Theoretical Computer Science 412 (2011) 3822–3837

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Generalized Besicovitch and Weyl spaces: Topology, patterns, and sliding block codes

Silvio Capobianco

Institute of Cybernetics at Tallinn University of Technology, Akadeemia tee 21, 12618 Tallinn, Estonia

ARTICLE INFO

Keywords:
Pseudo-distance
Quotient topology
Cellular automaton
Sliding block code
Orphan pattern

ABSTRACT

The Besicovitch and Weyl topologies on the space of configurations take a point of view completely different from the usual product topology; as such, the properties of the former are much different from that of the latter. The one-dimensional case has already been the subject of thorough studies; we extend it to greater dimensions, and also to more general groups.

1. Introduction

The space $\mathcal{C} = S^{\mathbb{Z}}$ of bi-infinite words on a finite alphabet S is usually given the *product topology*, where a sequence $\{w_n\}_{n\in\mathbb{N}}$ converges if and only if, for every $z \in \mathbb{Z}$, the sequence $w_n(z)$ is ultimately constant. This is the "natural", in the sense of "most straightforward given the setting", way of constructing a topology; moreover, it has several advantages, *e.g.*, it is compact and induced by a metric, which are very desirable properties when dealing with *dynamics* on such spaces. However, it also displays several "side effects" that may be deemed unpleasant in some contexts: one for all, an action as simple as the *shift* – $\sigma(c)$'s value at x is the same as c's value at $x+1$ – displays a *chaotic behavior*, in the sense that it satisfies the three conditions in Devaney's definition of chaos [11, Section 1.8].

To overcome this inconvenience, Cattaneo et al. [9] took inspiration from the work of Besicovitch on almost periodic functions [1]. Their *Besicovitch pseudo-distance* $d_{\mathcal{B}}$ is the point of view of an observer who can only see a fixed, finite portion of the space, which however *becomes larger and larger* until each point of the space is reached. In a subsequent work, Blanchard et al. [3] also define a second pseudo-distance $d_{\mathcal{W}}$, this time named after Weyl and his own work. The new point of view is that of an observer who not only enlarges the window, but also *moves it around the whole space*.

The key idea of the two papers, is to consider the *quotient spaces* $\mathcal{B}$ and $\mathcal{W}$, where two configurations $c, c' \in \mathcal{C}$ are identified iff $d_{\mathcal{B}}(c, c') = 0$ (for $\mathcal{B}$) or $d_{\mathcal{W}}(c, c') = 0$ (for $\mathcal{W}$). The shift then becomes an *isometry* of each of the new spaces: which is only the first of a series of differences between these and $\mathcal{C}$. Noteworthily, both $\mathcal{B}$ and $\mathcal{W}$ are pathwise connected, while each connected component of $\mathcal{C}$ is a singleton; and while each sequence in the latter has a convergent subsequence, this does not happen in the former, not even inside some neighborhood of some point. But the most striking feature of the new spaces is that one-dimensional *cellular automata* (CA) do induce transformations on $\mathcal{B}$ and $\mathcal{W}$, *i.e.*, are well defined on equivalence classes, and also, the properties of the induced transformations are, in several cases, linked to those of the original CA. Just to name two instances of these phenomena, equicontinuous CA induce equicontinuous maps [3, Proposition 7], and a CA is surjective *if and only if* it induces surjective transformations of $\mathcal{B}$ and $\mathcal{W}$ [3, Proposition 6].

In this paper, which is ideally a continuation of [3] and is based on our previous works [8,7,5], we display some of our findings in the search for extensions of the results by Blanchard et al. in the broader context of *finitely generated groups*. This includes all of the usual d-dimensional grids – and in particular, the plane – but also grids with more complicated geometries, such as the *uniform 4-tree*. Besicovitch and Weyl pseudo-distances (namely, $d_{\mathcal{B}}$ and $d_{\mathcal{W}}$) are then defined through *exhaustive*

E-mail address: silvio@cs.ioc.ee.

0304-3975/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2011.02.020

sequences of finite sets that grow to fill the whole space, thus taking the role of the symmetric, centered "windows". In this setting, it is already known from [8] that several properties of $d_{\mathcal{B}}$ – notably, invariance by translation – depend on those of the associate sequence; in the same work, we had provided sufficient conditions for translation-invariance of $d_{\mathcal{B}}$, that do hold in "natural d-dimensional setting", *i.e.*, when the underlying group is $\mathbb{Z}^d$ and the sequence is formed by the disks linked to classical subsets such as the von Neumann or the Moore neighborhoods. In fact, though the methods we employ may be applied for a very broad range of situations, our main focus is on d-dimensional grids, which are of great interest in several branches of theoretical computer science, *e.g.*, automata theory and image recognition.

One of our original targets was the study of the topologies of the quotient spaces under these more general definition and conditions. Some preliminary results had been presented in the short paper [7] at the Automata 2009 conference: in particular, we reproduced some classical results – such as topological dimension and non-density of the family of periodic configurations – in a more general setting, provided some special conditions (satisfied by $\mathbb{Z}^d$ and the Moore disks) are met. What initially gave us a shock instead, was the *completeness* of the Besicovitch metric space, which occurs *in every case we contemplate*: whatever choice we make for the underlying set and the exhaustive sequence, every sequence $\{c_n\}_{n\in\mathbb{N}}$ which is Cauchy for the Besicovitch distance – *i.e.*, "stays ultimately packed arbitrarily tight" – admits an object c such that the set where c_n differs from c becomes more and more "sparse". The reason for this phenomenon shall be clear later on in the paper.

On this subject, the present paper offers an improvement of the result in [7]: we prove not only that several "classical" properties are satisfied in arbitrary dimension – which is not entirely surprising: after all, $S^{\mathbb{Z}^d}$ is notoriously a *Cantor space* regardless of d and S – but also provide a general condition for *homeomorphism* between "general" and 1D Besicovitch spaces. (We regret not having found similar results for Weyl spaces.) In particular, d-dimensional Besicovitch spaces are homeomorphic to the "classical" one.

Another target was the identification of some kind of *local* principles – such as those that characterize CA – in the setting of Besicovitch and Weyl spaces as well. Since both pseudo-distances *erase finite differences* – *i.e.*, configurations that differ on finitely many points are identified – it is pointless to talk, *e.g.*, about single occurrences of patterns; however, it is still meaningful to consider *the density of a set*—which is a concept that turns out to be *dual* to these pseudo-distances. This was the subject of [5], which we presented at the SOFSEM 2010 conference: the Besicovitch (resp., Weyl) density of the set of occurrences of a pattern in a configuration only depends on the Besicovitch (resp, Weyl) equivalence class of the configuration.

Finally, we consider cellular automata on these spaces. About these, we remarked that several results proved in previous papers [8] also hold for the more general class of *sliding block codes* (SBC)—where the *source* alphabet can differ from the *target* alphabet. In this respect, other than provide new results, we improve some we had given in [8], extending them to arbitrary SBC in either Besicovitch or Weyl spaces: in particular, we extend to SBC a theorem that characterizes surjective CA in the three topologies, and was an extension of the one given in [2] for 1D CA. To get that result, we exploit a well-known fact, known as the *orphan pattern principle*: a CA – and more in general a SBC – has a *global configuration* without a predecessor, if and only if it has a *local pattern* without a predecessor. Now we move even further, and prove a variant of this principle that holds directly in the Besicovitch and Weyl spaces, and makes reference to the density of the set of the occurrences.

This is not everything that can be proved about the Besicovitch and Weyl spaces, and surely is not all that *we* want to prove about them. Because of this, throughout the paper, we point out some facts and conjecture that we would like to see proved or disproved: we will state them as *challenges*, for us and for anyone who wants to go through them.

The paper is organized as such. Section 2 provides a background. Section 3 is dedicated to the definition of the Besicovitch and Weyl spaces in the general context, and the study of their properties. Section 4 focuses on patterns and their occurrences from the point of view of these quotient spaces. Section 5 deals with sliding block codes in the new setting, and some of their properties. Conclusions and acknowledgments follow.

2. Background

2.1. Topology

We presume the reader to be familiar with the usual concepts of general topology [16]. A topological space is *Lindelöf* if every open cover admits a countable subcover. A metric space is *totally bounded* if, for every $r > 0$, the covering made of disks of radius r has a finite subcover.

Fact 1. *Let (X, d) be a metric space.*

1. *The following are equivalent.*
   (a) *(X, d) is compact.*
   (b) *(X, d) is sequentially compact.*
   (c) *(X, d) is complete and totally bounded.*
2. *If (X, d) is either separable or Lindelöf, then it is second countable. In particular, for a complete metric space the following hold:*

$$\begin{array}{ccccc} \textit{compact} & \Rightarrow & \textit{Lindelöf} & \Rightarrow & \textit{second countable} \\ \Updownarrow & & & & \Updownarrow \\ \textit{totally bounded} & & & & \textit{separable} \end{array}$$

The *Cantor set* is obtained by starting from the interval $[0, 1]$ and iteratively removing the *open* middle thirds of each interval remaining from the previous step. This space is thus made of precisely those $x \in [0, 1]$ that have a base-3 expansion without 1's: it is compact, totally disconnected, and perfect. A metric that induces the topology of the Cantor space is given by $d(x, y) = 2^{-n}$, where the nth subdivision step is the first to send x and y into different segments.

A *pseudo-distance* on a set X is a map $d : X \times X \to [0, +\infty)$ satisfying all of the axioms for a distance, except $d(x, y) > 0$ for every $x \neq y$. If d is a pseudo-distance on X, then $x_1 \sim x_2$ iff $d(x_1, x_2) = 0$ is an equivalence relation, and $d(\kappa_1, \kappa_2) = d(x_1, x_2)$ with $x_i \in \kappa_i$ is a distance on $X/\sim$. The pair (X, d) is also called a pseudo-metric space.

Let (X, d_X) and (Y, d_Y) be pseudo-metric spaces. A function $f : X \to Y$ is *Lipschitz continuous* w.r.t. (d_X, d_Y) if there exists $L > 0$ such that

$$d_Y(f(x_1), f(x_2)) \leq L \cdot d_X(x_1, x_2) \quad \forall x_1, x_2 \in X. \tag{1}$$

2.2. Group theory

Let G be a group. The identity element of G shall be indicated by 1_G, or 1 if no ambiguity is possible. If the group is abelian (*e.g.*, $\mathbb{Z}^d$) we may indicate its operation as $+$, and its identity element as 0. Product and inverse are extended to subsets of G element-wise. We write $H \leq G$ if H is a subgroup of G. The classes of the equivalence relation on G defined by $x\rho y$ iff $xy^{-1} \in H$ are called the *right cosets* of H. If U is a *set of representatives* of the right cosets of H (one representative per coset) then $(h, u) \mapsto hu$ is a bijection between $H \times U$ and G. The number of right cosets of H is called the *index* of H in G, written $[G : H]$.

Let $f_1, f_2 : \mathbb{N} \to [0, +\infty)$. We write $f_1(n) \preccurlyeq f_2(n)$ if there exist $n_0 \in \mathbb{N}$ and $C, \beta > 0$ such that $f_1(n) \leq C \cdot f_2(\beta n)$ for all $n \geq n_0$; we write $f_1(n) \approx f_2(n)$ if $f_1(n) \preccurlyeq f_2(n)$ and $f_2(n) \preccurlyeq f_2(n)$ Observe that, if either f_i is a polynomial, the choice $\beta = 1$ is always allowed.

If $E \subseteq G$ is finite and nonempty, the *closure* and *boundary* of $X \subseteq G$ w.r.t. E are the sets $X^{+E} = \{g \in G : gE \cap X \neq \emptyset\} = XE^{-1}$ and $\partial_E X = X^{+E} \setminus X$, respectively. In general, $X \not\subseteq X^{+E}$ unless $1_G \in E$. $V \subseteq G$ is a *set of generators* for G if each $g \in G$ can be written as a (finite!) word on $V \cup V^{-1}$; equivalently, if the *Cayley graph* $(G, \mathcal{E}_V)$, where $\mathcal{E}_V = \{(x, xz) : x \in G, z \in V \cup V^{-1}\}$, is connected. A group is *finitely generated* (briefly, f.g.) if it has a finite set of generators (briefly, FSOG). The *distance* between g and h w.r.t. V is their distance in the graph $(G, \mathcal{E}_V)$; the *length* of $g \in G$ w.r.t. V is its distance from 1_G. The *disk* of center g and radius r w.r.t. V will be indicated by $D_{r,V}(g)$; we will omit g if equal to 1_G, and V if irrelevant or clear from the context. Observe that $D_r(g) = gD_r$, and that $(D_{n,V})^{+D_{r,V}} = D_{n+r,V}$. For the rest of the paper, we will only consider f.g. infinite groups.

The *growth function* of G w.r.t. V is $\gamma_S(n) = |D_{n,V}|$. It is well known [12] that $\gamma_V(n) \approx \gamma_{V'}(n)$ for any two FSOG V, V'. G is of *sub-exponential growth* if $\gamma_V(n) \preccurlyeq \lambda^n$ for all $\lambda > 1$; G is of *polynomial growth* if $\gamma_V(n) \approx n^k$ for some $k \in \mathbb{N}$. Observe that, if $G = \mathbb{Z}^d$, then $\gamma_V(n) \approx n^d$.

A sequence $\{X_n\}$ of *finite* subsets of G is *exhaustive* if $X_n \subseteq X_{n+1}$ for every $n \in \mathbb{N}$ and $\bigcup_{n\in\mathbb{N}} X_n = G$. $\{D_n\}$ is an exhaustive sequence. An exhaustive sequence is *amenable* [10,12,14,18] if $\lim_{n\to\infty} |\partial_E X_n|/|X_n| = 0$ for every finite $E \subseteq G$; a group is amenable if it has an amenable sequence.

Fact 2 (See [12]). *Let G be a f.g. group and V a* FSOG *for G; let $\mathcal{X} = \{D_{n,V}\}_{n\geq 0}$.*

1. *G is of polynomial growth iff $\mathcal{X}$ is amenable.*
2. *G is of sub-exponential growth iff $\mathcal{X}$ contains an amenable subsequence.*

Definition 1. Let $U, W \subseteq G$ be nonempty. A (U, W)-*net* is a set $N \subseteq G$ such that the sets xU, $x \in N$, are pairwise disjoint, and $NW = G$.

Any subgroup $H \leq G$ is a (U, U)-net for any set U of representatives of its right cosets, because $G = \bigcup_{h\in H} hU$ is a partition of G.

Proposition 1. *Let G be a group.*

1. *For every nonempty $U \subseteq G$ there exists a (U, UU^{-1})-net.*
2. *If N is a (U, W)-net and $\phi(x) \in xU$ for every $x \in N$, then $\phi(N)$ is a $(\{1_G\}, U^{-1}W)$-net.*

Proof. For point 1, let $\mathcal{P}$ be the set of subsets $M \subseteq G$ s.t. the xU, $x \in M$, are pairwise disjoint: $\mathcal{P}$ is partially ordered by set inclusion, and the union of a chain in $\mathcal{P}$ is still in $\mathcal{P}$. If $M \in \mathcal{P}$ and some $g \in G$ does not have the form $g = xu_1u_2^{-1}$ for $x \in M$ and $u_1, u_2 \in U$, then $M \cup \{g\} \in \mathcal{P}$; thus, a maximal element of $\mathcal{P}$ (which exists by Zorn's lemma) must be a (U, UU^{-1})-net.

For point 2, since the xU are pairwise disjoint, the $\phi(x)$ for $x \in N$ are pairwise distinct, *i.e.*, the $\phi(x)\{1_G\}$ are pairwise disjoint. Also, since each $g \in G$ has the form $g = xw = \phi(x)u^{-1}w$ where $x \in N$, $w \in W$, $\phi(x) = xu$, $u \in U$, also $\phi(N)(U^{-1}W) = G$. □

2.3. Symbolic dynamics

If $2 \leq |S| < \infty$ and G is a f.g. group, the space $\mathcal{C} = S^G$ of *configurations* of G over S, endowed with the product topology, is homeomorphic to the Cantor set. If $E \subseteq G$ is finite, a *pattern* over S of *support* E is a map $p \in S^E$. For $g \in G$, the *translation by* g is the transformation $\sigma_g : \mathcal{C} \to \mathcal{C}$ defined by

$$\sigma_g(c)(h) = c(g \cdot h) \quad \forall h \in G. \tag{2}$$

Observe that $\sigma = \{\sigma_g\}_{g \in G}$ is a (right) action of G on $\mathcal{C}$, called the *natural action*. For $G = \mathbb{Z}$ and $g = +1$, the translation $c \mapsto c^{+1}$ is the shift map. An *occurrence* of a pattern p in a configuration c is a point $g \in G$ such that $c^g|_{\operatorname{supp} p} = p$; the set of occurrences of p in c is indicated as $\operatorname{occ}(p, c)$. From now on, we will often write c^g for $\sigma_g(c)$.

A *sliding block code* (briefly, SBC) over G is a triple $\mathcal{K} = \langle S, T, \mathcal{N}, f \rangle$, where the *source alphabet* S and the *target alphabet* T are finite and have at least two elements each, the *neighborhood index* $\mathcal{N} \subseteq G$ is finite and nonempty, and the *local function* f maps $S^{\mathcal{N}}$ into T. A SBC with $S = T$ is called a *cellular automaton* (briefly, CA). The map $F_{\mathcal{K}} : S^G \to T^G$ defined by

$$(F_{\mathcal{K}}(c))(g) = f\left(c^g\big|_{\mathcal{N}}\right) \tag{3}$$

is the *global function* of $\mathcal{K}$.

Fact 3 (*Hedlund's Theorem*)**.** *Let $F : S^G \to T^G$ for a f.g. group G and two alphabets S, T. The following are equivalent.*

1. *F is continuous in the product topology and commutes with translations, i.e., $(F(c))^g = F(c^g)$ for every $g \in G$, $c \in S^G$.*
2. *There exists a* SBC *$\mathcal{K} = \langle S, T, \mathcal{N}, f \rangle$ such that $F_{\mathcal{K}} = F$.*

$\mathcal{K}$ is said to be injective, surjective, and so on, if $F_{\mathcal{K}}$ is.

Fact 4. *Let $\mathcal{K} = \langle S, T, \mathcal{N}, f \rangle$ be a* SBC*. If $F_{\mathcal{K}}$ is bijective, then there exists a* SBC *$\mathcal{K}'$ s.t. $F_{\mathcal{K}'} = F_{\mathcal{K}}^{-1}$.*

Let $\mathcal{K} = \langle S, T, \mathcal{N}, f \rangle$ be a SBC over the group G. A pattern p over T is an *orphan* for $\mathcal{K}$ if $F_{\mathcal{K}}(c)$ has no occurrence of p for every $c \in S^G$.

Fact 5 (*Orphan Pattern Principle; cf. [19]*)**.** *A* SBC *is surjective iff it has no orphan pattern.*

3. The Besicovitch and Weyl distances

As we have said in the introduction, the Besicovitch and Weyl points of view corresponds to series of observations of the space of configurations through a sequence of finite, but growing, windows. The idea at the basis of the definitions is a very well grounded concept in computer science. From now on, we write

$$\Delta(\phi, \psi) = \{x \in X \mid \phi(x) \neq \psi(x)\} \tag{4}$$

whenever $\phi, \psi : X \to Y$.

Definition 2. Let X and Y be sets, let $U \subseteq X$, let $\phi, \psi : X \to Y$. The *Hamming distance between ϕ and ψ relative to U* is

$$H_U(\phi, \psi) = |U \cap \Delta(\phi, \psi)|. \tag{5}$$

In general, (5) is a pseudo-distance unless $U = X$. If X is a metric space and U is a disk of radius r, we may write H_r instead of H_{D_r}. Also, if $X = G$ is a f.g. group and U is a disk of radius r w.r.t. a FSOG V, we may write $H_{r,V}$ instead of $H_{D_{r,V}}$.

If X is countable and $\mathcal{X} = \{X_n\}$ is an exhaustive sequence for X, we can use the Hamming distances relative to the X_n's to define a new pseudo-distance on Y^X.

Definition 3. Let X be a countable set and let $\mathcal{X} = \{X_n\}_{n \in \mathbb{N}}$ be an exhaustive sequence for X. Let $\phi, \psi : X \to Y$ with Y an arbitrary set. The *Besicovitch pseudo-distance between ϕ and ψ w.r.t. $\mathcal{X}$* is the quantity

$$d_{\mathcal{B},\mathcal{X}}(\phi, \psi) = \limsup_{n \in \mathbb{N}} \frac{H_{X_n}(\phi, \psi)}{|X_n|}. \tag{6}$$

The quotient space $\mathcal{B}_{\mathcal{X}}^{X,Y} = Y^X / \sim_{\mathcal{B},\mathcal{X}}$, where $\phi \sim_{\mathcal{B},\mathcal{X}} \psi$ iff $d_{\mathcal{B},\mathcal{X}}(\phi, \psi) = 0$, is called the *Besicovitch space* induced by $\{X_n\}$.

The quantity (6) corresponds to the following approximation. Each X_n is a *window* through which the observer sees the functions: he/she can never see the *whole* space, but the portion under his/her eyes *grows with time*. What he/she then does, is to compute a *probability* of the event "$f(x) \neq g(x)$, given x is under the window". The upper limit of this sequence of probabilities is a pseudo-distance.

If $X = G$ is a group, our observer has another possibility.

Definition 4. Let G be a countable group and let $\mathcal{X} = \{X_n\}_{n\in\mathbb{N}}$ be an exhaustive sequence for G. Let $\phi, \psi : G \to Y$ with Y an arbitrary set. The *Weyl pseudo-distance between f and g w.r.t. $\mathcal{X}$* is the quantity

$$d_{\mathcal{W},\mathcal{X}}(\phi, \psi) = \limsup_{n\in\mathbb{N}} \left(\frac{1}{|X_n|} \sup_{g\in G} H_{gX_n}(\phi, \psi) \right). \tag{7}$$

The quotient space $\mathcal{W}_{\mathcal{X}}^{G,Y} = Y^G / \sim_{\mathcal{W},\mathcal{X}}$, where $\phi \sim_{\mathcal{W},\mathcal{X}} \psi$ iff $d_{\mathcal{W},\mathcal{X}}(\phi, \psi) = 0$, is called the *Weyl space* induced by $\{X_n\}$.

Definition 4 mirrors the point of view of an observer that sees the world through a set of enlarging windows, but also *centers each window at every possible point*. Since the windows are finite, the supremum in parentheses is actually a maximum. Also, since moving the window in one direction is the same as moving the space in the *opposite* direction – *i.e.*, $\Delta(\phi^g, \psi^g) = g^{-1}\Delta(\phi, \psi)$ – we get for free

$$d_{\mathcal{W},\mathcal{X}}(\phi, \psi) = \limsup_{n\in\mathbb{N}} \left(\frac{1}{|X_n|} \max_{g\in G} H_{X_n}(\phi^g, \psi^g) \right). \tag{8}$$

In the case $\mathcal{X} = \{D_{r,V}\}_{r\geq 0}$ for some FSOG V, we will replace the subscript $\mathcal{X}$ with V. So, for instance, $\mathcal{B}_{\mathcal{X}}$ will become $\mathcal{B}_V$.

It is critical to keep in mind that, in general, neither $d_{\mathcal{B}}$ nor $d_{\mathcal{W}}$ are continuous, as real-valued functions, with respect to the product topology.

Example 1. Let $G = \mathbb{Z}$, $S = \{0, 1\}$, $X_n = [-n, \ldots, n]$. (We shall refer to this as the *standard case* through the rest of the paper.) Fix $c \in S^{\mathbb{Z}}$, and let $c_k(x) = c(x)$ if and only if $|x| \leq k$. Then $\lim_{k\to\infty} c_k = c$ in the product topology. However, $d_{\mathcal{B}}(c_k, c) = d_{\mathcal{W}}(c_k, c) = 1$ for every $k \in \mathbb{N}$.

Regarding notations, we follow the usual convention of removing superscripts and/or subscripts when irrelevant or clear from the context. Moreover, since the case $X_n = D_{n,V}$ is rather common, we will often write $d_{\mathcal{B},V}$ and $d_{\mathcal{W},V}$ instead of $d_{\mathcal{B},\{D_{n,V}\}}$ and $d_{\mathcal{W},\{D_{n,V}\}}$.

Observe that, by definition, if $\phi, \psi : G \to S$, then $d_{\mathcal{B},\mathcal{X}}(\phi, \psi) \leq d_{\mathcal{W},\mathcal{X}}(\phi, \psi)$ for any two $\phi, \psi : G \to Y$. As a consequence, $\sim_{\mathcal{W},\mathcal{X}}$ *saturates* $\sim_{\mathcal{B},\mathcal{X}}$, *i.e.*, every equivalence class of $\sim_{\mathcal{B},\mathcal{X}}$ is union of equivalence classes of $\sim_{\mathcal{W},\mathcal{X}}$: the points of $\mathcal{W}_{\mathcal{X}}^{G,S}$ are "thinner" than those of $\mathcal{B}_{\mathcal{X}}^{G,S}$, and the latter is *coarser-grained* than the former.

Because of their definition, which actually relies on families of sets, the Besicovitch and Weyl pseudo-distances may be seen from another point of view.

Definition 5. Let $\mathcal{X}$ be an exhaustive sequence for a set X. Let $U \subseteq X$. The *Besicovitch upper density* of U w.r.t. $\mathcal{X}$ is the quantity

$$\operatorname{dens\,sup}_{\mathcal{B},\mathcal{X}} U = \limsup_{n\in\mathbb{N}} \frac{|U \cap X_n|}{|X_n|}. \tag{9}$$

Definition 5 is "dual" to Definition 3 in the following sense. By construction, $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) = \operatorname{dens\,sup}_{\mathcal{B},\mathcal{X}} \Delta(c_1, c_2)$. On the other hand, if $y_0, y_1 \in Y$ are distinct, then $\operatorname{dens\,sup}_{\mathcal{B},\mathcal{X}} U = d_{\mathcal{B},\mathcal{X}}(\chi_U, \chi_\emptyset)$, where $\chi_U(x)$ is y_1 if $x \in U$, and y_0 otherwise. A similar "dualism" exists between Definition 4 and

Definition 6. Let $\mathcal{X}$ be an exhaustive sequence for a group G. Let $U \subseteq G$. The *Weyl upper density* of U w.r.t. $\mathcal{X}$ is the quantity

$$\operatorname{dens\,sup}_{\mathcal{W},\mathcal{X}} U = \limsup_{n\in\mathbb{N}} \sup_{g\in G} \frac{|U \cap gX_n|}{|X_n|} = \limsup_{n\in\mathbb{N}} \sup_{g\in G} \frac{|gU \cap X_n|}{|X_n|}. \tag{10}$$

According to these "dualisms", we have a simple, approximated characterization of these quotient spaces: $d_{\mathcal{B}}(c_1, c_2)$ is "small" if $\Delta(c_1, c_2)$ is "sparse"; $d_{\mathcal{W}}(c_1, c_2)$ is "small" if $\Delta(c_1, c_2)$ is "sparse without large chunks". Symmetrically, we can consider the corresponding *lower densities* $\operatorname{dens\,inf}_{\mathcal{B},\mathcal{X}} U$ and $\operatorname{dens\,inf}_{\mathcal{W},\mathcal{X}} U$, defined as the lower limits of the respective quantities.

It is already known that several properties of $d_{\mathcal{B}}$ (and possibly $d_{\mathcal{W}}$) depend on the properties of $\mathcal{X}$. One of this, when the base space is a group, is *invariance by translations*: using the fact that the size of $[-n-1, \ldots, n+1]^d \setminus [-n, \ldots, n]^d$ is a polynomial in n of degree $d-1$, it is easy to check that $d_{\mathcal{B}}(c_1^x, c_2^x) = d_{\mathcal{B}}(c_1, c_2)$ for any $c_1, c_2 : \mathbb{Z}^d \to Q$ and $x \in \mathbb{Z}^d$. Indeed, the main reason for the definition of $d_{\mathcal{B}}$ in [9] was precisely to find a pseudo-distance which is translation-invariant and allows definition of cellular automata. In [8] we provide a condition on $\mathcal{X}$ sufficient for $d_{\mathcal{B},\mathcal{X}}$ to be translation-invariant: as a consequence of that, if $G = \mathbb{Z}^d$ and $\mathcal{X} = \{D_{n,V}\}_{n\geq 0}$ for some FSOG V, then $d_{\mathcal{B},\mathcal{X}}$ is translation-invariant. However, this is not always the case.

Example 2 (*Cf.* [8])**.** Let $G = \mathbb{F}_2$ be the free group on two generators a, b; let X_n be the ball of radius n. Let $c_0(g) = 0$ for all $g \in G$; let $c(g) = 1$ if $g = aw$ as a reduced word, 0 otherwise. Then $d_{\mathcal{B},\mathcal{X}}(c_0, c) = 1/4$ but $d_{\mathcal{B},\mathcal{X}}(c_0^a, c^a) = 3/4$; cf. Fig. 1.

The next result is thus rather surprising.

Theorem 2. *Let X and Y be sets and let $\mathcal{X}$ be an exhaustive sequence for X. Then $(\mathcal{B}_{\mathcal{X}}^{X,Y}, d_{\mathcal{B},\mathcal{X}})$ is a complete metric space.*

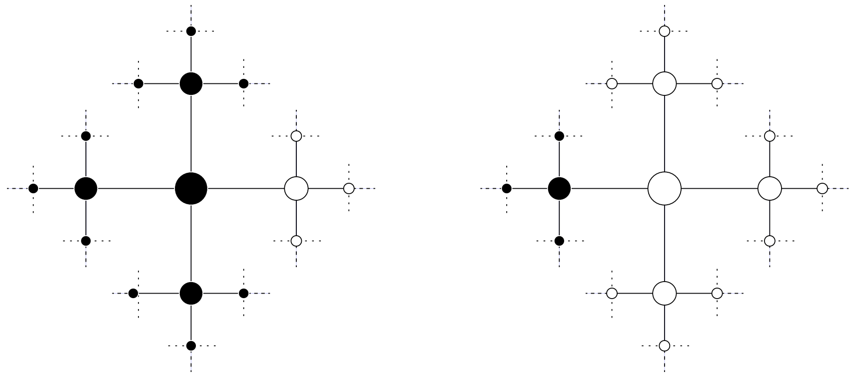

Fig. 1. On the left: configuration c is 1 on the right subtree (indexed by the generator a) and 0 elsewhere. On the right: configuration c^a is 0 on the left subtree and 1 elsewhere.

Proof. It is sufficient to prove the thesis when X is infinite. The following proof is modeled on that of [3, Proposition 2]. Let $\{c_k\}_{k\in\mathbb{N}} \subseteq Y^X$ satisfy Cauchy condition w.r.t. $d_{\mathcal{B},\mathcal{X}}$, *i.e.*,

$$\forall \varepsilon > 0\, \exists n_\varepsilon \mid \forall n > n_0 \forall p > 0 . d_{\mathcal{B},\mathcal{X}}(c_{n+p}, c_n) < \varepsilon . \tag{11}$$

Our plan is to show that it has a convergent subsequence. Then the original sequence shall converge to the same limit, because a Cauchy sequence has at most one limit point.

Choose $\{k_m\}$ so that $d_B(c_{k_m}, c_{k_{m+1}}) < 2^{-m-1}$ for all m. Let $\{\lambda_m\}$ satisfy the following properties:

1. $|X_{\lambda_{m+1}}| \geq 2 \cdot |X_{\lambda_m}|$ for every $m \in \mathbb{N}$.
2. $\sup_{n\geq\lambda_m} \frac{H_{X_n}(c_{k_m}, c_{k_{m+1}})}{|X_n|} \leq 2^{-m}$ for every $m \in \mathbb{N}$.

Note that property 2 implies $H_{X_n}(c_{k_m}, c_{k_{m+p}}) \leq |X_n| \cdot 2^{-m} \cdot \frac{1-2^{-p}}{1-2^{-1}} \leq |X_n| \cdot 2^{1-m}$ for all $p \geq 1, n \geq \lambda_{m+p}$. Call $\Delta_m = X_{\lambda_{m+1}} \setminus X_{\lambda_m}$. Put

$$c(x) = \begin{cases} c_{k_m}(x) & \text{if } x \in \Delta_m, \\ \text{arbitrary} & \text{if } x \in X_{\lambda_0}. \end{cases} \tag{12}$$

Given $n > \lambda_m$, there is exactly one $M \geq m$ s.t. $n \in \{\lambda_M + 1, \ldots, \lambda_{M+1}\}$. Then by property 2

$$\begin{aligned} H_{X_n}(c_{k_m}, c) &= H_{X_{\lambda_m}}(c_{k_m}, c) + \sum_{i=m+1}^{M-1} H_{\Delta_i}(c_{k_m}, c_{k_i}) + H_{X_n \setminus X_{\lambda_M}}(c_{k_m}, c_{k_M}) \\ &\leq |X_{\lambda_m}| + 2 \sum_{i=m+1}^{M-1} |X_{\lambda_{i+1}}| 2^{-m} + |X_n| \cdot 2^{1-m}. \end{aligned}$$

By property 1, $\sum_{i=m+1}^{M-1} |X_{\lambda_{i+1}}| \leq |X_{\lambda_M}| \cdot \sum_{j=1}^{M-m-1} 2^{-j} \leq 2|X_{\lambda_M}|$, so that $\frac{H_{X_n}(c_{k_m}, c)}{|X_n|} \leq \frac{|X_{\lambda_m}|}{|X_n|} + 6 \cdot 2^{-m}$ for all $n \geq \lambda_m$. Hence, $d_{B,\mathcal{X}}(c_{k_m}, c) \leq 3 \cdot 2^{1-m}$. $\square$

Theorem 2 is surprising in that it is true whatever X, Y, and $\mathcal{X}$ are. The reason why this happens, is that the original proof is based on three properties which hold whatever the ingredients above are. The first property, is just the characterization of $d_{\mathcal{B}}$: c_1 and c_2 are "near" iff $\Delta(c_1, c_2)$ is "sparse", the "sparseness" being measured according to the sequence $\mathcal{X}$, and this is true in any case. The second property, is shared by Cauchy sequences in any metric space: they have at most one limit point. The third and last, is that $\mathcal{X}$ is exhaustive, so that $|X_n|$ is unbounded.

We remark that $\mathcal{W}$ is not complete even when $G = \mathbb{Z}$ and $X_n = [-n, \ldots, n]$: see [3] for a full proof. This is because $\mathcal{W}_{\mathcal{X}}$, which we recall being finer-grained than $\mathcal{B}_{\mathcal{X}}$, has "too many" points, hence also "too many" Cauchy sequences.

Challenge 1. Check that $\mathcal{W}$ is never complete.

In general, the classes of $\sim_{\mathcal{B}}$ and $\sim_{\mathcal{W}}$ depend on the choice of $\{X_n\}$. However, if the group "does not grow too fast" and the X_n's are disks, then all the FSOG for G determine the same *notion of convergence* for $d_{\mathcal{B}}$ and $d_{\mathcal{W}}$.

Lemma 3. *Let G be a group of polynomial growth of order d. Let V, V' be* FSOG *for G. There exist $C, \beta, n_0 > 0$ such that, for any $c_1, c_2 \in \mathcal{C}$, $n > n_0$*

$$\frac{H_{n,V'}(c_1, c_2)}{\gamma_{V'}(n)} \leq C \cdot \frac{H_{\beta n,V}(c_1, c_2)}{\gamma_V(\beta n)} . \tag{13}$$

Proof (*Sketch*). Put $C = \alpha_1 \beta^d / \alpha_2$, where $\gamma_V(n) \leq \alpha_1 n^d$ and $\alpha_2 n^d \leq \gamma_{V'}(n)$ for every n large enough, and $D_{1,V'} \subseteq D_{\beta,V}$. □

Theorem 4. *Let G be a group of polynomial growth. If $\lim_{k\to\infty} d_{\mathcal{B},V}(c_k, c) = 0$ for some* FSOG *V, then $\lim_{k\to\infty} d_{\mathcal{B},V}(c_k, c) = 0$ for every* FSOG *V. In particular, if $d_{\mathcal{B},V}(c_1, c_2) = 0$ for some V, then $d_{\mathcal{B},V}(c_1, c_2) = 0$ for all V.*

The above remain true if $d_{\mathcal{B},V}$ is replaced with $d_{\mathcal{W},V}$.

Proof. Let V be a FSOG for G such that $\lim_{k\to\infty} d_{\mathcal{B},V}(c_k, c) = 0$. Let V' be another FSOG for G. Choose C and β so that (13) is satisfied.

Given $\varepsilon > 0$, let k_ε be such that $d_{\mathcal{B},V}(c_k, c) < \frac{\varepsilon}{4C}$ for all $k > k_\varepsilon$. Fix $k > k_\varepsilon$ and choose $n_{\varepsilon,k}$ such that

$$\frac{H_{\beta n,V}(c_k, c)}{\gamma_V(\beta n)} < \frac{\varepsilon}{2C} \quad \forall n > n_{\varepsilon,k}. \tag{14}$$

By (13), $\frac{H_{n,V'}(c_k,c)}{\gamma_{V'}(n)} < \frac{\varepsilon}{2}$ for all $n > n_{\varepsilon,k}$, so that $\varepsilon > \limsup_n \frac{H_{n,V'}(c_k,c)}{\gamma_{V'}(n)} = d_{\mathcal{B},V'}(c_k, c)$. This happens for every $k > k_\varepsilon$, thus $\limsup_k d_{\mathcal{B},V'}(c_k, c) < \varepsilon$. This happens for every $\varepsilon > 0$, thus $\lim_{k\to\infty} d_{\mathcal{B},V'}(c_k, c) = 0$.

A proof for $d_{\mathcal{W},V}$ is obtained by replacing (14) with

$$\frac{H_{\beta n,V}(c_k^g, c^g)}{\gamma_V(\beta n)} < \frac{\varepsilon}{2C} \quad \forall g \in G \; \forall n > n_{\varepsilon,k}. \quad \square \tag{15}$$

Corollary 5. *As V varies in the class of* FSOG *for $\mathbb{Z}^d$, $\mathcal{B}_V^{\mathbb{Z}^d,S}$ remains the same and so does $\mathcal{W}_V^{\mathbb{Z}^d,S}$. In particular, V can be freely chosen as either the von Neumann or Moore neighborhood, according to which is more convenient.*

We recall that the Moore neighborhood is

$$\mathcal{M}_d = \left\{ z \in \mathbb{Z}^d \mid |z_i| \leq 1 \; \forall i \in \{1, \ldots, d\} \right\}. \tag{16}$$

Challenge 2. Prove Theorem 4 for groups of *sub-exponential* growth.

To end the section with a final remark, we observe that, in the case of $\mathbb{Z}^d$, the Weyl distance is a straightforward extension of the one defined in [3].

Theorem 6. *If V is a* FSOG *for $\mathbb{Z}^d$ then*

$$d_{\mathcal{W},V}(c_1, c_2) = \lim_{n\to\infty} \max_{z\in\mathbb{Z}^d} \frac{H_{\{z,\ldots,z+n-1\}^d}(c_1, c_2)}{n^d}. \tag{17}$$

Proof. Because of [4, Theorem 5 and Corollary 6], Equation (17) holds for $V = \mathcal{M}_d$. Apply Theorem 4. □

3.1. The topology of the Besicovitch space

We know from [3] the topological properties of the classical Besicovitch space $\mathcal{B}^{\mathbb{Z},\{0,1\}}$; and we know that they are much different from those of $S^{\mathbb{Z}}$. We suspect that a similar phenomenon happens more in general; and, to find support to this hypothesis, we take into account another construction from the same paper.

Definition 7. The *unilateral Besicovitch space* on the alphabet S is the quotient $\mathcal{B}^{\mathbb{N},S}$ of $S^{\mathbb{N}}$ – the set of *infinite words* over S – with respect to the pseudo-distance

$$d_{\mathcal{B},\mathbb{N}}(w_1, w_2) = \limsup_n \frac{|\{x \in \{0, \ldots, n-1\} | w_1(x) \neq w_2(x)\}|}{n}. \tag{18}$$

For $S = \{0, 1\}$ – and, actually, for arbitrary finite S – the properties of $\mathcal{B}^{\mathbb{N},S}$ are known from [3], and are the same as those of $\mathcal{B}^{\mathbb{Z},S}$.

Proposition 7. *$\mathcal{B}^{\mathbb{N},S}$ is complete, arcwise connected, and infinite-dimensional. $\mathcal{B}^{\mathbb{N},S}$ is neither separable nor locally compact.*

The reason why we use $\mathcal{B}^{\mathbb{N},S}$ instead of $\mathcal{B}^{\mathbb{Z},S}$, is that it allows a simpler way to *enumerate* the elements of the sets in an exhaustive sequence.

Let X be a countable set and let $\mathcal{X} = \{X_n\}_{n\in\mathbb{N}}$ an exhaustive sequence for X. Consider a bijection $\iota_{\mathcal{X}} : \mathbb{N} \to X$ that satisfies the following property: if $\iota_{\mathcal{X}}(m) \in X_k$ and $\iota_{\mathcal{X}}(n) \in X_{k+p} \setminus X_k$, then $m < n$. In other words, $\iota_{\mathcal{X}}$ first counts the elements of X_1, then the remaining elements of X_2, then the remaining elements of X_3, and so on. Call $\kappa_{\mathcal{X}}$ the inverse of $\iota_{\mathcal{X}}$.

Given an alphabet S, we lift $\iota_{\mathcal{X}}$ to a bijection from $S^{\mathbb{N}}$ to S^X, which we also label $\iota_{\mathcal{X}}$ for convenience, via the straightforward definition

$$\iota_{\mathcal{X}}(c)(x) = c(\kappa_{\mathcal{X}}(x)). \tag{19}$$

Observe that the inverse of (19) is precisely the lifting of $\kappa_{\mathcal{X}}$ to a bijection from S^X to $S^{\mathbb{N}}$: we are then justified to call $\kappa_{\mathcal{X}}$ this new function. It is straightforward to check that $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) \leq d_{\mathcal{B},\mathbb{N}}(\kappa_{\mathcal{X}}(c_1), \kappa_{\mathcal{X}}(c_2))$.

Theorem 8. *If* $\limsup_{k\in\mathbb{N}} |X_{k+1}|/|X_k| = L < \infty$, *then* $\iota_{\mathcal{X}}$ *is Lipschitz continuous w.r.t.* $(d_{\mathcal{B},\mathbb{N}}, d_{\mathcal{B},\mathcal{X}})$ *with Lipschitz constant L, and* $\kappa_{\mathcal{X}}$ *is Lipschitz continuous w.r.t.* $(d_{\mathcal{B},\mathcal{X}}, d_{\mathcal{B},\mathbb{N}})$ *with Lipschitz constant L. In particular,* $\iota_{\mathcal{X}}$ *and* $\kappa_{\mathcal{X}}$ *are homeomorphisms between the Besicovitch spaces* $\mathcal{B}^{\mathbb{N},S}$ *and* $\mathcal{B}^{X,S}_{\mathcal{X}}$. *Moreover, if* $L = 1$, *then they are isometries.*

Proof. For every $n \in \mathbb{N}$ there exists $k \in \mathbb{N}$ such that $|X_k| \le n < |X_{k+1}|$. Let $c_1, c_2 \in S^X$: by definition of $\kappa_{\mathcal{X}}$,

$$H_{X_k}(c_1, c_2) \le H_{\{0,\dots,n-1\}}(\kappa_{\mathcal{X}}(c_1), \kappa_{\mathcal{X}}(c_2)) \le H_{X_{k+1}}(c_1, c_2).$$

Consequently,

$$\frac{|X_k|}{n} \cdot \frac{H_{X_k}(c_1, c_2)}{|X_k|} \le \frac{H_{\{0,\dots,n-1\}}(\kappa_{\mathcal{X}}(c_1), \kappa_{\mathcal{X}}(c_2))}{n} \le \frac{H_{X_{k+1}}(c_1, c_2)}{|X_{k+1}|} \cdot \frac{|X_{k+1}|}{n}.$$

But because of the choice of n and k, and since $1/L = \liminf_{k\in\mathbb{N}} |X_k|/|X_{k+1}|$, for any $\varepsilon > 0$, if n is large enough then $|X_k|/n \in [|X_k|/|X_{k+1}|, 1] \subseteq [(1/L) - \varepsilon, 1]$ and $|X_{k+1}|/n \in [1, |X_{k+1}|/|X_k|] \subseteq [1, L + \varepsilon]$. Thus,

$$\frac{1}{L} \cdot d_{\mathcal{B},\mathcal{X}}(c_1, c_2) \le d_{\mathcal{B},\mathbb{N}}(\kappa_{\mathcal{X}}(c_1), \kappa_{\mathcal{X}}(c_2)) \le L \cdot d_{\mathcal{B},\mathcal{X}}(c_1, c_2).$$

Rightmost inequality tells that $\kappa_{\mathcal{X}}$ is Lipschitz continuous with Lipschitz constant L. Leftmost inequality, rewritten in the form $d_{\mathcal{B},\mathcal{X}}(\iota_{\mathcal{X}}(w_1), \iota_{\mathcal{X}}(w_2)) \le L \cdot d_{\mathcal{B},\mathbb{N}}(w_1, w_2)$ with $w_i = \kappa_{\mathcal{X}}(c_i)$, tells that $\iota_{\mathcal{X}}$ is Lipschitz continuous with Lipschitz constant L. □

Corollary 9. *Let V be a* FSOG *for G. Then* $\mathcal{B}^{G,S}_V$ *is homeomorphic to* $\mathcal{B}^{\mathbb{N},S}$.

Proof. If $X_k = D_{k,V}$ for some FSOG V then $|X_{k+1}| \le (1 + 2|V|) \cdot |X_k|$. □

Corollary 10. *Let* $\mathcal{X}$ *be a sequence of disks on* $\mathbb{Z}^d$ *and let* $\mathcal{B} = \mathcal{B}^{\mathbb{Z}^d,S}_{\mathcal{X}}$.

1. $\mathcal{B}$ *is complete.*
2. $\mathcal{B}$ *is arcwise connected.*
3. $\mathcal{B}$ *is infinite-dimensional, hence has the power of continuum.*
4. $\mathcal{B}$ *is not separable, hence not second countable, hence not Lindelöf, hence not compact, hence not totally bounded.*
5. $\mathcal{B}$ *is not locally compact.*

Corollary 9 states that, in the case of the Euclidean groups, the topology does not change with the dimension. The methods employed, however, do not (seem to) allow to be used for the Weyl space, because in general, ι_{χ} does not "behave well" with respect to translations: we cannot be sure that the translate of the image is the image of the translate.

Challenge 3. Prove that $\mathcal{W}^{\mathbb{Z}^d,S}_V$ is homeomorphic to $\mathcal{W}^{\mathbb{Z},S}_V$.

3.2. Periodic configurations in $\mathcal{B}$ *and* $\mathcal{W}$

A noteworthy property of the Besicovitch distance on $S^{\mathbb{Z}}$, other than invariance by translations, is that it is positive on distinct periodic configurations: each class of Besicovitch (and, consequently, Weyl) equivalence has *at most one* periodic representative. But there is more than this: contrary to the case of the product topology, periodic configurations are not dense in either $\mathcal{B}$ or $\mathcal{W}$. In our pursuit of proving (or disproving!, no one ever said that everything has to be the same) similar results of general groups G, the first question to ask ourselves is: *what does it mean*, for a configuration, to be periodic?

Let us start with $G = \mathbb{Z}$. We know that $c \in S^{\mathbb{Z}}$ is periodic if it is the juxtaposition of infinitely many copies of some $w \in S^*$: that is, if there is some number k (*e.g.*, the length of w) such that $c(i + kj) = c(i)$ for every $i, j \in \mathbb{Z}$. But $c(i + kj) = c^{kj}(i)$ because of our definition: which means that c is periodic of period k if it is *invariant by action of* $k\mathbb{Z}$. And here we have a hint of the second "ingredient" we need for the "recipe" of periodicity: the *index* of $k\mathbb{Z}$ in $\mathbb{Z}$ is k, which is a finite value.

Example 3. In the standard case, let $c(i) = i \mod 2$. Then $c^k = c$ if and only if k is an even integer. The stabilizer of c is the subgroup $2\mathbb{Z}$ of $\mathbb{Z}$, which has the set $\{0, 1\}$ of representatives of its right cosets, and so has index 2.

Let us now switch from the line to the plane. We are at ease with configurations that are made of rectangular *pictures*; but what about the one in Fig. 2? Our intuition tells us that it has full right of being called "periodic": after all, it displays a scheme that repeats itself.

A definition of periodicity that holds for arbitrary groups and does not depend on "the shape of the period", is the following.

Definition 8. Let G be a group and let $c : G \to Y$ for some set Y. The *stabilizer* of c is the subgroup $\mathrm{St}(c) = \{g \in G \mid c^g = c\}$. c is *periodic* if $[G : \mathrm{St}(c)] < \infty$. c is H-*periodic* if $H \le \mathrm{St}(c)$ and $[G : H] < \infty$. We call $\mathrm{Per}(G, Y)$ the set of periodic configurations in Y^G.

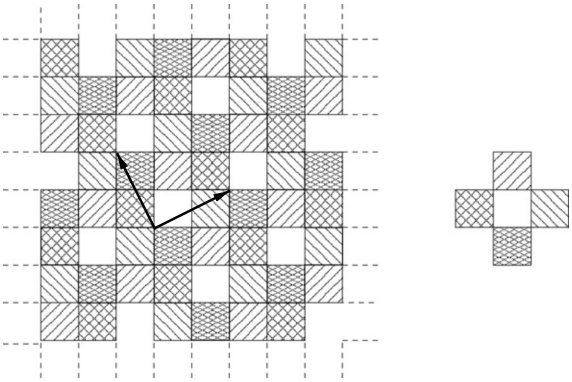

Fig. 2. A 2D periodic configuration which displays a repetition of a cross-shaped pattern. The arrows displayed are generators of the stabilizer.

Definition 8 extends to arbitrary groups the usual definition of periodicity on $\mathbb{Z}$. Moreover, since intersection of groups of finite index has finite index, two periodic configurations have a common *period shape*. We stress that $\mathrm{St}(c)$ is not necessarily a normal subgroup.

Definition 9. A group G is *residually finite* (briefly, r.f.) if the intersection of all its subgroups of finite index is trivial.

$\mathbb{Z}^d$ is r.f. for every d, because the subgroup $H_{j,k} = \{(x_1, \ldots, x_d) \mid x_j \in k\mathbb{Z}\}$ has index k in $\mathbb{Z}^d$, and $x \in \bigcap_{j,k} H_{j,k}$ iff each coordinate of x is divisible by all integers—which is only possible if it is zero. Free groups are also r.f., although the proof is more complicated.

Proposition 11. *Let G be a group, Y a finite set, $|Y| \geq 2$. Let $\mathcal{C} = Y^G$ with the product topology.*

1. *If G is not r.f. then* $\mathrm{Per}(G, Y)$ *is not dense in* $\mathcal{C}$.
2. *If G is both r.f. and f.g. then* $\mathrm{Per}(G, Y)$ *is dense in* $\mathcal{C}$.

Proof (*Sketch; cf. [13, Lemma 2.3.2 and Theorem 2.3.3 and 2.3.4]*). If G is not r.f., then there is some $g \neq 1$ common to all the stabilizers, so that any c s.t. $c(1) \neq c(g)$ is not the limit of a sequence of periodic configurations. If G is f.g., then the disks of finite radius are finite, and if it is also r.f., then for any $n \geq 0$ there exists $H_n \leq G$ of finite index such that $H_n \cap D_n = \{1\}$. □

As a consequence of Proposition 11, periodic d-dimensional configurations and periodic configurations on the free group $\mathbb{F}_2$ form a dense set in the product topology. While the first is immediate to check, the other one is much less intuitive to visualize.

We now want to check what happens to periodic configurations under the Besicovitch or Weyl equivalence. Shall different periodic configurations belong to different classes? And if they do, what is the *reason* that makes them behave as such?

Lemma 12 (*Lemma 3.10 of [8]*). *Let $\mathcal{X}$ be amenable and N be a (U, W)-net with $|U|, |W| < \infty$. Then* $\mathrm{dens\,inf}_{\mathcal{B},\mathcal{X}} N \geq 1/|W|$ *and* $\mathrm{dens\,sup}_{\mathcal{B},\mathcal{X}} N \leq 1/|U|$.

Proof. It is safe to suppose $1 \in U \cap W$. Every $x \in X_n$ is $x = \nu_1 u$ for *at most* one $\nu_1, u \in N \times U$ and $x = \nu_2 w$ for *at least* one $\nu_2, w \in N \times W$: these imply $\nu_1 \in X_n^{+U}$ and $\nu_2 \in X_n^{+W}$, so that

$$|U| \cdot |N \cap X_n^{+U}| \leq |X_n| \leq |W| \cdot |N \cap X_n^{+W}|.$$

Since $1 \in U$, $N \cap X_n = (N \cap X_n^{+U}) \setminus (N \cap \partial_U X_n)$; similar for W. Hence,

$$|U| \cdot |N \cap X_n| \leq |X_n| - |U| \cdot |N \cap \partial_U X_n| \tag{20}$$

and

$$|X_n| - |W| \cdot |N \cap \partial_W X_n| \leq |W| \cdot |N \cap X_n| \tag{21}$$

By dividing (20) by $|U| \cdot |X_n|$, (21) by $|W| \cdot |X_n|$, and merging, we get

$$\frac{1}{|W|} - \frac{|N \cap \partial_W X_n|}{|X_n|} \leq \frac{|N \cap X_n|}{|X_n|} \leq \frac{1}{|U|} - \frac{|N \cap \partial_U X_n|}{|X_n|}. \tag{22}$$

Since $\mathcal{X}$ is amenable, the terms with the minus sign in (22) vanish for $n \to \infty$. □

Theorem 13. *Let $\{X_n\}$ be an amenable sequence for G. Let c_1 and c_2 be distinct periodic configurations. Then*

$$d_{\mathcal{B},\{X_n\}}(c_1, c_2) \geq \frac{1}{[G : \mathrm{St}(c_1) \cap \mathrm{St}(c_2)]^2} > 0. \tag{23}$$

Proof. Let U be a set of representatives of the right cosets of $H = \mathrm{St}(c_1) \cap \mathrm{St}(c_2)$ in G: then $|U| = [G : H] < \infty$ and $c_1(u) \neq c_2(u)$ for some $u \in U$. Let $\phi(x) = xu$ for every $x \in H$: then $\phi(H)$ is a $(\{1_G\}, U^{-1}U)$-net as explained in Section 2.2, and $c_1(g) \neq c_2(g)$ for all $g \in \phi(H)$. By Lemma 12, $d_{\mathcal{B},\{X_n\}}(c_1, c_2) \geq \text{dens sup}_{\{X_n\}}\phi(H) \geq 1/|U^{-1}U| \geq 1/|U|^2$. □

As a consequence of Theorem 13, if $\mathcal{X}$ is amenable, then each element of $\mathcal{W}_\mathcal{X}$ has at most a single periodic representative, which is also unique in the corresponding element of $\mathcal{B}_\mathcal{X}$.

Corollary 14. *Let V be a* FSOG *for $\mathbb{Z}^d$. The set of (classes of) periodic configurations is not dense in either $\mathcal{B}_V^{\mathbb{Z}^d,S}$ or $\mathcal{W}_V^{\mathbb{Z}^d,S}$.*

Proof. The countable set of (classes of) periodic configurations is not dense in $\mathcal{B}_V^{\mathbb{Z}^d,S}$ because the space is not separable, and is not dense in $\mathcal{W}_V^{\mathbb{Z}^d,S}$ because the latter space is finer-grained than the former. □

Example 4. Consider the sequence $X_n = \{-n, \ldots, n-1\}^d$. It is straightforward to check that $d_{B,\{X_n\}}(c_1, c_2) = d_{B,\mathcal{M}_d}(c_1, c_2)$ whatever c_1 and c_2 are.

Let $s_0, s_1 \in S$ with $s_0 \neq s_1$. Consider $c \in S^{\mathbb{Z}^d}$ defined by

$$c(x) = \begin{cases} s_0 & \text{if } x_1 < 0, \\ s_1 & \text{if } x_1 \geq 0. \end{cases} \tag{24}$$

Let $V = \{v_1, \ldots, v_d\}$ be the standard set of generators for $\mathbb{Z}^d$, *i.e.*, $(v_i)_j = \delta_{i,j}$. Let c' be a periodic configuration. Since $\mathrm{St}(c')$ is a subgroup of finite index in the f.g. group $\mathbb{Z}^d$, it has itself a finite set Σ of generators. But any $\sigma \in \Sigma$ can be rewritten as a linear combination of the v_i's, hence $\mathrm{St}(c')$ has a subgroup of finite index $H = \langle LV \rangle$ for some $L > 0$, and c' may be seen as being H-periodic.

Now, for any $x = (x_1, \ldots, x_d) \in X_{mL}$ with $x_1 \geq 0$, let $x' = (x_1 - mL, \ldots, x_d)$. Then either $c'(x) \neq c(x)$, or $c'(x') \neq c(x')$, or both. Thus, $\frac{H_{X_{mL}}(c,c')}{(2mL)^d} \geq \frac{1}{2}$ for any m, so that $d_{W,\mathcal{M}_d}(c, c') \geq d_{B,\mathcal{M}_d}(c, c') \geq 1/2$.

Having "good" dense subsets is useful in the study of dynamical systems, since several properties hold for the whole space once they are proved on a dense set. The product space has the periodic configurations; do $\mathcal{B}$ and $\mathcal{W}$ have some as well? A partial answer was provided in [2].

Definition 10. A 1D configuration $c : \mathbb{Z} \to S$ is *Toeplitz* if for every $x \in \mathbb{Z}$ there exists $p = p(x) > 0$ such that $c(x+kp) = c(x)$ for every $k \in \mathbb{Z}$.

In other words, a Toeplitz configuration is one in which every pattern occurs periodically, *though the size of the period needs not be bounded*. Thus, any periodic configuration is Toeplitz, but not vice versa. Blanchard et al. [2] prove that 1D Toeplitz configurations are dense in $\mathcal{B}$ – which, incidentally, implies that they are *un*countably many – but not in $\mathcal{W}$.

Challenge 4. Prove that Toeplitz d-dimensional configurations are dense in $\mathcal{B}_V$.

Challenge 5. Define Toeplitz configurations on more general groups, and check whether they are dense in $\mathcal{B}$.

Incidentally, we observe another consequence of Lemma 12, with which we close the section.

Theorem 15. *Let $\mathcal{X}$ be an amenable sequence for G. Then $(\mathcal{B}_\mathcal{X}, d_{\mathcal{B},\mathcal{X}})$ is a perfect metric space.*

Proof. Let $c \in \mathcal{C}$, $\varepsilon > 0$. Let $E \subseteq G$ be finite and $\varepsilon \cdot |E| > 1$. Let N be a (E, EE^{-1})-net. Let $c_\varepsilon \in \mathcal{C}$ satisfy $c_\varepsilon(g) = c(g)$ iff $g \notin N$. Then $d_{\mathcal{B},\mathcal{X}}(c, c_\varepsilon) = \text{dens sup}_{\mathcal{B},\mathcal{X}}N \in \left[1/|EE^{-1}|, 1/|E|\right] \subseteq (0, \varepsilon)$. □

4. Patterns in the Besicovitch and Weyl spaces

Consider a pattern p and a configuration c. We can compute the Besicovitch and Weyl upper and lower densities of the set $\mathrm{occ}(p, c)$ of the occurrences of p in c. But before embarking in this task, we want to have at least a hint of the *information* that these quantities can provide.

The most naive check that we can do is: Suppose $\text{dens sup}_{\mathcal{B},\mathcal{X}}\mathrm{occ}(p, c) = 0$. Can we infer that there exists c' such that $d_{\mathcal{B},\mathcal{X}}(c, c') = 0$ and p does not occur in c'? The answer, as are the answers to most of the naive guesses, is negative.

Example 5. In the standard case, let $c \in S^\mathbb{Z}$ such that $c(i) = 1$ iff $i \geq 0$. Choose $E = \{0, 1\}$ and define $p : E \to S$ as $p(i) = i$: then $\text{dens sup}_{\mathcal{B},\mathcal{X}}(p, c) = 0$; However, any c' s.t. $d_{\mathcal{B},\mathcal{X}}(c, c') = 0$ must have at least one occurrence of p. Indeed, if $\mathrm{occ}(p, c') = \emptyset$, then either $c'(x) = 0$ for all $x \in \mathbb{Z}$, or $c'(x) = 1$ for all $x \in \mathbb{Z}$, or there exists $y \in \mathbb{Z}$ such that $c'(x) = 0$ iff $x \geq y$. In the first two cases, $d_{\mathcal{B},\mathcal{X}}(c, c') = \frac{1}{2}$; in the third one, $c'(x) \neq c(x)$ for all x such that $|x| > |y|$, so that $d_{\mathcal{B},\mathcal{X}}(c, c') = 1$.

So, let us suppose $\text{dens sup}_{\mathcal{B},\mathcal{X}}\mathrm{occ}(p, c_1) = 0$ and $\text{dens sup}_{\mathcal{B},\mathcal{X}}\mathrm{occ}(p, c_2) > 0$. What can we infer about $d_{\mathcal{B},\mathcal{X}}(c_1, c_2)$? Intuitively, it cannot be zero: the two evaluations suggest that there is a set of positive upper measure made of occurrences of p in c_2 that are not occurrences of p in c_1; which in turn should suggest that the two configurations differ on a set of positive upper measure. And here is the question we should always ask ourselves: *Is our intuition correct?*

The answer is welcome.

Theorem 16. *Let* $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) = 0$*. For every pattern* p*,* $\text{dens sup}_{\mathcal{B},\mathcal{X}}\text{occ}(p, c_1) = \text{dens sup}_{\mathcal{B},\mathcal{X}}\text{occ}(p, c_2)$ *and* $\text{dens inf}_{\mathcal{B},\mathcal{X}}\text{occ}(p, c_1) = \text{dens inf}_{\mathcal{B},\mathcal{X}}\text{occ}(p, c_2)$*. These hold for Weyl pseudo-distance and densities as well.*

Proof. Let $D_i = \text{dens sup}_{\mathcal{B},\mathcal{X}}\text{occ}(p, c_i)$. Suppose, for the sake of contradiction, $D_1 > D_2$. Let $\delta > 0$ and let $\{n_k\} \subseteq \mathbb{N}$ be a strictly increasing sequence such that $|\text{occ}(p, c_1) \cap X_{n_k}| \geq (D_2 + \delta) |X_{n_k}|$ for all $k \in \mathbb{N}$. On the other hand, for all k large enough,

$$|\text{occ}(p, c_2) \cap X_{n_k}| < \left(D_2 + \frac{\delta}{2}\right) |X_{n_k}| \ . \tag{25}$$

Let E be the support of p. There are at least $\left\lfloor \frac{1}{2}\delta|X_{n_k}| \right\rfloor$ points $g \in G$ such that $gE \subseteq X_{n_k}$, $c_1^g\big|_E = p$, and $c_2^g\big|_E \neq p$: hence, $H_{X_{n_k}}(c_1, c_2) \geq \left\lfloor \frac{1}{2|E|}\delta|X_{n_k}| \right\rfloor$ for every $k \in \mathbb{N}$. This implies $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) \geq \frac{\delta}{2|E|}$, against the hypothesis that $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) = 0$. The case $D_1 < D_2$ is analogous, as is the proof for lower densities.

The proof in the Weyl case is similar, replacing (25) with

$$|\text{occ}(p, c_2) \cap g_k X_{n_k}| < \left(D_2 + \frac{\delta}{2}\right) |X_{n_k}| \tag{26}$$

for suitable $g_k \in G$ s.t. $|\text{occ}(p, c_1) \cap g_k X_{n_k}| \geq (D_2 + \delta) |X_{n_k}|$ for every $k \in \mathbb{N}$. □

Observe that Theorem 16 holds if $\mathcal{B} = \mathcal{B}_{\mathcal{X}}^{X,Y}$ for any two sets X, Y, and if $\mathcal{W} = \mathcal{W}_{\mathcal{X}}^{G,Y}$ for any group G and set Y. The argument is purely counting, and does not take into account any particular set size or group structure.

Theorem 16 states that the Besicovitch upper and lower densities of the occurrences of a pattern in a configuration are preserved by the Besicovitch equivalence, and similarly for the Weyl case. This allows us to speak of the upper density of a pattern in a point of the Besicovitch or Weyl space.

The condition of Theorem 16 is sufficient, but not necessary.

Example 6. In the standard case, let $c_1(x) = 1$ iff $x \geq 0$, $c_2(x) = 1$ iff $x < 0$. It is straightforward to check that any density of any pattern is the same for c_1 and c_2. However, $d_{\mathcal{B},V}(c_1, c_2) = 1$.

Theorem 16 gives us confidence to check other pattern-related properties. Of these, the *higher block codes* come to our mind as a standard tool from symbolic dynamics. We recall that, given an infinite word w on an alphabet S, its k-higher block transform is the word $w^{[k]}$ on the alphabet S^k such that $w_i^{[k]} = (w_i, w_{i+1}, \ldots, w_{i+k-1})$ for every $i \in \mathbb{Z}$. In other words, to construct $w^{[k]}$, we take "snapshots" of w through the "window" $\{0, \ldots, k-1\}$. We may extend this procedure to arbitrary groups and windows, as follows.

Definition 11. Let $E \subseteq G$ s.t. $|E| < \infty$ and $1_G \in E$. The *E-shaped block transform* (briefly, E-SBT) of $c : G \to S$ is the configuration $c^{[E]} : G \to S^E$ defined as $c^{[E]}(g) = c^g|_E$, that is,

$$(c^{[E]}(g))(e) = c(ge) \ \forall g \in G, e \in E \ . \tag{27}$$

The value of $c^{[E]}$ at g is the set of values of c on a set shaped as E and based on g. This is what is done in the 1D case, where the Nth higher block code is obtained by taking $E = \{0, \ldots, N-1\}$ (cf. [15, Section 1.4]).

The construction of higher block transforms commutes with translations. In fact, let $g \in G$: then for every $h \in G, e \in E$

$$\left((c^{[E]})^g(h)\right)(e) = (c^{[E]}(gh))(e) = c(ghe) = c^g(he) = ((c^g)^{[E]}(h))(e),$$

so that $\left(c^{[E]}\right)^g = (c^g)^{[E]}$. This commutation property is satisfied even on non-commutative groups, because translations operate via *left* multiplication, while E-SBT operate via *right* multiplication.

As we might have guessed, neither the Besicovitch nor the Weyl distance are preserved in the passage to E-shaped block transform.

Example 7. In the standard case, let $N = 2$ (*i.e.*, $E = \{0, 1\}$), $c_1(x) = 0$ for all x, $c_2(x) = x \mod 2$. Then $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) = \frac{1}{2}$ but $d_{\mathcal{B},\mathcal{X}}(c_1^{[E]}, c_2^{[E]}) = 1$.

However, it looks reasonable that $d_{\mathcal{B}}$ and $d_{\mathcal{W}}$ are not changed "too much" in the passage from two configurations to their E-SBT. And it is so, provided that the sequence $\mathcal{X}$ has a property of the kind *"the orange grows faster than the peel"*.

Theorem 17. *Let* $\mathcal{X}$ *be an exhaustive sequence for* G*. Let* $c_1, c_2 \in S^G$*. Let* E *be a finite subset of* G *such that* $1_G \in E$*.*

1. $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) \leq d_{\mathcal{B},\mathcal{X}}\left(c_1^{[E]}, c_2^{[E]}\right)$.
2. *If* $\mathcal{X}$ *is amenable then*

$$d_{\mathcal{B},\mathcal{X}}\left(c_1^{[E]}, c_2^{[E]}\right) \leq |E| \cdot d_{\mathcal{B},\mathcal{X}}(c_1, c_2). \tag{28}$$

3. *In particular, if* $\mathcal{X}$ *is amenable then the following are equivalent:*

(a) $d_{\mathcal{B},\mathcal{X}}(c_1, c_2) = 0$.
(b) $d_{\mathcal{B},\mathcal{X}}\left(c_1^{[E]}, c_2^{[E]}\right) = 0$ *for some finite* E *s.t.* $1_G \in E$.
(c) $d_{\mathcal{B},\mathcal{X}}\left(c_1^{[E]}, c_2^{[E]}\right) = 0$ *for all finite* E *s.t.* $1_G \in E$.

The same hold with $d_{\mathcal{W},\mathcal{X}}$ *in place of* $d_{\mathcal{B},\mathcal{X}}$.

Proof. If $c_1(x) \neq c_2(x)$ then $c_1^{[E]}(x) \neq c_2^{[E]}(x)$ as well, from which point 1 follows easily.

For point 2, given $U \subseteq G$, to each $x \in U$ s.t. $c_1^{[E]}(x) \neq c_2^{[E]}(x)$ correspond no more than $|E|$ points $y \in UE$ s.t. $c_1(y) \neq c_2(y)$, *i.e.*,

$$H_U(c_1^{[E]}, c_2^{[E]}) \leq |E| \cdot H_{UE}(c_1, c_2). \tag{29}$$

But $UE = U^{+E^{-1}} = U \sqcup \partial_{E^{-1}}U$ because $1_G \in E$, hence $H_{UE}(c_1, c_2) \leq H_U(c_1, c_2) + |\partial_{E^{-1}}U|$. From this, (29), and the fact that $\{X_n\}$ is amenable follows the thesis.

The corresponding statements for $d_{\mathcal{W},\mathcal{X}}$ can be proved similarly, replacing (29) by

$$H_U((c_1^{[E]})^g, (c_2^{[E]})^g) \leq |E| \cdot H_{UE}(c_1^g, c_2^g) \quad \forall g \in G, \tag{30}$$

allowed because translation and E-SBT commute. $\square$

5. Sliding block codes on Besicovitch and Weyl spaces

At this point, our knowledge of the Besicovitch and Weyl space is sufficiently broad to stimulate us towards studying the behavior of cellular automata – and, more in general, sliding block codes – on the new environment. For this, we can rely on the results in the classical case [9,3,2] and the previous findings in the broader context [8].

The first thing that we want, is that SBC are well defined on $\mathcal{B}$ and $\mathcal{W}$. In [8, Theorem 3.7] we prove that any CA is Lipschitz continuous w.r.t. $d_{\mathcal{B},\mathcal{X}}$ – which is a sufficient condition for well-definition – provided $\mathcal{X}$ is either amenable or a sequence of disks. The argument extends straightforwardly not only to arbitrary SBC, but also to the Weyl distance: for the convenience of the reader, we rewrite it here.

Theorem 18. *Let* G *be a f.g. group and let* $\mathcal{K} = \langle S, T, \mathcal{N}, f\rangle$ *be a* SBC *over* G.

1. *If* $\mathcal{X}$ *is amenable, then* $F_{\mathcal{K}}$ *is Lipschitz continuous w.r.t.* $(d_{B,\mathcal{X}}, d_{B,\mathcal{X}})$ *and* $(d_{W,\mathcal{X}}, d_{W,\mathcal{X}})$*, with* $L = |\mathcal{N} \cup \{1_G\}|$.
2. *If* $X_n = D_{n,V}$ *for all* n *for some* FSOG V*, and* $\mathcal{N} \subseteq D_{r,V}$*, then* $F_{\mathcal{K}}$ *is Lipschitz continuous w.r.t.* $(d_{B,\mathcal{X}}, d_{B,\mathcal{X}})$ *and* $(d_{W,\mathcal{X}}, d_{W,\mathcal{X}})$*, with* $L = (\gamma_V(r))^2$.

Proof. We give the proof for the Weyl pseudo-distance; the argument for the Besicovitch pseudo-distance is similar, and can be found (for CA) in [8].

First, observe that, if $X \subseteq G$ and $\mathcal{N} \subseteq E$, then

$$H_X(F_{\mathcal{K}}(c_1), F_{\mathcal{K}}(c_2)) \leq |E| \cdot H_{X^{+E}}(c_1, c_2).$$

Moreover, since $\mathcal{A}$ is a CA, $(F_{\mathcal{K}}(c))^g = F_{\mathcal{K}}(c^g)$ for every $g \in G, c \in \mathcal{C}$.

If $\mathcal{X}$ is amenable, put $E = \mathcal{N} \cup \{1_G\}$. Then for every $g \in G$

$$\begin{aligned} H_{X_n}((F_{\mathcal{K}}(c_1))^g, (F_{\mathcal{K}}(c_2))^g) &= H_{X_n}(F_{\mathcal{K}}(c_1^g), F_{\mathcal{K}}(c_2^g)) \\ &\leq |E| \cdot H_{X_n^{+E}}(c_1^g, c_2^g) \\ &\leq |E| \cdot \left(H_{X_n}(c_1^g, c_2^g) + |\partial_E X_n|\right), \end{aligned}$$

thus

$$\sup_{g \in G} H_{X_n}((F_{\mathcal{K}}(c_1))^g, (F_{\mathcal{K}}(c_2))^g) \leq |E| \cdot \left(|\partial_E X_n| + \sup_{g \in G} H_{X_n}(c_1^g, c_2^g)\right),$$

from which the thesis since $\mathcal{X}$ is amenable.

If $X_n = D_{n,V}$, put $E = D_{r,V}$ with $\mathcal{N} \subseteq D_{r,V}$. Then for every $g \in G$

$$H_{n,V}((F_{\mathcal{K}}(c_1))^g, (F_{\mathcal{K}}(c_2))^g) = H_{n,V}(F_{\mathcal{K}}(c_1^g), F_{\mathcal{K}}(c_2^g)) \leq \gamma_V(r) \cdot H_{n+r,V}(c_1^g, c_2^g),$$

and since $\gamma_S(n+r) \leq \gamma_S(n)\gamma_S(r)$, for all $n \in \mathbb{N}$

$$\frac{H_{n,V}((F_{\mathcal{K}}(c_1))^g, (F_{\mathcal{K}}(c_2))^g)}{\gamma_V(n)} \leq (\gamma_V(r))^2 \cdot \frac{H_{n+r,S}(c_1^g, c_2^g)}{\gamma_V(n+r)}.$$

Point 2 is then achieved similarly to point 1. $\square$

Theorem 18 provides sufficient conditions for $F_{\mathcal{K}} : S^G \to T^G$ to induce transformations $F_{\mathcal{B}} : \mathcal{B}_{\mathcal{X}}^{G,S} \to \mathcal{B}_{\mathcal{X}}^{G,T}$ and $F_{\mathcal{W}} : \mathcal{W}_{\mathcal{X}}^{G,S} \to \mathcal{W}_{\mathcal{X}}^{G,T}$ through

$$F_{\mathcal{B}}([c]_{\mathcal{B}}) = [F_{\mathcal{K}}(c)]_{\mathcal{B}} \tag{31}$$

and

$$F_{\mathcal{W}}([c]_{\mathcal{W}}) = [F_{\mathcal{K}}(c)]_{\mathcal{W}}. \tag{32}$$

Corollary 19. *Let G be a group of polynomial growth (e.g., $\mathbb{Z}^d$), let V be a* FSOG *for G (e.g., $\mathcal{M}_d$) and let $\mathcal{X} = \{D_{n,V}\}_{n\geq 0}$. Then (31) and (32) are well defined.*

In the context of Besicovitch spaces on $\mathbb{Z}^d$ w.r.t. some FSOG, being translation-invariant and Lipschitz continuous in the Besicovitch distance are necessary conditions for being induced by a CA. However, they are not sufficient.

Example 8 (*Example 7 of [3]*)**.** In the standard case, let $T = \{0, 1, \perp\} = S \cup \{\perp\}$. Consider the function $F : T^{\mathbb{Z}} \to T^{\mathbb{Z}}$ defined as follows:

- If $c(x) = \perp$ then $(F(c))(x) = \perp$.
- If $c(x) \in S$, let $l = \sup\{z \leq x - 1 \mid c(z) \in S\}$, $r = \inf\{z \geq x + 1 \mid c(z) \in S\}$, with the conventions $\sup \emptyset = -\infty$, $\inf \emptyset = +\infty$, and let $\mathcal{I} = \{x, l, r\} \cap \mathbb{Z}$. Then $(F(c))(x) = \sum_{i\in\mathcal{I}} c(i) \mod 2$.

F is clearly translation-commuting. Moreover, the value of c at a point can influence the value of $F(c)$ at no more than three points: consequently, F satisfies (1) w.r.t. $(d_{\mathcal{B},\mathcal{X}}, d_{\mathcal{B},\mathcal{X}})$ and $(d_{\mathcal{W},\mathcal{X}}, d_{\mathcal{W},\mathcal{X}})$, with $L = 3$. However, F is not continuous in the product topology, and thus not a CA global function. (Intuitively, F needs an *unbounded* neighborhood.) Indeed, let $c(0) = 0$ and $c(x) = \tau$ if $x \neq 0$; let $c_k(x) = c(x)$ if $x \neq k$, $c_k(k) = 1$. Then $d_V(c_k, c) = 2^{-k}$ but $d_V(F(c_k), F(c)) = 1$.

In a recent paper [17], Müller and Spandl give a characterization of those transformations of either $\mathcal{B}$ or $\mathcal{W}$ which may be induced by one-dimensional CA.

Proposition 20 (*Müller and Spandl, 2009*)**.** *Let $X_n = [-n, \dots, n]$. Let $\mathcal{B} = \mathcal{B}_{\mathcal{X}}^{\mathbb{Z},S}$ for some finite alphabet S of the form $S = \mathbb{Z}/m\mathbb{Z} = \{0, \dots, m-1\}$. Let $f : \mathcal{B} \to \mathcal{B}$ satisfy the following properties:*

1. *$f \circ \sigma_{\mathcal{B}} = \sigma_{\mathcal{B}} \circ f$, where $\sigma_{\mathcal{B}}$ is the operator on $\mathcal{B}$ induced by the shift.*
2. *$f(\mathrm{Per}(\mathbb{Z}, S)) \subseteq \mathrm{Per}(\mathbb{Z}, S)$.*
3. *There exists $C > 0$ such that for every $A \subseteq \mathbb{Z}$, $c_1, c_2 \in S^{\mathbb{Z}}$*

$$d_{\mathcal{B}}(f([\rho_A(c_1)]_{\mathcal{B}}), f([\rho_A(c_2)]_{\mathcal{B}})) \leq C \cdot d_{\mathcal{B}}([\rho_{A^{+r}}(c_1)]_{\mathcal{B}}, [\rho_{A^{+r}}(c_2)]_{\mathcal{B}}), \tag{33}$$

where $\rho_A(c)$ coincides with c on A, and is 0 otherwise.

Then there exists a cellular automaton $\mathcal{A}$ of radius r whose global function satisfies $d(f([c]_{\mathcal{B}}), [F_{\mathcal{A}}(c)]_{\mathcal{B}}) = 0$ for all $c \in S^{\mathbb{Z}}$.

The same holds for the Weyl space $\mathcal{W}_{\mathcal{X}}^{S,\mathbb{Z}}$, with $d_{\mathcal{W}}$ in place of $d_{\mathcal{B}}$.

Proposition 20 is the best possible result for one-dimensional CA, in the following sense. Since $\mathcal{B}$ and $\mathcal{W}$ are quotient spaces, there might well be several functions $F : S^{\mathbb{Z}} \to S^{\mathbb{Z}}$ that induce the same $f : \mathcal{B}^{\mathbb{Z},S} \to \mathcal{B}^{\mathbb{Z},S}$: therefore, given f, we can at most hope to find a CA whose global function is indistinguishable from f on $\mathcal{B}$. Proposition 20 provides a sufficient (and necessary) condition for this to happen.

Now, to get a CA, we need local definability at any point and commutation with translations. Müller and Spandl prove that point 3 captures *local* definability; however, point 1 is *not enough* to recover commutation with translations, or (which is the same given the previous point) *uniform* definability at each point, because to have that, the existence of an invariant dense set (in the product topology!) is required, which is what point 2 states.

Challenge 6. Prove Proposition 20 for SBC on r.f. and/or amenable groups.

5.1. Properties of sliding block codes on $\mathcal{B}$ and $\mathcal{W}$

As we have said in the introduction, a key feature of $\mathcal{B}$ and $\mathcal{W}$ in the 1D case is that properties of CA are linked to properties of induced maps. In this subsection, we will see several such properties; some apply to arbitrary SBC, while others require that we are dealing with CA.

A first consequence of Theorem 18 is the preservation of *equicontinuity*. Recall that, given a function $f : X \to X$, the *k-th iterate of f* is defined as $f^{(0)}(x) = x$ and $f^{(k+1)}(x) = f(f^{(k)}(x))$ for every $x \in X$.

Definition 12. Let d be a pseudo-distance on X and $f : X \to X$ be a function. f is *equicontinuous on X w.r.t. d* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that, if $d(x_1, x_2) < \delta$, then $d(f^{(k)}(x_1), f^{(k)}(x_2)) < \varepsilon$ for every $k \in \mathbb{N}$.

In other words, $f : X \to X$ is equicontinuous if the sequence of its iterates is continuous on X uniformly w.r.t. the order of iteration.

Theorem 21. *Let $\mathcal{A} = \langle S, \mathcal{N}, f\rangle$ be a* CA *on G. Suppose $F_{\mathcal{A}}$ is equicontinuous in the product topology. Also suppose that at least one of the following is satisfied:*

1. *$\mathcal{X}$ is amenable.*
2. *$\mathcal{X} = \{D_{n,V}\}_{n\geq 0}$ for some* FSOG *V.*

Then all the iterates of $F_{\mathcal{B}}$ and $F_{\mathcal{W}}$ are Lipschitz continuous with the same constant L.

Proof. Let $\mathcal{A}$ be a CA equicontinuous in the product topology. By the considerations above, whatever the FSOG V is, there exists ρ_0 such that, if $c_1|_{D_{\rho_0}} = c_2|_{D_{\rho_0}}$, then $F_{\mathcal{A}}^{(k)}(c_1)(1_G) = F_{\mathcal{A}}^{(k)}(c_2)(1_G)$ for every $k \in \mathbb{N}$. Consequently, any iterate $F_{\mathcal{A}}^{(k)}$ is the global evolution function of some CA of the form $\langle Q, D_{\rho_0}, f_k\rangle$.

We now apply Theorem 18. If the hypotheses are satisfied, then an upper bound for $d_{\mathcal{W},\mathcal{X}}(F_{\mathcal{A}}^{(k)}(c_1), F_{\mathcal{A}}^{(k)}(c_2))$ is $(\gamma_S(\rho_0))^2 \cdot d_{\mathcal{W},\mathcal{X}}(c_1, c_2)$, and an upper bound for $d_{\mathcal{B},\mathcal{X}}(F_{\mathcal{A}}^{(k)}(c_1), F_{\mathcal{A}}^{(k)}(c_2))$ is $(\gamma_S(\rho_0))^2 \cdot d_{\mathcal{B},\mathcal{X}}(c_1, c_2)$. These hold whatever the iteration k is. □

Challenge 7 (*Cf. [3]*)**.** Find links between dynamical properties of $F_{\mathcal{A}}$, $F_{\mathcal{B}}$ and $F_{\mathcal{W}}$ under the hypotheses of Theorem 18. Keep in mind that, by [2, Theorem 4 and Corollary 3], if $G = \mathbb{Z}$ then $F_{\mathcal{B}}$ and $F_{\mathcal{W}}$ cannot be transitive.

The second property that we are going to consider, is injectivity. This is a fairly interesting field, because it is well known that an injective d-dimensional CA is surjective. This phenomenon seems to be widespread, as we are going to see.

Definition 13 (*cf. [20]*)**.** A group G is called *surjunctive* if, for every alphabet S, every injective continuous map $F : S^G \to S^G$ which commutes with translations is surjective.

We stress that Definition 13 does not require G to be f.g. Currently, no example of a non-surjunctive group is known; on the other hand, amenable groups (*e.g.*, $\mathbb{Z}^d$) and residually finite groups (*e.g.*, $\mathbb{F}_2$ and $\mathbb{Z}^d$) have been proved to be surjunctive.

We also stress that Definition 13 cannot be extended from CA to SBC. In fact, if S is a proper subset of T, then the embedding of S^G into T^G is an injective SBC which is not surjective.

Theorem 22. *Let $\mathcal{A}$ be a* CA *and let $F = F_{\mathcal{A}}$ be its global map. Suppose G is surjunctive and $\mathcal{X}$ is either amenable or a sequence of disks. If F is injective then $F_{\mathcal{B}}$ and $F_{\mathcal{W}}$ are injective.*

Proof. Since F is injective and G is surjunctive, F is surjective. By Fact 4 (cf. [6, Corollary 3.11]) $P = F^{-1}$ is the global map of some CA, and is Lipschitz continuous w.r.t. $d_{\mathcal{B}}$ and $d_{\mathcal{W}}$ by Theorem 18.

Suppose then $[c_1]_{\mathcal{B}} \neq [c_2]_{\mathcal{B}}$. Put $\chi_i = F(c_i)$: then $P_{\mathcal{B}}([\chi_1]_{\mathcal{B}}) = [P(\chi_1)]_{\mathcal{B}} \neq [P(\chi_2)]_{\mathcal{B}} = P_{\mathcal{B}}([\chi_2]_{\mathcal{B}})$. Since $P_{\mathcal{B}}$ is a function, $F_{\mathcal{B}}([c_1]_{\mathcal{B}}) = [\chi_1]_{\mathcal{B}} \neq [\chi_2]_{\mathcal{B}} = F_{\mathcal{B}}([c_2]_{\mathcal{B}})$. A similar argument holds for $F_{\mathcal{W}}$. □

The third property that is relevant to us, is surjectivity. This is very well characterized in the product topology through the orphan pattern principle: which is what we enforce to prove the following extension of the main result of [8].

Theorem 23. *Let G be an amenable f.g. group and let $\mathcal{K} = \langle S, T, \mathcal{N}, f\rangle$ be a* SBC *on G. If $\mathcal{X}$ contains an amenable subsequence, then the following are equivalent.*

1. *$F_{\mathcal{K}}$ is surjective.*
2. *For every $c_T \in T^G$ there exists $c_S \in S^G$ such that $d_{\mathcal{B},\mathcal{X}}(F_{\mathcal{K}}(c_S), c_T) = 0$.*
3. *For every $c_T \in T^G$ there exists $c_S \in S^G$ such that $d_{\mathcal{W},\mathcal{X}}(F_{\mathcal{K}}(c_S), c_T) = 0$.*

Proof. (Sketch; cf. [8, Theorem 3.11]) It is straightforward that point 1 implies point 3, which implies point 2; so, let us suppose, for the sake of contradiction, that $\mathcal{K}$ has an orphan pattern p. Let $R = r + n$ where $\mathcal{N} \subseteq D_r$ and $\operatorname{supp} p \subseteq D_n$, and let N be a (D_R, D_{2R})-net.; define $c(g)$ as $p(x^{-1}g)$ if $g \in D_R(x)$ for some $x \in N$, and arbitrary otherwise. Then c and $F_{\mathcal{K}}(c')$ differ at least on the points of a $(\{1_G\}, D_{3R})$-net, and $d_{\mathcal{B}}(c, F_{\mathcal{K}}(c')) \geq 1/\gamma(3R)$ because of Lemma 12. □

Corollary 24. *Let G be a group of polynomial growth (e.g., $\mathbb{Z}^d$), let V be a* FSOG *for G (e.g., $\mathcal{M}_d$) and let $\mathcal{X} = \{D_{n,V}\}_{n\geq 0}$. For an arbitrary* SBC *$\mathcal{K} = \langle S, T, \mathcal{N}, f\rangle$ on G, the following are equivalent.*

1. *$F_{\mathcal{K}} : S^G \to T^G$ is surjective.*
2. *$F_{\mathcal{B}} : \mathcal{B}_V^{G,S} \to \mathcal{B}_V^{G,T}$ is surjective.*
3. *$F_{\mathcal{W}} : \mathcal{W}_V^{G,S} \to \mathcal{W}_V^{G,T}$ is surjective.*

5.2. An orphan pattern principle for $\mathcal{B}$ and $\mathcal{W}$

At this point we have what we need to make an attempt towards a characterization of surjective SBC which refers to the set of the occurrences of a pattern from the point of view of either the Besicovitch or Weyl upper density. Instead of linking surjectivity to single occurrences of patterns (which are meaningless in our quotient topologies) we take into account the *density* of the set of occurrences of given patterns—which, by Theorem 16, is a property of *the equivalence class* of a configuration.

Theorem 25. *Let G be an amenable f.g. group and let $\mathcal{K} = \langle S, T, \mathcal{N}, f \rangle$ be a* SBC *on G. If $\mathcal{X}$ contains an amenable subsequence, then the following are equivalent.*

1. *For every $c_T \in T^G$ there exists $c_S \in S^G$ such that $d_{\mathcal{B},\mathcal{X}}(c_T, F_{\mathcal{K}}(c_S)) = 0$. (Thus, $F_{\mathcal{K}}$ is surjective by Theorem 23.)*
2. *For every finite $E \subseteq G$ and every $p : E \to T$ there exists $c_S \in S^G$ such that* $\text{dens sup}_{\mathcal{B},\mathcal{X}}\text{occ}(p, F_{\mathcal{K}}(c_S)) > 0$.

The same hold with $d_{\mathcal{W}}$ and $\text{dens sup}_{\mathcal{W}}$ *in place of $d_{\mathcal{B}}$ and* $\text{dens sup}_{\mathcal{B}}$.

Proof. Let $p : E \to T$. Suppose $D_R \supseteq E$. Let N be a (D_R, D_{2R})-net. Define $c_T \in T^G$ as

$$c_T(g) = \begin{cases} p(x^{-1}g) & \text{if } \exists x \in N \mid x^{-1}g \in E, \\ \text{arbitrary} & \text{otherwise.} \end{cases} \tag{34}$$

Then each $x \in N$ is an occurrence of p in c, so that $\text{dens sup}_{\mathcal{W},\mathcal{X}}\text{occ}(p, c_T) \geq \text{dens sup}_{\mathcal{B},\mathcal{X}}\text{occ}(p, c_T) \geq \text{dens sup}_{\mathcal{B},\mathcal{X}} N \geq 1/\gamma(2R)$. If $d_{\mathcal{B},\mathcal{X}}(c_T, F_{\mathcal{K}}(c_S)) = 0$ for some $c_S \in S^G$, then $\text{dens sup}_{\mathcal{B},\mathcal{X}}\text{occ}(p, F(c_S)) \geq 1/\gamma(2R)$ as well. The previous statement holds if d_B and dens sup_B are replaced by d_W and $\text{dens sup}_{\mathcal{W}}$. The thesis then follows by Theorem 23. □

We stress that the hypotheses of Theorems 23 and 25 do not, as far as we know, ensure that F_B and F_W are well defined. They are if G is of sub-exponential growth (*e.g.*, $\mathbb{Z}^d$) and $\mathcal{X}$ is a sequence of disks (*e.g.*, $X_n = \{-n, \ldots, n\}^d$).

6. Conclusion

We have arrived at the end of our journey through two most noteworthy translation-invariant pseudo-distances for CA spaces The related field of studies is relatively young, but nevertheless very appealing, because of both the numerous changes of spatial properties *due to* the change of viewpoint (cf. Corollary 14) and the preservation of CA dynamics and properties *regardless of* the change of viewpoint. The new spaces seem thus to provide a useful framework in the study of dynamics on symbolic spaces and the properties of SBC global functions. Several of the results presented here belong to this thread of research.

Many more questions arise about these spaces. We think not only about their topologies and the CA dynamics on them, but also the identification of "good" subsets, *e.g.*, dense subspaces that may come out "handy" also in the studies of the topics above. In this respect, Challenges 4 and 5 seem especially appealing to us, since Toeplitz configurations seem to be a straightforward extension of periodic ones.

Our (not at all hidden) hope is that the compendium of results presented in this paper can raise interest in the subject and attract more researchers towards this field. Should any of the challenges we have issued be collected and overcome, that would be one of the greatest reward for the work we have done here.

Acknowledgements

This research was supported by the European Regional Development Fund (ERDF) through the Estonian Centre of Excellence in Computer Science (EXCS). The author thanks Tommaso Toffoli, Patrizia Mentrasti, Tarmo Uustalu, and all the people who provided him with their suggestions and encouragement. The author also thanks the anonymous referees for their thorough review and helpful suggestions.

References

[1] A.S. Besicovitch, Almost Periodic Functions, Cambridge University Press, 1932.
[2] F. Blanchard, J. Cervelle, E. Formenti, Some results about the chaotic behavior of cellular automata, Theor. Comp. Sci. 349 (2005) 318–336.
[3] F. Blanchard, E. Formenti, P. Kurka, Cellular automata in Cantor, Besicovitch, and Weyl topological spaces, Complex Systems 11 (2) (1999) 107–123.
[4] S. Capobianco, Multidimensional cellular automata and generalization of Fekete's lemma, Disc. Math. Theor. Comp. Sci. 10 (3) (2008) 95–104.
[5] S. Capobianco, On pattern density and sliding block code behavior for the Besicovitch and Weyl pseudo-distances, in: Jan van Leeuwen, et al. (Eds.), Proc. of 36th Int. Conf. on Current Trends in Theory and Practice of Computer Science, SOFSEM 2010 (Špindlerův Mlýn, Jan. 2010), in: Lecture Notes in Computer Science, vol. 5901, 2010, pp. 259–270. doi:10.1007/978-3-642-11266-9_22.
[6] S. Capobianco, On the induction operation for shift subspaces and cellular automata as presentations of dynamical systems, Inform. Comput. 207 (11) (2009) 1169–1180.
[7] S. Capobianco, Some notes on Besicovitch and Weyl distances over higher-dimensional configurations, in: Pedro P.B. de Oliveira, Jarkko Kari (Eds.), Proceedings of Automata 2009: 15th International Workshop on Cellular Automata and Discrete Complex Systems, Universidade Presbiteriana Mackenzie, São Paulo, SP, Brazil, 2009. Short papers section, 2009, pp. 300–308.
[8] S. Capobianco, Surjunctivity for cellular automata in Besicovitch spaces, J. Cell. Autom. 4 (2) (2009) 89–98.
[9] G. Cattaneo, E. Formenti, L. Margara, J. Mazoyer, A shift-invariant metric on $S^{\mathbb{Z}}$ inducing a non-trivial topology, Lect. Not. Comp. Sci. 1295 (1997) 179–188.
[10] T.G. Ceccherini-Silberstein, A. Machì, F. Scarabotti, Amenable groups and cellular automata, Ann. Inst. Fourier, Grenoble 42 (1999) 673–685.
[11] R. Devaney, An Introduction to Chaotic Dynamical Systems, Westview Press, 2003.
[12] P. de la Harpe, Topics in Geometric Group Theory, The University of Chicago Press, 2000.
[13] F. Fiorenzi, Cellular automata and finitely generated groups, Doctoral Thesis, Sapienza Università di Roma, 2000.

[14] F. Fiorenzi, Cellular automata and strongly irreducible shifts of finite type, Theor. Comp. Sci. 299 (2003) 477–493.
[15] D. Lind, B. Marcus, An Introduction to Symbolic Dynamics and Coding, Cambridge University Press, 1995.
[16] J. Munkres, Topology, Prentice Hall, 2000.
[17] J. Müller, C. Spandl, A Curtis-Hedlund-Lyndon theorem for Besicovitch and Weyl spaces, Theor. Comp. Sci. 410 (2009) 3606–3615.
[18] I. Namioka, Følner's Condition for Amenable Semi Groups, Math. Scand. 15 (1962) 18–28.
[19] T. Toffoli, S. Capobianco, P. Mentrasti, When—and how—can a cellular automaton be rewritten as a lattice gas? Theor. Comp. Sci. 403 (2008) 71–88.
[20] B. Weiss, Sofic groups and dynamical systems, Sankhyā: Indian J. Stat. 62 (2000) 350–359.