damage spreading and μ-sensitivity in cellular automata, in: J. van Leeuwen, O. Watanabe, M. Hagiya, P.D. Mosses, T. Ito (Eds.), IFIP TCS 2000, Lecture Notes in Computer Science, Vol. 1872, Springer, Berlin, 2000, pp. 226–242.
[25] E.F. Moore, Machine models of self-reproduction, Proc. Symp. Appl. Math. 14 (1962) 13–33.
[26] J. Myhill, The converse to Moore's garden-of-eden theorem, Proc. Amer. Math. Soc. 14 (1963) 685–686.
[27] M. Nasu, Textile Systems for Endomorphisms and Automorphisms of the Shift, Memoirs of the American Mathematical Society, Vol. 546, American Mathematical Society, Providence, RI, 1995.
[28] W. Parry, Symbolic dynamics and transformations of the unit interval, Trans. Amer. Math. Soc. 122 (1966) 366–378.
[29] J. Silver, Counting the number of equivalence classes of Borel and coanalytic equivalence relations, Ann. Math. Logic 19 (1980) 1–28.
[30] S. Wolfram, Theory and Applications of Cellular Automata, World Scientific, Singapore, 1986.