Cellular automata and substitutions in topological spaces defined via edit distances

Firas Ben Ramdhane¹,² · Pierre Guillon¹

Accepted: 24 May 2023 / Published online: 18 August 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
The Besicovitch pseudometric is a shift-invariant pseudometric over the set of infinite sequences, that enjoys interesting properties and is suitable for studying the dynamics of cellular automata. It corresponds to the asymptotic behavior of the Hamming distance on longer and longer prefixes. Though dynamics of cellular automata were already studied in the literature, we propose the first study of the dynamics of substitutions. We characterize those that yield a well-defined dynamical system as essentially the uniform ones. We also explore a variant of this pseudometric, the Feldman–Katok pseudometric, where the Hamming distance is replaced by the Levenshtein distance. Like in the Besicovitch space, cellular automata are Lipschitz in this space, but here also all substitutions are Lipschitz. In both spaces, we discuss equicontinuity of these systems, give a number of examples, and generalize our results to the class of dill maps, that embeds both cellular automata and substitutions.

Keywords Besicovitch topology · Cellular automata · Substitutions · Dill maps · Noncompact spaces · Topological dynamics

1 Introduction

In Blanchard et al. (1997) were studied the dynamics of cellular automata in the spaces of sequences endowed with the Besicovitch pseudometric, which is defined as the asymptotics of the Hamming distance over prefixes of the sequences. This corresponds to the $d$-metric defined for ergodic purposes in Ornstein (1974), Feldman (1976), and independently (Katok 1977), proposed to replace the Hamming distance by the Levenshtein distance from Levenshtein (1966), and get the $f$-metric, which is useful in Kakutani equivalence theory. The Levenshtein distance depends on the minimum number of edit operations (deletion, insertion, substitution) required to change one word into another word.

It is extensively used for information theory, linguistics, word algorithmics, statistics.... One can read some properties of the pseudometric in Ornstein et al. (1982, Chapter 2), and a nice history of this notion in Kwietniak and Łacka (2017). The recent (García-Ramos and Kwietniak 2020) can be seen as presenting a nice picture of those systems for which the identity map from the Cantor space into the Feldman–Katok space is a topological factor map, after a similar task has been achieved for the Besicovitch space in García-Ramos (2017). Here, we adopt a complementary point of view, by considering the dynamics within the space itself. Though this task on the Besicovitch space has concerned mainly cellular automata so far, relaxing the pseudometric to edit space allows to naturally consider a larger class of systems, that also includes substitutions: the so-called dill maps.

In Sect. 2, we will introduce some basic vocabulary of symbolic dynamical systems, including dill maps, and define the Besicovitch and Feldman–Katok spaces. In Sect. 3, we study dill maps over the Besicovitch space, give a sufficient and necessary conditions for them to induce a well-defined dynamical system over this space, and give some examples of behaviours. In Sect. 4, we do the same within the Feldman–Katok space.
2 Definitions and basic results

The aim of this section is to introduce some concepts and basic notations in symbolic dynamics, that will be used throughout this paper, and to introduce some symbolic dynamical objects and topological spaces.

We start with some terminology in word combinatorics. We fix once and for all an alphabet $A$ of finitely many letters (it will be precisely in each example, but general in our statements). A finite word over $A$ is a finite sequence of letters in $A$; it is convenient to write a word as $u = u[0..|u|]$ to express $u$ as the concatenation of the letters $u_0, u_1, \ldots, u_{|u|-1}$, with $|u|$ representing the length of $u$, i.e., the number of letters appearing in $u$, and $[0, |u|] = \{0, \ldots, |u| - 1\}$. The unique word of length 0 is the empty word denoted by $\lambda$. The number of occurrences of some subalphabet $B \subset A$ within a finite word $u$ is denoted by $|u|_B$. A configuration $x = x_0x_1x_2\ldots$ over $A$ is the concatenation of infinitely many letters from $A$. The set of all finite (resp. infinite) words over $A$ is denoted by $A^*$ (resp. $A^\omega$), and $A^n$ is the set of words of length $n \in \mathbb{N}$ over $A$.

2.1 Symbolic dynamics

Let us now introduce some basic notions in symbolic dynamics. Most classically, the set $A^\omega$ is endowed with the product topology of the discrete topology on each copy of $A$. The topology defined on $A^\omega$ is metrizable, corresponding to the Cantor distance $d_C$ and defined as follows:

$$d_C(x, y) = 2^{-\min\{n \in \mathbb{N} \mid x_n \neq y_n\}}, \forall x \neq y \in A^\omega,$$

and $d_C(x, x) = 0, \forall x \in A^\omega$.

This space, called the Cantor space, is compact, totally disconnected and perfect.

A (topological) dynamical system is a pair $(X_d, F)$ where $F$ is a uniformly continuous map of a metric space $X_d = (X, d)$ to itself, i.e.,

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in X, d(x, y) < \delta \Rightarrow d(F(x), F(y)) < \varepsilon.$$

Note that in the literature, $X_o$ is usually assumed compact, but in this article we will focus on a more general setting, in which many known results cannot be applied. Recall that, if $(X_d, F)$ is uniformly continuous implies that $(X_d, F^t)$ is uniformly continuous also. When $X_d$ is understood from the context, we may omit it. In particular, the shift dynamical system is the pair $(A^\omega, \sigma)$, where $\sigma$ is the shift map, defined for all $x \in A^\omega$ by $\sigma(x)_i = x_{i+1}$, for $i \in \mathbb{N}$.

We can now introduce some topological properties of a dynamical system $(X_d, F)$. We say that $x \in X$ is a fixed point if $F(x) = x$; it is periodic if $F^t(x) = x$ for some $t > 0$. The map $F$ is $\varepsilon$-Lipschitz, for $\varepsilon > 0$, if $d(F(x), F(y)) \leq \varepsilon d(x, y)$ for all $x, y \in X$. It is clear that, if $F$ is Lipschitz, then $F$ is uniformly continuous. A point $x \in X$ is an equicontinuity point of $(X_d, F)$ if:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall y \in X, d(x, y) < \delta \Rightarrow \exists \tau \in \mathbb{N}, d(F^\tau(x), F^\tau(y)) < \varepsilon.$$

A dynamical system $(X_d, F)$ is equicontinuous if:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in X, d(x, y) < \delta \Rightarrow \exists \tau \in \mathbb{N}, d(F^\tau(x), F^\tau(y)) < \varepsilon.$$

Note that if $F$ is $\varepsilon$-Lipschitz, then $F^t$ is $\varepsilon^t$-Lipschitz. It is then clear that if $F$ is 1-Lipschitz, then $F$ is equicontinuous (and it is actually an equivalence, up to equivalent distance, as seen for instance in Kûrka (2003, Proposition 2.41)). A dynamical system $(X_d, F)$ is sensitive if:

$$\exists \varepsilon > 0, \forall x \in X, \exists \delta > 0, \exists y \in X, d(x, y) < \delta \text{ but } \exists \tau \in \mathbb{N}, d(F^\tau(x), F^\tau(y)) > \varepsilon.$$

A dynamical system $(X_d, F)$ is (positively) expansive if:

$$\exists \varepsilon > 0, \forall x \neq y \in X, \exists \tau \in \mathbb{N}, d(F^\tau(x), F^\tau(y)) > \varepsilon.$$

As examples of dynamical systems, we will be interested in this paper by cellular automata, substitutions, and in general dill maps. For more details in the case of the Cantor space, we can refer to [7], [3] and Kûrka (2003).

2.2 Cellular automata

Definition 1 A cellular automaton (CA) with diameter $\theta$ is a map $F : A^N \to A^N$, such that there exists a map called local rule $f : A^0 \to A$ such that for all $x \in A^N$ and all $i \in \mathbb{N}$: $F(x)_i = f(x_{i+[i, i+1]})$.

Example 2

1. The shift is the CA with diameter $\theta = 2$ and local rule $f$ defined by $f(u[0, 1]) = u_1$ for all $u_0, u_1 \in A$.

2. Let $A = \{a, b\}$. The Xor is the CA with diameter $\theta = 2$ and local rule $f$ defined by: $f(aa) = f(bb) = a$ and $f(ab) = f(ba) = b$.

3. Let $A = \{a, b\}$. The Min is the CA with diameter $\theta = 2$ and local rule $f$ defined by: $f(aa) = f(ab) = f(ba) = a$ and $f(bb) = b$.

In the Cantor space, an elegant characterization of cellular automata was given by Curtis, Hedlund and Lyndon in Hedlund (1969) as follows: A function $F : A^N \to A^N$ is a cellular automaton if and only if it is continuous with respect to the Cantor metric and shift-equivariant (i.e., $F(\sigma(x)) = \sigma(F(x))$), for all $x \in A^N$.

2.3 Substitutions

Recall that $f(t) = \lim_{t \to \infty} f(t)$ if $\lim_{t \to \infty} f(t) = 0$, $f(t) = \lim_{t \to \infty} f(t)$ if there is $\varepsilon > 0$ such that for every sufficiently
large \( t \in \mathbb{N} \), we have \( f(t) \leq \varphi g(t) \), and \( f(t) = \Theta_{t-\infty}(g(t)) \) if both \( f(t) = \Theta_{t-\infty}(g(t)) \) and \( g(t) = \Omega_{t-\infty}(f(t)) \).

**Definition 3**

1. A substitution \( \tau \) is a nonerasing homomorphism of monoid \( A^* \), i.e., \( \tau^{-1}(\lambda) = \{ \lambda \} \) and \( \tau(uv) = \tau(u)\tau(v) \), for all \( u, v \in A^* \).
2. \( \tau \) yields a dynamical system, denoted by \( \tau \), and defined over \( \mathbb{N} \) by:
   \[ \tau(z) = \tau(z_0)\tau(z_1)\tau(z_2)\tau(z_3) \ldots, \forall z \in A^\mathbb{N}. \]
3. The lower norm \( |\tau| \) and upper norm \( \|\tau\| \) of \( \tau \) are defined by:
   \[ |\tau| = \min \{|\tau(a)| \mid a \in A\} \quad \text{and} \quad \|\tau\| = \max \{|\tau(a)| \mid a \in A\}. \]

   We say that \( \tau \) is uniform if \( |\tau| = \|\tau\| \).
4. Thanks to Berthé and Rigo (2010, Theorem 4.7.15), one knows that there exists a nonempty subalphabet \( A_t^+ \subset A \) such that for every letter \( b \in A \),
   \[ |\tau^t(b)| = \Theta_{t-\infty}(|\tau^t|) \quad \text{if} \quad b \in A_t^+, \quad \text{and} \quad |\tau^t(b)| = \Theta_{t-\infty}(|\tau^t|) \quad \text{if} \quad b \in A_t^- = A \setminus A_t^+. \]
5. Following (Pansiot 1984), we say that \( \tau \) is quasi-uniform if
   \[ A_t^+ = A, \text{ i.e., for any letter} \quad a \in A, \quad |\tau^t(a)| = \Theta_{t-\infty}(|\tau^t|). \]

**Remark 4** Quasi-uniform substitutions include uniform substitutions of course, as well as irreducible substitutions, i.e., \( \tau \) such that for every letters \( a, b \in A \), \( b \) appears in \( \tau^t(a) \), for some iteration \( t \in \mathbb{N} \). This comes from the observation that, in general, if \( b \) appears in \( \tau^t(a) \), then \( |\tau^t(b)| = O_{t-\infty}(|\tau^t|) \) (in particular \( b \in A_t^+ \Rightarrow a \in A_t^+ \)).

The matrix of \( \tau \) is defined as \( M(\tau) = (M(\tau))_{ab} \) such that \( M(\tau)_{ab} \) is the number \( |\tau(a)|_b \) of occurrences of \( b \) in \( \tau(a) \). When it is reducible, \( A_t^+ \) can be seen from its irreducible components, i.e., the blocks in the triangular form of the matrix \( M(\tau) \).

If the spectral radius \( \rho \) is strictly greater than 1, then Perron-Frobenius theory (see for example Kürka 2003, Theorem A.72) establishes that \( \|\tau^t\| = \Theta_{t-\infty}(\rho^t) \), and for every \( a \in A_t^- \), \( |\tau^t(a)| = \Theta_{t-\infty}(\rho^t) \) for some \( \rho_- < \rho \).

**Example 5** Let \( A = \{a, b\} \).

1. The Thue-Morse substitution defined over \( A \) by:
   \[ \tau : \quad a \mapsto ab \quad b \mapsto ba \quad M(\tau) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad M(\tau)^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \]
   This is an irreducible uniform substitution.
2. The Fibonacci substitution defined over \( A \) by:
   \[ \tau : \quad a \mapsto ab \quad b \mapsto a \quad M(\tau) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad M(\tau)^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \]
   This is an irreducible non-uniform substitution:
   \[ |\tau| = 1 < 2 = \|\tau\|, \quad \text{and} \quad \|\tau\| \text{ is the } n^{th} \text{ Fibonacci number,} \]
   \[ \text{which is } \Theta(\rho^n), \text{ where } \rho \text{ is the golden ratio.} \]
3. The doubling substitution defined over \( A \) by:
   \[ \tau : \quad a \mapsto aa \quad M(\tau) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad b \mapsto bb \]
   This is a uniform reducible substitution: \( \{a\} \) and \( \{b\} \) are two disjoint invariant subalphabets.
4. A uniform substitution \( \tau \) is Teplitz if there exists \( i \in [0, \|\tau\|[ \) such that for all \( a, b \in A \), \( \tau(a)_t = \tau(b)_t \). An example is the Cantor substitution, defined over \( A = \{a, b\} \) by:
   \[ \tau : \quad a \mapsto aba \quad M(\tau) = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \quad b \mapsto bbb \]
   This is a reducible uniform substitution: \( \{b\} \) is an invariant subalphabet.
5. For \( k \in \mathbb{N} \), consider the substitution:
   \[ \tau : \quad a \mapsto a^kb \quad M(\tau) = \begin{bmatrix} k & 1 \\ 0 & 1 \end{bmatrix} \]
   If \( k > 0 \), then this substitution is not quasi-uniform: \( A_t^+ = \{a\} \) and \( A_t^- = \{b\} \). If \( k > 1 \), then \( |\tau^t(a)| = |\tau^t| = \Theta(k^t) \), whereas if \( k = 1 \), then \( |\tau^t(a)| = |\tau^t| = t + 1 \).

The following lemma states that the fast-growing orbits of letters must involve many fast-growing letters.

**Lemma 6** Let \( \tau \) be any substitution with spectral radius strictly greater than 1, and \( a \in A_t^+ \). Then
   \[ |\tau^t(a)|_{A_t^+} = \Theta_{t-\infty}(\rho^t). \]

**Proof** Since the spectral radius is \( \rho > 1 \), Remark 4 gives \( \beta, \beta_- > 0 \) such that for every \( t \in \mathbb{N} \) and every letter \( a \in A_t^+, \beta \rho^t \leq |\tau^t(a)| \leq \beta \rho^t \), and for every letter \( a \in A_t^- \), \( |\tau^t(a)| \leq \beta \rho^t \). Now let \( s, t \in \mathbb{N} \), and let us write:
   \[ \tau^{t+s}(a) = \tau^t(\tau(s)_{(a)}) \cdot \tau^{t+s}(a_{(t)}) \cdots \tau^{t+s}(a_{(s)}) \].

One can now bound the length:
\[
|\tau^{t+s}(a)| = \sum_{\tau^t(a)_{(a)} \in A_t^+} |\tau^t(\tau(s)_{(a)})| + \sum_{\tau^t(a)_{(a)} \in A_t^-} |\tau^t(\tau(s)_{(a)})| \\
\leq |\tau^t(a)_{(a)}| \rho^s + (|\tau^t(a)| - |\tau^t(a)_{(a)}|) \beta \rho^t \\
\leq |\tau^t(a)_{(a)}| \rho^s (\rho^t - \rho_-^t) + x^t \beta \rho^t \rho_-^t.
\]

Since \( |\tau^{t+s}(a)| \geq \beta \rho^{t+s} \), we obtain that:
   \[ |\tau^t(a)|_{A_t^+} \geq \rho^s \frac{\beta - x^t (\frac{\rho}{\rho_-})^t}{\rho(1 - (\frac{\rho}{\rho_-})^t)}. \]

When \( s \) grows, this fraction converges to \( \frac{\beta}{\rho} \) \( \square \)
2.4 Dill maps

The dill maps were defined in Salo and Törmä (2015), and generalize both substitutions and CA. Here we give a simple definition, which is equivalent to Salo and Törmä (2015, Definition 2).

**Definition 7**
- A dill map $F$ with diameter $\theta \in \mathbb{N} \setminus \{0\}$ is a dynamical system over the set of configurations such that there exists a local rule $f : A^0 \rightarrow A^+$ satisfying:
  \[
  \forall x \in A^N, F(x) = f(x_{[0,\theta]}^0) f(x_{[1,\theta+1]}^1) f(x_{[2,\theta+2]}^2) \cdots .
  \]
- The lower norm $|f|$ and the upper norm $\|f\|$ of a dill map $F$ with diameter $\theta$ and local rule $f$ are defined by:
  \[
  |f| = \min \{ |f(u)| \mid u \in A^0 \} \quad \text{and} \quad \|f\| = \max \{ |f(u)| \mid u \in A^0 \}.
  \]
- We extend the local rule into a self-map $f^* : A^* \rightarrow A^*$ by:
  \[
  f^*(u) = f(u_{[0,\theta]}^0) f(u_{[1,\theta+1]}^1) \cdots f(u_{[\ell,\ell+\theta]}^\ell),
  \]
  for $u$ such that $|u| \geq \theta$ and $f^*(u) = \lambda$ if $|u| < \theta$.
- We also consider the cocycle $s^p_m = \sum_{i \leq n} f(x_{[i,i+\theta]}^i)$ for $n \in \mathbb{N}$ and $x \in A^N$, which represents the position in which one can read the image at offset $n$:
  \[
  F(s^p(x)) = \sigma^p(F(x)).
  \]
- If $\|f\| = |f|$, then we say that $F$ is uniform. In that case, the cocycle $s^p_m$ does not depend on the configuration $x$, and we note it $s^p$.

When it is clear from the context, we may identify a dill map with its local rule.

**Remark 8**
1. The substitutions are the dill maps with diameter $\theta = 1$.
2. The cellular automata are the uniform dill maps with $|f| = \|f\| = 1$.
3. The composition of a substitution $\tau$ and a cellular automaton local rule $f$ with diameter $\theta$ is a dill map local rule $\tau \circ f$ with diameter $\theta$. Actually, every dill map is the composition of a substitution and a shift homomorphism (which is like a cellular automaton, but allowing to change the alphabet).

**Example 9** Let $f$ be the local rule of the Xor CA and $\tau$ be the Fibonacci substitution. Then $\tau \circ f$ is a local rule of a dill map with diameter 2 and defined as follows:
\[
\tau \circ f: \begin{array}{ccl}
(aa, bb) & \mapsto & ab \\
(ba, ab) & \mapsto & a
\end{array}
\]

**Remark 10** For all $x \in A^N$, $m \rightarrow s^p$ is a one-to-one function.

**Proof** For $n \leq m \in \mathbb{N}$ such that $x^i_{s^p_m} = x^i_{s^p_n}$ we have:
\[
\sum_{i \leq m} |f(x_{i+i+\theta}^i)| = \sum_{i < n} |f(x_{i+i+\theta}^i)| = \sum_{n \leq i \leq m} |f(x_{i+i+\theta}^i)| = 0
\]
Then for all $n < i \leq m$, we have $|f(x_{i+i+\theta}^i)| = 0$. Since $f$ is nonerasing, $n = m$. □

Similarly to the case of cellular automata, we give a characterization of dill maps à la Hedlund; it is quite classical but we are not aware of a reference with this exact statement, so we include it for completeness. Recall that $\mathbb{N}$ can be naturally endowed with the discrete topology.

**Theorem 11** A function $F : A^N \rightarrow A^N$ is a dill map if and only if it is continuous over the Cantor space and there exists a continuous map $s : A^N \rightarrow \mathbb{N}$ such that for all $x \in A^N$, $F(s(x)) = \sigma^{s(x)}(F(x))$.

**Proof** Let $F$ be a dill map with diameter $\theta$ and local rule $f$. For $x \in A^N$, $\varepsilon = 2^{-\theta}$ with $p \in \mathbb{N} \setminus \{0\}$, we take $m = \min \{i \in \mathbb{N} \mid s^i \geq p \}$. Then $\delta = 2^{-\min(m+p)}$ and $y \in A^N$ such that $d_C(x, y) \leq \delta$ we have $x_{[0,m+\theta]} = y_{[0,m+\theta]}$. Then $f^*(x_{[0,m+\theta]}) = f^*(y_{[0,m+\theta]})$. Hence $F(x)_{[0,p]} = F(y)_{[0,p]}$. So, $d_C(F(x), F(y)) \leq 2^{-\theta} = \varepsilon$. In conclusion, $F$ is continuous.

Now let us define $s(x) = s^1_x = |f(x_{[0,\theta]}^0)|$ for all $x \in A^N$. Let $x \in A^N$ and $\varepsilon > 0$. For $y \in A^N$ such that $d_C(x, y) < 2^{-\theta}$ we have $x_{[0,\theta]} = y_{[0,\theta]}$. Then $f(x_{[0,\theta]}) = f(y_{[0,\theta]})$ and hence $|f(x_{[0,\theta]})| = |f(y_{[0,\theta]})|$. So $|s(x) - s(y)| = 0$.

In conclusion, $s$ is continuous, and it satisfies $F(s(x)) = \sigma^{s(x)}(F(x))$.

Assume now that there exists a continuous map $s : A^N \rightarrow \mathbb{N}$ such that $F(s(x)) = \sigma^{s(x)}(F(x))$ for every $x \in A^N$. We can write $A^N = \bigcup_{i \in \mathbb{N}} s^{-1} \{ \{i\} \}$. For all $x \in A^N$, since $\{n\}$ is clopen and $s$ is continuous, $s^{-1} \{ \{i\} \}$ is also open. On the other hand, $(A^N, d_C)$ is compact, so that $A^N = \bigcup_{i \in \mathbb{N}} s^{-1} \{ \{i\} \}$ for some finite set $I \subset \mathbb{N}$. In other words, for every $x \in A^N$, $s(x) \leq \max I$. Since $F$ is continuous, for $\varepsilon = 2^{-\max I}$, there exists $\theta \in \mathbb{N}$ such that for all $y \in A^N$ verifying $d_C(x, y) < 2^{-\theta}$, we have $d_C(F(x), F(y)) < \varepsilon$. Hence for all $x \in A^N$ and all $y \in A^N$ with $x_{[0,\theta]} = y_{[0,\theta]}$, we have $F(x)_{[0,s(x)]} = F(y)_{[0,s(y)]}$. So one can define a map $f : A^0 \rightarrow A^*$ such that $f(x_{[0,\theta]}) = F(x)_{[0,s(x)]}$, for all $x \in A^N$. On the other hand, for $x \in A^N$ and $j \in \mathbb{N}$ we have:
3 Dill maps in the Besicovitch space

3.1 The Besicovitch space

In this subsection, we recall the definition and topological properties of the Besicovitch space.

We can endow the set $A^n$ of words, for $n \in \mathbb{N}$, with some distance, i.e., some application from $A^n \times A^n$ to $\mathbb{R}^+$ satisfying: separation, symmetry, triangle inequality. The prototypical example is the Hamming distance denoted by $d_H$.

It is usually defined as the number of differences between two finite words of the same length. Let us present a definition in terms of edit distance, that is the number of operations of a specific kind to transform a word into another one.

Definition 12

- The substitution operation $S_j^k$ at position $j \in \{0, |u|\}$, for $a \in A$, is defined over finite word $u \in A^*$ as follows:
  $$S_j^k(u) = u_0 \cdots u_{j-1} au_{j+1} \cdots u_{|u|-1}.$$  
- Between two finite words with the same length $u$, $v$, we define the Hamming distance:
  $$d_H(u, v) = \min \left\{ m \in \{0, |u|\} \mid \exists j_1 < \cdots < j_m, (a_{j})_{j \leq m} \in A^m, S_{j_1}^{a_{j_1}} \circ \cdots \circ S_{j_m}^{a_{j_m}}(u) = v \right\}.$$

Of course, the definition would be completely equivalent by allowing substitution operations to be performed on both $u$ and $v$, and is simply the number of differences, letterwise. One should be careful, in this article, not to confuse the substitution operations (used only in this subsection) with the general substitutions from Sect. 2.3.

Remark 13

Note that this distance is additive i.e., for all $u, u', v, v'$ such that $|u| = |u'|$ and $|v| = |v'|$,
$$d_H(uu', vv') = d_H(u, u') + d_H(v, v').$$

Definition 14

The Besicovitch pseudometric, denoted by $d_B$, is defined as follows:

$$d_B(x, y) = \limsup_{\ell \to \infty} \frac{d_H(x_{[0, \ell]}, y_{[0, \ell]})}{\ell}, \forall x, y \in A^\mathbb{N}.$$  

It is easy to verify that this is a pseudometric: it is symmetric, zero over diagonal pairs, and satisfies the triangular inequality. On the other hand, it is not a distance since we can find two different configurations between which the pseudometric is worth zero (for instance, we can take two configurations with finitely many of differences). Hence, it is relevant to quotient the space of configurations by the equivalence of zero distance, in order to get a separated topological space:

Definition 15

- The relation $x \sim_{d_B} y$ if and only if $d_B(x, y) = 0$, is an equivalence relation.
- The quotient space $A^\mathbb{N} / \sim_{d_B}$ is a topological space, called the Besicovitch space, denoted $X_{d_B}$.
- We denote by $x_{d_B}$ the equivalence class of $x \in A^\mathbb{N}$ in the quotient space.
- Any map $F : A^\mathbb{N} \to A^\mathbb{N}$ such that $d_B(F(x), F(y)) = 0$ for all $x, y \in A^\mathbb{N}$, induces a well-defined map $F_{d_B} : X_{d_B} \to X_{d_B}$ over the Besicovitch space.

According to Blanchard et al. (1997), the Besicovitch space is path-connected, complete, and infinite-dimensional, but it is neither separable nor locally compact.

Proposition 16 (Cattaneo et al. 1997, Proposition 3) If $x / = y \in A^\mathbb{N}$ are periodic configurations, then $0 < d_B(x, y) \leq d_H(x, y)$.

Remark 17

Note that for $k \neq k' \in \mathbb{N}$, if $\sigma^k \sim_{d_B} \sigma^{k'}$ (resp. $\sigma^k \sim_{d_B} \sigma^{k'}$) if and only if $k = k'$. Indeed, suppose that $k \neq k'$. Let $u \in A^*$ such that $|u| \geq \max\{k, k'\}$ and $u_k \neq u_{k'}$.

For $x = u^\mathbb{N}$, it is clear that $\sigma^k(x)_0 = u_k \neq u_{k'} = \sigma^{k'}(x)_0$.

Hence, since $\sigma^k(x)$ and $\sigma^{k'}(x)$ are $|u|$-periodic, and thanks to Proposition 16 we obtain:
$$0 < \frac{1}{|u|} \leq d_H(\sigma^k(x), \sigma^{k'}(x)) \leq d_B(\sigma^k(x), \sigma^{k'}(x)).$$

Which contradict the fact that $\sigma^k$ and $\sigma^{k'}$ are in the same equivalence class.

3.2 Lipschitz property of dill maps

It is known since (Blanchard et al. 1997) that every cellular automaton induces a (well-defined) Lipschitz function over the Besicovitch space.
Müller and Spandl (2009, Theorem 13) goes further, by establishing a characterization à la Hedlund of cellular automata in the Besicovitch space by three conditions: shift invariance, a condition in terms of uniform continuity and a condition in terms of periodic configurations.

Some dill maps, on the contrary, are not well-defined.

**Example 18** The Fibonacci substitution is not well-defined over the Besicovitch space $X_{b^\infty}$.

Even worse, for every $x \in \{a, b\}^\mathbb{N}$ such that $x_{b^\infty} \neq (b^\infty)_{b^\infty}$, altering simply the first letter will induce a shift in the substituted word. Indeed, if $x_0 = b$, then:

$$d_H(\tau(S^\infty_0(x)), \tau(x)) = d_H(a\tau(x), \tau(x)).$$

Symmetrically, if $x_0 = a$, then:

$$d_H(\tau(S^\infty_0(x)), \tau(x)) = d_H(\sigma(\tau(x)), \tau(x)).$$

In both cases, the pseudometric is at least half the frequency $d_H(b^\infty, a)$ of $a$ in $x$. For example, $d_H(ab^\infty, ba^\infty) = 0$ but $d_H(ab^\infty, ba^\infty) = d_H((ab)^\infty, (ba)^\infty) = 1$. On the other hand, for all $x \in \{a, b\}^\mathbb{N}$, $d_H(\tau(x), \tau(b^\infty)) \leq d_H(x, b^\infty)$ (frequency of $a$ in $x$). So $b^\infty$ is the only continuity point.

In this section, we characterize dill maps which induce a well-defined function over this space.

Let us denote, for a uniform dill map $F$ with local rule $f$ and diameter $\theta$:

$$d_f^- = \max \{ d_H(f(u), f(v)) | u, v \in A^0 \}$$

and

$$d_f^+ = \min \{ d_H(f(u), f(v)) | u \neq v \in A^0 \}. $$

**Proposition 19** Let $F$ be a uniform dill map with diameter $\theta$ and local rule $f$. Then for all $x, y \in A^\mathbb{N}$:

$$d_f^+ \cdot d_H(x, y) \leq d_H(F(x), F(y)) \leq \frac{\theta d_f^+}{|f|} \cdot d_H(x, y).$$

**Proof** We start by the first inequality. Let $x, y \in A^\mathbb{N}$ and $\ell \in \mathbb{N}$. Let $D = \{i \in \mathbb{N} | x_i \neq y_j \}$ and $D' = \{i \in \mathbb{N} | F(x)_i \neq F(y)_i \}$. If $i \in D$, then $|D' \cap \|i\|f, (j + 1)\|f|\| \geq d_f^-$ for each $j \in \mathbb{N}$, which implies that $|D' \cap \|i - \ell\|f, \ell\|f|\| \geq 0d_f^-$, so provided that $i \geq \ell$. Hence:

$$d_H(F(x)_{\ell}, F(y)_{\ell}) \leq \frac{1}{\ell} \sum_{i=0}^{\ell} |D' \cap \|i - \ell\|f, \ell\|f|\|$$

$$\geq |D \cap \|0, \ell\|f|\|$$

$$\geq |d_H(x_{0, \ell}, y_{0, \ell})| - |D \cap \|0, \ell\|f|\|$$

$$\geq d_H(x_{0, \ell}, y_{0, \ell})d_f^- - \theta d_f^-.$$

Then:

$$d_f^+ \cdot d_H(x_{0, \ell}, y_{0, \ell}) \leq \frac{1}{\ell} \sum_{i=0}^{\ell} |D' \cap \|i - \ell\|f, \ell\|f|\|$$

$$\leq |D \cap \|0, \ell\|f|\|$$

$$\geq |d_H(x_{0, \ell}, y_{0, \ell})| - |D \cap \|0, \ell\|f|\|$$

$$\geq d_H(x_{0, \ell}, y_{0, \ell})d_f^- - \theta d_f^-.$$
In particular, the Cantor and Thue-Morse substitutions are well-defined over this space (we will discuss them in the next subsection).

We now reach necessary and sufficient conditions for dill maps to induce well-defined dynamical systems over this space.

**Theorem 20** Let $F$ be a dill map with diameter $\theta$ and local rule $f$. Then the following statements are equivalent:

1. $F_{\partial_H}$ is well-defined.
2. $F$ is $\frac{\partial d_f}{|f|}$–Lipschitz with respect to $\partial_H$.
3. $F$ is either constant or uniform.

**Proof** Implication 3⇒2 follows from Proposition 19. Implication 2⇒1 is clear from the definition of Lipschitz function.

Let us prove implication 1⇒3. Assume that $F$ is non-uniform, i.e., there are two words $u$ and $v$ of equal length such that $|f^*(u)| \neq |f^*(v)|$. One can assume that their longest common suffix has length $\theta - 1$. Indeed, otherwise let $u \in A$, $u' = u_{[|u| - \theta + 1, |u|]}a^{d_0 - 1}$ and $v' = v_{[|v| - \theta + 1, |v|]}a^{d_0 - 1}$; one can note that $f^*(ua^{d_0 - 1}) = f^*(u)f^*(a)$ and $f^*(va^{d_0 - 1}) = f^*(v)f^*(a)$, so that either $|f^*(ua^{d_0 - 1})| \neq |f^*(va^{d_0 - 1})|$, or $|f^*(u')| = |f^*(va^{d_0 - 1})| - |f^*(u)| \neq |f^*(va^{d_0 - 1})| - |f^*(v)| = |f^*(v')|$, and both of these pairs of words share a common suffix of length at least $\theta - 1$. Assume also without loss of generality that $k = |f^*(u)| - |f^*(v)| > 0$.

- First assume that there exist $w \in A^*$ and $i \in \mathbb{N}$ such that $f^*(w) \neq f^*(w)'_i$. By our previous assumption, we know that for $z = w^\infty$ and

$$w' = u_{[|u| - \theta + 1, |u|]}z_{[0, \theta]}$$

we have: $F(uz) = f^*(u)f^*(w')F(z)$, so that for every $i \in \mathbb{N}$,

$$F(uz)[f^*(u)]|f^*(w')| + f^*(w') = f^*(w)_i.$$

On the other hand:

$$F(vz)[f^*(u)]|f^*(w')| + f^*(w') = f^*(w)_i + k$$

We deduce that $\partial_H(F(uz), F(vz)) \geq \frac{1}{|f^*(z_{[0, \theta]})|}$. Since $|u| = |v|$, we know $\partial_H(uz, vz) = 0$, so that $F$ is not well-defined over the quotient space.

- Otherwise, for all $w \in A^*$, $i \in [0, |f^*(w')| - k]$, we have $f^*(w)_i = f^*(w)'_i$. Consider any fixed $u \in A^{k + \theta}$, and $x \in A^\mathbb{N}$. Our assumption gives that $F(x)$ is $k$-periodic, thus $F(x) = f^*(x_{[0, k + \theta]})_{[0, \theta]}^\infty$. Assuming that $F_{\partial_H}$ is well-defined, we get $\partial_H(F(x), F(x_{[0, k + \theta]}) = 0$. According to Proposition 16, we can deduce that $F(x) = f^*(x_{[0, k + \theta]})_{[0, \theta]}^\infty$. Gathering the two equalities from above, we get that $F(x) = f^*(x_{[0, k + \theta]})_{[0, \theta]}^\infty$. Hence, $F$ is constant.

In the particular case of uniform substitutions, $\partial d_f = d_f^+ \leq |f|$, which allows the following.

**Corollary 21** A substitution $\tau$ yields a well-defined dynamical system $\tau_{\partial_H}$ over $X_{\partial_H}$ if and only if it is $1$-Lipschitz with respect to $\partial_H$.

### 3.3 Equicontinuity

The following derives directly from Proposition 19 (and completeness of the Besicovitch space).

**Corollary 22** Let $F$ be a uniform dill map with diameter $\theta$ and local rule $f$.

1. If $\theta d_f^+ \leq |f|$, then $F_{\partial_H}$ is equicontinuous.

   For example, for every uniform substitution $\tau$, $\tau_{\partial_H}$ is equicontinuous.

2. If $\theta d_f^+ < |f|$, then $F_{\partial_H}$ is contracting. So according to the Banach fixed point theorem (Banach 1922), every orbit converges to a unique fixed point. For example, for every uniform substitution $\tau$ such that $d_f^+ < |\tau|$, $\tau_{\partial_H}$ is contracting.

3. For every uniform substitution $\tau$ such that $d_f^+ = |\tau|$ (which means that the substitution is everywhere marked: any two images have no letter in common), $\tau_{\partial_H}$ is an isometry.

The case of cellular automata is quite well understood, since a characterization for contraction and isometry have been given in Salo and Törnä (2012). Let us give some examples of substitutions.

**Example 23** Let $\tau$ be a Toeplitz substitution. By definition and by Corollary 22, $\tau_{\partial_H}$ is contracting, so that all orbits converge towards a unique fixed point: the class for $\sim_{\partial_H}$ of the usual fixed points of the substitution (which is unique if for all $a, b \in A$, $\tau(a)_0 = \tau(b)_0$, but may not be otherwise, like for the Cantor substitution).

**Example 24** On the contrary, the Thue-Morse substitution is an isometry, thanks to Item 3 of Corollary 22. In particular, if $\Sigma_1$ is the orbit closure of the two fixed points, then for every $x \notin \Sigma_1$, the pseudometric $\partial_H(\tau(x), \Sigma_1)$ is constant, positive, so that our intuition that orbits converge towards $\Sigma_1$, though justified in the Cantor space, is completely false in the Besicovitch space.
Remark 25 The behaviors from Example 23 and 24 give an essentially full picture of what can occur for substitutions. Indeed, if there exists \( p \in \mathbb{N} \) such that \( d_\tau^n < |\tau|^p \), then \( \tau \) is contracting; consequently, for every \( t \in \mathbb{N} \) the diameter of \( \tau'(A^N) \) is bounded by that of \( \tau^{|t|} \), which is bounded by \( \left( \frac{|d_\tau^n|}{|t|} \right)^{|t|} \); so all orbits of \( \tau \) converge to a unique fixed point.

If, on the other hand, for every \( t \in \mathbb{N} \), \( d_\tau^t = |\tau| \), this means that there exists a subalphabet \( A_t \in A \) containing at least two letters, such that \( a \neq b \in A_t \Rightarrow d(H(\tau(a), \tau(b))) = |\tau|^t \). Hence, for \( t \in \mathbb{N} \) and \( A_t \), the maximal subalphabet satisfying the previous condition, it is not difficult to see that \( A_{t+1} \subseteq A_t \), and since \( A_t \) contains at least two letters for all \( t \in \mathbb{N} \), then the subalphabet \( A' = \bigcap_{t \in \mathbb{N}} A_t \) contains at least two letters. Finally, since \( A' \) is preserved by \( \tau \), then the restriction of \( \tau \) to \( A'^N \) is an isometry (because Proposition 19 remains true when the minimum and maximum are taken over a subalphabet).

The links between dynamical properties in the Cantor space and in the Besicovitch space appeared for cellular automata in Blanchard et al. (1997), Formenti and Kůrka (2009): in particular, sensitivity in the Besicovitch space implies sensitivity in the Cantor space, and equicontinuity in the Cantor space implies equicontinuity in the Bescovitch space. Nevertheless, unlike for cellular automata, there exist dill maps which are equicontinuous in the Cantor space but not in the Besicovitch space.

Example 26 Consider the dill map \( F \) with diameter 2 defined over the alphabet \( A = \{a, b\} \) by the following local rule: \( f(bb) = bbb \), and \( f(u_0u_1) = u_0u_1 \) if \( u_0u_1 \neq bb \). This dill map is 1-Lipschitz in the Cantor space (because it either preserves or doubles the common prefix), and hence it is equicontinuous. On the contrary, \( F_{bb} \) is not well-defined over the Besicovitch space since it is neither constant nor uniform (thanks to Theorem 20).

Some weak robustness properties of cellular automata from Formenti and Kůrka (2009), though, can be generalized to dill maps, like in the following statement.

Proposition 27 Let \( F \) be a uniform dill map and \( m \in \mathbb{N} \). Then we have the following:
1. If \( F_{b_1} \) is sensitive, then \( \sigma_{b_1}^m \circ F_{b_1} \) is sensitive, too.
2. If \( x \) is an equicontinuity point for \( F_{b_1} \), then it is an equicontinuity point for \( \sigma_{b_1}^m \circ F_{b_1} \), too.
3. If \( F_{b_1} \) is equicontinuous, then \( \sigma_{b_1}^m \circ F_{b_1} \) is equicontinuous, too.

Proof The key to prove the three statements is the following:

\[
\forall n \in \mathbb{N}, (\sigma^m \circ F)^n = \sigma^{|t|-1} \circ F^n.
\]

Let us prove this by induction on \( n \in \mathbb{N} \). The case \( n = 0 \) is obvious. Suppose that it is true for some \( n \).

\[
(\sigma^m \circ F)^{n+1} = (\sigma^m \circ F) \circ (\sigma^m \circ F)^n
= \sigma^m \circ F \circ (\sigma^{|t|-1} \circ F^n)
= \sigma^m \circ \sigma^{|t|-1} \circ F \circ F^n,
\]

which is the next step of the induction hypothesis. The last equality comes from the fact that \( F \) is uniform, so that \( \delta^m_{\tau} = |\tau| = \|\tau\| \).

Now we can deduce the proof of the statements: for all \( x, y \in A^N \) and for all \( m, n \in \mathbb{N} \):

\[
\delta_H((\sigma^m \circ F)^n(x), (\sigma^m \circ F)^n(y))
= \delta_H(\sigma^{|t|-1} \circ (F^n(x), \sigma^{|t|-1} \circ (F^n(y)))
= \delta_H(F^n(x), F^n(y)) \quad (\text{since } \delta_H \text{ is shift-invariant}).
\]

It was known that the cellular automata suit well in the Besicovitch pseudometric, and we have seen in this section that it is also the case for uniform substitutions. But we proved that this is not true for non-uniform substitutions. In the next section, we consider another topological space, in which both cellular automata and all substitutions are well-defined over this space.

4 Dill maps in the Feldman–Katok space

4.1 The Feldman–Katok space

Another classical edit distance is the Levenshtein distance (Levenshtein 1966). Instead of allowing to edit finite words only via substitution operations (like for the Hamming distance), we now allow to edit using deletions.

Definition 28

- The deletion operation \( D_j \) at position \( j \in [0, |u|] \) is defined over word \( u \in A^* \) as follows: \( D_j(u) = u_0u_1\ldots u_{j-1}u_{j+1}\ldots u_{|u|-1} \), for all \( u \in A^* \).
- The Levenshtein distance \( d_L \) is defined over \( u, v \in A^* \) as follows:

\[
d_L(u, v) = \frac{1}{2} \min \left\{ m + m' \left| \exists j_1 < \cdots < j_m, f_1 < \cdots < f_{m'} \right. \left. < f_{m'} \right) \circ D_{j_1} \circ \ldots \circ D_{j_m} \circ (u) = D_{f_1} \circ \ldots \circ D_{f_{m'}}(v) \right\}.
\]

Most frequently, we will consider the distance between two words of equal length, so that the result is an integer,
and can be defined as the minimal length \( m \) of two sequences \( D_{j_1} \ldots D_{j_n} \) and \( D_{j_1'} \ldots D_{j_n'} \) such that

\[
D_{j_1} \circ \ldots \circ D_{j_n}(u) = D_{j_1'} \circ \ldots \circ D_{j_n'}(v).
\]

The distance \( d_L(u, v) \) can also be defined as \( \frac{|u|+|v|}{2} - \ell \), where \( \ell \) is the length of the longest common (possibly non-contiguous) subword between \( u \) and \( v \).

Several variants exist in the literature:

- One may want to remove factor \( \frac{1}{2} \) in the definition, to make the definition look more natural. Nevertheless, the two points above, as well as the next two remarks, motivate our definition. Anyway, the two distances \( d_L \) and \( 2d_L \) are equivalent.
- If one allows two edition operations, insertion and deletion, the purpose could be that it can be defined by performing all operations only on one of the two words. The two distances are here exactly equal because an insertion on one side corresponds to a deletion on the other side. Manipulations are a little more technical because one has to deal with as many insertion operations as there are letters in the alphabet.
- If one additionally allows the substitution operation from Definition 12, with weight 1, then again the two obtained distances are equal, because a substitution corresponds to a sequence of an insertion and a deletion.
- If one gives the same weights to the substitution and deletion operations, then one gets an equivalent distance (bounded between \( d_L \) and \( 2d_L \)).

**Example 29** Let \( A = \{a, b\} \).

1. For \( u = ababab \) and \( v = bababa \), we have: \( d_L(u, v) = 1 \). Indeed, \( D_0(u) = babab \) (we delete the letter of index 0 in \( u \)), then we delete the last letter in the end of the word \( v \) and we find \( D_0(u) = babab = D_5(v) \). For the sake of comparison, note that \( d_H(u, v) = 6 \).

2. For \( u = aaaa \) and \( v = aaaa \), we have \( d_L(u, v) = \frac{1}{2} \) since it is enough to delete the last letter of \( v \).

**Remark 30** For every \( u, v \in A^* \), we have:

\[
\frac{|u|+|v|}{2} \leq d_L(u, v) \leq \frac{|u|+|v|}{2}.
\]

**Proof** The upper bound comes from the trivial edition sequence producing:

\[
D_1 \circ D_2 \circ \ldots \circ D_{|u|}(u) = \lambda = D_1 \circ \ldots \circ D_{|v|}(v).
\]

On the other hand, if

\[
D_{j_1} \circ D_{j_2} \circ \ldots \circ D_{j_n}(u) = D_{j_1} \circ D_{j_2} \circ \ldots \circ D_{j_n}(v),
\]

then

\[
|D_{j_1} \circ D_{j_2} \circ \ldots \circ D_{j_n}(u) - D_{j_1} \circ D_{j_2} \circ \ldots \circ D_{j_n}(v)| = \left| D_{j_1} \circ D_{j_2} \circ \ldots \circ D_{j_n}(v) \right|.
\]

Hence, \( |u| - m = |v| - m' \). Then we can conclude that

\[
\frac{|u|+|v|}{2} \leq \frac{|u|+|v|}{2} = d_L(u, v).
\]

**Remark 31** The Hamming distance is an upper bound for the Levenshtein distance, i.e., for all words \( u, v \in A^* \) such that \( |u| = |v| \),

\[
d_L(u, v) \leq d_H(u, v).
\]

**Proof** Let \( d_H(u, v) = m \). Then there exist \( j_1 < \ldots < j_m \) such that for all \( h \in [1, m] \), \( u_{j_h} \neq v_{j_h} \). If we delete \( u_{j_h} \) and \( v_{j_h} \) for all \( h \in [1, m] \), then we find \( D_{j_1} \circ \ldots \circ D_{j_m}(u) = D_{j_1} \circ \ldots \circ D_{j_m}(v) \). Hence: \( d_L(u, v) \leq \frac{m}{2} = d_H(u, v) \). \( \square \)

**Proposition 32** The Levenshtein distance is subadditive, i.e., for all words \( u, u', v, v' \),

\[
d_L(uu',vv') \leq d_L(u,v) + d_L(u',v').
\]

**Proof** Consider words \( u, u', v, v' \), and \( m, m', n, n' \) such that:

\[
D_{j_1} \circ \ldots \circ D_{j_m}(u) = D_{j_1'} \circ \ldots \circ D_{j_{m'}'}(v) \]

for some minimal edition sequences \( j_1 < \ldots < j_m < |u| \), \( j_1' < \ldots < j_{m'} < |v| \), \( n_1 < \ldots < n_{n'} < |u| \) and \( f_1 < \ldots < f_{n'} < |v| \), so that \( d_L(u, v) = \frac{m+n}{2} \) and \( d_L(u', v') = \frac{m'+n'}{2} \). By concatenating the two previous edited words, we obtain:

\[
D_{j_1} \circ \ldots \circ D_{j_m} \circ D_{j_{m+1}} \circ \ldots \circ D_{j_{m+n}}(uu') = D_{j_1'} \circ \ldots \circ D_{j_{m'}} \circ D_{j_{m'+1}} \circ \ldots \circ D_{j_{m'+n'}}(vv').
\]

Therefore \( d_L(uu', vv') \leq \frac{m+n+m'+n'}{2} = d_L(u, v) + d_L(u', v') \). \( \square \)

**Remark 33** Let \( u \in A^* \) and \( v \in B^* \). Thanks to the characterization of \( d_L \) in terms of longest common subword, we have

\[
d_L(u, v) = d_L(u', v) + \frac{1}{2} |u|_{A \Delta B},
\]

where \( u' \) is the subword of all letters in \( B \) from \( u \). In particular, if \( u \) and \( v \) have equal length, then \( d_L(u, v) \geq |u|_{A \Delta B} \).

This is because every deletion sequence must at least delete these letters and as many letters of \( v \).

If, besides, \( |B| = 1 \) (i.e., \( v = a^{|u|} \) for some \( a \)), then there is equality:
Following the idea behind the Besicovitch pseudometric, we define a pseudometric associated to the Levenshtein distance as follows:

**Definition 34** The Feldman–Katok pseudometric is:

\[ d_L(x, y) = \lim_{\ell \to \infty} \sup_{\ell} \frac{d_L(x[0, \ell], y[0, \ell])}{\ell}, \forall x, y \in A^N. \]

This pseudometric is often denoted \( \tilde{d} \) in the literature, but we keep our notation to emphasize the similarity with \( d_H \). Like the Besicovitch pseudometric, it is a pseudometric but not a distance. Actually, all pairs at Besicovitch pseudometric 0 are at Feldman–Katok pseudometric 0. More generally, the following inequality derives from Remark 31.

**Remark 35** For all \( x, y \in A^N \) we have: \( d_L(x, y) \leq d_H(x, y) \).

Here too, it is natural to quotient the space of configurations by the equivalence of zero distance; we obtain a metric space, called the Feldman–Katok space. Following the idea behind the Besicovitch pseudometric, it is a pseudometric but not a distance. Actually, all pairs at Besicovitch pseudometric 0 are at Feldman–Katok pseudometric 0. More generally, the following inequality derives from Remark 31.

**Definition 36** For every periodic configurations \( x, y \in A^N \) such that \( x[0, \ell] = y[0, \ell] \), we keep our notation to emphasize the similarity with \( d_H \). Like the Besicovitch pseudometric, it is a pseudometric but not a distance. Actually, all pairs at Besicovitch pseudometric 0 are at Feldman–Katok pseudometric 0. More generally, the following inequality derives from Remark 31.

**Definition 37** For every periodic configurations \( x, y \in A^N \), if \( d_L(x, y) = 0 \) then there exists \( k \in [0, p] \) such that \( x = \sigma^k(y) \).

**Proof** Let \( p \) be a common period for \( x \) and \( y \). According to Remark 36, there exist \( i, j \in \mathbb{N} \) such that \( x[i, i+p] = y[j, j+p] \).

Since both \( x \) and \( y \) are \( p \)-periodic for the shift, one can write, for every \( n \in \mathbb{N} \),

\[ x_n = x_{n-\text{mod}p+i} = y_{n-\text{mod}p+j} = y_{i+j+n} \]

Hence \( x = \sigma^{i+j+n} \).

One of the motivation to study the Besicovitch space is that the shift is an isometry over this space. In the Feldman–Katok space, this is still true, but even more than this: the shift is exactly the identity.

**Proposition 38** The shift over the Feldman–Katok space is the identity map.

**Proof** For \( x \in A^N \) and \( \ell \in \mathbb{N} \), we have \( d_L(x[0, \ell], \sigma(x)[0, \ell]) = d_L(x[0, \ell], x[1, \ell+1]) \leq 1 \): simply delete the first letter of \( x[0, \ell] \) and the last letter of \( \sigma(x)[0, \ell] \), to obtain \( x[1, \ell] \) in both cases. Hence:

\[ d_L(x, \sigma(x)) = \limsup_{\ell \to \infty} \frac{d_L(x[0, \ell], x[1, \ell+1])}{\ell} \leq \limsup_{\ell \to \infty} \frac{1}{\ell} = 0. \]

Since every equivalence class is invariant by the shift, dynamical systems over this space can be considered as acting on shift orbits.

Let us now see that the shifts are the only dill maps in the class of the identity.

**Theorem 39** For every dill map \( F, F \in \text{id}_{b_L} \iff \exists k \in \mathbb{N}, F = \sigma^k \).

**Proof** Proposition 38 proves that if \( F = \sigma^k \) for some \( k \in \mathbb{N} \setminus \{0\} \) then \( F \in \text{id}_{b_L} \).

Conversely, let \( F \) be a dill map with diameter \( \theta \in \mathbb{N} \setminus \{0\} \) and local rule \( f \) such that \( F \in \text{id}_{b_L} \). Let \( u \in A^* \) be such that for all \( w \in A^\theta \), \( w, \varnothing, u[0, \varnothing] = d^{\theta-1} \), for some \( a \in A \), \( u_0 \neq a \) and \( u[\varnothing] \neq a \). Let \( x = (ua^\theta)^\infty \) for some \( n \) strictly larger than both \( |u| \) and \( |f^*(ua^{\theta-1})| - |u| \). According to Proposition 37, there exists \( k \in [0, |u| + n] \) such that:

\[
F(x) = (f^*(ua^{\theta-1}))^\infty = \sigma^k(x) = \left( (ua^\theta)^{k+1} \right) \text{ if } k \leq |u|; \]
\[
(\sigma^{|u|+|k|}u_{a^{\theta-1}})^\infty \text{ if } k \geq |u|.
\]

Thanks to the assumption that \( u_0 \) is the first letter different from \( a \), and that there is no factor \( a^\theta \) in \( u \), we get the aperiodicity property of \( ua^\theta \): the shortest period of \( x \) is \( |u| + n \). Since \( F(x) \) has the same periods and \( |f^*(ua^{\theta-1})| \) is one of them, we deduce:
Cellular automata and substitutions... 519

Firstly, if \( q > 1 \), then by our assumption that 
\[ f^s(ua_0^q) = f^s(a_0^q) \]
ends with \( id \), hence it is not monochromatic. Hence \( q = 1 \), i.e.,
\[ |ua^q| = |f^s(ua^q)| = \sum_{i=0}^{n} f((ua)^{i-1}) |j, j+1\rangle. \]

Since \( f \) is nonrasing, we get that for every \( i \in [0, |u| + n] \),
\[ f((ua)^{i-1}) = 1, \]
and since \( u \) contains all words of \( A^0 \), we deduce that \( F \) is a CA.

Secondly, if \( \theta < k < |u| \), then \( f(a)^{k} = f^s(a^k) \)
starts with \( a \) but contains \( a_0 \) at position \( n - k + \delta \).
Similarly, if \( k > |u| \), then \( f(a)^{k} = f^s(a^k) \)
starts with \( a \) but contains \( a_0 \) at position \( n - k + |u| \).
In both cases, we have contradicted its monochromaticity.

We have proven that \( k \leq \theta \). Since \( u \) contains all words of \( A^0 \), we can conclude then, \( f(v) = v_k \), for all \( v \in A^0 \), which is exactly the local rule of \( \sigma^k \).

**Corollary 40** For every dill map \( F, F \in id_{b_k} \iff F \in id_{b_k} \iff \exists k \in \mathbb{N}, F = \sigma^k. \)

**Proof** If \( F \in id_{b_k} \), then \( b_k(x, F(x)) = 0 \), for all \( x \in A^n \). Hence \( b_k(x, F(x)) = 0 \) for all \( x \in A^n \) since \( b_k(x, F(x)) = b_k(x, F(x)) \), and thus, \( F \in id_{b_k} \). According to Theorem 39, there exists \( k \in \mathbb{N} \) such that \( F = \sigma^k \). On the other hand, if \( F = \sigma^k \) for some \( k \in \mathbb{N} \), then for all \( x \in A^n \) we have, \( b_k(x, F(x)) = 0 \) thanks to Proposition 38. Hence \( F \in id_{b_k} \).

It is proven in Salo and Törmä (2012) that any shift-commuting continuous map of the Besicovitch space which is equivalent to a CA is actually a CA. We give the corresponding result for the particular case of shifts.

**Corollary 41** If \( k \in \mathbb{N} \), then a dill map \( F \) is in the class \( \sigma^k_{b_k} \)
if and only if \( F = \sigma^k \).

**Proof** If \( F \in \sigma^k_{b_k} \), then \( F \in id_{b_k} \) (thanks to Remark 35 and Theorem 39). Then, according to Corollary 40 there exists \( k' \in \mathbb{N} \) such that \( F = \sigma^{k'} \). Hence, \( \sigma^k \sim b_k \sigma^k \). Therefore, thanks to Remark 17 we obtain that \( k = k' \). And thus, \( F = \sigma^k \).

### 4.3 Lipschitz property of dill maps

Now, we aim at proving that, unlike in the Besicovitch space, all dill maps are well-defined in the Feldman–Kátok space.

The following notion gives the idea of a dill map which transforms a deletion into a bounded number of deletions.

**Definition 42** A dill map \( F \) with local rule \( f \) is \((M, M')\)-bounded-deletion-spreading (BDS) for some \( M, M' \in \mathbb{N} \) if for every \( u \in A^+ \) and every \( j \in [0, |u|] \):
\[
d_L(f^s(D_j(u)), f^s(u)) \leq M + \frac{|f^s(u)| - |f^s(D_j(u))|}{2}, \]
\[
d_L(f^s(D_j(u)), f^s(u)) \leq M' - \frac{|f^s(u)| - |f^s(D_j(u))|}{2}. \]

Note that this property does not depend on the choice of a local rule, since the definition involves the application of \( f \) over arbitrarily long words.

**Proposition 43** Let \( F \) be a dill map with local rule \( f \) and diameter \( 0 \). Then \( F \) is \((0, |f|, (0 - 1) |f|)\)-BDS.

**Proof** Let \( u \in A^+ \) and \( j \in [0, |u|] \). By definition of \( f^s \), we have the following:
\[
f^s(u) = f^s(u_{0, j - 0|j}^j f^s(w) f^s(u_{j+1, |u|}) \]
and,
\[
f^s(D_j(u)) = f^s(u_{0, j - 0|j}^j f^s(w') f^s(u_{j+1, |u|}) \]
where \( w = f^s(u_{j-0, j+0|j}) \) and \( w' = f^s(D_j(u))_{j-0, j+0|j} \). By subadditivity and Remark 30, \( d_L(f^s(u), f^s(D_j(u))) \leq d_L(w, w') \leq \frac{|w| + |w'|}{2} \).

Obviously,
\[
|w| = |f^s(u_{j-0, j+0|j})| \leq |f| \quad \text{and} \quad |w'| = |f^s(D_j(u))_{j-0, j+0|j}| \leq (0 - 1) |f| .
\]
Moreover:\n\[
|f^s(u)| - |f^s(D_j(u))| = |w| - |w'| \]. \]

Hence,
\[
d_L(f^s(u), f^s(D_j(u))) \leq \frac{|w| + |w'|}{2} = |w| + \frac{|w| - |w'|}{2} \leq \frac{\theta |f| + |\Delta f^s(D_j(u))|}{2}.
\]

Also,
\[
d_L(f^s(u), f^s(D_j(u))) \leq \frac{|w| + |w'|}{2} = |w'| + \frac{|w| - |w'|}{2} \leq \frac{\theta |f| + |\Delta f^s(D_j(u))|}{2} .
\]

**Lemma 44** Let \( F \) be a \((M, M')\)-BDS dill map for some \( M, M' \in \mathbb{N} \) and local rule \( f \). Then for all \( l \in \mathbb{N} \) and \( u, v \in A^+ \), we have:
\[ d_L(f^*(u), f^*(v)) \leq (M + M')d_L(u, v) - \frac{|f^*(u)| - |f^*(v)|}{2}. \]

**Proof** Consider words \( u, v \), and \( m \in \mathbb{N} \) such that:
\[ D_{j_1} \cdots D_{j_m}(u) = D_{i_1} \cdots D_{i_m}(v), \]
for some minimal edit sequences \( j_1 < \cdots < j_m \leq |u| \) and \( i_1 < \cdots < i_m \leq |v| \), so that \( d_L(u, v) = m \). By the triangular inequality, one gets:
\[ d_L(f^*(u), f^*(v)) \leq \sum_{k=1}^{m} d_L(f^*(D_{j_k} \cdots D_{j_1}(u)), f^*(D_{i_k} \cdots D_{i_1}(v))) + \sum_{k=1}^{m} d_L(f^*(D_{j_k} \cdots D_{j_1}(u)), f^*(D_{i_k} \cdots D_{i_1}(v))). \]

Now our two assumptions allow to write:
\[ d_L(f^*(u), f^*(v)) \leq \sum_{k=1}^{m} \left( M + \frac{f^*(D_{j_k} \cdots D_{j_1}(u)) - f^*(D_{i_k} \cdots D_{i_1}(u))}{2} \right) + \sum_{k=1}^{m} \left( M + \frac{f^*(D_{i_k} \cdots D_{i_1}(v)) - f^*(D_{j_k} \cdots D_{j_1}(v))}{2} \right) \]
\[ \leq Mm + \frac{|f^*(u)| - |f^*(v)|}{2} + Mm + \frac{|f^*(u)| - |f^*(v)|}{2} \]
\[ \leq (M + M')m + \frac{|f^*(u)| - |f^*(v)|}{2}. \]

**Lemma 45** Let \( F \) be a \((M, M')\)-BDS dill map with local rule \( f \) and diameter \( \theta \). Let \( x \) be such that for every \( i \in \mathbb{N} \), \( |f(x[i, i+\theta]|) \geq L \), for some \( L > 0 \). Then for every \( y \in A^\mathbb{N} \), \( d_L(F(x), F(y)) \leq \frac{M + M'}{L} d_L(x, y) \).

**Proof** Let \( x, y \in A^\mathbb{N} \) and \( \ell \in \mathbb{N} \). Consider the largest \( k \in \mathbb{N} \) such that \( |f^*(x[0, \ell]|) \leq \ell \). Then \( F(x[0, \ell]) \) can be written \( f^*(x[0, \ell])w \) for some \( w \) of length less than \( \|f\| \). Note that \( \ell \geq \sum_{i=0}^{k-\theta} |f(x[i, i+\theta]|) \geq (k - \theta + 1)L \). Proposition 32 gives the following:
\[ d_L(F(x[0, \ell]), F(y[0, \ell])) \leq d_L(f^*(x[0, \ell]), f^*(y[0, \ell])) + d_L(w, F(y)[f^*(y[0, \ell])]). \]

Note that the previous inequality still holds if \( |f^*(y[0, \ell]|) \geq \ell \), in which case the second term is \( d_L(w, \lambda) = |w|/\ell \). Otherwise,
\[d_L(\tau(x), \tau(y)) \leq \frac{\|x\|}{L}d_L(x, y).\]

3. In particular, any dill map with local rule \(f\) and 
   diameter \(\theta\) is \((2L-1)\frac{\|f\|}{\|f\|}\)-Lipschitz.

   For example, any substitution \(\tau\) yields a \(\|\cdot\|\)-Lipschitz dynamical system.

\[d_L(\tau'(x), \tau'(y)) \leq \frac{\|x\|}{2\|\tau'\|}d_L(x, y) = \frac{1}{2}d_L(x, y).\]

So all iterates \(\tau^n\) are Lipschitz with a uniform coefficient: \(\tau^n\) is equicontinuous.

2. Consider any configuration \(x \in (A^-)^N\). Let \(k \in \mathbb{N} \setminus \{0\}\), 
   and \(y\) defined by \(y_i = x_i\) for every \(i \not\in k\mathbb{N}\), and \(y_i\) be any letter
   from \(A_i^+\), if \(i \in k\mathbb{N}\). Note that \(d_L(x, y) = d_H(x, y) = \frac{1}{2}\). From
   Remark 33, if \(u \in (A_i^-)^N\), then for every \(v \in A^+\) with equal 
   length, \(d_L(u, v) \geq \frac{1}{2}i [i \in [0, |v|]\{v_i \in A_i^+\}]\). Remark 4
   gives that \(A_i^+\) is stable by \(\tau\). Hence, for every \(u \in (A_i^-)^N\)
   and \(v \in A^+\) such that \(|u| = |\tau'(v)|\), \(d_L(u, \tau'(v))\) is at least 
   \(|\tau'(v)|A_i^+\). Also, from Lemma 6, there exists \(z > 0\) (which 
   does not depend on \(k\)) such that for every \(i \in \mathbb{N}\) and sufficiently
   large \(t \in \mathbb{N}\), \(|\tau'(y_{ik})|A_i^+ \geq z|\tau'(y_{ik})|\). So, for every \(m \in \mathbb{N}\),
   and sufficiently large \(t \in \mathbb{N}\),
   \[d_L(\tau'(x)|_{[0,|\tau'(y_{0, mk})|]}, \tau'(y)|_{[0, |\tau'(y_{0, mk})|]}) \geq |\tau'(y)|_{[0, |\tau'(y_{0, mk})|]}\]
   \[= \sum_{i=0}^{m-1} |\tau'(y_{ik})|A_i^+ \geq z \sum_{i=0}^{m-1} |\tau'(y_{ik})|.\]

On the other hand:
\[|\tau'(y)|_{[0, |\tau'(y_{0, mk})|]} = \sum_{i=0}^{m-1} \left(|\tau'(y_{ik})| + |\tau'(y_{[k, (i+1)k]})|\right).\]

Overall, we get:
\[d_L(\tau'(x)|_{[0, |\tau'(y_{0, mk})|]}, \tau'(y)|_{[0, |\tau'(y_{0, mk})|]}) \geq 1 \frac{1}{2} + k \cdot \frac{a_{i \rightarrow \infty}(\|\tau'\|)}{\Theta_{i \rightarrow \infty}(\|\tau'\|)}.\]

This converges to \(z > 0\), independently of \(k \in \mathbb{N}\). Since \(k\) was taken arbitrary, \(y\) is arbitrarily close to \(x\), 
so that \(x\) is not an equicontinuity point.

\[\square\]

Moreover, from the proof above, remark that if there
exists a maximal terminal component, i.e., some letter \(a \in A_i^+\) such that for every \(t \in \mathbb{N}\), \(\tau'(a)\) contains only letters 
from \(A_i^+\), then \(a\) can be taken equal to 1. This proves some 
extreme form of non-equi-continuity: whatever the precision 
\(\frac{1}{2}\) with which one initially measures \(x\), there is a neighboring configuration \(y\) whose orbit will be nearly 
maximally distant to that of \(x\).
Example 49  Consider the following substitution $\tau$:

$\tau : a \mapsto ab$

$\quad b \mapsto b \quad M(\tau) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

Then $A_+^\tau = \{a\}$ and $A_-^\tau = \{b\}$ are both nontrivial. Yet, the system is asymptotically nilpotent: every orbit converges to the class of $b^\infty$ (even in the Besicovitch space). In particular all configurations are equicontinuous, even $b^\infty \in A_-^\tau$.

Here, the spectral radius is 1 and the growth is linear.

We believe that a substitution yields an equicontinuous system in the Feldman–Katok space if and only if it either is quasi-uniform, or admits a unique terminal component, which is a single vertex.

Apart from the subalphabet argument from Remark 33, it is quite hard to prove lower bounds for the class of example of a dill map without equicontinuity configuration.

Example 50  The Fibonacci substitution $\tau$ is irreducible, so $\tau$ is equicontinuous in the Feldman–Katok space, though it admits almost no equicontinuity point in the Besicovitch space (see Example 18).

Example 51  Let $F$ be the Xor CA. Then neither $F_{d_1}$ nor $F_{d_2}$ is equicontinuous. Indeed, let us prove that $x = a^\infty$ is a non-equicontinuity point. Let $k \in \mathbb{N}$, and $y = (a^{2^k-1}b)^\infty$.

Then $d_L(x, y) \leq d_H(x, y) \leq 2^{-k}$. A classical induction on $k$ gives that $F^{2^k-1}(y) = b^\infty$ (for more details see Kůrka 2003, Example 5.6). Hence:

$d_H(F^{2^k-1}(x), F^{2^k-1}(y)) = d_L(F^{2^k-1}(x), F^{2^k-1}(y)) = 1$.

Figure 1 illustrates that $b^\infty$ is also a non-equicontinuity point.

Example 52  Let $\tau$ be a substitution defined as follows:

\[ \tau : a \mapsto b \quad M(\tau) = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} \quad M(\tau)^2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}. \]

Hence, $M(\tau)$ is irreducible (though not primitive: no iteration yields a positive matrix, unlike all of our irreducible examples so far). So, according to Corollary 48, $\tau_{b_1}$ is equicontinuous. More precisely, $\tau$ is actually the doubling substitution, proven to be 1-Lipschitz by Corollary 46.

Besides, it can be shown that the latter is even an isometry: the longest common subword of $\tau^2(u)$ and $\tau^2(v)$ is always obtained by doubling a common subword of $u$ and $v$.

Though quasi-uniform substitutions behave smoothly in our spaces, other substitutions may be more pathological.

Example 53  Let $\tau$ be a substitution defined over $A = \{a, b\}^\mathbb{N}$ as follows:

$\tau : a \mapsto a \quad M(\tau) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad M(\tau)^n = \begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix}$.

Hence there are two components, $\{a\}$ and $\{b\}$, and $A_-^\tau = \{b\}$. Then, according to Theorem 47, $b^\infty$ is an equicontinuity point of $\tau_{b_1}$ and $a^\infty$ is not, as illustrated in Figs. 2 and 3 (where $b$ is represented in black, and $a$ in red).

One natural question is whether, in the Feldman–Katok space, the orbit of a dill map always converges towards its (classical) limit set (the well-studied substitutive subshift). On the one hand, in the Besicovitch space, this is not always the case (see Example 24).

On the other hand, the limit set of any primitive substitutions (endowed with the shift map) is a particular finite-rank system, hence, as explained in García-Ramos and Kwietniak (2020, Sect. 6.1), a particular topologically loosely Kronecker system. From García-Ramos and Kwietniak (2020, Theorem 1.1), this implies that it is a
singleton in the Feldman–Katok space. The tools involved may be useful to understand our question.

### 4.5 Expansivity

It was proven in Blanchard et al. (1997) that there is no expansive CA over the Besicovitch space. By a similar method, we prove that there is no expansive CA over the Feldman–Katok space. Note that it does not derive directly from the corresponding result in the Besicovitch space, because some configurations from different Besicovitch classes could be in the same Feldman–Katok class, hence not concerned by the expansivity property.

**Lemma 54** If \( x = a^{(x)} b^{(y)} \) with \( a \neq b \), then:

\[
\delta(x, a^\infty) = \delta_H(x, a^\infty) = \limsup_{n \to \infty} \frac{\sum_{i=0}^{n-1} \beta_i}{\sum_{i=0}^{n-1} (x_i + \beta_i)}.
\]

**Proof** Thanks to Remark 33, we have \( \delta_L(x, a^\infty) = \delta_H(x, a^\infty) \).

For \( n \in \mathbb{N} \), let \( \ell_n = \sum_{i=0}^{n-1} (x_i + \beta_i) \). We have:

\[
\delta_H(x^{[0,\ell_n]}, a^{\ell_n^\infty}) = \sum_{i=0}^{n-1} \beta_i.
\]

Then:

\[
\limsup_{n \to \infty} \frac{\sum_{i=0}^{n-1} \beta_i}{\sum_{i=0}^{n-1} (x_i + \beta_i)} = \limsup_{n \to \infty} \frac{\delta_H(x^{[0,\ell_n]}, a^{\ell_n^\infty})}{\ell_n}.
\]

\[
\leq \limsup_{n \to \infty} \frac{\delta_H(x^{[0,\delta]}, a^{\delta^\infty})}{\ell} = \delta_H(x, a^\infty).
\]

On the other hand, for \( \ell \in \mathbb{N} \), there exist \( n = n(\ell) \in \mathbb{N} \) and \( \delta \in [0, x_n + \beta_n] \) such that \( \ell = \sum_{i=0}^{n-1} (x_i + \beta_i) + \delta \). If \( \delta \leq x_n \), then:
On the other hand, for 
\[ a \neq a_i \text{ and } a_{i+1} = f(a_i^t) \text{ for every } t \in \mathbb{N}. \]

The previous remark, applied in parallel, gives, if \( b_i \) is any letter different from \( a_i \), that \( F^t(x) \) has at least as many occurrences of \( a_i \) as \( a_i^{(x_i \rightarrow b_i)_2} b_i^{(y_i \rightarrow b_i)_2} \).

\[ (a_i^{(x_i \rightarrow b_i)_2} b_i^{(y_i \rightarrow b_i)_2})_m = a_i \Rightarrow F^t(a_i^{(x_i \rightarrow b_i)_2} b_i^{(y_i \rightarrow b_i)_2})_m = a_i. \]

In other words,
\[ d_H(F^t(a_i^{(x_i \rightarrow b_i)_2} b_i^{(y_i \rightarrow b_i)_2}), a_i^\infty) \leq d_H(a_i^{(x_i \rightarrow b_i)_2} b_i^{(y_i \rightarrow b_i)_2}, a_i^\infty). \]

As a consequence:
\[ d_L(F^t(x), F^t(y)) \leq d_H(F^t(a_i^{(x_i \rightarrow b_i)_2} b_i^{(y_i \rightarrow b_i)_2}), a_i^\infty) \]
\[ \leq d_H(a_i^{(x_i \rightarrow b_i)_2} b_i^{(y_i \rightarrow b_i)_2}, a_i^\infty) \]
\[ \leq \limsup_{n \to \infty} \frac{\sum_{i=0}^{n-1} \beta_i + n \theta}{\sum_{i=0}^{n-1} (x_i + \beta_i)} \text{ (by shift-invariance)} \]
\[ \leq \limsup_{n \to \infty} \frac{\sum_{i=0}^{n-1} \beta_i + \frac{n \theta}{p}}{\sum_{i=0}^{n-1} (x_i + \beta_i)} \]
\[ \leq \frac{1}{1 + p} \cdot \epsilon. \]

The orbits of these two distinct configurations stay very close forever. Hence, \( F \) is not expansive. \( \qed \)

Other dynamical properties would be interesting to study but seem much harder to tackle because of the lack of tools to compute general lower bounds for the Feldman–Katok pseudometric. This problem is known to have high algorithms complexity.

### 5 Conclusion and perspectives

In this paper, we studied CA, substitutions, and in general dill maps over two nontrivial topological spaces (the Besicovitch space and the Feldman–Katok space). Those spaces were constructed using two pseudometrics depending on two different edit distances over finite words (the Hamming distance and the Levenshtein distance). It was known since (Cattaneo et al. 1997; Blanchard et al. 1997) that the Besicovitch space is a suitable playground to study the dynamics of cellular automata. Here we saw that the same can be said for the Feldman–Katok space with respect to the dynamics of dill maps. In this space, the shift is equal to the identity, there are no expansive CA, every substitution is well-defined and admits at least one equicontinuous point.
This makes it natural to suggest a general definition, using any distance $d$ over the set of finite words:

**Definition 56** We define the centered pseudometric, denoted by $d_C$, as follows:

$$d_C(x, y) = \lim_{\ell \to \infty} \sup_{0 \leq \ell \leq \ell' \leq |x|} \frac{|d_{ij}(x_{[0, \ell]}, y_{[0, \ell]})|}{\max_{u, v \in A^*} d(u, v)}, \forall x, y \in A^N.$$

Following another point of view, a similar pseudometric, known as Weyl pseudometric, also based on $d_M$, measures the density of differences between two given sequences in arbitrary segments of a given length. A general definition, based on any distance $d$ over finite words, would become the following:

**Definition 57** We define the sliding pseudometric, denoted by $d_S$, as follows:

$$d_S(x, y) = \lim_{\ell \to \infty} \max_{k \in \mathbb{N}} \frac{d_{ij}(x_{[k, k + \ell]}, y_{[k, k + \ell]})}{\max_{u, v \in A^*} d(u, v)}, \forall x, y \in A^N.$$

The Weyl space shares many properties with the Besicovitch space; one of the main differences though is that it is not complete, according to Downarowicz and Iwanik (1988); in terms of dynamics, an open question is whether there exists an expansive cellular automaton over the Weyl space. A relevant question is now the following: which properties of distance $d$ make dill maps well-defined in the corresponding pseudometric space?

Generalizations exist of the Besicovitch pseudometric over groups (see for instance Łacka and Straszak 2016; Capobianco et al. 2020). An interesting work would be to generalize more of these pseudometrics to this setting. Let us replace $(\mathbb{N}, +)$ by any monoid $(\mathbb{M}, \cdot)$. A pattern with finite support $U \subset \mathbb{M}$ is some coloring $u \in A^U$. If $U = V \cdot g$, then the translate by $g \in \mathbb{M}$ of a pattern $u \in A^U$ is the pattern $\sigma^g(u)$ defined on support $V$ such that $\sigma^g(u)_v = u_{vg}$. Let $G$ be a (say finite) set of right-cancelable elements, that is, $i \cdot g = j \cdot g \Rightarrow i = j$ for every $g \in G$. The deletion $D_j^g$ at position $j \in \mathbb{M}$ with respect to $g \in G$ is the function mapping any pattern $u$ with support $U$ into the pattern $v$ defined over support $V = U \setminus \{jg^k | k \in \mathbb{N}\} \cup \{jg^{k+1} | k \in \mathbb{N}, g^{k+1} \in U \setminus \{j\}\}$ by $v_i = u_i$ if $i \in U \setminus \{jg^k | k \in \mathbb{N}\}$ and $v_{jg^k} = u_{jg^{k+1}}$ otherwise. By the cancelability property, $|V| = |U \setminus \{j\}|$, so that iterating $|U|$ deletions can exhaust the subset. Now one can consider, as a generalization of the Levenshtein distance, the following: the distance $d_M(u, v)$ between patterns $u \in A^U$ and $v \in A^V$ is the minimal half-number $\frac{m+n}{2}$ of deletions such that $D_{j_1}^{g_1} \circ D_{j_2}^{g_2} \circ \cdots \circ D_{j_m}^{g_m}(u)$ and $D_{j_1}^{g_1} \circ D_{j_2}^{g_2} \circ \cdots \circ D_{j_m}^{g_m}(v)$ have a common translate, where $j_1, \ldots, j_m, g_1, \ldots, g_m \in \mathbb{N}$ and $g_1, \ldots, g_m, g'_1, \ldots, g'_m \in G$. From the cardinality remark above, this distance is at most $\frac{|U|+|V|}{2}$. The Feldman–Katok-like pseudometric over configurations from $A^M$, endowed with a spanning sequence $(F_n)_{n \in \mathbb{N}}$ of finite subsets of $\mathbb{M}$, would then be:

$$d_M = \lim_{n \to \infty} \frac{d_M(x_{F_n}, y_{F_n})}{|F_n|}.$$ 

From the previous remark, $d_M$ is between 0 and 1. To expect the pseudometric to enjoy nice properties (like shift-invariance, for example), one should probably assume some Folner-like condition (without any, pathologies are known to appear for the Besicovitch distance Capobianco et al. 2020).

Once any distance $d$ over finite patterns is fixed, a generalisation of Definition 56 is relevant for space $A^M$, where $\mathbb{M}$ is any space endowed with a spanning sequence $(F_n)$, and a generalisation of Definition 57 is also meaningful, with the additional assumption that $\mathbb{M}$ is a monoid, so as to be able to slide the windows.

**Acknowledgements** We thank the anonymous referee for many valuable comments and corrections. We also thank Ville Salo, with whom the idea of dill maps was discussed, Dominik Kwietniak and Felipe Garcia-Ramos, for pointing to the bibliography about the Feldman–Katok space.

**Declarations**

**Competing Interests** The authors have no relevant financial or non-financial interests to disclose.

**References**

Banach S (1922) Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. Fundam Math 3:133–181

Blanchard F, Formenti E, Kürka P (1997) Cellular automata in the Cantor, Besicovitch, and Weyl topological spaces. Complex Syst 11:107–123

Berthé V, Rigo M, eds (2010) Combinatorics, automata, and number theory., volume 135 of Encyclopedia of Mathematics and Its Applications. Cambridge: Cambridge University Press

Cattaneo G, Formenti E, Mazoyer J (1997) A shift-invariant metric on $S^2$ inducing a non-trivial topology. In: International Symposium on Mathematical Foundations of Computer Science, pages 179–188. Springer

Capobianco S, Guillon P, Noûs C (2020) A characterization of amenable groups by besicovitch pseudodistances. In: Zenil H (ed) Cellular automata and discrete complex systems. Springer International Publishing, Cham, pp 99–110

Downarowicz T, Iwanik A (1988) Quasi-uniform convergence in compact dynamical systems. Stud Math 89(1):11–25

Feldman J (1976) New K-automorphisms and a problem of Kakutani. Israel J Math 24(1):16–38
Formenti E, Kurka P (2009) Dynamics of cellular automata in non-compact spaces
García-Ramos F (2017) Weak forms of topological and measure theoretical equicontinuity: relationships with discrete spectrum and sequence entropy. Ergodic Theory Dynam Syst 37(4):1211–1237
García-Ramos F, Kwietniak D (2020) On topological models of zero entropy loosely Bernoulli systems. arXiv:2005.02484
Hedlund GA (1969) Endomorphisms and automorphisms of the shift dynamical system. Math Syst Theory 3(4):320–375
Katok AB (1977) Monotone equivalence in Ergodic theory. Math USSR-Izvestiya 11(1):99–146
Kwietniak D, Łacka M (2017) Feldman-Katok pseudometric and the GIKN construction of nonhyperbolic ergodic measures. arXiv:1702.01962
Kurka P (2003) Topological and symbolic dynamics. Société mathématique de France Paris
Levenshtein V (1966) Binary codes capable of correcting deletions, insertions, and reversals. In: Soviet physics doklady, volume 10, pages 707–710
Łacka M, Straszak M (2016) Quasi-uniform Convergence in Dynamical Systems Generated by an Amenable Group Action. arXiv:1610.09675
Müller J, Spandl C (2009) A Curtis-Hedlund-Lyndon theorem for Besicovitch and Weyl spaces. Theoret Comput Sci 410(38–40):3606–3615
Ornstein DS (1974) Ergodic theory, randomness, and dynamical systems. Number 5 in Yale mathematical monographs. Yale University Press, Yale
Ornstein DS, Rudolph DJ, Weiss B (1982) Equivalence of measure preserving transformations. Number 262 in Memoirs of the AMS. AMS
Pansiot J-J (1984) Complexité des facteurs des mots infinis engendrés par morphismes itérés. In Automata, languages and programming, volume 172 of Lecture Notes in Computer Science, pages 380–389, Antwerp. Springer Berlin
Pytheas FN, Berthé V, Ferenczi S, Mauduit C, Siegel A eds (2002) Substitutions in dynamics, arithmetics and combinatorics, volume 1794 of Lecture Notes in Mathematics. Berlin: Springer
Saló V, Törnä I (2012) Geometry and dynamics of the Besicovitch and Weyl spaces. In Developments in language theory. 16th international conference, DLT 2012, Taipei, Taiwan, August 14–17, 2012. Proceedings, pages 465–470. Berlin: Springer
Saló V, Törnä I (2015) Block maps between primitive uniform and Pisot substitutions. Ergodic Theory Dyn Syst 35(7):2292–2310

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.