Ergod. Th. & Dynam. Sys., (2024), **44**, 408–431
doi:10.1017/etds.2023.16

Mixing properties of erasing interval maps

DARIO CORONA and ALESSANDRO DELLA CORTE

Mathematics Division, School of Sciences and Technology, University of Camerino, Camerino 62032, Italy
(e-mail: dario.corona@unicam.it, alessandro.dellacorte@unicam.it)

(*Received* 27 *August* 2022 *and accepted in revised form* 2 *February* 2023)

Abstract. We study the measurable dynamical properties of the interval map generated by the model-case erasing substitution ρ, defined by

$$\rho(00) = \text{empty word}, \quad \rho(01) = 1, \quad \rho(10) = 0, \quad \rho(11) = 01.$$

We prove that, although the map is singular, its square preserves the Lebesgue measure and is strongly mixing, thus ergodic, with respect to it. We discuss the extension of the results to more general erasing maps.

Key words: mixing, substitutions, combinatorics on words
2020 Mathematics Subject Classification: 37A25, 28D05, 37A05 (Primary); 68R15 (Secondary)

1. *Introduction*

Dynamical systems defined by means of substitutions represent a well-developed research field, with rich interactions with ergodic theory, spectral analysis, chaos theory and number theory. In the usual meaning, substitutions are rules for replacing symbols or words on a given alphabet by non-empty words. The iterated application of the substitution rule to a given finite or infinite word determines a discrete dynamical system (see [**13**, **19**] for useful reference works on substitutive dynamics).

In the recent past, some attention has been devoted to the more general case in which also the empty word is allowed as an output. These maps are called *erasing substitutions*, and the corresponding morphism, induced by concatenation on $\{0, 1\}^\infty$, has a dense set of discontinuities in the standard product topology.

Some works within theoretical computer science have addressed the problem of extending classical results on substitutions to the erasing case (see, for instance, [**10**, **11**, **20**]). On a more specifically dynamical ground, the study of maps generated by the action of erasing substitutions on the binary expansion of reals was begun in [**5**, **6**]. In our opinion, the main interest of these systems arises from the combination of very simple defining rules with rich properties and dynamical behavior. This makes them a good model case for the investigation of dynamical properties of densely discontinuous maps, which are attracting increasing attention in the last years (see, for instance, [15, 21, 22]).

The present paper extends the results of [5, 6] by widening the perspective from purely topological to measurable dynamical properties, specifically mixing.

The paper is organized as follows: in §2, we define the erasing interval maps, give some background concepts and prove that the iterates of even order of the model-case erasing map preserve the Lebesgue measure; in §3, we prove that the above map is strongly mixing, thus weakly mixing and ergodic, with respect to the Lebesgue measure; in §4, we discuss some possible generalization of our results to other erasing interval maps.

2. *Erasing interval maps*

We start by defining what we mean herein by erasing substitution. The notation used throughout the paper is summarized in Appendix A.

Definition 2.1. We call *simple substitution* a map $\sigma : \{0, 1\} \to \{0, 1\}^*$. For every integer $k \geq 2$, we call a map $\sigma : \{0, 1\}^k \to \{0, 1\}^*$ a *k-block substitution*. By *erasing k-block substitution* we mean a k-block substitution σ such that there exists a unique $w \in \{0, 1\}^k$ verifying $\sigma(w) = \epsilon$. We indicate the unique word w whose σ-image is the empty word by the symbol w_ϵ.

We say that $w \in \{0, 1\}^\infty$ is a *binary expansion* for $x \in [0, 1]$ if

$$x = (0.w)_2 := \sum_{i=1}^{|w|} 2^{-i} w_i. \tag{2.1}$$

Note that the previous equation applies whether the length of w is finite or infinite. For $x \in (0, 1]$, we indicate by $w_x \in \{0, 1\}^\omega$ the unique infinite binary expansion of x not ending with 0^∞.

Let us denote the concatenation of words multiplicatively. Assuming $\sigma(\epsilon) = \epsilon$, the k-block substitution σ can be extended by concatenation to a map over $(\bigcup_{n\in\mathbb{N}_0}\{0, 1\}^{nk}) \cup \{0, 1\}^\omega$ by setting, for every w belonging to this set,

$$\sigma(w) = \prod_{i=0}^{\frac{|w|}{k}-1} \sigma(w_{ki+1}w_{ki+2} \ldots w_{ki+k}),$$

where the concatenation index has to be intended to be up to ∞ if $w \in \{0, 1\}^\omega$.

We are now ready to introduce what we call the *erasing interval maps*, which are maps from [0, 1] to itself generated by the symbolic action of an erasing substitution on the binary expansion of reals.

Definition 2.2. Given an erasing substitution σ, the interval map generated by it, indicated by f_σ, is defined as follows:

$$f_\sigma(x) := \begin{cases} \displaystyle\sum_{h=1}^{|\sigma(w_x)|} \frac{(\sigma(w_x))_h}{2^h} = (0.\sigma(w_x))_2 & \text{if } x \in (0, 1] \text{ and } w_x \neq w_\epsilon^\infty, \\ 0 & \text{if } w_x = w_\epsilon^\infty \text{ or } x = 0. \end{cases} \tag{2.2}$$

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

In the following, the main objects of our analysis are the model-case erasing block substitution

$$\rho : \{0, 1\}^2 \to \{0, 1\}^*$$

defined as follows:

$$\rho(00) = \epsilon, \quad \rho(01) = 1, \quad \rho(10) = 0, \quad \rho(11) = 01, \tag{2.3}$$

and the corresponding erasing interval map $f_\rho \colon [0, 1] \to [0, 1]$, defined using (2.2).

Some properties of this map were already investigated in [6], where it was denoted by R. We point out that the map R was a slightly different object than f_ρ. Indeed, R maps 0 to $2/3$ and this makes its combinatorial dynamical properties more uniform. In the present paper, this slight discrepancy makes no difference as we are interested in measurable dynamical properties concerning absolutely continuous measures, which are unaffected by what happens at single points.

Let us now consider the following map $\rho_e : \{0, 1\}^\infty \to \{0, 1\}^\infty$:

$$\rho_e : \begin{cases} 0 \mapsto \epsilon, & \\ 1 \mapsto 0 & \text{in odd positions,} \\ 1 \mapsto 1 & \text{in even positions.} \end{cases} \tag{2.4}$$

As the action of ρ_e coincides with that of ρ on the words of even or infinite length, it can be taken as an extension of ρ. In the following, we write simply $\rho(w)$ instead of $\rho_e(w)$ for words of any length.

The definition of ρ given through (2.4) shows why ρ can be considered the model-case erasing substitution. Indeed, ρ is defined using the simplest non-trivial alphabet, namely $\{0, 1\}$, and acts by sending the digit 0 to the empty word and the digit 1 to 0 or 1 according to the parity of its position in the original word, which is arguably the simplest possible choice.

For the reader convenience, we recall here the main results of the two aforementioned papers, namely [5, 6], because we are now ready to formulate them more precisely:

- in [6] it has been proven that f_ρ is Devaney chaotic, exhibits distributional chaos of type 1 (in the sense of [1]), has infinite topological entropy and uncountably many periodic cycles for every order; moreover, it is Borel-singular and not bi-measurable, because there are sets of Lebesgue measure zero whose f_ρ-image is not measurable;
- in [5], Li–Yorke and Devaney chaos, as well as infinite topological entropy, have been established for maps generated by a more general class of erasing substitutions.

As the results that we are going to show will concern the even-order iterates of f_ρ, and f_ρ^2 is equal to f_{ρ^2} Lebesgue-almost everywhere (as we prove in Lemma 3.5), a legitimate question is whether ρ^2 itself can be rewritten as a block substitution. However, the following lemma shows that this is not the case, even asking only that ρ^2 coincide with a block substitution on *infinite* words.

LEMMA 2.3. *The map $\rho^2 \colon \{0, 1\}^\omega \to \{0, 1\}^\infty$ is not expressible as a block substitution. That is, there is no k-block substitution ς, with $k \geq 2$, such that $\rho^2(w) = \varsigma(w)$ for every $w \in \{0, 1\}^\omega$.*

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

Proof. Suppose, towards a contradiction, that there exists $k \in \mathbb{N}$ such that ρ^2 coincides with the k-block substitution ς when both are extended to a morphism over $\{0, 1\}^\omega$. Note that, because $\rho^2(0^k) = \epsilon$, the equality $\rho^2(0^\infty) = \varsigma(0^\infty)$ implies $\varsigma(0^k) = \epsilon$. Moreover, it should be

$$\rho^2(w) = \varsigma(w) \quad \text{for every } w \in \{0, 1\}^k, \tag{2.5}$$

because

$$\varsigma(w) = \varsigma(w0^\infty) = \rho^2(w0^\infty) = \rho^2(w).$$

Set now $a = 0^{k-1}$ and take $b \in \{0, 1\}^k$ such that $\varsigma(b) \neq \epsilon$. Such a word b must exist, because otherwise the block substitution ς would send all words to the empty word. Take also $c \in \{0, 1\}^\omega$.

Recalling (2.5), and because, in particular, $\varsigma(1a) = \rho^2(1a) = \epsilon$, we have

$$\begin{aligned}\rho^2(0abc) &= \varsigma(0abc) = \varsigma(0^k)\varsigma(b)\varsigma(c) = \varsigma(1a)\varsigma(b)\varsigma(c) = \varsigma(1abc)\\ &= \rho^2(1abc).\end{aligned} \tag{2.6}$$

Note now that

$$\rho^2(0abc) = \rho(\rho(0^k)u\rho(c)) = \rho(u)w, \tag{2.7}$$

where:

- u is equal to $\rho(b)$ or its bitwise negation $\widetilde{\rho(b)}$, depending on the parity of k;
- w is equal to $\rho^2(c)$ or $\widetilde{\rho^2(c)}$, depending on the parity of $|u|$.

On the other hand, we have

$$\rho^2(1abc) = \rho(\rho(1a)u\rho(c)) = \rho(0u\rho(c)) = \widetilde{\rho(u)}\widetilde{w} = \widetilde{\rho(u)w}, \tag{2.8}$$

where u, v, w are defined similarly as before.

By equation (2.6), the right-hand sides of equations (2.7) and (2.8) must be equal, so it should be

$$\rho^2(0abc) = \rho^2(1abc) = \epsilon,$$

which is absurd because, again by the chain of equalities in equation (2.6), both have $\varsigma(b) \neq \epsilon$ as a subword. □

The previous result tells us that we cannot hope to straightforwardly apply to ρ^2 the tools developed in [5], so that we must study this map and the corresponding interval map $(f_\rho)^2$ in itself.

As an immediate consequence of [6, Proposition 5.3], the f_ρ-preimages of single points has Lebesgue measure zero. However, the map f_ρ is not very well-behaved with respect to the Lebesgue measure, being a singular map. To show this, let us define the set of points whose binary expansion does not contain the word 00:

$$M := \{x \in [0, 1] : w_x \text{ does not have pairs of consecutive 0s}\}.$$

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

As the subword 00 has no effect in the ρ-image of any word, it is clear that $f_\rho(M) = [0, 1]$. On the other hand, the Lebesgue measure of M is 0, because none of its points can be in $\mathcal{N}_2$, namely in the set of normal real numbers in base 2 (see [18] for a reference work on normal numbers). Moreover, in [6, Proposition 5.2] it has also been proven that $f_\rho(\mathcal{N}_2) \subset [0, 1] \setminus \mathcal{N}_2$. Therefore, f_ρ displaces all the Lebesgue-mass in the null set $[0, 1] \setminus \mathcal{N}_2$, whereas at the same time it maps the null set M to the whole interval $[0, 1]$. Quite surprisingly, it does so in such a way that in two iterations 'things are set right again'. More precisely, we have the following result.

THEOREM 2.4. *The interval map f_ρ^2 preserves the Lebesgue measure.*

The proof of Theorem 2.4 is given later, after having proven some needed lemmas. Let us first look at some numerical computations to get a sense of what is going on. The behavior of the even and odd iterates of f_ρ is graphically shown by the numerical results presented in Figure 1, where the substantial difference between the odd and even iterates of f_ρ can be appreciated. As already recalled, f_ρ is singular and, therefore, it cannot preserve any absolutely continuous measure. The graphs corresponding to the odd iterates can give some hint on a singular-continuous measure possibly preserved by it. Whether it exists is an open question for the authors. The self-similarities shown by the plots corresponding to the odd iterates are linked to the self-similarity properties of the graph of f_ρ, which are investigated in some depth in §4 of [6].

Before giving the proof of Theorem 2.4, let us introduce some useful notation.

For every $w \in \{0, 1\}^*$, we indicate by $[w]$ the real cylinder set corresponding to w, that is the dyadic subinterval of $[0, 1]$ whose points have a binary expansion that starts with w:

$$[w] = \{x \in [0, 1] : x = (0.wv)_2 \text{ for some } v \in \{0, 1\}^\infty\}. \tag{2.9}$$

We remark that, with this definition, $[\epsilon] = [0, 1]$ and all real cylinder sets are closed intervals whose endpoints are dyadic rationals.

For $w \in \{0, 1\}^*$, we indicate by $\langle\langle w\rangle\rangle$ the (countable) set of all distinct words obtained by inserting (in any position, but not as a suffix) any finite number of subwords of type 00 in w. More formally, for $w = w_1 w_2 \ldots w_k \in \{0, 1\}^k$, a word v is in $\langle\langle w\rangle\rangle$ if it can be written as

$$v = (00)^{h_1} w_1 (00)^{h_2} w_2 \ldots (00)^{h_k} w_k, \tag{2.10}$$

for some choice of the k non-negative integers $h_1, \ldots, h_k$. Thus, for instance,

$$\langle\langle 1\rangle\rangle = \{w \in \{0, 1\}^* : w = 0^{2k}1,\ k \in \mathbb{N}_0\},$$

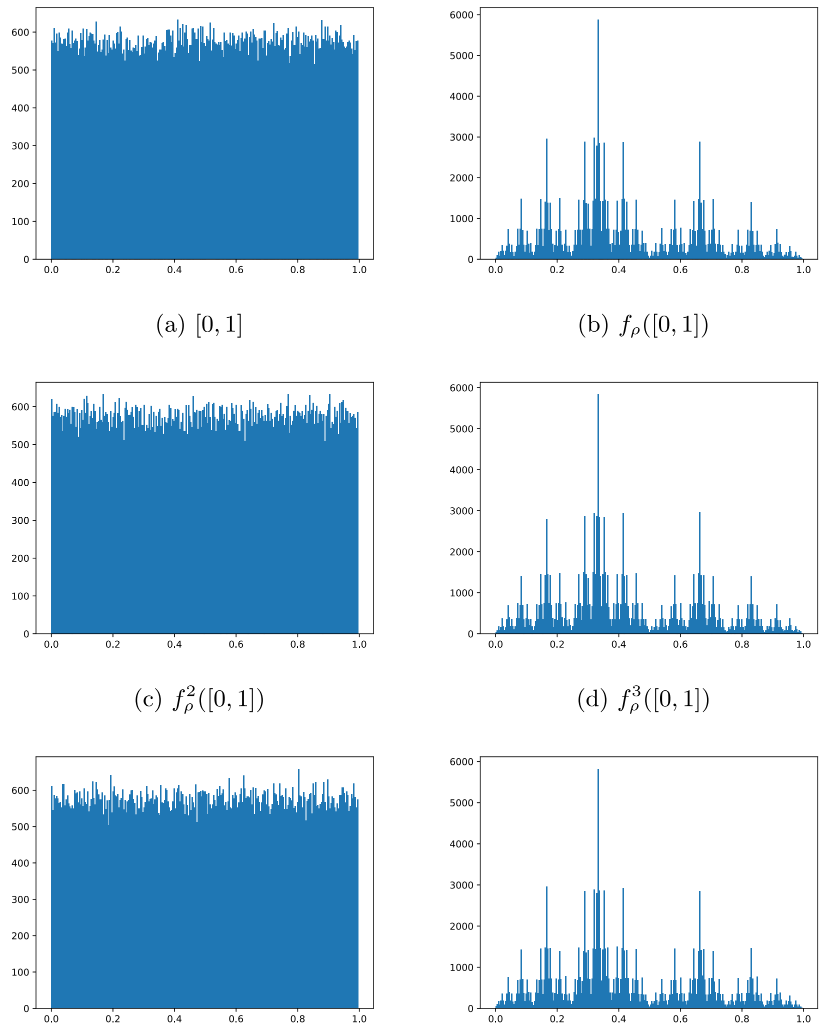
and

$$\langle\langle 101\rangle\rangle = \{w \in \{0, 1\}^* : w = 0^{2k}10^{2h}01,\ k, h \in \mathbb{N}_0\}.$$

For every $w \in \{0, 1\}^*$, we indicate by $[[w]]$ the countable union of the real cylinder sets characterized by words in $\langle\langle w\rangle\rangle$, namely

$$[[w]] = \bigcup_{v \in \langle\langle w\rangle\rangle} [v]. \tag{2.11}$$

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press



(a) $[0, 1]$ (b) $f_\rho([0, 1])$

(c) $f_\rho^2([0, 1])$ (d) $f_\rho^3([0, 1])$

(e) $f_\rho^4([0, 1])$ (f) $f_\rho^5([0, 1])$

FIGURE 1. Numerical results on the first five iterates of $f_\rho([0, 1])$. The graphs show how many, out of approximately 150,000 randomly selected points from $[0, 1]$, occupy the subintervals of length 2^{-8} at each iterate.

We prove now a technical result which shows an element of regularity in ρ^2 that is missing in ρ.

LEMMA 2.5. *Let $u, v, w \in \{0, 1\}^*$ be such that*

$$\rho(u) = v \quad \textit{and} \quad \rho(v) = w,$$

with both u and v not containing 00 *as a subword and ending with* 1. *Then* $|u| = 2|w|$.

Proof. We proceed by induction on the length of w. Consider first the two words, 0 and 1, having length 1.

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

It is straightforward to check that 01 and 11 are the only two words such that:

(1) their ρ^2-images are respectively 0 and 1;

(2) both they and their ρ-images do not have 00 as a subword;

(3) they and their ρ-images end with 1.

This settles the case of words with length 1.

Suppose now to have established the result for every word of length k. Fix $w \in \{0, 1\}^k$ and let u and v have the properties indicated in the statement. As u has even length, we have

$$\rho(u01) = v1, \quad \rho(v1) = wx,$$

where x is 0 or 1 according to $|v|$ being even or odd, respectively. Moreover, we have

$$\rho(u11) = v01, \quad \rho(v01) = wx,$$

where x is 0 or 1 according to $|v|$ being odd or even, respectively. The words $u01$ and $u11$ are such that:

(1) their ρ^2-images are respectively $w0$ and $w1$ (if $|v|$ is even) or respectively $w1$ and $w0$ (if $|v|$ is odd);

(2) both they and their ρ-images do not have 00 as a subword;

(3) they and their ρ-images end with 1.

It is easy to check that they are the unique words with these properties. As $|u| = 2|w|$ by the inductive assumption, we have $|u01| = |u11| = 2|wx|$, so we have established the result for the two words $w0$ and $w1$ whatever the parity of v is, which completes the induction procedure. □

Under the assumptions of Lemma 2.5, we call v and u the *minimal preimages* of order 1 and 2 of w under ρ, respectively.

Before giving the proof of Theorem 2.4, we need another result.

LEMMA 2.6. *For every $w \in \{0, 1\}^*$, let $W \subset \{0, 1\}^*$ be defined as follows:*

$$W = \{q \in \{0, 1\}^* : \rho^2(q) = w \text{ and both } q \text{ and } \rho(q) \text{ do not end in } 0\}. \tag{2.12}$$

Then, $|q|$ is even for every $q \in W$. Moreover, for every $q, q' \in W$ such that $q \neq q'$, $[q] \cap [q']$ contains at most one point.

Proof. As a first step, let us show that $|q|$ is even for every $q \in W$ by a contradiction argument. If $|q|$ is odd, then q can be written as $q = v1$, with $|v|$ even. As a consequence, $\rho(q) = \rho(v)0$, so it ends in 0 and $q \notin W$ by definition.

Now, let us consider $q, q' \in W$ such that $q \neq q'$. Looking again for a contradiction, let us assume that $[q] \cap [q']$ contains more than one point. This implies that one of them is strictly contained in the other. Without loss of generality, let us assume that $[q'] \subsetneq [q]$, so q is a prefix of q', namely there exists $v \in \{0, 1\}^+$ (hence $v \neq \epsilon$) such that

$$q' = qv.$$

As both $\rho^2(q)$ and $\rho^2(q') = w$, we have that $\rho^2(v) = \epsilon$. Moreover, because both $|q|$ and $|q'|$ are even, $|v|$ is even too. Hence, there exists $v' \in \{0, 1\}^+$ such that $|v'|$ is odd and

$$v = v'1.$$

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

As a consequence, $\rho(v) = \rho(v')1$ and

$$\rho^2(v) = \rho^2(v')a,$$

where a is 0 or 1 depending on the parity of $|\rho(v')|$. Therefore, $|\rho^2(v)| \geq |a| = 1 > 0 = |\epsilon|$, so $\rho^2(v) \neq \epsilon$, that is a contradiction. □

Now, we are ready to prove Theorem 2.4.

Proof of Theorem 2.4. In the following computations, it is useful to recall that the number of ways in which one can arrange m indistinguishable objects in n distinguishable sites is given by

$$\binom{m+n-1}{m}.$$

In our case, the objects will be subwords of type 00 and the sites can be thought of as the places 'at the left of the digits 1' in a given word. Moreover, we are going to use the following equality:

$$\sum_{m=0}^{\infty} \binom{m+n-1}{m} 4^{-m} = \left(\frac{4}{3}\right)^n. \tag{2.13}$$

We proceed by induction.

(1) First, we check the result in the case of the real cylinder set $[0] = [0, 1/2]$. Its f_ρ-preimage is the set $[[1]] = \bigcup_{k=0}^{\infty}[0^{2k}1]$. Therefore, the f_ρ^2-preimage of $[0]$ is the set

$$[[01]] \cup \left(\bigcup_{k=0}^{\infty} [[(10)^{2k+1}11]] \right).$$

Let us calculate its Lebesgue measure. We have

$$m([[01]]) = 4^{-1} \sum_{j=0}^{\infty} 4^{-j} = \frac{1}{3}, \tag{2.14}$$

and, using also (2.13), we have also

$$m([[(10)^{2k+1}11]]) = 4^{-2k-2} \sum_{j=0}^{\infty} \left(\binom{j+2k+2}{j} 4^{-j} \right) = \frac{4}{3^{2k+3}}.$$

Thus, we obtain

$$m\left(\bigcup_{k=0}^{\infty} [[(10)^{2k+1}11]] \right) = \sum_{k=0}^{\infty} \frac{4}{3^{2k+3}} = \frac{1}{6}, \tag{2.15}$$

which leads to $m(f_\rho^{-2}([0])) = m([0]) = 1/2$, as claimed.

(2) From the previous step, it immediately follows that $m(f_\rho^{-2}([1])) = m([0, 1] \setminus f_\rho^{-2}([0])) = 1/2$.

(3) Having established the claim for real cylinder sets associated to words of length $N = 1$, let us fix $N \geq 1$ and suppose that for every $w \in \{0, 1\}^N$ we have $m(f_\rho^{-2}([w])) = m([w])$. We are going to prove that both $[w0]$ and $[w1]$ have f_ρ^2-preimages having Lebesgue measure $2^{-|w|-1}$, which suffices to prove the theorem.

Let $P = \{p \in \{0, 1\}^* : \rho^2(p) = w$ and both p and $\rho(p)$ do not end in $0\}$. By Lemma 2.6 the length of every $p \in P$ is even and $[p] \cap [p']$ consists of at most one point for every $p, p' \in P$ such that $p \neq p'$. Using the induction hypothesis we obtain

$$m\bigg(\bigcup_{p\in P}[p]\bigg) = \sum_{p\in P} 2^{-|p|} = 2^{-|w|}.$$

Let v and u be the minimal preimages of w under ρ of order 1 and 2, respectively, so that $\rho(u) = v$, $\rho(v) = w$ and neither u or v have 00 as a subword. Let $a \in \{0, 1\}$. The minimal ρ-preimage of wa is of type $v1$ or $v01$ depending on a and on the parity of the length of v. Let us analyze the two cases separately.

(i) The minimal ρ-preimage of wa is of type $v01$. The minimal ρ-preimage of $v01$ is $u11$, because by Lemma 2.5 u has even length. Set

$$Q = \{q \in \{0, 1\}^* : \rho^2(q) = wa \text{ and both } q \text{ and } \rho(q) \text{ do not end in } 0\}.$$

Every word in Q is of type pr where $r \in \langle\langle 1(01)^{2k}1\rangle\rangle$ for some integer $k \geq 0$, because $|p|$ is even and this implies that the subword $(01)^{2k}$ in r goes to the empty word in two iterations of ρ. Therefore, recalling (2.13), $m(f_\rho^{-2}([wa]))$ can be written as follows:

$$\begin{aligned}
m(f_\rho^{-2}([wa])) &= \sum_{p\in P}\bigg(m([p]) \sum_{j,k=0}^{\infty}\binom{j+2k+1}{j}2^{-2-2j-4k}\bigg)\\
&= 2^{-|w|}\sum_{j,k=0}^{\infty}\binom{j+2k+1}{j}2^{-2-2j-4k}\\
&= 2^{-|w|}\sum_{k=0}^{\infty}\frac{4}{9^{k+1}}\\
&= 2^{-|w|-1}.
\end{aligned}$$

(ii) The minimal ρ-preimage of wa is of type $v1$. The minimal ρ-preimage of $v1$ is $u01$, because again by Lemma 2.5 u has even length. Let P and Q be defined as previously. Arguing as before (and recalling that $|p|$ is even), every word in Q is of type pr' where $r' \in [[01]]$ or $r' \in [[1(01)^{2k-1}1]]$ for some integer $k \geq 1$. Therefore, $m(f_\rho^{-2}([wa]))$ can be computed as follows:

$$\begin{aligned}
m(f_\rho^{-2}([wa])) &= \sum_{p\in P} m([p])\bigg(\sum_{j=0}^{\infty}2^{-2-2j} + \sum_{j=0}^{\infty}\sum_{k=1}^{\infty}\binom{j+2k}{j}2^{-2j-4k}\bigg)\\
&= 2^{-|w|}\bigg(\frac{1}{3}+\frac{4}{3}\sum_{k=1}^{\infty}\frac{1}{9^k}\bigg)\\
&= 2^{-|w|-1}.
\end{aligned}$$

As a consequence, the inductive step is proven and this concludes the proof. □

3. *Dynamical properties of f_ρ^2*

In defining the map f_ρ, we exploited the fact that, using (2.4), the erasing block substitution ρ can be extended to a morphism on $\{0,1\}^\infty$. However, we remark that ρ does not commute with the operation of concatenation of words. Indeed, for any $v \in \{0,1\}^*$ and $w \in \{0,1\}^\infty$ with w containing at least one digit 1, the word equality

$$\rho(vw) = \rho(v)\rho(w)$$

holds if and only if $|v|$ is even, whereas for $|v|$ odd we have

$$\rho(vw) = \rho(v)\widetilde{\rho}(w),$$

where $\widetilde{\rho}$ is defined as follows:

$$\widetilde{\rho} : \begin{cases} 0 \mapsto \epsilon, & \\ 1 \mapsto 1 & \text{in odd positions,} \\ 1 \mapsto 0 & \text{in even positions.} \end{cases} \tag{3.1}$$

By Lemma 2.5 for every finite binary word w, there is $k \in \mathbb{N}$ such that $\rho^k(w) = \epsilon$, because the length of subsequent iterates of ρ applied to w has to strictly decrease every two steps of iteration. This means that ρ is completely erasing in the sense of [5] (a formal definition of this concept will be given in §4). The following result is then an immediate consequence of [5, Theorem 4], the key point here being that both ρ and $\widetilde{\rho}$ are surjective as maps on $\{0,1\}^\omega$.

PROPOSITION 3.1. *The map f_ρ is topologically exact (otherwise said locally eventually onto), that is, for every non-empty open set $A \subset [0,1]$, there exists a positive integer n such that $f_\rho^n(A) = [0,1]$. Consequently, f_ρ^2 is topologically exact too.*

In the definition given in [16], a function $f : [0,1] \to [0,1]$ is called *turbulent* if there exist two compact subintervals $I_1, I_2 \subset [0,1]$ with at most one point in common such that

$$I_1 \cup I_2 \subset f(I_1) \cap f(I_2).$$

PROPOSITION 3.2. *The function f_ρ^2 is turbulent.*

Proof. It is enough to observe that $f_\rho^2([00]) = f_\rho^2([10]) = [0,1]$. □

In [4] it has been proven that every turbulent interval map whose graph is a connected G_δ set has positive topological entropy. This does not apply to f_ρ, because its graph is totally disconnected; however, it has been proven that its topological entropy is infinite (both results are proven in [6]).

A map f preserving the measure μ is said to be *exact* with respect to μ if $\lim_{n\to\infty} \mu(f^n(A)) = 1$ for every set of positive measure A. For a measure-preserving transformation, topological exactness and exactness are scarcely related. For instance, in [2] it is proven that the typical continuous non-invertible Lebesgue-preserving interval map is topologically exact, whereas the set of strongly mixing maps is of first category.

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

In the following, we prove that $([0, 1], m, f_\rho^2)$ is a strong mixing, which is a weaker property than exactness with respect to the preserved measure. Let us recall the definition of strong mixing.

Definition 3.3. Let $(X, \mathcal{M}, \mu)$ be a measure space and $f : X \to X$ a measure-preserving map. Then f is said *strongly mixing* if for every $A, B \in \mathcal{M}$ the following holds:

$$\lim_{n\to\infty} \mu(A \cap f^{-n}(B)) = \mu(A)\mu(B). \tag{3.2}$$

Our main result is the following.

THEOREM 3.4. *The map f_ρ^2 is strongly mixing.*

The fact that f_ρ^2 is strongly mixing looks surprising, at first. Roughly speaking, if the Lebesgue measure is considered, the strongly mixing property requires that the preimages of increasing order of every measurable set tend to 'spread' uniformly on [0, 1]. Now, the ρ-preimage of a real cylinder set $[w]$ is obtained in two steps: finding the minimal preimage v of w and then generating the set $[[v]]$. The latter is obtained inserting pairs of consecutive 0s in v, an operation which seems to be non-uniform, favoring instead a global transfer of Lebesgue-mass towards the left of the interval. However, this rough intuition is wrong because the minimal preimage of 00 is 101 (inserted in such a way that the 1s are at odd-indexed places) and at the second iteration for the construction of the ρ^2-preimage the Lebesgue-mass can move again towards the right of the interval. Nevertheless, a perfect uniformity of finite-order ρ^2-preimages cannot be expected in general and it will be achieved only through a limit process (see the proof of Lemma 3.8 for more details).

To prove the strongly mixing property of f_ρ^2 we do not need to work with all the Lebesgue measurable subset of [0, 1]. Let us indicate by $\mathcal{I}_d$ the set of all subintervals of [0, 1] (open, closed, half-open) consisting of points x whose binary expansion begins with a given $w \in \{0, 1\}^*$, plus the empty set and the degenerate cases in which the interval becomes a single dyadic rational point. More precisely, every element of $\mathcal{I}_d$ has one of the following forms for some $w \in \{0, 1\}^*$:

$$\{x \in [0, 1] : (0.w0^\infty)_2 \le x \le (0.w1^\infty)_2\};$$

$$\{x \in [0, 1] : (0.w0^\infty)_2 < x \le (0.w1^\infty)_2\};$$

$$\{x \in [0, 1] : (0.w0^\infty)_2 \le x < (0.w1^\infty)_2\};$$

$$\{x \in [0, 1] : (0.w0^\infty)_2 < x < (0.w1^\infty)_2\};$$

$$\{(0.w0^\infty)_2\},\ \{1\},\ \emptyset.$$

We call the elements of $\mathcal{I}_d$ the *dyadic intervals*.

It is immediate to check that $\mathcal{I}_d$ is a semi-algebra (see [**14**]), that is:

- $\emptyset \in \mathcal{I}_d$;
- $J \in \mathcal{I}_d$ implies that $[0, 1] \setminus J$ is a finite union of pairwise disjoint members of $\mathcal{I}_d$;
- $\mathcal{I}_d$ is closed with respect to finite intersection.

Therefore (see, for instance, [12, p. 52]), to establish that f_ρ^2 is strongly mixing it is enough to prove that

$$\lim_{n\to\infty} m(I \cap f_\rho^{-2n}(J)) = m(I)m(J) \tag{3.3}$$

holds for every $I, J \in \mathcal{I}_d$. As Theorem 2.4 ensures that

$$m(f_\rho^{-2}(x)) = 0 \quad \text{for every } x \in [0, 1], \tag{3.4}$$

it is enough to check that (3.3) holds for non-degenerate closed intervals, which we can identify with real cylinder sets $[w]$ such that w is a finite, possibly empty binary word. This will be done in the following. However, before addressing the proof of Theorem 3.4, it is convenient to prove some lemmas.

In the following, we usually construct the preimages of a dyadic interval $[w]$ under the action of $(f_\rho)^2$ by finding the words v such that $\rho^2(v) = w$. However, it is not always true that f_{ρ^2} coincides with $(f_\rho)^2$. For instance, let

$$x = 0.1111(10)^\infty,$$

so that

$$f_\rho(x) = 0.\rho(w_x) = 0.0101(0)^\infty = 0.0101.$$

As a consequence, $w_{f_\rho(x)} = 0100(1)^\infty$, so we have

$$f_\rho^2(x) = f_\rho(0.w_{f_\rho(x)}) = 0.(10)^\infty \neq 0.11 = 0.\rho^2(x).$$

The next lemma ensures, however, that f_{ρ^2} coincides with $(f_\rho)^2$ Lebesgue-almost everywhere on [0, 1], as $f_\rho^2(x) \neq f_{\rho^2}(x) = 0.\rho^2(w_x)$ only if $\rho(w_x)$ ends with 0^∞, so that $w_{f_\rho(x)} \neq \rho(w_x)$.

LEMMA 3.5. *The set*

$$S = \{x \in [0, 1] : f_\rho^2(x) \neq (0.\rho^2(w_x))_2\} \tag{3.5}$$

has Lebesgue measure 0.

Proof. If $f_\rho(x) \notin \mathbb{Q}_2$ and $x \neq 0$, then, by (2.2),

$$f_\rho^2(x) = (0.\rho(w_{f_\rho(x)}))_2 = (0.\rho^2(w_x))_2.$$

It is then sufficient to prove that $m(f_\rho^{-1}(\mathbb{Q}_2)) = 0$. As $|\mathbb{Q}_2| = \aleph_0$, this follows from the fact that the f_ρ-preimages of single points has Lebesgue measure zero (see [6, Proposition 5.3]). □

LEMMA 3.6. *For every finite word $w \in \{0, 1\}^*$ such that $|w|$ is even and every dyadic interval J, we have*

$$m([w] \cap f_\rho^{-2}(J)) = 2^{|\rho^2(w)|-|w|} m([\rho^2(w)] \cap J). \tag{3.6}$$

Proof. We divide the proof in different cases. Indeed, because both $[\rho^2(w)]$ and J are dyadic intervals, their intersection could either contain at most one point, or be equal to one of the two sets.

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

(i) $[\rho^2(w)] \cap J$ is the empty set or it contains exactly one point. In this case, recalling Lemma 3.5, we have that $[w] \cap f_\rho^{-2}(J)$ is a Lebesgue-null set, hence (3.6) trivially holds as $0 = 0$. Indeed, a point x is in $[w] \cap f_\rho^{-2}(J)$ if and only if $w_x = wq$ for some $q \in \{0, 1\}^\infty$ and $f_\rho^2(x) = 0.\rho^2(wq) = 0.\rho^2(w)\widetilde{q} \in J$, (with $\widetilde{q}$ equal to either $\rho^2(q)$ or $\widetilde{\rho}(\rho(q))$), and this is not possible by assumption.

(ii) $[\rho^2(w)] \subset J$. In this case, $[\rho^2(w)] \cap J = [\rho^2(w)]$, thus the word $w_J \in \{0, 1\}^*$ that characterizes J is a prefix of $\rho^2(w)$, namely there exists $v \in \{0, 1\}^*$ (eventually $v = \epsilon$ if $[\rho^2(w)] = J$) such that

$$\rho^2(w) = w_J v.$$

As a consequence, by Lemma 3.5, $[w] \cap f_\rho^{-2}(J) = [w]$ except for a Lebesgue-null set. Indeed, for Lebesgue-almost every $x \in [w]$ we have that

$$f_\rho^2(x) = 0.\rho^2(w_x) = 0.\rho^2(wq) = 0.\rho^2(w)\widetilde{q} = 0.w_J v\widetilde{q} \in [w_J] = J,$$

hence, $x \in f_\rho^{-2}(J)$. Therefore, in this case (3.6) holds because $m([w]) = 2^{-|w|}$ and $m([\rho^2(w)]) = 2^{-|\rho^2(w)|}$.

(iii) $J \subsetneq [\rho^2(w)]$. In this case, $\rho^2(w)$ is a prefix of w_J and there exists $v \in \{0, 1\}^+$ (so v is different from ϵ) such that

$$w_J = \rho^2(w)v.$$

Therefore, we have

$$m([\rho^2(w)] \cap J) = m([\rho^2(w)v]) = 2^{-|\rho^2(w)|}m(v). \tag{3.7}$$

Let us define the set $V \subset \{0, 1\}^+$ as

$$V = \{q \in \{0, 1\}^+ : \rho^2(wq) = \rho^2(w)v\},$$

so that, again neglecting a Lebesgue-null set by Lemma 3.5,

$$[w] \cap f_\rho^{-2}(J) = \bigcup_{q \in V} [wq].$$

As w is even by hypothesis, $\rho(wq) = \rho(w)\rho(q)$ for every $q \in V$. Now, if $\rho(w)$ is even we have $\rho^2(wq) = \rho^2(w)\rho^2(q)$. Otherwise, if $\rho(w)$ is odd we have $\rho^2(wq) = \rho^2(w)\widetilde{\rho}(\rho(q))$. Hence, we have either

$$V = \rho^{-2}(v) \quad \text{or} \quad V = \rho^{-1}(\widetilde{\rho}^{-1}(v)). \tag{3.8}$$

By Theorem 2.4, both f_ρ^2 and $f_{\widetilde{\rho}} \circ f_\rho = (1 - x) \circ f_\rho^2$ preserve the Lebesgue measure. Therefore, in both cases we obtain

$$m\bigg(\bigcup_{q \in V} [q]\bigg) = m([v]).$$

As for all $q_1, q_2 \in V$ with $q_1 \neq q_2$ the sets $[wq_1]$ and $[wq_2]$ are disjoint, Lemma 3.5 implies that we have

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

$$m([w] \cap f_\rho^{-2}(J)) = \sum_{q \in V} m([wq]) = 2^{-|w|} \sum_{q \in V} m(q) = 2^{-|w|} m([v]). \tag{3.9}$$

Combining (3.7) and (3.9) we have that (3.6) holds. □

COROLLARY 3.7. *For every dyadic interval J and every word w such that |w| is even, the following holds:*

$$m([w] \cap f_\rho^{-2n}(J)) = 2^{|\rho^2(w)|-|w|} m([\rho^2(w)] \cap f^{-2n+2}(J)) \quad \textit{for all } n \geq 1. \tag{3.10}$$

Proof. If $n = 1$, then (3.10) reduces to (3.6). If $n > 1$, then, neglecting the intersection with the Lebesgue-null set S defined in (3.5), $f_\rho^{-2n+2}(J)$ can be written as a countable union of disjoint dyadic intervals $(W_k)_{k \in \mathbb{N}}$, namely

$$f_\rho^{-2n+2}(J) = \bigcup_{k=1}^{\infty} W_k.$$

Hence, applying (3.6) to every W_k, we obtain

$$\begin{aligned} m([w] \cap f_\rho^{-2n}(J)) &= \sum_{k=1}^{\infty} m([w] \cap f_\rho^{-2}(W_k)) \\ &= 2^{|\rho^2(w)|-|w|} \sum_{k=1}^{\infty} m([\rho^2(w)] \cap W_k) \\ &= 2^{|\rho^2(w)|-|w|} m([\rho^2(w)] \cap f^{-2n+2}(J)). \end{aligned}$$

□

LEMMA 3.8. *For every dyadic interval $J \subset [0, 1]$ we have*

$$\lim_{n \to \infty} m([0] \cap f_\rho^{-2n}(J)) = \tfrac{1}{2} m(J) = m([0]) m(J) \tag{3.11}$$

and, consequently,

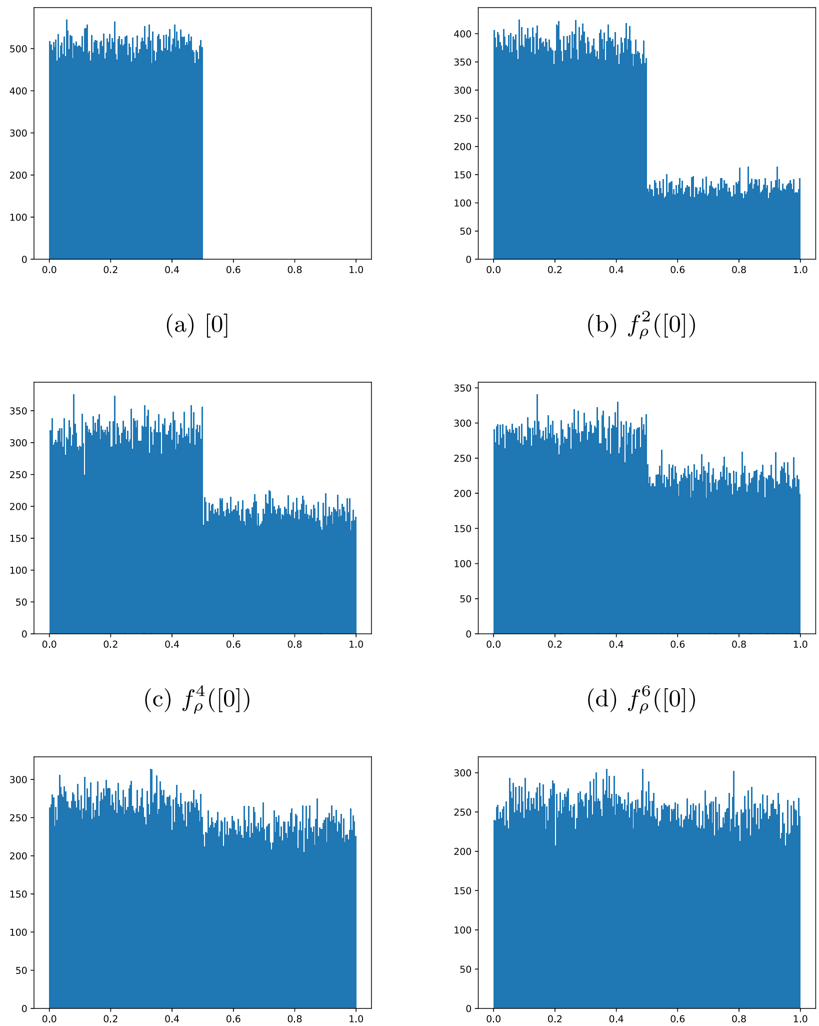
$$\lim_{n \to \infty} m([1] \cap f_\rho^{-2n}(J)) = \tfrac{1}{2} m(J) = m([1]) m(J). \tag{3.12}$$

The numerical results presented in Figure 2 show the mixing property ensured by Lemma 3.8. Indeed, it can be seen how a randomly selected set of points in [0] spreads uniformly to the whole interval [0, 1] under the action of subsequent iterations of f_ρ^2.

Proof of Lemma 3.8. If $J = [0, 1]$ or $J = \emptyset$, then the thesis immediately follows from $f_\rho^{-2}(J) = J$.

Let us consider a dyadic interval W different from [0, 1] and $\emptyset$. Hence, neglecting degenerate cases of Lebesgue measure zero, W is characterized by a finite word w of length greater or equal than 1, namely $W = [w]$ with $1 \leq |w| < \infty$. Hence, either $w = 0v$ or $w = 1v$, for some $v \in \{0, 1\}^*$. As a first step, let us consider $w = 0v$. To obtain $f_\rho^{-2}(W)$, we can follow the same construction given in the first step of the proof of Theorem 2.4.

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press



(a) $[0]$ (b) $f_\rho^2([0])$

(c) $f_\rho^4([0])$ (d) $f_\rho^6([0])$

(e) $f_\rho^8([0])$ (f) $f_\rho^{10}([0])$

FIGURE 2. Numerical results on the first five iterates of $f_\rho^2([0])$. The graphs show how many, out of approximately 150,000 randomly selected points from $[0, 1/2]$, occupy the subintervals of length 2^{-8} at each iterate.

Hence, recalling Lemma 3.5, for Lebesgue almost every point x in $f_\rho^{-2}(W)$, we have that $w_x = aqr$, where

$$a \in \left(\langle\langle 01 \rangle\rangle \bigcup \left(\bigcup_{k=1}^{\infty} \langle\langle (10)^{2k+1} 11 \rangle\rangle \right) \right) = \rho^{-2}(0),$$

$q \in \{0, 1\}^+$ is such that $\rho^2(aq) = 0v$ and r is a general infinite binary word. As $|a|$ is even for every $a \in \rho^{-2}(0)$, by Lemma 3.6 we obtain

$$m([a] \cap f_\rho^{-2}(W)) = \frac{2m(W)}{2^{|a|}} \quad \text{for all } a \in \rho^{-2}(0). \tag{3.13}$$

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

To compute $m([0] \cap f_\rho^{-2}(W))$, it suffices to apply the previous equation for every $a \in \rho^{-2}(0)$ that has 0 as first letter. Let A_1 be the set of words in $\langle\langle 01 \rangle\rangle$, A_2 the set of words in $\langle\langle (10)^{2k+1} 11 \rangle\rangle$ and A_2^i the subset of A_2 of words that start with $i = 0, 1$. Using both (2.14) and (3.13), and recalling Lemma 3.5, we obtain

$$\sum_{a \in A_1} m([a] \cap f_\rho^{-2}(W)) = 2m(W) \sum_{a \in A_1} \frac{1}{2^{|a|}} = \frac{2}{3} m(W). \tag{3.14}$$

Using (2.15), we have that

$$\sum_{a \in A_2} \frac{1}{2^{|a|}} = \frac{1}{6}.$$

As

$$\sum_{a \in A_2} \frac{1}{2^{|a|}} = \frac{4}{3} \sum_{a \in A_2^1} \frac{1}{2^{|a|}} \quad \text{and} \quad \sum_{a \in A_2^0} \frac{1}{2^{|a|}} = \frac{1}{3} \sum_{a \in A_2^1} \frac{1}{2^{|a|}},$$

we obtain

$$\sum_{a \in A_2^0} m([a] \cap f_\rho^{-2}(W)) = 2m(W) \sum_{a \in A_2^0} \frac{1}{2^{|a|}} = \frac{1}{12} m(W), \tag{3.15}$$

where we have neglected again a set of measure zero by Lemma 3.5.

Using (3.14) and (3.15), we have

$$\begin{aligned} m([0] \cap f_\rho^{-2}(W)) &= \sum_{a \in A_1} m([a] \cap f_\rho^{-2}(W)) + \sum_{a \in A_2^0} m([a] \cap f_\rho^{-2}(W)) \\ &= \frac{2}{3} m(W) + \frac{1}{12} m(W) = \frac{3}{4} m(W). \end{aligned}$$

By the generality of $W = [0v]$, we conclude that, for every $v \in \{0, 1\}^*$,

$$m([0] \cap f_\rho^{-2}(W)) = \tfrac{3}{4} m(W) \quad \text{if } W = [0v]. \tag{3.16}$$

Using a similar procedure, for every $v \in \{0, 1\}^*$ we obtain that

$$m([0] \cap f_\rho^{-2}(W)) = \tfrac{1}{4} m(W) \quad \text{if } W = [1v]. \tag{3.17}$$

Now we are ready to prove (3.11) for every dyadic interval J different from $[0, 1]$ and $\emptyset$. In the following, we neglect systematically the intersection of the subsets of $[0, 1]$ under consideration with the Lebesgue-null set S defined in (3.5). The idea is to exploit (3.16) and (3.17) to see that $m([0] \cap f_\rho^{-2n}(J))$ converges to $m(J)/2$ as $n \to \infty$. Let us define the sequence $(y_n)_{n \in \mathbb{N}} \subset [0, 1]$ as follows:

$$y_n = \frac{m([0] \cap f_\rho^{-2n}(J))}{m(J)}.$$

As J is a dyadic interval different from $[0, 1]$ and $\emptyset$, we have that y_0 is either 0 or 1. For every $n \geq 1$, $f_\rho^{-2n}(J)$ is a countable union of dyadic intervals. More precisely, there are

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

two infinite countable sets of disjoint dyadic intervals, say $(W_i^0)_{i\in\mathbb{N}}$ and $(W_i^1)_{i\in\mathbb{N}}$, such that

$$[0]\cap f_\rho^{-2n}(J)=\bigcup_{i=0}^{\infty} W_i^0 \quad \text{and} \quad [1]\cap f_\rho^{-2n}(J)=\bigcup_{i=0}^{\infty} W_i^1.$$

As a consequence, we can write

$$f_\rho^{-2(n+1)}(J)=\bigcup_{i=0}^{\infty} f_\rho^{-2}(W_i^0)\ \cup\ \bigcup_{i=0}^{\infty} f_\rho^{-2}(W_i^1),$$

and we have also

$$\sum_{i=0}^{\infty} m(W_i^0)=y_n m(J) \quad \text{and} \quad \sum_{i=0}^{\infty} m(W_i^1)=(1-y_n)m(J).$$

For every W_i^0 we can apply (3.16), so we obtain

$$\sum_{i=0}^{\infty} m([0]\cap f_\rho^{-2}(W_i^0))=\frac{3}{4}\sum_{i=0}^{\infty} m(W_i^0)=\frac{3}{4}y_n m(J). \tag{3.18}$$

Similarly, for every W_i^1 we can apply (3.17), so we have

$$\sum_{i=0}^{\infty} m([0]\cap f_\rho^{-2}(W_i^1))=\frac{1}{4}\sum_{i=0}^{\infty} m(W_i^1)=\frac{1}{4}(1-y_n)m(J). \tag{3.19}$$

As a consequence, summing (3.18) and (3.19), we obtain

$$\begin{aligned} y_{n+1}m(J) &= m([0]\cap f_\rho^{-2(n+1)}(J)) \\ &= \sum_{i=0}^{\infty} m([0]\cap f_\rho^{-2}(W_i^0))+\sum_{i=0}^{\infty} m([0]\cap f_\rho^{-2}(W_i^1)) \\ &= \frac{3}{4}y_n m(J)+\frac{1}{4}(1-y_n)m(J)=\left(\frac{1}{4}+\frac{1}{2}y_n\right)m(J). \end{aligned} \tag{3.20}$$

Therefore, $(y_n)_{n\in\mathbb{N}}$ can be computed using the following discrete dynamical system

$$y_{n+1}=\frac{1}{4}+\frac{1}{2}y_n,$$

thus, we obtain

$$\lim_{n\to\infty} y_n=\frac{1}{2},$$

from which (3.11) follows. □

Now, we are ready to give a proof of the strongly mixing property of f_ρ^2.

Proof of Theorem 3.4. As discussed previously, it is enough to show that (3.3) holds for every pair of non-degenerate dyadic intervals $I, J\subseteq[0,1]$. The proof is given by an induction argument on the length of the word w that characterizes the interval I (namely $I=[w]$).

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

By Theorem 2.4, we have that

$$m([\epsilon]\cap f_\rho^{-2n}(J)) = m(J) = m([\epsilon])m(J)$$

for every $n \in \mathbb{N}$, hence the equality holds for $n \to \infty$. Using also Lemma 3.8, we have that (3.3) holds for every I that is characterized by a word w such that $|w| \le 1 = 2^0$.

Let us show that if (3.3) holds for every I that is characterized by a word w such that $|w| \le 2^n$, then it holds also for every $I = [w]$ with $|w| \le 2^{n+1}$.

As a first step, let w be a word of length 2^{n+1}, so $I = [w]$ and $m(I) = 2^{-|w|}$. It follows immediately from Lemma 2.5 that $|\rho^2(w)| \le |w|/2 = 2^n$. By Corollary 3.7, in particular by applying (3.10), and using the induction hypothesis we obtain

$$\begin{aligned}\lim_{n\to\infty} m([w]\cap f_\rho^{-2n}(J)) &= 2^{|\rho^2(w)|-|w|}\lim_{n\to\infty} m([\rho^2(w)]\cap f_\rho^{-2n+2}(J))\\ &= 2^{|\rho^2(w)|-|w|}\frac{m(J)}{2^{|\rho^2(w)|}} = \frac{m(J)}{2^{|w|}} = m(I)m(J),\end{aligned}$$

so (3.3) holds.

As a second step, let I be a dyadic interval that is characterized by a word w of length between 2^n and 2^{n+1}. As I can be obtained as a finite union of intervals of the type $I_j = [w_j]$ with $|w_j| = 2^{n+1}$ for every j, then (3.3) holds. This concludes the inductive step, so we are done. □

In the proof of Lemma 3.6 we had to consider the composition $f_{\widetilde{\rho}} \circ f_\rho$ (see (3.8)), and we exploited the fact that it preserves the Lebesgue measure. A graphic illustration of this fact is shown in Figure 3.

In the proof of Lemma 3.6 we could neglect the interval maps generated by the other compositions of ρ and $\widetilde{\rho}$, that is, $f_\rho \circ f_{\widetilde{\rho}}$ and $f_{\widetilde{\rho}}^2$. However, in order to investigate further dynamical properties of the model-case erasing interval map, such as exactness, rate of mixing and metric entropy, it seems relevant to study also these compositions. The Lebesgue measure is not preserved by them. Instead, the numerical results in Figure 4 and in Figure 5 suggest respectively that:

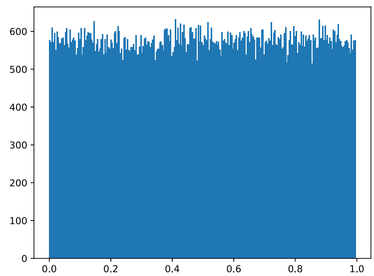
- $f_{\widetilde{\rho}}^2$ preserves the probability measure supported and uniformly distributed on the set [1];
- $f_\rho \circ f_{\widetilde{\rho}}$ preserves a probability measure μ uniformly distributed on both [0] and [1] and such that $\mu([0]) > \mu([1])$; most likely, $\mu([0]) = 2/3$ and $\mu([1]) = 1/3$.

We conclude this section with some thoughts on the entropy of the map f_ρ^2. As we already said, the topological entropy of f_ρ (thus, of f_ρ^2) is infinite. The link between topological entropy and metric entropy, in a compact metric space, is expressed by the well-known *variational principle*:

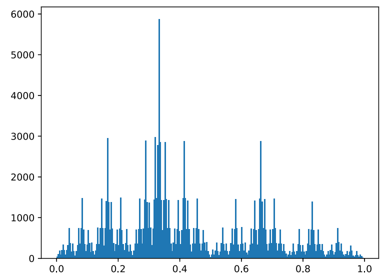
$$h(f) = \sup\{h_\mu(f) : \mu \text{ is an } f\text{-invariant Borel measure}\}, \tag{3.21}$$

where $h(f)$ is the topological entropy and $h_\mu(f)$ is the metric entropy with respect to μ. Classically, the variational principle is stated assuming that the map f is continuous (the elegant proof by M. Misiurewicz is given, for instance, in [17]).

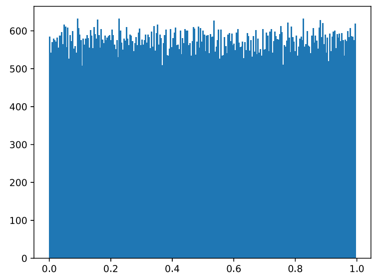
https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press



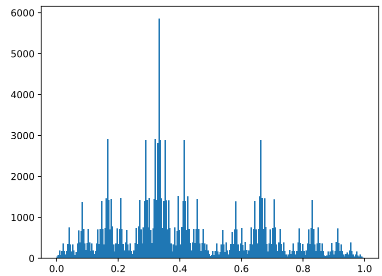
(a) $[0, 1]$



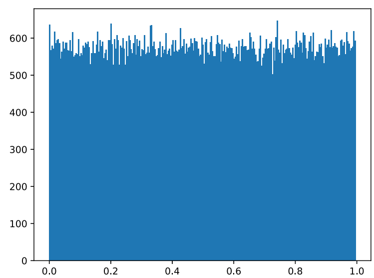
(b) $f_\rho([0, 1])$



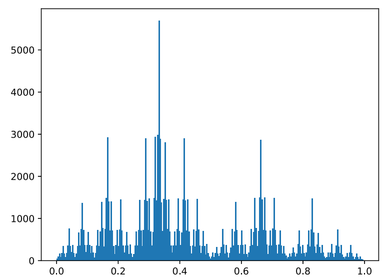
(c) $f_{\widetilde{\rho}}(f_\rho([0, 1]))$



(d) $f_\rho(f_{\widetilde{\rho}}(f_\rho([0, 1])))$



(e) $(f_{\widetilde{\rho}} \circ f_\rho)^2([0, 1])$



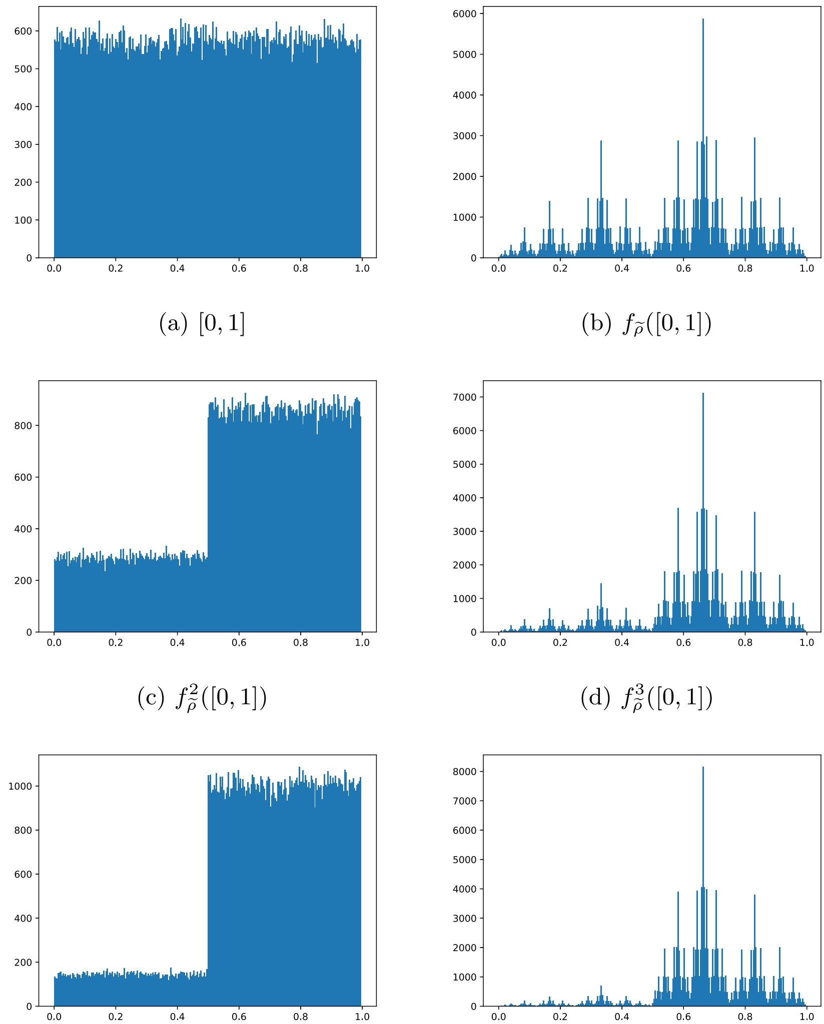
(f) $f_\rho\big((f_{\widetilde{\rho}} \circ f_\rho)^2([0, 1])\big)$

FIGURE 3. Numerical results on the first iterates of $f_{\widetilde{\rho}} \circ f_\rho$. The graphs show how many, out of approximately 150,000 randomly selected points from [0, 1], occupy the subintervals of length 2^{-8} at each iterate.

A remarkable result of the already cited (and, in our opinion, somewhat underestimated) work [4] is that the variational principle holds in far greater generality. In fact, if we use the metric definition of topological entropy given by Bowen and Dinaburg [3, 7], Theorem 3.18 in [4] shows that the variational principle holds for completely general self-maps of a compact metric space.

Another relatively recent development, whose application to our model-case erasing map looks interesting, is the theory of entropy structure due to Downarowicz (see [8]), which studies how the entropy emerges when improving the topological resolution power,

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press



(a) $[0,1]$ (b) $f_{\widetilde{\rho}}([0,1])$

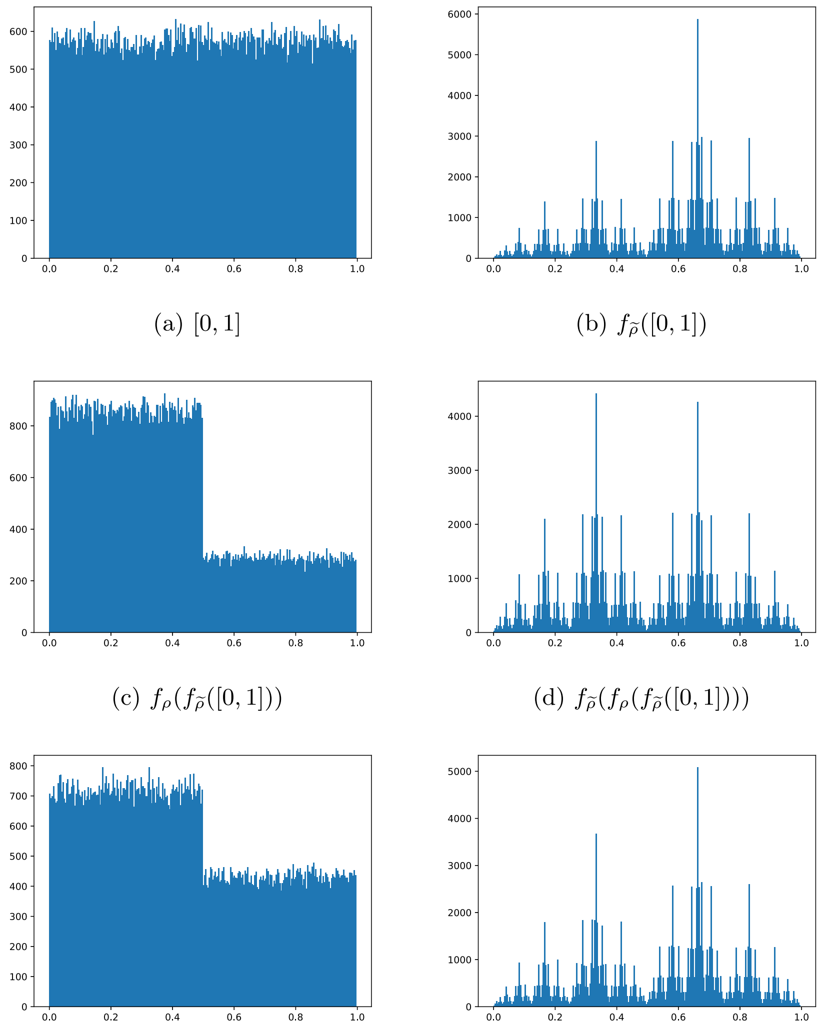
(c) $f^2_{\widetilde{\rho}}([0,1])$ (d) $f^3_{\widetilde{\rho}}([0,1])$

(e) $f^4_{\widetilde{\rho}}([0,1])$ (f) $f^5_{\widetilde{\rho}}([0,1])$

FIGURE 4. Numerical results on the first iterates of $f_{\widetilde{\rho}}$. The graphs show how many, out of approximately 150,000 randomly selected points from [0, 1], occupy the subintervals of length 2^{-8} at each iterate.

and allows to define some powerful dynamical invariants. The theory can be applied to discontinuous maps through the passage to an extension of the system (preserving the entropy of invariant measures) in which the map is continuous. This is a viable approach if the set of discontinuities has measure zero for all invariant measures of positive metric entropy (see [9], especially Chs. 6–9), which seems like a reasonable hope in the case of f^2_ρ, whereas things are more difficult for f_ρ.

The investigation of invariant measures for f^2_ρ (which, we recall, has a dense set of discontinuities), other than the Lebesgue measure, seems particularly interesting in view of the aforementioned results.



(a) $[0, 1]$ (b) $f_{\widetilde{\rho}}([0, 1])$

(c) $f_\rho(f_{\widetilde{\rho}}([0, 1]))$ (d) $f_{\widetilde{\rho}}(f_\rho(f_{\widetilde{\rho}}([0, 1])))$

(e) $(f_\rho \circ f_{\widetilde{\rho}})^2([0, 1])$ (f) $f_{\widetilde{\rho}}\big((f_\rho \circ f_{\widetilde{\rho}})^2([0, 1])\big)$

FIGURE 5. Numerical results on the first iterates of $f_\rho \circ f_{\widetilde{\rho}}$. The graphs show how many, out of approximately 150,000 randomly selected points from [0, 1], occupy the subintervals of length 2^{-8} at each iterate.

4. *More general erasing maps*

In this section we address the problem of generalizing the results proven up to now to a larger class of interval maps generated by erasing substitutions.

We start by defining more formally a concept already introduced at the beginning of §3.

Definition 4.1. We say that the k-block substitution σ is *alternating* if there exist k simple substitutions $\sigma_1, \sigma_2, \ldots, \sigma_k$ such that the map σ_e defined below is an extension of σ on $\{0, 1\}^\infty$:

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

$$\sigma_e(u) = \prod_{j=0}^{n-1} \left(\prod_{i=1}^{k} \sigma_i(u_{i+jk}) \right) \prod_{i=1}^{m} \sigma_i(u_{i+nk}), \tag{4.1}$$

where $|u| = nk + m$ ($n, m \in \mathbb{N}_0$, $m < k$) and the first (last) product has to be taken as empty if $n = 0$ ($m = 0$).

The substitution ρ is indeed alternating, as shown by (2.4) and subsequent comments.

In [5], erasing substitutions have been classified in a hierarchy according to the 'strength' of their erasing character. In particular, a k-block substitution σ is defined as follows.

(1) *Completely erasing* if it is alternating (in the sense of Definition 4.1) and, for every $w \in \{0, 1\}^*$, there is $n \in \mathbb{N}$ such that

$$\sigma^n(w) = \epsilon. \tag{4.2}$$

The smallest positive integer n verifying (4.2), indicated by $\epsilon(w)$, is called *vanishing order* of w. The topological entropy of f_σ depends on the asymptotic behavior of the vanishing order of words of increasing length.

(2) *Boundedly erasing* if σ is completely erasing and $\epsilon(\cdot)$ is bounded over $\{0, 1\}^*$.

In [5] the concept of *strongly erasing* substitution was also introduced, an intermediate property between being erasing and completely erasing, but it is not relevant herein.

Boundedly erasing substitutions can be interpreted as an extreme, uninteresting case, because the dynamics of boundedly erasing maps is almost trivial (see [5, Lemma 4.1]). Completely erasing substitutions, instead, show the richest topological dynamical properties. Thus, it is natural to try to establish the strong mixing property for the class of interval maps generated by completely erasing substitutions. However, things do not look easy.

First, understanding whether f_σ is always Borel singular if σ is completely erasing seems not trivial. In [6], the singularity of the map f_ρ was established through a quite *ad hoc* argument connected to normal real numbers in base 2. The argument has a combinatorial nature and it boils down to showing that the f_ρ-image of a 2-normal real number cannot be 2-normal by upper/lower bounding the asymptotic frequencies of blocks of length 8 (see [6, Proposition 5.2]). None of that can be straightforwardly extended to a generic completely erasing substitution.

The existence of an absolutely continuous measure which is preserved by f_σ^2 under the sole hypothesis that σ is completely erasing looks also far from trivial to establish. The reasoning in Theorem 2.4 is based on the possibility of counting explicitly the number of *distinguishable* insertions of the pair 00 in a given binary word which does not have any 00 subword. To achieve this, it was enough to consider the insertions made at the left of every digit 1 in the original word.

In the case of general completely erasing substitutions, indicating as usual by w_ϵ the word mapped to the empty word, this amounts to counting the distinguishable insertions of suitable circular permutations of w_ϵ in the original word w. The difficulty, here, lies in the identification of indistinguishable cases. For instance, let us consider the completely erasing k-block substitution σ and the case in which we have just one insertion of (a suitable circular permutation of) the word mapped to ϵ in a given word $u \in \{0, 1\}^n$.

https://doi.org/10.1017/etds.2023.16 Published online by Cambridge University Press

In this case, we should check how many of the following word (non-)equalities hold for $0 \leq h \leq k$, $0 \leq j \leq k$ and m, m' such that $mk + h + 1 \leq n$ and $m'k + j + 1 \leq n$:

$$u_1 \ldots u_{mk+h}(w_\epsilon)_{h+1} \ldots (w_\epsilon)_k (w_\epsilon)_1 \ldots (w_\epsilon)_h u_{mk+h+1} \ldots u_n$$
$$\neq u_1 \ldots u_{m'k+j}(w_\epsilon)_{j+1} \ldots (w_\epsilon)_k (w_\epsilon)_1 \ldots (w_\epsilon)_j u_{m'k+j+1} \ldots u_n.$$

where by $(w_\epsilon)_i$ we mean the ith digit of the word w_ϵ.

The problem becomes much more intricate when we consider multiple insertions of the circular permutations of w_ϵ. This would ultimately lead to a countable family of word equalities, and it seems unlikely that this approach will prove suitable to achieve the result. Probably, in order to study the existence of absolutely continuous invariant measures for (the square of) general erasing maps, new ideas have to be used.

Acknowledgements. The authors thank Tomasz Downarowicz, Marco Lenci and Riccardo Piergallini for some useful exchanges. D. Corona thanks the GNAMPA INdAM—Italian National Institute of High Mathematics for partial support.

A. *Appendix. Notation*

$\mathbb{N}$	the set of positive integers
$\mathbb{N}_0$	the set of non-negative integers
$\mathbb{Q}$	the set of rational numbers
$\mathbb{Q}_2$	the dyadic rationals in $[0, 1]$, that is the set $[0, 1] \cap \{n/2^k : n, k \in \mathbb{N}_0\}$
$\mathbb{R}$	the set of real numbers
ϵ	the empty word
$\{0, 1\}^*$	the set of all finite words on the alphabet $\{0, 1\}$
$\{0, 1\}^+$	the set of all finite non-empty words on the alphabet $\{0, 1\}$
$\{0, 1\}^\omega$	the set of all infinite words on the alphabet $\{0, 1\}$
$\{0, 1\}^\infty$	the set $\{0, 1\}^* \cup \{0, 1\}^\omega$
$\{0, 1\}^{\infty+}$	the set $\{0, 1\}^\infty \setminus \epsilon$
w_k	the kth digit of the non-empty word w
$\lvert w\rvert$	the length of the finite word w, that is, the non-negative integer n if $w = w_1 \ldots w_n$ and 0 if $w = \epsilon$
$\widetilde{w}$	the bitwise negation of w, namely $\widetilde{w}_k = 1 - w_k$ for every $k = 1, \ldots, \lvert w\rvert$
$w_{(n)}, w^{(n)}$	a word indexed by the non-negative integer n
$\prod_{i=1}^{n} w_{(i)}$	the concatenation $w_{(1)} \ldots w_{(n)}$
$\prod_{i=1}^{\infty} w_{(i)}$	the infinite concatenation $w_{(1)} w_{(2)} \ldots$
$\{0, 1\}^{\leq n}$	the set of all finite words on the alphabet $\{0, 1\}$ with length less than or equal to n
$\{0, 1\}^{\geq n}$	the set of all infinite words and of all finite words on the alphabet $\{0, 1\}$ with length greater than or equal to n
w^n	the concatenation of n copies of the finite word w
w^∞	the concatenation of infinitely many copies of the finite word w
$(0.w)_2$	for $w \in \{0, 1\}^{\infty+}$, the real number $\sum_{i=1}^{\lvert w\rvert} 2^{-i} w_i \in [0, 1]$
$\mathcal{N}_2$	the subset of $[0, 1]$ consisting of real numbers which are normal in base 2
σ	the symbol used for a generic erasing block substitution (see Definition 2.1)

w_ϵ	the unique word that is mapped to the empty word by an erasing block substitution
w_x	for $x \in (0, 1]$, the unique element of $\{0, 1\}^\omega$ which is a binary expansion for x and does not end with 0^∞
$[w]$	the real cylinder set generated by $w \in \{0, 1\}^*$ (see (2.9))
$\langle\langle w \rangle\rangle$	the set of all distinct words obtained by inserting any finite number of subwords of type 00 in w (see (2.10))
$[[w]]$	the union of dyadic intervals characterized by words in $\langle\langle w \rangle\rangle$ (see (2.11))
f_σ	the interval map generated by the erasing substitution σ (see (2.2))
ρ	the model-case erasing block substitution (see (2.3))
$\mathcal{I}_d$	the semi-algebra of dyadic subintervals of [0, 1]

References

[1] F. Balibrea, J. Smıtal and M. Štefánková. The three versions of distributional chaos. *Chaos Solitons Fractals* **23**(5) (2005), 1581–1583.
[2] J. Bobok and S. Troubetzkoy. Typical properties of interval maps preserving the Lebesgue measure. *Nonlinearity* **33**(12) (2020), 6461.
[3] R. Bowen. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* **153** (1971), 401–414.
[4] M. Čiklová. Dynamical systems generated by functions with connected G_δ graphs. *Real Anal. Exchange* **30**(2) (2005), 617–638.
[5] A. Della Corte and M. Farotti. Dynamics of interval maps generated by erasing substitutions. *J. Differential Equations* **323** (2022), 86–112.
[6] A. Della Corte, S. Isola and R. Piergallini. The simplest erasing substitution. *Nonlinearity* **35** (2022), 4317–4356.
[7] E. I. Dinaburg. A correlation between topological entropy and metric entropy. *Dokl. Akad. Nauk* **190**(1) (1970), 19–22.
[8] T. Downarowicz. Entropy structure. *J. Anal. Math.* **96**(1) (2005), 57–116.
[9] T. Downarowicz. *Entropy in Dynamical Systems (New Mathematical Monographs, 18)*. Cambridge University Press, Cambridge, 2011.
[10] F. Durand. Cobham's theorem for substitutions. *J. Eur. Math. Soc. (JEMS)* **13**(6) (2011), 1799–1814.
[11] F. Durand and M. Rigo. Syndeticity and independent substitutions. *Adv. Appl. Math.* **42**(1) (2009), 1–22.
[12] M. Einsiedler and T. Ward. *Ergodic Theory*. Springer, London, 2013, pp. 4–5.
[13] N. P. Fogg, V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel (Eds.). *Substitutions in Dynamics, Arithmetics and Combinatorics*. Springer, Berlin, 2002.
[14] S. Gentili. Semi-algebra and algebra of sets. *Measure, Integration and a Primer on Probability Theory. (UNITEXT, 125)*. Springer, Cham, 2020, pp. 121–129.
[15] A. Loranty, R. J. Pawlak and M. Terepeta. On almost continuous functions and peculiar points. *Eur. J. Math.* **5**(1) (2019), 106–115.
[16] R. Pawlak. On the entropy of Darboux functions. *Colloq. Math.* **2** (2009), 116.
[17] K. Petersen. *Ergodic Theory (Cambridge Studies in Advanced Mathematics, 2)*. Cambridge University Press, Cambridge, 1983.
[18] M. Queffelec. Old and new results on normality. *IMS Lecture Notes Monogr. Ser.* **48** (2006), 225–236.
[19] M. Queffélec. *Substitution Dynamical Systems–Spectral Analysis (Lecture Notes in Mathematics, 1294)*. Springer, Berlin, 2010.
[20] D. Reidenbach and J. C. Schneider. Restricted ambiguity of erasing morphisms. *Theoret. Comput. Sci.* **412**(29) (2011), 3510–3523.
[21] T. H. Steele. Dynamics of Baire-1 functions on the interval. *Eur. J. Math.* **5**(1) (2019), 138–149.
[22] T. H. Steele. Dynamics of Baire-2 functions on the interval. *Topology Appl.* **265** (2019), 106821.