LONDON MATHEMATICAL SOCIETY
EST. 1865

IOP Publishing

PAPER

The simplest erasing substitution

View the article online for updates and enhancements.

You may also like

The simplest erasing substitution

Alessandro Della Corte, Stefano Isola* and Riccardo Piergallini

Scuola di Scienze e Tecnologie, Università di Camerino, Italy

E-mail: stefano.isola@unicam.it

Abstract

In this work, we begin the study of a new class of dynamical systems determined by interval maps generated by the symbolic action of erasing substitution rules. We do this by discussing in some detail the geometric, analytical, dynamical and arithmetic properties of a particular example, which has the virtue of being arguably the simplest and that at the same time produces interesting properties and new challenging problems.

(Some figures may appear in colour only in the online journal)

Introduction

Substitutive dynamical systems are a widely studied and quite well-understood class (see for instance [17, 33]). In this context, it is usual to intend the term *substitution* in a rather specific sense. Namely, a substitution S is typically understood as a rule replacing every element from a given alphabet by a finite word on the same alphabet, in order to extend S by concatenation to a morphism over all the finite or infinite words. The morphic nature of S means that the action of the substitution on a certain symbol within a word w is independent of its position in w. Moreover, most of the times no symbol is mapped by S to the empty word, that is the substitution is assumed to be *non-erasing*.

In recent years, erasing substitutions have been taken in some consideration within combinatorics of words (see for instance [15, 34]), mostly in the context of theoretical computer science, while little attention seems to be devoted to the analytical and dynamical properties of real maps generated by the symbolic action of erasing substitutions. The shift from symbolic

spaces to the continuous real context, that is from zero-dimensional spaces to one-dimensional ones, has relevant consequences on the richness of the properties involved and the problems that seems natural to consider. Lately, topological dynamics questions initially considered only in case of continuous maps, such as the structure of ω-limit sets, topological entropy, the presence of Devaney, Li-Yorke or distributional chaos and mixing properties, have been addressed in a more general context. For instance: Darboux interval maps of Baire class 1 are considered in [21, 39], general Darboux interval maps are investigated from the point of view of the connections between transitivity, turbulence and topological entropy in [29, 30], interval maps with G_δ connected graph are studied in [8, 9], while in [20] an adjustment of the concept of Devaney chaos is proposed for discontinuous maps. In this connection, the study of the dynamics of real maps generated by erasing block substitutions appears as a natural development, as the objects constructed in this way fall into a category which is in a way the direct generalisation of the ones cited above, that is Baire class 1, generally not Darboux functions. The systematic investigation of the dynamical properties of these maps has just begun; for instance, generic interval maps of Baire class 1 are studied in [18, 36, 37].

Our aim is the investigation of an interval map generated by what is arguably the simplest erasing substitution rule, which we will indicate by ρ. We are led to the choice of ρ as follows. We start selecting the smallest non-trivial alphabet, that is $\{0, 1\}$. The simplest way to get an erasing substitution is to map one of the symbols, say 0, to the empty word. Then, we must make a choice on how to transform the other symbol 1. Clearly, to obtain a non-trivial map, we cannot replace every 1 with a fixed word, independently of its position. Making once again the simplest choice, we distinguish between odd and even positions, replacing one case with the symbol 0 and the other with the symbol 1.

Even in this model-case, the real map $R : [0, 1] \to [0, 1]$ generated by the symbolic action of ρ on binary expansions, presents interesting properties and challenging problems, which involve, among others, measure-theoretic, dimension-theoretic, topological, dynamical and arithmetical aspects. As we will see better in section 2, from an operational point of view the map is hardly manageable, as to obtain the first m digits of the nth iterate of its action at point x we must typically provide x with accuracy of order $\approx 2^{-2^{mn}}$. This entails that it is almost impossible to investigate in a purely numerical fashion objects as ω-limit sets and attractors. Moreover, as we will see, there is an uncountable and dense subset of [0, 1] exhibiting extreme distributional chaos, which means that the qualitative relative behaviour of sets of points, even arbitrarily close, cannot be extrapolated from finite samplings of the dynamics of any time length.

It is this combination of simplicity of the object and variety of related properties which in our opinion confers interest to the proposed subject.

The main properties of the map R, defined in sections 2 and 3, can be summarised as follows.

(a) The graph of R is a totally disconnected subset of the plane, whose Hausdorff dimension is between 1 and $\log_2 3$ (section 4).
(b) R is a singular Borel map of Baire class 1, it is not Darboux, and its mean value is the (admittedly a bit surprising) rational number $3/7$ (section 5).
(c) The closure of the fibers and of the fixed point set of R are null-measure Cantor sets, with Hausdorff dimension belonging to $[1/2, \log_2 \varphi]$ (sections 5 and 6).
(d) There are two rational R-cycles of order two which attract (in a finite time) every rational, moreover R has uncountably many periodic points for any given period, it exhibits Devaney and (uniform) DC1 chaos, it is topologically mixing, it has infinite topological entropy and every point in [0, 1] is a full entropy point (section 6).

(e) The combinatorially simplest periodic points of odd period (including period 1) are transcendental numbers and there are uncountably many periodic points that are transcendental as well for every given period (section 7).

For the reader's convenience, a list of the most used symbols is given at the end of the paper.

1. The general setting

Let $\mathcal{A} = \mathbb{N}^\omega$ be the set of all infinite sequences of non-negative integers, and $\mathcal{B} = \{0, 1\}^\omega \subset \mathbb{N}^\omega$ be the subset of all infinite binary sequences, both endowed with the metric given by

$$d(s, t) = 1/2^{\min\{k \geqslant 1 \ | \ s_k \neq t_k\}},$$

for every $s = (s_1, s_2, \ldots, s_n, \ldots)$ and $t = (t_1, t_2, \ldots, t_n, \ldots)$ such that $s \neq t$.

Then, $\mathcal{B}$ splits as the union of two dense subspaces

$$\mathcal{B} = \mathcal{C} \cup \mathcal{C}',$$

where $\mathcal{C}$ consists of all the infinite binary sequences which are not eventually 0, while $\mathcal{C}'$ consists of all the infinite binary sequences which are not eventually 1.

There is a one-Lipschitz homeomorphism $\eta : \mathcal{A} \to \mathcal{C}$ defined by

$$a = (a_1, a_2, \ldots, a_n, \ldots) \stackrel{\eta}{\longmapsto} \eta(a) = (0^{a_1}, 1, 0^{a_2}, 1, \ldots, 0^{a_n} 1, \ldots)$$

$$= (\underbrace{0, \ldots, 0}_{a_1}, 1, \underbrace{0, \cdots, 0}_{a_2}, 1, \cdots, \underbrace{0, \cdots, 0}_{a_n}, 1, \cdots) \in \mathcal{C} \quad (1)$$

for every $a = (a_1, a_2, \cdots, a_n, \cdots) \in \mathcal{A}$.

For the sake of convenience, we introduce the notation

$$\langle a \rangle^b = \eta(2a + \eta^{-1}(b)) \tag{2}$$

with $a \in \mathcal{A}$ and $b \in \mathcal{C}$. If $a = (a_1, a_2, \cdots, a_n, \cdots)$ then $\langle a \rangle^b$ is obtained from b by inserting 0^{2a_k} immediately before the kth occurrence of 1 in b. In particular, $\langle a \rangle^{1^\infty} = \langle a \rangle^{(1,1,\cdots)} = \eta(2a)$.

There is also a two-Lipschitz map $\xi : \mathcal{B} \to [0, 1]$ defined by

$$b = (b_1, b_2, \cdots, b_n, \cdots) \stackrel{\xi}{\longmapsto} \xi(b) = \sum_{n=1}^{\infty} \frac{b_n}{2^n} = 0.b_1 b_2 \cdots b_n \cdots \in [0, 1] \tag{3}$$

for every $b = (b_1, b_2, \cdots, b_n, \cdots) \in \mathcal{B}$.

The map ξ splits as the union of two bijective maps

$$\xi|_{\mathcal{C}} : \mathcal{C} \to (0, 1] \quad \text{and} \quad \xi|_{\mathcal{C}'} : \mathcal{C}' \to [0, 1).$$

Then, we have the inverse maps

$$\beta = (\xi|_{\mathcal{C}})^{-1} : (0, 1] \to \mathcal{C} \quad \text{and} \quad \beta' = (\xi|_{\mathcal{C}'})^{-1} : [0, 1) \to \mathcal{C}', \tag{4}$$

giving the binary expansions of the real numbers in $[0, 1]$, meaning the infinite sequences of binary digits in their binary expressions, completed by infinitely many 0's in the case of the finite binary expressions of dyadic rationals.

We emphasise the difference between a *binary expression* $0.x_1 x_2 \ldots$ of $x \in [0, 1]$, which can be finite (for dyadic rationals) or infinite, and a *binary expansion* of x which is always

an infinite sequence in $\mathcal{B}$. Namely, for a dyadic rational $x \in (0,1]$ admitting the finite binary expression $0.x_1 x_2 \ldots x_n$, we have $\beta'(x) = (x_1, x_2, \ldots, x_n, 0, 0, \ldots) \in \mathcal{C}'$, while we have $\beta(x) = (x_1, x_2, \ldots, x_{k-1}, 0, 1, 1, \ldots) \in \mathcal{C}$ if $k \leqslant n$ is the index of the last occurrence of 1 in $0.x_1 x_2 \ldots x_n$.

Now, let $\mathbb{D} \subset \mathbb{Q}$ denote the set of all dyadic rationals, and put

$$\mathcal{D} = \mathbb{D} \cap [0,1] \quad \text{and} \quad \mathcal{Q} = \mathbb{Q} \cap [0,1].$$

Notice that the two maps β and β' coincide and are continuous on $(0,1) \backslash \mathcal{D}$, while they differ on the set $\mathcal{D} \backslash \{0,1\}$, where β is only left-continuous and β' is only right-continuous.

Finally, we introduce the further notations $b_{|k}$ for the kth truncation of an infinite binary sequence b, and $[w]$ for the cylinder consisting of all the infinite binary sequences having the finite prefix w. Namely,

$$b_{|k} = (b_1, b_2, \ldots, b_k) \in \{0,1\}^k \quad \text{for every } b = (b_1, b_2, \ldots, b_n, \ldots) \in \mathcal{B}, \tag{5}$$

$$[w] = \{(w,b) \mid b \in \mathcal{B}\} \subset \mathcal{B} \quad \text{for every } w = (w_1, w_2, \ldots, w_k) \in \{0,1\}^k. \tag{6}$$

Then, for every $b \in \mathcal{B}$ and $k \geqslant 0$ we can express the open ball $B(b, 1/2^k) \subset \mathcal{B}$ as $[b_{|k}]$.

In the following, we identify the infinite binary sequence $b = (b_1, b_2, \ldots)$ with the infinite binary word $b = b_1 b_2 \cdots$, and we write $x = 0.b$ for (the corresponding binary expression of) $x = \xi(b)$. An analogous convention is adopted for finite binary sequences. Thus, the notations introduced in the equations (2), (5) and (6) make sense also in the context of binary words, and hence in the context of binary expressions/expansions of real numbers in $[0,1]$.

2. The substitution rule

Let $\{0,1\}^\infty$ be the set of all the finite or infinite binary words, including the empty word ϵ. For a word $w \in \{0,1\}^\infty$, we indicate by $|w|$ its (possibly infinite) length and by $|w|_1$ the (possibly infinite) number of 1's occurring in it.

Consider the map $\rho : \{0,1\}^\infty \to \{0,1\}^\infty$ digit-wise defined by the alternating substitution rule

$$\rho : \begin{cases} 0 \mapsto \epsilon \\ 1 \mapsto 0 & \text{in odd positions} \\ 1 \mapsto 1 & \text{in even positions} \end{cases}. \tag{7}$$

In other words, for any word $b = b_1 b_2 \cdots \in \{0,1\}^\infty$ the word $\rho(b)$ is obtained from b by deleting all the $b_n = 0$ and replacing each $b_n = 1$ by 0 if n is odd and by 1 if n is even.

We remark that the parity condition in the above definition makes ρ a non-morphic map with respect to the concatenation. Indeed, for any $v, w \in \{0,1\}^\infty$ with v a finite word and $|w|_1 > 0$, the equality $\rho(vw) = \rho(v)\rho(w)$ holds only if $|v|$ is even, while for $|v|$ odd we have $\rho(vw) = \rho(v)\widetilde{\rho}(w)$, where $\widetilde{\rho}$ is the complementary substitution rule

$$\widetilde{\rho} : \begin{cases} 0 \mapsto \epsilon \\ 1 \mapsto 1 & \text{in odd positions} \\ 1 \mapsto 0 & \text{in even positions} \end{cases}. \tag{8}$$

For the sake of convenience, we introduce the notation

$$
\rho_k = \begin{cases} \rho & \text{if } k \text{ is an even integer} \\ \widetilde{\rho} & \text{if } k \text{ is an odd integer} \end{cases},
\tag{9}
$$

and we write ρ_v to mean $\rho_{|v|}$. In this way, for any v and w as above we have

$$
\rho(vw) = \rho(v)\rho_v(w).
\tag{10}
$$

We observe that, for binary words of even or infinite length, the action of ρ coincides with that of the block substitution rule

$$
\tau : \begin{cases} 00 \mapsto \epsilon \\ 01 \mapsto 1 \\ 10 \mapsto 0 \\ 11 \mapsto 01 \end{cases}.
\tag{11}
$$

This interpretation of the substitution rule ρ enables us to estimate the *vanishing order*

$$
n_\epsilon(w) = \min\{k > 0 \mid \rho^k(w) = \epsilon\}
\tag{12}
$$

of a finite binary word w under the action of ρ in terms of $|w|$.

Proposition 2.1.   *If $w \neq \epsilon$ is a finite binary word then $n_\epsilon(w) \leqslant 2\lfloor \log_2 |w| \rfloor + 2$.*

Proof.   We proceed by induction on $|w| \geqslant 1$, based on the trivial case of $|w| = 1$. So, let $|w| > 1$ and observe that we can assume $|w|$ even, because of the obvious equality $n_\epsilon(w) = n_\epsilon(w0)$. In this case, we have $|\rho^2(w)| = |\rho(\tau(w))| = |\tau(w)|_1$. Looking at the contribution to $|\tau(w)|_1$ of each single pair, according to (11), we immediately get $|\rho^2(w)| \leqslant |w|/2$. Then, by the inductive hypothesis,

$$
n_\epsilon(w) \leqslant n_\epsilon(\rho^2(w)) + 2 \leqslant 2\lfloor \log_2 |\rho^2(w)| \rfloor + 4 \leqslant 2\lfloor \log_2(|w|/2) \rfloor + 4 = 2\lfloor \log_2 |w| \rfloor + 2.
$$

$\square$

As it can be easily realised, the vanishing order $n_\epsilon(w)$ of a word of length $n \geqslant 1$ can assume any value between 1 and the upper bound established by the proposition. In particular, the extremal values are attained for the words 0^n and 1^n, which vanish exactly after 1 and $2\lfloor \log_2 n \rfloor + 2$ iterations of ρ, respectively.

Up to the identification between binary sequences and binary words, we have $\mathcal{C} \subset \mathcal{B} \subset \{0,1\}^\infty$. Moreover, it can be easily seen that $\rho^{-1}(\mathcal{B}) = \mathcal{C}$. Then, it makes sense to consider the restriction $\rho|_{\mathcal{C}} = \tau|_{\mathcal{C}} : \mathcal{C} \to \mathcal{B}$.

Now, consider the substitution rule

$$
\overline{\tau} : \begin{cases} 0 \mapsto 10 \\ 1 \mapsto 01 \end{cases},
\tag{13}
$$

and let $\overline{\tau} : \{0,1\}^\infty \to \{0,1\}^\infty$ be the corresponding digit-wise generated map. Then, the image $\overline{\tau}(w)$ has even or infinite length for every $w \in \{0,1\}^\infty$. Moreover, $\tau \circ \overline{\tau} = \mathrm{id}_{\{0,1\}^\infty}$ and $\overline{\tau}(w) \in \mathcal{C}$ for every $w \in \mathcal{B}$. It immediately follows that the map $\rho|_{\mathcal{C}}$ is surjective.

Actually, $\overline{\tau}(w)$ is not the 'simplest' element in $(\rho|_{\mathcal{C}})^{-1}(w)$ for $w \in \mathcal{B}$. In fact, $\overline{\tau}(w)$ can contain pairs of consecutive 0's and these can be deleted without changing the image under

$\rho|_{\mathcal{C}}$. The reason is that $\rho(0) = \epsilon$ and that deleting/inserting subwords of even length does not change the parity of the subsequent positions in the word.

In the following we indicate by $\mathcal{S} \subset \mathcal{C}$ the set of the 'simplest' infinite binary words, meaning those which do not contain any pair of consecutive 0's.

In order to directly construct the 'simplest' element in $(\rho|_{\mathcal{C}})^{-1}(w)$, the unique one which belongs to $\mathcal{S}$, we consider the section $\sigma : \{0, 1\}^\infty \to \{0, 1\}^\infty$ digit-wise generated by the rule

$$\sigma : \begin{cases} x \mapsto 1 & \text{if } x \text{ is preceded by } 1 - x \text{ in the word or it is 0 as the first digit} \\ x \mapsto 01 & \text{if } x \text{ is preceded by } x \text{ in the word or it is 1 as the first digit} \end{cases}. \tag{14}$$

Given $w \in \{0, 1\}^\infty$, $\sigma(w)$ is actually given by a true substitution rule (in the usual sense) applied to the sequence of first differences of the word $1w$ obtained prepending 1 to w. Moreover, $\sigma(b) \in \mathcal{S} \subset \mathcal{C}$ for every $b \in \mathcal{B}$, and hence $\sigma|_{\mathcal{B}} : \mathcal{B} \to \mathcal{C}$ is a section of $\rho|_{\mathcal{C}}$.

Proposition 2.2. *For every $b \in \mathcal{B}$ the fiber $(\rho|_{\mathcal{C}})^{-1}(b)$ contains uncountably many elements. In fact,*

$$(\rho|_{\mathcal{C}})^{-1}(b) = \{\langle a \rangle^{\sigma(b)} \mid a \in \mathcal{A}\}, \tag{15}$$

and hence $\sigma(b)$ is the unique element of $(\rho|_{\mathcal{C}})^{-1}(b)$ belonging to $\mathcal{S}$ and it is the maximum of $(\rho|_{\mathcal{C}})^{-1}(b)$ with respect to the lexicographic order.

Proof. We only need to prove equation (15), since the rest of the statement immediately follows from it. A straightforward induction on n gives $\rho|_{\mathcal{C}}(\sigma(b))_n = b_n$ for every $n \geqslant 1$, which implies that $\rho(\sigma(b)) = b$. On the other hand, by deleting pairs of 0's as discussed above, we get $\rho(\langle a \rangle^{\sigma(b)}) = b$ for every $a \in \mathcal{A}$. Conversely, let any $v \in (\rho|_{\mathcal{C}})^{-1}(b)$ be written as $v = \eta(c)$, thanks to (1). The nth occurrence of 1 in v is in position $k_n = c_1 + \cdots + c_n + n$. Since the parity of k_n is determined by $\rho(v)_n = b_n$, induction on n shows that also the parity of c_n is determined for every $n \geqslant 1$. So, if $\sigma(b) = \eta(d)$ then we have $d_n = 0, 1$ and $c_n \equiv d_n \mod 2$ for every $n \geqslant 1$. This yields $c = 2a + d$ for a suitable $a \in \mathcal{A}$, and hence, according to (2), $v = \eta(c) = \eta(2a + d) = \langle a \rangle^{\sigma(b)}$. $\square$

We conclude this section by considering some continuity properties of the maps $\rho|_{\mathcal{C}} : \mathcal{C} \to \mathcal{B}$ and $\sigma|_{\mathcal{B}} : \mathcal{B} \to \mathcal{C}$ with respect to the metric induced by the inclusions $\mathcal{C} \subset \mathcal{B} \subset \mathcal{A}$.

Proposition 2.3. *The map $\rho|_{\mathcal{C}} : \mathcal{C} \to \mathcal{B}$ is continuous but not uniformly continuous. Moreover, $\lim_{v \to b} \rho(v)$ does not exist for any $b \in \mathcal{B} \setminus \mathcal{C}$. So, $\rho|_{\mathcal{C}}$ cannot be continuously extended to any subset of $\mathcal{B}$ larger than $\mathcal{C}$.*

Proof. For any $b \in \mathcal{C}$ and $n \geqslant 1$, let $k_n = |\rho(b_{|n})| = |b_{|n}|_1$. Then, $\rho(B(b, 1/2^n)) = \rho([b_{|n}]) \subset [\rho(b_{|n})] = [\rho(b)_{|k_n}] = B(\rho(b), 1/2^{k_n})$, and hence the continuity of $\rho|_{\mathcal{C}}$ at b immediately follows from the fact that $k_n \to \infty$ for $n \to \infty$. On the other hand, if ρ were uniformly continuous then it would be possible to extend it to the whole space $\{0, 1\}^\infty$, and this would contradict the second part of the statement, which we are going to prove.

Let $b \in \mathcal{B} \setminus \mathcal{C}$. Then, b is eventually 0, and hence it can be written as $b = v0^\infty$. Consider the sequence $(b_n = v0^n1^\infty)_{n \geqslant 1} \subset \mathcal{C}$. We have, $\lim_{n \to \infty} b_n = b$, while the limit $\lim_{n \to \infty} \rho(b_n)$ cannot exist, being $\rho(b_{2k+1}) = \rho(b_1) \neq \rho(b_2) = \rho(b_{2k+2})$ for every $k \geqslant 1$. $\square$

Proposition 2.4. *The map $\sigma|_{\mathcal{B}} : \mathcal{B} \to \mathcal{C}$ is uniformly continuous, in fact it is one-Lipschitz.*

Proof. Since $|\sigma(v)| \geqslant k$ for every $v \in \{0, 1\}^k$, we have that $d(b, b') = 1/2^k \Rightarrow b_{|k} = b'_{|k} \Rightarrow \sigma(b)_{|k} = \sigma(b')_{|k} \Rightarrow d(\sigma(b), \sigma(b')) \leqslant 1/2^k$, for every $b, b' \in \mathcal{B}$. $\square$

3. The real map

We define the map $R : [0, 1] \to [0, 1]$ as follows

$$R(x) = \begin{cases} 2/3 & \text{if } x = 0 \\ \xi(\rho(\beta(x))) & \text{if } x \in (0, 1] \end{cases}. \tag{16}$$

For $x \neq 0$ this amounts to say that $R(x)$ admits a binary expansion which is the transform under ρ of the unique binary expansion of x in $\mathcal{C}$. In formulas, if $x = 0.b$ with $b \in \mathcal{C}$ then

$$R(x) = R(0.b) = 0.\rho(b).$$

The image of 0 could be chosen somewhat arbitrarily. This particular choice is convenient for later purposes.

The continuity properties we have seen in the previous section for the maps involved in the definition of R give us the next proposition.

Proposition 3.1. *The map R is continuous everywhere but on the set $\mathcal{D}\backslash\{1\}$, where it is only left-continuous. Hence, R is a Borel function of Baire class 1, i.e. it is a point-wise limit of continuous functions.*

Proof. The first assertion immediately derives from the continuity of ξ and $\rho|_{\mathcal{C}}$ (proposition 2.3), and the fact that β is continuous everywhere but on the set $\mathcal{D}\backslash\{1\}$, where it is only left-continuous. Then, the second assertion follows from the fact that R has countably many discontinuities (see [41, theorem 11.8] or [22, chapter three, section 34, paragraph VII]). $\square$

It is worth remarking that the right-discontinuity of R on $\mathcal{D}\backslash\{1\}$ is not simply due to the right-discontinuity of β on that set, but it is instead also related to the fact that $\rho|_{\mathcal{C}}$ is not continuously extendable, as stated in proposition 2.3. In fact, as we will see in proposition 4.2, for $x \in \mathcal{D}\backslash\{0, 1\}$ the right limit $\lim_{y \searrow x} R(y)$ does not exist and even more

$$R(x) \notin \left[\liminf_{y \searrow x} R(y), \limsup_{y \searrow x} R(y) \right].$$

For example:

$$R(1/2) = 2/3, \quad \text{while } \liminf_{x \searrow 1/2} R(x) = 0 \quad \text{and} \quad \limsup_{x \searrow 1/2} R(x) = 1/2,$$

$$R(1/4) = 1/3, \quad \text{while } \liminf_{x \searrow 1/4} R(x) = 1/2 \quad \text{and} \quad \limsup_{x \searrow 1/4} R(x) = 1.$$

In particular, R is not a Darboux function, meaning that it does not satisfy the intermediate value property. Actually, it is not Darboux from the right for every $x \in \mathcal{D}\backslash\{1\}$ (see corollary 4.4). Thus, by Darboux's theorem, the integral function $\int_0^x R(t)\mathrm{d}t$ is differentiable everywhere but on the set $\mathcal{D}\backslash\{1\}$, where it is only left-differentiable.

By a classical characterisation of Baire class 1 functions given by Lebesgue [23] (see also [4, section 4.4] or [22, chapter two, section 31, paragraph II]), we know that for every $\varepsilon > 0$ there exists a countable closed covering $\{C_n\}_{n \geqslant 1}$ of $[0, 1]$ such that the oscillation of R on each C_n is less than ε. We recall that the oscillation of a real function f on a subset S of its domain is defined as $O_f(S) = \sup_{x,y \in S} |f(x) - f(y)|$.

Before going on, we want to provide an explicit construction of a covering $\{C_n\}_{n \geqslant 1}$ as above. First, we need to estimate the oscillation of R on the dyadic intervals as follows.

For every $x \in \{0, 1\}^n$ and $n \geqslant 1$, we consider the cylinder $[x] \subset \mathcal{B}$ and the interval $\xi([x]) = [0.x0^\infty, 0.x1^\infty] \subset [0, 1]$, and observe that

$$R(\xi([x])) = \begin{cases} [0, 1] & \text{if } x = 0^n \\ \xi(\{\rho(x_{|k-1}01^\infty)\} \cup [\rho(x)]) & \text{if } x \neq 0^n \end{cases}, \tag{17}$$

where k is the index of the last occurrence of 1 in x. In particular, in the latter case we have $R(\xi([x])) \subset \xi([\rho(x_{|k-1})])$, which implies that

$$O_R(\xi([x])) \leqslant 1/2^h \quad \text{with } h = |\rho(x_{|k-1})| = |x|_1 - 1. \tag{18}$$

Now we can start with our construction. Given $\varepsilon > 0$, we choose a positive integer ℓ such that $1/2^{\ell-1} < \varepsilon$, and consider the set $D_\ell \subset \mathcal{D}$ consisting of all dyadic rational x whose binary expansion $\beta'(x)$ contains at most ℓ 1's, that is $|\beta'(x)|_1 \leqslant \ell$. Then, the ordered set $(D_\ell, \geqslant)$ is isomorphic to the countable ordinal $\omega^\ell + 1$, and hence D_ℓ has Cantor–Bendixson rank $\ell + 1$. Moreover, considering D_ℓ with the standard order in $\mathbb{R}$, we have $\min D_\ell = 0$ and $\max D_\ell = 0.1^\ell = \xi(1^\ell)$, and for every $x \in D_\ell \backslash \{0\}$, the immediate predecessor of x in D_ℓ, that is the largest element of D_ℓ smaller than x, is $p(x) = \xi(\beta(x)_{|k}0^\infty)$, where k denotes the index of the ℓth 1 in $\beta(x)$. Keeping the same notations, we put

$$C_x = \begin{cases} \xi([1^\ell]) = [\xi(1^\ell 0^\infty), \xi(1^\ell 1^\infty)] = [0.1^\ell, 1] & \text{if } x = 1 \\ \xi([\beta(x)_{|k}]) = [\xi(\beta(x)_{|k}0^\infty), \xi(\beta(x)_{|k}1^\infty)] = [p(x), x] & \text{if } x \in D_\ell \backslash \{0\}. \\ \{0\} & \text{if } x = 0 \end{cases}$$

Then, $\{C_x\}_{x \in D_\ell \cup \{1\}}$ is a countable closed covering of $[0, 1]$ with the wanted oscillation bound. Indeed, in all cases, we have $O_R(C_x) \leqslant 1/2^{\ell-1} < \varepsilon$ by (18).

Concerning the fibers, it immediately follows from (16) that

$$R^{-1}(y) \cap (0, 1] = \beta^{-1}(\rho^{-1}(\xi^{-1}(y))) = \xi(\rho^{-1}\{\beta(y), \beta'(y)\})) \tag{19}$$

for every $y \in [0, 1]$. This actually coincides with $R^{-1}(y)$ if $y \neq 2/3$, while 0 must be added to get $R^{-1}(y)$ if $y = 2/3$.

Taking into account equation (15) and the properties of the maps β and β' we have seen at the end of section 1, we have

$$R^{-1}(y) = \begin{cases} \xi(\langle\!\langle \sigma(\beta(y)) \rangle\!\rangle) & \text{if } y \in [0, 1] \backslash (\mathcal{D} \cup \{2/3\}) \text{ or } y = 1 \\ \xi(\langle\!\langle \sigma(\beta(y)) \rangle\!\rangle) \cup \{0\} & \text{if } y = 2/3 \\ \xi(\langle\!\langle \sigma(\beta(y)) \rangle\!\rangle) \cup \xi(\langle\!\langle \sigma(\beta'(y)) \rangle\!\rangle) & \text{if } y \in \mathcal{D} \backslash \{0, 1\} \\ \xi(\langle\!\langle \sigma(\beta'(y)) \rangle\!\rangle) & \text{if } y = 0 \end{cases}, \tag{20}$$

where the following notation is used for $b \in \mathcal{C}$

$$\langle\!\langle b \rangle\!\rangle = \{\langle a \rangle^b \mid a \in \mathcal{A}\}. \tag{21}$$

Since $\xi|_{\mathcal{C}}$ is injective, the fiber $R^{-1}(y)$ contains uncountably many elements for every $y \in [0,1]$. Moreover, according to proposition 2.2, any fiber $R^{-1}(y)$ admits a maximum element

$$S(y) = \begin{cases} \xi(\sigma(\beta(y))) & \text{if } y \in [0,1]\backslash\mathcal{D} \text{ or } y = 1 \\ \max\{\xi(\sigma(\beta(y))), \xi(\sigma(\beta'(y)))\} & \text{if } y \in \mathcal{D}\backslash\{0,1\} \\ \xi(\sigma(\beta'(y))) & \text{if } y = 0 \end{cases} . \quad (22)$$

The map $S : [0,1] \to [0,1]$ defined by the equation above is a null measure section for R, as we will see in proposition 5.1. Here, we limit ourselves to consider the following continuity properties of it.

Proposition 3.2. *The map S is continuous everywhere but on the set $\mathcal{D}\backslash\{0,1\}$, where it is either left- or right-continuous, with different existing left- and right-limits. Hence, S is a Borel function of Baire class 1.*

Proof. The first assertion easily follows from the continuity properties of $\xi, \sigma|_{\mathcal{B}}$ (proposition 2.4), β and β', and the fact that, for every $y = (2k+1)/2^n \in \mathcal{D}\backslash\{0,1\}$, we have

$$|\xi(\sigma(\beta(y))) - \xi(\sigma(\beta'(y)))| \leqslant 2d(\sigma(\beta(y)), \sigma(\beta'(y)))$$
$$\leqslant 2d(\beta(y), \beta'(y)) = 1/2^{n-1}.$$

Then, we can apply [41] as above to get the second assertion. □

Recalling how ρ and $\widetilde{\rho}$ act on the binary digits of x we readily deduce the next proposition.

Proposition 3.3. *The map R satisfies the functional equations for every $x \in (0,1]$*

$$R\left(\frac{x}{2}\right) = 1 - R(x), \quad (23)$$

$$R\left(\frac{x+1}{2}\right) = \frac{1-R(x)}{2}. \quad (24)$$

Proof. For every $x \in (0,1]$, we have

$$R\left(\frac{x}{2}\right) = \xi\left(\rho\left(\beta\left(\frac{x}{2}\right)\right)\right) = \xi(\rho(0\beta(x))) = \xi(\widetilde{\rho}(\beta(x))) = 1 - \xi\rho(\beta(x)))$$
$$= 1 - R(x),$$

$$R\left(\frac{x+1}{2}\right) = \xi\left(\rho\left(\beta\left(\frac{x+1}{2}\right)\right)\right) = \xi(\rho(1\beta(x))) = \xi(0\widetilde{\rho}(\beta(x)))$$
$$= \frac{1 - \xi\rho(\beta(x)))}{2} = \frac{1-R(x)}{2}.$$

□

Notice that the relations (23) and (24) are verified by any map defined by means of a generalised substitution of type (7), when odd-indexed 1's go into a word w and even-indexed 1's go into its complementary word $\widetilde{w}$.

By using the relations (23) and (24), one can get a countable family of functional equations, one for each word $w \in \{0, 1\}^\infty$. Namely, if w is a finite binary word such that $|w| = n$, $\rho(w) = v$ and $|v| = m$, for every $x \in (0, 1]$ we have

$$R \left(\frac{x}{2^n} + \sum_{i=1}^{n} \frac{w_i}{2^i} \right) = \sum_{i=1}^{m} \frac{v_i}{2^i} + \frac{1}{2^m} \left(\frac{1 + (-1)^{n+1}}{2} + (-1)^n R(x) \right).$$

Moreover, from (23) and (24) it easily follows that

$$R(x) = \begin{cases} 2R \left(x + \dfrac{1}{2} \right) \text{ if } x \in \left(0, 1/2 \right] \\ \dfrac{1}{2}R \left(x - \dfrac{1}{2} \right) \text{ if } x \in \left(1/2, 1 \right] \end{cases}. \tag{25}$$

4. Graph properties

In this section we consider some properties of the map R related to the geometry of its graph $G = \{(x, R(x)) | x \in [0, 1]\} \subset [0, 1]^2$, which is depicted in figure 1 below.

We remark that there is *a priori* no reason to think that the dots in the picture represent a numerical approximation of points of the graph of R, as numerical computation is of course based on (finite) binary expressions of dyadic rational numbers, whereas ρ acts always on the infinite representation and is discontinuous precisely on dyadic rationals. However, the picture still represents a numerical approximation of the graph of R for a somewhat deeper reason, that is because we can interpret it as the plot of an element of a sequence of step functions, constant on cylinder sets characterized by the same finite prefix, converging to R. That this 'numerical' convergence is mathematically meaningful is indeed ensured by proposition 3.1, which establishes that R can be point-wise limit of *continuous* functions.

The functional relations (23) and (24) induce an interesting recursive structure on G.

Consider, the transformations $T_0, T_1 : [0, 1]^2 \to [0, 1]^2$ defined by the equations

$$T_0(x, y) = \left(\frac{x}{2}, 1 - y \right) \quad \text{and} \quad T_1(x, y) = \left(\frac{x + 1}{2}, \frac{1 - y}{2} \right).$$

Let $\mathcal{T}$ be the operator which associates to each subset $E \subset [0, 1]^2$ the subset

$$\mathcal{T}(E) = T_0(E) \cup T_1(E).$$

Of course, $\mathcal{T}$ preserves compactness. Moreover, (23) and (24) imply the equality $\mathcal{T}(H) = H$ for $H = G - \{(0, 2/3)\}$. Indeed, for every $x \in (0, 1]$, we have

$$T_0(x, R(x)) = \left(\frac{x}{2}, 1 - R(x) \right) = \left(\frac{x}{2}, R \left(\frac{x}{2} \right) \right)$$

and

$$T_1(x, R(x)) = \left(\frac{x + 1}{2}, \frac{1 - R(x)}{2} \right) = \left(\frac{x + 1}{2}, R \left(\frac{x + 1}{2} \right) \right).$$

Therefore,
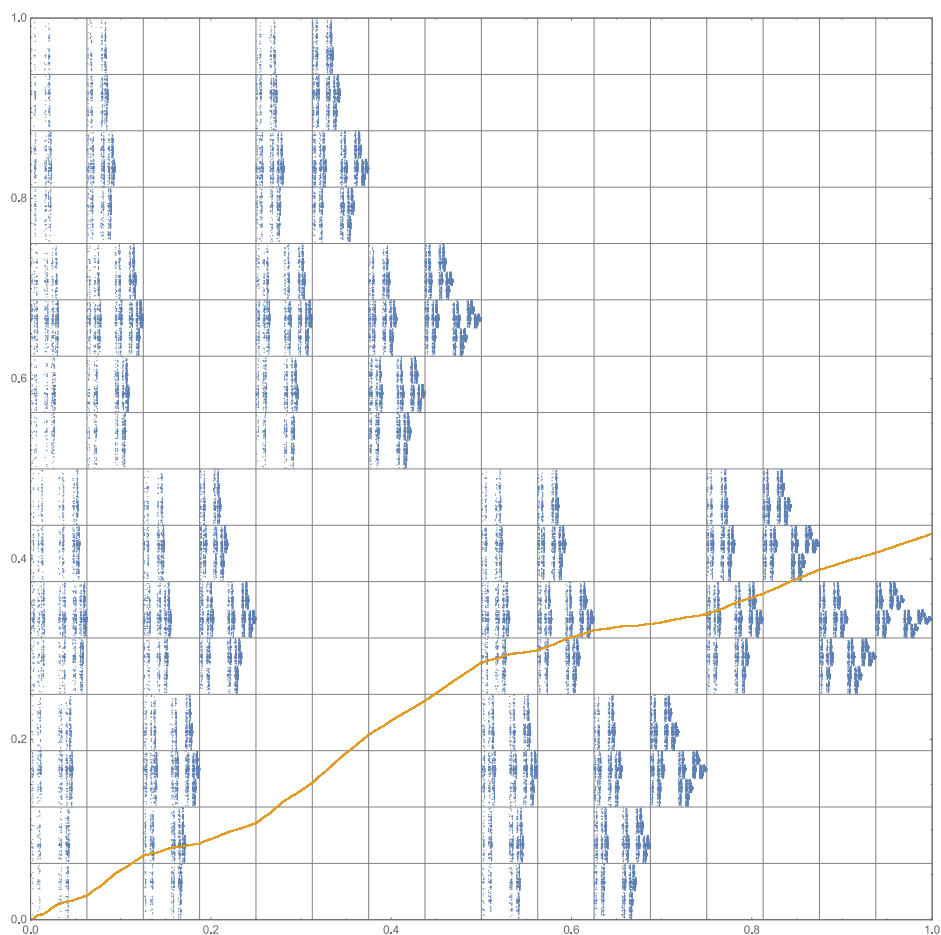
Figure 1. The graph of the map R (blue) and of its integral function (orange) (color figure online).

$$T_0(H) = H \cap ([0, 1/2] \times [0, 1]) \quad \text{and} \quad T_1(H) = H \cap ([1/2, 1] \times [0, 1]),$$

which gives $\mathcal{T}(H) = H$.

We inductively define $K_n \subset [0, 1]^2$ for $n \geqslant 0$, by putting

$$K_0 = [0, 1]^2 \quad \text{and} \quad K_{n+1} = \mathcal{T}(K_n) = T_0(K_n) \cup T_1(K_n),$$

where

$$T_0(K_n) = K_{n+1} \cap ([0, 1/2] \times [0, 1]) \quad \text{and} \quad T_1(K_n) = K_{n+1} \cap ([1/2, 1] \times [0, 1]).$$

(26)

The trivial inclusion $K_1 \subset K_0$ implies that $K_{n+1} \subset K_n$ for every $n \geqslant 0$. So, the K_n's form a decreasing sequence of compact subspaces of $[0, 1]^2$ converging to

$$K_\infty = \cap_{n \geqslant 0} K_n \subset [0, 1]^2.$$

In order to describe the sequence of the K_n's from a different point of view, we consider the maps $t_0, t_1 : [0, 1] \rightarrow [0, 1]$ defined by

$$t_0(y) = 1 - y \quad \text{and} \quad t_1(y) = \frac{1 - y}{2},$$

which give the action of T_0 and T_1 on the second coordinate.

Moreover, we consider the intervals

$$I_0 = [0, 1] \quad \text{and} \quad I_x = t_{x_1} \circ t_{x_2} \circ \cdots \circ t_{x_n}([0, 1]) \quad \text{for } x \in \mathcal{D} \setminus \{1\}$$

where $0.x_1 x_2 \cdots x_n$ is a finite binary expression of x. Since the terminal 0's in the binary word $x_1 x_2 \cdots x_n$ do not affect the interval I_x, this is well-defined, depending only on x.

We point out that the interval I_x satisfies the identities

$$t_0(I_x) = I_{x/2} \quad \text{and} \quad t_1(I_x) = I_{(x+1)/2}. \tag{27}$$

Then, by reasoning inductively as above, the interval I_x for $x = 0.x_1 \cdots x_n$ can be shown to coincide with the cylinder generated by the word $\rho(x_1 \cdots x_n)$, that is

$$I_x = [\rho(x_1 \cdots x_n)]. \tag{28}$$

For every $n \geqslant 0$, we can decompose K_n as a union of rectangles as follows

$$K_n = \cup_{k=0,\ldots,2^n-1} [k/2^n, (k+1)/2^n] \times I_{k/2^n}.$$

This equality is trivially true for $n = 0$, while it can be proved by induction for $n > 0$ by taking into account of (27), from which

$$T_0(K_n) = \cup_{k=0,\ldots,2^n-1} [k/2^{n+1}, (k+1)/2^{n+1}] \times I_{k/2^{n+1}}$$

and

$$T_1(K_n) = \cup_{k=2^n,\cdots,2^{n+1}-1} [k/2^{n+1}, (k+1)/2^{n+1}] \times I_{k/2^{n+1}}.$$

Then, we can obtain K_{n+1} from K_n by replacing each rectangle

$$[k/2^n, (k+1)/2^n] \times I_{k/2^n}$$

by the union

$$([k/2^n, (2k+1)/2^{n+1}] \times I_{k/2^n})$$
$$\cup ([(2k+1)/2^{n+1}, (k+1)/2^n] \times I_{(2k+1)/2^{n+1}}).$$

Remark 4.1. Passing from K_n to K_{n+1} amounts to split each rectangle $[k/2^n, (k+1)/2^n] \times I_{k/2^n}$ into four congruent closed sub-rectangles and remove the right-top or right-bottom one, depending on n being even or odd respectively, while keeping the union of the other three sub-rectangles.

Now, come back to H and G. The trivial inclusion $H \subset K_0$ and the equality $\mathcal{T}(H) = H$ imply that $H \subset K_n$ for every $n \geqslant 0$. On the other hand, as

$$\{0\} \times [0,1] = T_0(\{0\} \times [0,1]) \subset \mathcal{T}(\{0\} \times [0,1]),$$

we also have $\{0\} \times [0,1] \subset K_n$ for every $n \geqslant 0$. Hence, $G \subset (\{0\} \times [0,1]) \cup H$ implies $G \subset K_\infty$, and therefore $\mathrm{Cl}\, G \subset K_\infty$. We want to show that the reversed inclusion holds as well. In fact, we have the following slightly stronger result.

Proposition 4.2. *The following equalities hold*

$$K_\infty = \mathrm{Cl}\, G = G \cup (\cup_{x \in \mathcal{D} \setminus \{1\}} (\{x\} \times I_x)).$$

Moreover, for every $x \in \mathcal{D} \setminus \{0,1\}$,

$$R(x) \notin I_x = \left[\liminf_{\xi \searrow_* x} R(\xi), \limsup_{\xi \searrow_* x} R(\xi) \right].$$

In order to prove the proposition a technical lemma is needed.

Lemma 4.3. *For every $x \in (0,1)$ we have*

$$K_\infty \cap (\{x\} \times [0,1]) = \begin{cases} \{x\} \times (\{R(x)\} \cup I_x) \text{ with } R(x) \notin I_x & \text{if } x \in \mathcal{D} \\ \{(x, R(x))\} & \text{if } x \notin \mathcal{D} \end{cases}.$$

Proof. The case when $x \in \mathcal{D}$ immediately follows from the definition of K_∞ and the following statement: if $0.x_1 \cdots x_k$ is the shortest binary expression of x, then for every $n \geqslant k$

$$K_n \cap (\{x\} \times [0,1]) = \{x\} \times (I_x \cup J_{x,n})$$

where $(J_{x,n})_{n \geqslant k}$ is a non-increasing sequence of intervals such that $\cap_{n \geqslant k} J_{x,n} = \{R(x)\}$ with $R(x) \notin I_x$. We prove this statement by induction on $k \geqslant 1$.

If $k = 1$ then $x = 1/2$, and for every $n > 1$, we have

$$K_n \cap (\{1/2\} \times [0,1]) = T_0(K_{n-1} \cap (\{1\} \times [0,1])) \cup T_1(K_{n-1} \cap (\{0\} \times [0,1])),$$

so that, by recalling equation (26),

$$K_n \cap (\{1/2\} \times [0,1]) = T_0(\{1\} \times t_1^{n-1}([0,1])) \cup T_1(\{0\} \times [0,1])$$

$$= \{1/2\} \times (t_0 \circ t_1^{n-1}([0,1]) \cup [0,1/2]).$$

Then, $J_{1/2,n} = (t_0 \circ t_1^{n-1}([0,1]))_{n \geqslant 1}$ is a non-increasing sequence of intervals whose intersection is $t_0(R(1)) = R(1/2) = 2/3$. Moreover, $J_{1/2,n} \cap I_{1/2} = J_{1/2,n} \cap [0,1/2] = \emptyset$ for every n large enough, and hence $R(1/2) \notin I_{1/2}$.

Consider now the case of a dyadic rational $x = 0.x_1 \cdots x_k$ with $k > 1$. Let $x' = 0.x_2 \cdots x_k = 2x \pmod 1$ and assume, by the inductive hypothesis, that $K_n \cap (\{x'\} \times [0,1]) = \{x'\} \times (I_{x'} \cup J_{x',n})$ for every $n \geqslant k-1$, with $(J_{x',n})_{n \geqslant k-1}$ a non-increasing sequence of intervals whose intersection is $\{R(x')\}$ and $R(x') \notin I_{x'} = t_{x_2} \circ \cdots \circ t_{x_k}([0,1])$. Then, for every $n \geqslant k$ we have

$$K_n \cap (\{x\} \times [0,1]) = T_{x_1}(K_{n-1} \cap (\{x'\} \times [0,1])),$$

and hence

$$K_n \cap (\{x\} \times [0,1]) = \{x\} \times (t_{x_1}(I_{x'}) \cup t_{x_1}(J_{x',n-1})).$$

Here, $(J_{x,n} = t_{x_1}(J_{x',n-1}))_{n \geqslant k}$ is a non-increasing sequence of intervals whose intersection is $\{R(x)\} = \{t_{x_1}(R(x'))\}$, and $R(x) \notin I_x = t_{x_1}(I_{x'})$.

We are left to consider the case when $x \notin \mathcal{D}$. For every $n \geqslant 1$, let $x_{|n} = 0.x_1 \cdots x_n$ the nth truncation of the binary expression of x. We want to prove, by induction on $n \geqslant 1$, that

$$K_n \cap (\{x\} \times [0,1]) = \{x\} \times I_{x_{|n}}.$$

In fact, for $n = 1$ we have $I_{x_{|1}} = [0,1]$ if $x_1 = 0$ and $I_{x_{|1}} = [0,1/2]$ if $x_1 = 1$. In both cases, $K_1 \cap (\{x\} \times [0,1]) = \{x\} \times I_{x_{|1}}$ and $R(x) \in I_{x_{|1}}$. Moreover, for $n > 1$ we have

$$I_{x_n} = \begin{cases} I_{x_{|n-1}} & \text{if } x_n = 0 \\ I^+_{x_{|n-1}} & \text{if } x_n = 1 \quad \text{and} \quad n \text{ is even,} \\ I^-_{x_{|n-1}} & \text{if } x_n = 1 \quad \text{and} \quad n \text{ is odd} \end{cases}$$

where $I^+_{x_{|n-1}}$ and $I^-_{x_{|n-1}}$ denote the upper half and the lower half of $I_{x_{|n-1}}$, respectively. In all the cases, the equality $K_n \cap (\{x\} \times [0,1]) = \{x\} \times I_{x_{|n}}$ and the relation $R(x) \in I_{x_{|n}}$ follow by induction on n, taking into account that

$$0.x_1 \cdots x_{n-1} < x < 0.x_1 \cdots x_{n-1}1 \qquad \text{if } x_n = 0$$
$$0.x_1 \cdots x_{n-1}1 < x < 0.x_1 \cdots x_{n-1}1^\infty \quad \text{if } x_n = 1.$$

Then,

$$K_\infty \cap (x \times [0,1]) = \{x\} \times (\cap_{n \geqslant 1} I_{x_{|n}}) = \{(x, R(x))\}.$$

$\square$

Proof of proposition 4.2.   According to the lemma above, taking the union over $x \in [0,1)$ we immediately get the following equality

$$K_\infty = G \cup (\cup_{x \in \mathcal{D} \setminus \{1\}} (\{x\} \times I_x)), \tag{29}$$

with

$$G \cap (\cup_{x \in \mathcal{D} \setminus \{1\}} (\{x\} \times I_x)) \ = \{(0, 2/3)\}.$$

Equation (29) implies that $\mathrm{Cl}\, G \subset K_\infty$. Therefore, in order to conclude that $K_\infty = \mathrm{Cl}\, G$, it suffices to show that the inclusion $\{x\} \times I_x \subset \mathrm{Cl}\, G$ holds for every $x = 0.x_1 \cdots x_n \in \mathcal{D} \setminus \{1\}$. According to equation (28), for a generic point $(x, y) \in \{x\} \times I_x$ we have $y = 0.x_1 \cdots x_n v$, with $v \in \{0, 1\}^\infty$. Then, the point (x, y) can be arbitrarily approximated by a point $(x', R(x')) \in G$ with $x' = 0.x_1 \cdots x_n(00)^k w$, k sufficiently large and $w \in \{0, 1\}^\infty$ any binary word such that $\rho(w) = v$. This concludes the proof of the first part of the proposition.

To prove the second part of the proposition, we start by observing that the argument just exposed immediately entails that

$$I_x \subset \left[\liminf_{\xi \searrow x} R(\xi), \limsup_{\xi \searrow x} R(\xi) \right].$$

The opposite inclusion readily follows by equation (28). $\square$

Corollary 4.4.   *The map R is not Darboux from the right* (*in the sense of* [10]) *at any dyadic rational $x \in \mathcal{D} \setminus \{1\}$ and its graph G is totally disconnected.*
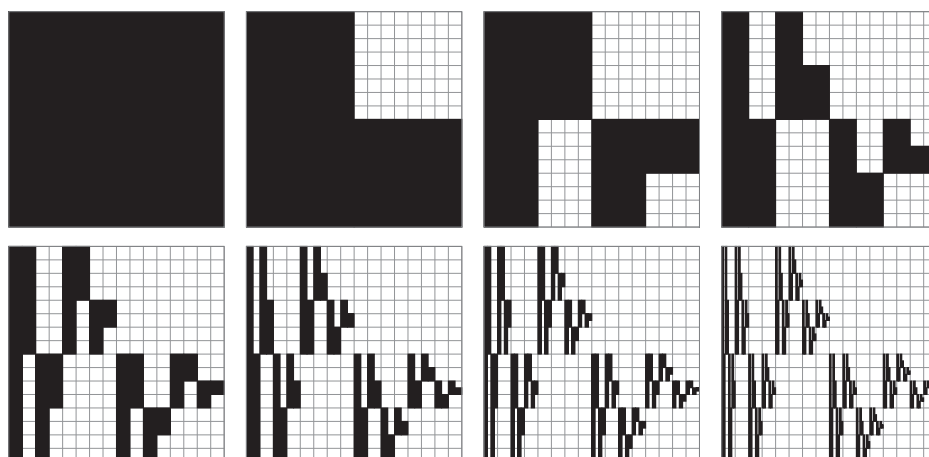
Figure 2. K_n for $n = 0, \cdots, 7$.

Proof. This is an immediate consequence of proposition 4.2. $\qquad\square$

Using the inclusion $G \subset K_\infty$, we can prove the next proposition.

Proposition 4.5. $\int_0^1 R(x)\mathrm{d}x = \frac{3}{7}$.

Proof. By remark 4.1, Area $K_n = \frac{3}{4}$ Area K_{n-1} for every $n > 0$. Then, taking into account that Area $K_0 = 1$, we get

$$\text{Area } K_n = \left(\frac{3}{4}\right)^n \tag{30}$$

for every $n \geqslant 0$. Since this vanishes for $n \to \infty$, we can approximate the integral of R by the integral A_n of the piecewise constant function R_n whose graph is given by the bottom edges of all the rectangles forming K_n. Now, $R_{2n+1} = R_{2n}$ for every $n \geqslant 1$, so we can write

$$\int_0^1 R(x)\mathrm{d}x = \lim_{n\to\infty} A_{2n}.$$

Looking at what happens inside each rectangle of K_{2n-2} when passing from R_{2n-2} to R_{2n} (compare K_0 and K_2 in figure 2), we have

$$A_{2n} - A_{2n-2} = \frac{3}{16} \text{ Area } K_{2n-2} = \frac{3}{16} \cdot \left(\frac{9}{16}\right)^{n-1}$$

Therefore, starting from $A_0 = 0$ we get

$$\lim_{n\to\infty} A_{2n} = \frac{3}{16}\sum_{n=0}^{\infty}\left(\frac{9}{16}\right)^n = \frac{3}{16} \cdot \frac{16}{7} = \frac{3}{7}.$$

$\qquad\square$

Finally, we can compute the Box dimension $\dim_B G$ and estimate the Hausdorff dimension $\dim_H G$ as follows.

Proposition 4.6. $1 \leqslant \dim_H G \leqslant \dim_B G = \dim_B K_\infty = \log_2 3.$

Proof.   The two inequalities derive from the projection of G onto $[0, 1]$ and from the general relation between the Hausdorff dimension and the box-dimension, respectively. The first equality comes from the invariance of the box-dimension under closure. The second equality follows from the fact that the boxes of the form $[k/2^n, (k+1)/2^n] \times [\ell/2^n, (\ell+1)/2^n]$ needed to cover K_∞ are exactly the ones contained in K_n, which are 3^n, as it is clear by equation (30). $\qquad\square$

5. Analytical properties

Let us start with some global properties of the map R with respect to the Lebesgue measure. We indicate by $\lambda(A)$ the standard Lebesgue measure of a measurable set $A \subset [0, 1]$.

In the following propositions we indicate by $\mathcal{N} \subset [0, 1]$ the subset of two-normal numbers, that is the numbers $x \in [0, 1]$ which admit a binary expansion where all the binary sequences of any given length $k \geqslant 1$ occur with the same asymptotic relative frequency $1/2^k$. In particular, $\mathcal{N}$ does not contain any rational number, and hence $\beta(x) = \beta'(x)$ for every $x \in \mathcal{N}$. We recall that $\mathcal{N}$ has full measure $\lambda(\mathcal{N}) = 1$, as proved in [2].

For any infinite binary word $b \in \mathcal{B}$ and any finite binary word $w \in \{0, 1\}^k$, we indicate by $f_n(w, b)$ the relative frequency of w among all the n subwords of $b_{|n+k-1}$ of length k, and by $f(w, b)$ the asymptotic relative frequency of w in b, if it exists. Namely, we put

$$f_n(w, b) = \frac{|\{i = 1, \ldots, n \mid w = b_i b_{i+1} \cdots b_{i+k-1}\}|}{n} \tag{31}$$

and

$$f(w, b) = \lim_{n \to \infty} f_n(w, b). \tag{32}$$

Then, to say that $x = 0.x_1 x_2 \cdots$ is normal means that $f(w, x_1 x_2 \cdots) = 1/2^k$ for every $w \in \{0, 1\}^k$ and for every $k > 1$. This is equivalent to the property that for every $k \geqslant 1$, if $(w_1, w_2, \ldots)$ is the decomposition of $x_1 x_2 \cdots$ into contiguous blocks of length k, then the asymptotic relative frequency of any $w \in \{0, 1\}^k$ in the sequence $(w_1, w_2, \ldots)$ is $1/2^k$ (see [2, 5, 7, 27, 31]).

Proposition 5.1.   *The map R is measurable but not bi-measurable. Moreover, the map S defined by equation* (22) *is a null measure section of R, being $S([0, 1]) \subset [0, 1]\backslash\mathcal{N}$.*

Proof.   The map R is almost everywhere continuous by proposition 3.1, hence it is measurable. Concerning the section S, for every $y \in [0, 1]$ the image $S(y)$ does not contain any pair of consecutive 0's in its binary expansion $\beta(S(y))$, which coincides with either $\sigma(\beta(y))$ or $\sigma(\beta'(y))$. Then, we have $S([0, 1]) \subset [0, 1]\backslash\mathcal{N}$, which implies that $\lambda(S([0, 1])) = 0$ by [2]. Therefore, every set $A \subset [0, 1]$ is the image $R(S(A))$ of the null measure set $S(A)$. In particular, this is true for any non-measurable subset of $[0, 1]$, like the Vitali set. Thus, R is not bi-measurable. $\qquad\square$

We remark that, in the light of the uncountability of all the fibers of R discussed in Section 3, part of proposition 5.1 could be also obtained as a direct consequence of a general theorem by Purves [32] stating that, if f is a bi-measurable map from a standard Borel space X to a Polish space Y, then at most countably many fibers of f can be uncountable.

Proposition 5.2. *The map R is a singular Borel function, being $R(\mathcal{N}) \subset [0,1] \backslash \mathcal{N}$.*

Proof. It is enough to prove the last inclusion, since it implies that $\lambda(R^{-1}([0,1]\backslash\mathcal{N})) = 1$ with $\lambda([0,1]\backslash\mathcal{N}) = 0$, and hence the singularity of R. So, given $x = 0.x_1x_2\cdots \in \mathcal{N}$ and $R(x) = y = 0.y_1y_2\ldots$, with $y_1y_2\cdots = \rho(x_1x_2\cdots)$, we have to prove that y is not normal. Notice that $x_1x_2\cdots = \beta(x) \in \mathcal{C}$ and $y_1y_2\cdots \in \mathcal{B}$ are both infinite binary words (the latter coincides with either $\beta(y)$ or $\beta'(y)$). Here, we think of the normality of x in terms of asymptotic equi-distribution of contiguous blocks in $x_1x_2\cdots$ of every given length, as discussed above.

By considering the blocks of length 2 in the word $x_1x_2\ldots$, and taking into account the substitution rule τ defined in (11) and the asymptotic equi-distribution of 01 and 10, guaranteed by the normality of x, we can immediately conclude that $R(x)$ is simply normal, meaning that 0 and 1 have the same asymptotic relative frequency $1/2$ in the word $y_1y_2\cdots$.

By arguing instead on the blocks of length 8 in the word $x_1x_2\cdots$, we will prove that

$$\lim_{n\to\infty}(f_n(01, y_1y_2\cdots) + f_n(10, y_1y_2\cdots)) > \lim_{n\to\infty}(f_n(00, y_1y_2\cdots)$$
$$+ f_n(11, y_1y_2\cdots)), \tag{33}$$

assuming that both the limits exist, otherwise $y = R(x)$ cannot be normal and we are done. This prevents the word $y_1y_2\cdots$ to satisfy the asymptotic equi-distribution of the binary subwords of length 2, and hence implies once again that $y = R(x)$ is not normal.

To prove (33), let $\overline{n} \geqslant 1$ be arbitrarily fixed. Since $f(0, x_1x_1\cdots) = f(1, x_1x_2\cdots) = 1/2$, for every $\delta > 0$ there exists $m \geqslant 1$ such that the length n of the word $\rho(x_1x_2\cdots x_{8m})$ satisfies the following properties

$$\frac{n}{8m} = \frac{|\rho(x_1x_2\cdots x_{8m})|}{8m} = \frac{|x_1x_2\cdots x_{8m}|_1}{8m} \in (1/2 - \delta, 1/2 + \delta) \quad \text{and} \quad n > \overline{n}$$

We want to show that for a suitable choice of m the following inequality holds

$$f_n(01, y_1y_2\cdots) + f_n(10, y_1y_2\cdots) > f_n(00, y_1y_2\cdots) + f_n(11, y_1y_2\cdots) + c,$$
$$\tag{34}$$

where $c > 0$ is a constant independent of n, which gives (33).

Let $(w_1, w_2, \ldots, w_m)$ be the decomposition of the word $x_1x_2\cdots x_{8m}$ in blocks of length 8, and $(v_1, v_2, \ldots, v_m)$ be the corresponding decomposition of the word $y_1y_2\cdots y_n$ in blocks of variable length $\leqslant 8$, with $v_i = \rho(w_i)$ for every $i = 1, \ldots, k$.

Based on the normality of x, we can assume m sufficiently large in such a way that the relative frequency of any binary word w of length 8 in the sequence $(w_1, w_2, \ldots, w_m)$ belongs to the interval $(1/2^8 - \delta, 1/2^8 + \delta)$. Then, we consider the pairs of consecutive digits in $y_1y_2\cdots y_n$, separating those occurring inside the blocks from those formed by the last digit of a block and the first digit of the next one (disregarding the empty blocks).

The pairs of the latter type are less than m. Concerning the pairs of the former type, direct inspection shows that in the images of all binary words of length 8 under ρ, the pairs 01 and 10 occur 541 times, while the pairs 00 and 11 occur 228 times. Therefore, inside the blocks $(v_1, v_2, \ldots, v_m)$ the number of occurrences of the pairs 01 and 10 is at least $541(1/2^8 - \delta)m$, while the number of occurrences of the pairs 00 and 11 at most $228(1/2^8 + \delta)m$. Putting all

together we have

$$f_n(01, y_1y_2\cdots) + f_n(10, y_1y_2\cdots) \geqslant \frac{541(1/2^8 - \delta)m}{n} \geqslant \frac{541(1/2^8 - \delta)}{8(1/2 + \delta)}, \tag{35}$$

$$f_n(00, y_1y_2\cdots) + f_n(11, y_1y_2\cdots) \leqslant \frac{228(1/2^8 + \delta)m + m}{n}$$

$$\leqslant \frac{228(1/2^8 + \delta) + 1}{8(1/2 - \delta)}. \tag{36}$$

At this point, it is enough to observe that (35) is greater than (36) for δ sufficiently small. $\qquad\square$

As a consequence of proposition 5.2, the map R does not preserve the Lebesgue measure. The Lebesgue measure of the inverse images of dyadic cylinders can be directly computed as follows.

According to equation (20), for every $y \in \{0, 1\}^n$ and $n \geqslant 1$, we have

$$R^{-1}(\xi([y]))\backslash\{0\} = \begin{cases} \cup_{b\in\mathcal{B}}\xi(\langle\!\langle\sigma(yb)\rangle\!\rangle) & \text{if } y = 1^n \\ \cup_{b\in\mathcal{B}}\xi(\langle\!\langle\sigma(yb)\rangle\!\rangle) \cup \xi(\langle\!\langle\sigma(y')\rangle\!\rangle) & \text{otherwise'} \end{cases} \tag{37}$$

where $y' = \beta(\xi(y) + 1/2^n)$.

Then, by putting $|\sigma(y)| = m$ and recalling that $|\sigma(y)|_1 = |y| = n$, we get

$$\lambda(\cup_{b\in\mathcal{B}}\xi(\langle\!\langle\sigma(yb)\rangle\!\rangle)) = \lambda(\{\langle a\rangle^{\sigma(yb)} \mid a \in \mathcal{A}, b \in \mathcal{B}\}) = \lambda(\{\langle a\rangle^{\sigma(y)}b \mid a \in \mathbb{N}^n, b \in \mathcal{B}\})$$

$$= \lambda(\cup_{a\in\mathbb{N}^n}[\langle a\rangle^{\sigma(y)}]) = \sum_{a\in\mathbb{N}^n}\lambda([\langle a\rangle^{\sigma(y)}]) = \sum_{a\in\mathbb{N}^n}1/2^{|\langle a\rangle^{\sigma(y)}|}$$

$$= \sum_{a\in\mathbb{N}^n}1/2^{(2a_1+\cdots+2a_n+m)} = \left(\sum_{a_1\geqslant0}1/2^{2a_1}\right)\cdots\left(\sum_{a_n\geqslant0}1/2^{2a_n}\right)/2^m$$

$$= (4/3)^n/2^m = 2^{2n-m}/3^n \leqslant (2/3)^n.$$

Moreover, by taking into account the inclusions

$$\xi(\langle\!\langle\sigma(y')\rangle\!\rangle) \subset \cup_{b\in\mathcal{B}}\xi(\langle\!\langle\sigma(y'_{|k}b)\rangle\!\rangle),$$

which hold for every $k \geqslant 0$, and by applying the above formula for the measure of the right-side terms, we obtain

$$\lambda(\xi(\langle\!\langle\sigma(y')\rangle\!\rangle)) = 0.$$

Therefore, based on (37), we have

$$\lambda(R^{-1}(\xi([y]))) = 2^{2n-|\sigma(y)|}/3^n \leqslant (2/3)^n \tag{38}$$

for every $y \in \{0, 1\}^n$ and $n \geqslant 1$.

We remark that, by generalising the above computation of the Lebesgue measure of the inverse images of dyadic cylinders, one could show that actually the map R cannot preserve any product measure μ_p with $0 < p < 1$, determined by $\mu_p(\xi([x])) = p^m(1-p)^{n-m}$ where $m = |x|_1$, for every $x \in \{0, 1\}^n$ and $n \geqslant 1$.

Moreover, proposition 5.2 also entails that no R-invariant probability measure can be absolutely continuous with respect to Lebesgue measure. Indeed, if there were such a measure

ν, then we would have $\nu([0,1]\setminus\mathcal{N}) = 0$, and hence $\nu(R^{-1}([0,1]\setminus\mathcal{N})) = 0$. Thus, the density function of ν would vanish almost everywhere on a set containing $\mathcal{N}$.

Now we pass to the topological and metric structure of the fibers of R, which we recall to be all uncountable, as a consequence of equation (20) in section 3. In particular, we will compute the Hausdorff dimension $\dim_H R^{-1}(y)$ of the fibers of the rational numbers $y \in \mathcal{Q}$ and estimate that of all the other fibers. Let us start with some preliminary results.

Proposition 5.3. *For every $y \in [0,1]$, the closure $\operatorname{Cl} R^{-1}(y)$ of the fiber $R^{-1}(y)$ is a null measure Cantor set in $[0,1]$, with a countable remainder $\operatorname{Cl} R^{-1}(y)\setminus R^{-1}(y) \subset \mathcal{D}$.*

Proof. First we prove the inclusion

$$\operatorname{Cl}\xi(\langle\!\langle b\rangle\!\rangle) \subset \xi(\langle\!\langle b\rangle\!\rangle) \cup \mathcal{D} \tag{39}$$

for every $b \in \mathcal{C}$, which implies that $\operatorname{Cl} R^{-1}(y)\setminus R^{-1}(y) \subset \mathcal{D}$ for every $y \in [0,1]$ by (20).

Given $x \in \operatorname{Cl}\xi(\langle\!\langle b\rangle\!\rangle)$, let $(x_n)_{n\geqslant 1} \subset \xi(\langle\!\langle b\rangle\!\rangle)$ be a sequence such that $\lim_{n\to\infty} x_n = x$. Then, by the definition of $\langle\!\langle b\rangle\!\rangle$ in (21), there exists a sequence $(a_n)_{n\geqslant 1} \subset \mathcal{A}$ such that $x_n = \xi(\langle a_n\rangle^b)$, and hence $\langle a_n\rangle^b = \beta(x_n)$, for every $n \geqslant 1$. If $x \in \mathcal{D}$ we are done, otherwise the continuity of β at x (see section 1) implies that $\lim_{n\to\infty}\langle a_n\rangle^b = \beta(x) \in \mathcal{C}$. If the sequence of the ith components $((a_n)_i)_{n\geqslant 1}$ is bounded for every $i \geqslant 1$, then by compactness we can replace the sequence $(a_n)_{n\geqslant 1}$ by a subsequence converging to $a \in \mathcal{A}$, and recalling that η is a homeomorphism, we have

$$\lim_{n\to\infty}\langle a_n\rangle^b = \lim_{n\to\infty}\eta(2a_n + \eta^{-1}(b)) = \eta(\lim_{n\to\infty} 2a_n + \eta^{-1}(b))$$
$$= \eta(2a + \eta^{-1}(b)) = \langle a\rangle^b,$$

and hence $x = \xi(\langle a\rangle^b) \in \xi(\langle\!\langle b\rangle\!\rangle)$. If instead some sequence $((a_n)_i)_{n\geqslant 1}$ is unbounded, let m be the minimum index $i \geqslant 1$ for which this happens. Replacing the sequence $(a_n)_{n\geqslant 1}$ by a suitable subsequence such that $(a_n)_{|m-1} = c$ for some $c \in \{0,1\}^{m-1}$ and every $n \geqslant 1$, we have

$$\beta(x) = \lim_{n\to\infty}\langle a_n\rangle^b = \langle c\rangle^{b_{|m-1}} 0^\infty,$$

which is absurd, being $\beta(x) \in \mathcal{C}$. So this case cannot occur, and we finished the proof of (39).

At this point, for any $y \in (0,1]$, taking into account that $\{y\} = \cap_{n\geqslant 1}\xi([\beta(y)_{|n}])$, and hence $R^{-1}(y) = \cap_{n\geqslant 1}R^{-1}(\xi([\beta(y)_{|n}]))$, we obtain $\lambda(\operatorname{Cl} R^{-1}(y)) = \lambda(R^{-1}(y)) = 0$ by (38). Similarly, we can obtain $\lambda(\operatorname{Cl} R^{-1}(0)) = \lambda(R^{-1}(0)) = 0$ starting from $\{0\} = \cap_{n\geqslant 1}\xi([0^n])$.

We are left to prove that $\operatorname{Cl} R^{-1}(y)$ is a Cantor set for every $y \in [0,1]$. To this end, it is enough to show that $\operatorname{Cl} R^{-1}(y)$ is topologically 0-dimensional and perfect (see [22, chapter four, section 45, paragraph II]). The 0-dimensionality immediately follows from the fact that $\operatorname{Cl} R^{-1}(y)$ cannot contain any interval, having a null measure. For the perfectness, we observe that every $x \in \operatorname{Cl} R^{-1}(y)\setminus R^{-1}(y)$ cannot be isolated. On the other hand, if $x \in R^{-1}(y)\setminus\{0\}$, then $x = \xi(\eta(c))$ for some $c \in \mathcal{C}$ (see (20)), and we have $x = \lim_{n\to\infty} x_n$ with the sequence $(x_n = \xi(\eta(c_n)))_{n\geqslant 1} \subset R^{-1}(y)$ defined by $(c_n)_i = c_i + 2\delta_{i,n}$ for any $i,n \geqslant 1$. Finally, 0 is a limit point for every fiber $R^{-1}(y)$, and in particular for $y = 2/3$. In fact, $0 = \lim_{n\to\infty} x_n$ where $(x_n)_{n\geqslant 1} \subset R^{-1}(y)$ is defined like above, starting with any $c \in \mathcal{C}$ such that $\xi(\eta(c)) \in R^{-1}(y)$ and putting $(c_n)_i = c_i + 2n\delta_{i,1}$ for any $i,n \geqslant 1$. $\qquad\square$

In order to compare different fibers of R, we introduce a partial order relation between infinite binary sequences $b, c \in \mathcal{C}$ as follows

$$b \succcurlyeq c \iff \text{there exists } m \geqslant 0 \quad \text{such that } |c_{|i}|_1 \leqslant |b_{|i}|_1 + m \quad \text{for every } i \geqslant 1.$$
(40)

We also introduce the induced equivalence relation:

$$b \approx c \iff b \succcurlyeq c \quad \text{and} \quad c \succcurlyeq b.$$
(41)

The above relations provide a kind of uniform control on the distribution of the 1's in the words. In particular, $b \succcurlyeq c$ and $b \approx c$ imply analogous relations $f(1, b) \geqslant f(1, c)$ and $f(1, b) = f(1, c)$ respectively, between the asymptotic relative frequencies of 1's, if the latter exist. But of course the opposite implications are false.

The next two lemmas provide our technical tool for the comparison of fibers. Here, we recall that $\mathcal{S}$ stands for the set of all infinite binary words without any pair of consecutive 0's.

Lemma 5.4. *For every $b, c \in \mathcal{C}$, the natural bijection $\phi_{b,c} : \langle\!\langle b \rangle\!\rangle \to \langle\!\langle c \rangle\!\rangle$ defined by*

$$\phi_{b,c}(\langle a \rangle^b) = \langle a \rangle^c$$

with $a \in \mathcal{A}$, is a homeomorphism. Moreover, $\phi_{b,c}$ is a Lipschitz map if $b \succcurlyeq c$ and $c \in \mathcal{S}$, and hence $\phi_{b,c}$ is a bi-Lipschitz map if $b \approx c$ and $b, c \in \mathcal{S}$.

Proof. Given any $b \in \mathcal{C}$, we consider the natural bijection $\phi_b : \mathcal{A} \to \langle\!\langle b \rangle\!\rangle$ defined by $\phi_b(a) = \langle a \rangle^b = \eta(2a + \eta^{-1}(b))$. This is a one-Lipschitz homeomorphism, being the composition of the isometric map $\mathcal{A} \to \mathcal{A}$ which sends a to $2a + \eta^{-1}b$ with the one-Lipschitz homeomorphism $\eta : \mathcal{A} \to \mathcal{C}$. Moreover, for every $a, a' \in \mathcal{A}$ such that $a_{|n} = a'_{|n}$ and $a_{n+1} \neq a'_{n+1}$ we put

$$h(a, a') = a_1 + \cdots + a_n + \min\{a_{n+1}, a'_{n+1}\},$$

and observe that the maximum index i for which $\langle a \rangle^b_{|i} = \langle a' \rangle^b_{|i}$ is given by $2h(a, a') + k_n^b$, where k_n^b denotes the index of the nth occurrence of 1 in b.

Then, for every $b, c \in \mathcal{C}$ the identity $\phi_{b,c} = \phi_c \circ \phi_b^{-1}$ implies that $\phi_{b,c}$ is a homeomorphism. Furthermore, if $b \succcurlyeq c$ and $c \in \mathcal{S}$ we can see that $\phi_{b,c}$ is a Lipschitz map as follows.

Let $m \geqslant 1$ such that $|c_{|i}|_1 \leqslant |b_{|i}|_1 + m$ for every $i \geqslant 1$. In particular, we have

$$|c_{|k_i^b}|_1 \leqslant |b_{|k_i^b}|_1 + m = i + m,$$

and therefore

$$k_i^b + 1 \leqslant k_{i+m+1}^c \leqslant k_i^c + 2m + 2$$
(42)

for every $i \geqslant 1$, where k_i^b denotes the index of the ith occurrence of 1 in b as above, and the last inequality derives from the fact that $c \in \mathcal{S}$ does not contain any pair of consecutive 0's.

Given any two different sequences $a, a' \in \mathcal{A}$, let n be the maximum index i such that $a_{|i} = a'_{|i}$. Then, by the above observation, we have

$$d(\langle a \rangle^b, \langle a' \rangle^b) = 1/2^{2h(a,a')+k_n^b+1} \quad \text{and} \quad d(\langle a \rangle^c, \langle a' \rangle^c) = 1/2^{2h(a,a')+k_n^c+1},$$

from which by (42) we obtain

$$\frac{d(\phi_{b,c}(\langle a \rangle^b), \phi_{b,c}(\langle a' \rangle^b))}{d(\langle a \rangle^b, \langle a' \rangle^b)} \leqslant 2^{2m+1}.$$

So, we can conclude that $\phi_{b,c}$ is a Lipschitz map with Lipschitz constant 2^{2m+1}. $\qquad \square$

Lemma 5.5. *For every $b, c \in \mathcal{C}$, the bijection $\psi_{b,c} = \xi \circ \phi_{b,c} \circ \beta|_{\xi(\langle\!\langle b \rangle\!\rangle)} : \xi(\langle\!\langle b \rangle\!\rangle) \to \xi(\langle\!\langle c \rangle\!\rangle)$ given by*

$$\psi_{b,c}\left(\xi(\langle a \rangle^b)\right) = \xi(\langle a \rangle^c)$$

with $a \in \mathcal{A}$, is a Lipschitz map if $b \succcurlyeq c$ and $b \in \mathcal{S}$, and hence $\psi_{b,c}$ is a bi-Lipschitz homeomorphism if $b \approx c$ and $b, c \in \mathcal{S}$.

Proof. In the light of the previous lemma, recalling that ξ is two-Lipschitz, it suffices to show that $\beta|_{\xi(\langle\!\langle b \rangle\!\rangle)}$ is a Lipschitz map for every $b \in \mathcal{S}$.

Take any two different real numbers $x = \xi(\langle a \rangle^b)$ and $x' = \xi(\langle a' \rangle^b)$ in $\xi(\langle\!\langle b \rangle\!\rangle)$, with $a_{|n} = a'_{|n}$ and $a_{n+1} \neq a'_{n+1}$. Without loss of generality, we assume that $a_{n+1} < a'_{n+1}$. Like in the proof of the previous lemma, we have $\langle a \rangle^b_{|2h(a,a')+k^b_n} = \langle a' \rangle^b_{|2h(a,a')+k^b_n}$. Denoting by w this word, we can write $\langle a \rangle^b = w1c$ and $\langle a' \rangle^b = w00c'$, and hence $x = 0.w1c$ and $x' = 0.w00c'$, for suitable $c, c' \in \mathcal{C}$. Then, we get

$$|x - x'| = 0.x1c - 0.x00c' \geqslant 0.x1 - 0.x01 = 0.0^{|w|+1}1 = 1/2^{|w|+2}$$

while

$$d(\beta(x), \beta(x')) = d(\langle a \rangle^b, \langle a' \rangle^b) = 1/2^{|w|+1}.$$

These inequalities immediately imply that $\beta|_{\xi(\langle\!\langle b \rangle\!\rangle)}$ is a two-Lipschitz map. $\qquad\square$

Now we focus on the 'rational fibers', that is the fibers $R^{-1}(y)$ with $y \in \mathcal{Q}$.

Proposition 5.6. *If $y \in \mathcal{Q}$ then $\xi(\sigma(\beta(y))), \xi(\sigma(\beta'(y))) \in \mathcal{Q}$ when defined, and so $S(y) \in \mathcal{Q}$.*

Proof. Since any binary expansion of a rational number is periodic, both $\beta(y)$ and $\beta'(y)$ have the form wp^∞, where w and p are finite binary of length $n \geqslant 0$ and $\ell \geqslant 1$, respectively. Then, recalling that σ can be interpreted as a true substitution rule applied to the first differences, $\sigma(\beta(y))$ and $\sigma(\beta'(y))$ have both the corresponding form vq^∞, with

$$v = \begin{cases} \sigma(w) & \text{if } w_n = p_\ell \\ \sigma(wp) & \text{if } w_n \neq p_\ell \end{cases} \quad \text{and} \quad q = \begin{cases} \sigma(p) & \text{if } p_\ell = 1 \\ \sigma'(p) & \text{if } p_\ell = 0 \end{cases}, \tag{43}$$

where we assume $w_n = 1$ for $n = 0$ and indicate by $\sigma'(p)$ the word $\sigma(p)$ with the first digit complemented. In all cases, we have a periodic binary sequence with a period q of length n, and by applying ξ we get a rational number in $\mathcal{Q}$. $\qquad\square$

The next proposition expresses the Hausdorff dimension $\dim_H R^{-1}(y)$ for $y \in \mathcal{Q}$ in terms of the density d of the 1's in the period q of $\sigma(\beta(y))$ or $\sigma(\beta'(y))$ as given by (43), that is

$$d = |q|_1 / |q|.$$

We recall that $\beta(y)$ and $\beta'(y)$ coincide for $y \in [0, 1] \setminus \mathcal{D}$, while only one of them is defined for $y = 0, 1$. On the other hand, for $y \in \mathcal{D} \setminus \{0, 1\}$ the period of $\beta(y)$ is 1 and the period of $\beta'(y)$ is 0, and hence in both cases $q = 01$ or $q = 10$ and then $d = 1/2$.

We also observe that actually d does not depend on the specific choice for the period q in $\sigma(\beta(y))$ or $\sigma(\beta'(y))$, in that d does not change if we choose a different starting digit for the

period of the word or consider as a period any multiple of a minimal one. In fact, d coincides with the asymptotic relative frequency $f(1, \sigma(\beta(y)))$ or $f(1, \sigma(\beta'(y)))$.

Proposition 5.7. *For every $y \in \mathcal{Q}$ we have $\dim_H R^{-1}(y) = -\log_2 t$, with t the unique real number in $(0, 1)$ verifying the equation*

$$t^2 + t^{1/d} = 1,$$

where $d = |q|_1/|q| \in (0, 1]$ is the density of the 1's in the period q of $\sigma(\beta(y))$ or $\sigma(\beta'(y))$.

Proof. According to (20) and (43), and recalling that $\dim_H(A \cup B) = \max\{\dim_H A, \dim_H B\}$ for every $A, B \subset [0, 1]$, we only need to compute $\dim_H \xi(\langle\!\langle vq^\infty \rangle\!\rangle)$ with $vq^\infty \in \mathcal{S}$. Moreover, thanks to lemma 5.5 and the bi-Lipschitz invariance of the Hausdorff dimension, we can limit ourselves to consider the case when $v = \epsilon$, since $vq^\infty \approx q^\infty$ and both vq^∞ and q^∞ belong to $\mathcal{S}$. In addition, by applying once again lemma 5.5, we can assume that the last digit of q is 1 without changing the $\approx$class. Indeed, q must contain some 1 and thus it can be written as $q = q'1q''$ for some (possibly empty) binary words q' and q'', whence $q^\infty = (q'1q'')^\infty = q'1(q''q'1)^\infty \approx (q''q'1)^\infty$.

Given $q = q_1 q_2 \cdots q_n \in \{0, 1\}^n$ with $q_n = 1$, we put $m = |q|_1$ and consider the family of contractive similarities

$$\mathcal{T} = \{T_a : \mathbb{R} \to \mathbb{R}\}_{a \in \mathbb{N}^m}$$

defined by

$$T_a(x) = x/2^{2(a_1 + \cdots + a_m) + n} + x_a,$$

where $x_a = 0.\langle a \rangle^q$, for every $a \in \mathbb{N}^m$. Such family $\mathcal{T}$ is clearly relatively compact in the space of all the contractive similarities of $\mathbb{R}$ with the topology of the uniform convergence over bounded sets.

In terms of binary expressions, since $2(a_1 + \cdots + a_m) + n = |\langle a \rangle^q|$, we have

$$T_a(0.x_1 x_2 \cdots) = 0.\langle a \rangle^q x_1 x_2 \cdots. \tag{44}$$

This allows us to derive in a straightforward way that

$$\xi(\langle\!\langle q^\infty \rangle\!\rangle) = \cup_{a \in \mathbb{N}^m} T_a(\xi(\langle\!\langle q^\infty \rangle\!\rangle)).$$

Indeed, the equality $\langle a \rangle^q q^\infty = \langle a0^\infty \rangle^{q^\infty}$ gives the inclusion $T_a(\xi(\langle\!\langle q^\infty \rangle\!\rangle)) \subset \xi(\langle\!\langle q^\infty \rangle\!\rangle)$ for every $a \in \mathbb{N}^m$. On the other hand, for every $x = \xi(\langle a \rangle^{q^\infty}) \in \xi(\langle\!\langle q^\infty \rangle\!\rangle)$ with $a \in \mathcal{A}$, we have $x = \xi(\langle a_1 a_2 \cdots a_m \rangle^q \langle a_{m+1} a_{m+2} \cdots \rangle^{q^\infty}) \in T_{a_{|m}}(\xi(\langle\!\langle q^\infty \rangle\!\rangle))$.

From (44), we also see that $T_a((0, 1)) \subset (0, 1)$ for every $a \in \mathbb{N}^m$, and $T_a((0, 1)) \cap T_{a'}((0, 1)) = \emptyset$ for every $a, a' \in \mathbb{N}^m$ with $a \neq a'$, that is the family $\mathcal{T}$ satisfies the so called 'open set condition'.

Then, we are in position to apply theorem 2.2 of [26] (see also [16, theorem 3.11]), in order to conclude that

$$\dim_H(\xi(\langle\!\langle b \rangle\!\rangle)) = \inf \left\{ s > 0 \,\middle|\, \sum_{a \in \mathbb{N}^m} r_a^s \leqslant 1 \right\}, \tag{45}$$

where $r_a = 1/2^{2(a_1 + \cdots + a_m) + n}$ denotes the similarity ratio of T_a, for every $a \in \mathbb{N}^m$.

Now, recalling that $d = m/n$, we have

$$\sum_{a \in \mathbb{N}^m} r_a^s = \sum_{a \in \mathbb{N}^m} 1/2^{(2(a_1 + \cdots + a_m) + n)s}$$

$$= \left(\sum_{a_1 \geqslant 0} 1/2^{2a_1 s} \right) \cdots \left(\sum_{a_m \geqslant 0} 1/2^{2a_m s} \right) /2^{ns}$$

$$= 2^{-ns}/(1 - 1/2^{2s})^m = (2^{s/d}(1 - 2^{-2s}))^{-m}.$$

Therefore, from (45) we deduce that $\dim_H(\xi(\langle\!\langle b \rangle\!\rangle))$ is given by the unique $s > 0$ where the monotonic real function $f(s) = 2^{s/d}(1 - 2^{-2s})$ assumes the value 1. Finally, the equation $f(s) = 1$ becomes $t^2 + t^{1/d} = 1$ as in the statement, by putting $s = -\log_2 t$.   □

We remark that actually the open set condition as formulated in the proof above also implies that the rational fibers of R have positive Hausdorff measure in the relative Hausdorff dimension (see [26, theorem 2.2]).

To conclude this section, we want to provide a common lower and upper bound for the Hausdorff dimension of all the fiber of R. To this end, we first apply proposition 5.7 to compute the Hausdorff dimension of the two 'extremal' rational fibers $R^{-1}(1)$ and $R^{-1}(1/3)$. We have

$$\beta(1) = 1^\infty \qquad \Rightarrow \sigma(\beta(1)) = (01)^\infty \Rightarrow d = 1/2 \Rightarrow \dim_H R^{-1}(1) = 1/2,$$
$$\beta(1/3) = (01)^\infty \Rightarrow \sigma(\beta(1/3)) = 1^\infty \Rightarrow d = 1 \quad \Rightarrow \dim_H R^{-1}(1/3) = \log_2 \varphi,$$

where $\varphi = (1 + \sqrt{5})/2$ is the golden ratio.

Proposition 5.8.   *For every $y \in [0, 1]$, the following holds*

$$1/2 \leqslant \dim_H R^{-1}(y) \leqslant \log_2 \varphi.$$

Proof.   Thanks to (20) and the formula $\dim_H(A \cup B) = \max\{\dim_H A, \dim_H B\}$ holding for every $A, B \subset [0, 1]$, it suffices to prove that $1/2 \leqslant \dim_H \xi(\langle\!\langle b \rangle\!\rangle) \leqslant \log_2 \varphi$ for every $b \in \mathcal{S}$. This immediately follows from lemma 5.5, taking into account the obvious relation $1^\infty \succcurlyeq b \succcurlyeq (01)^\infty$ and the fact that the Hausdorff dimension cannot increase under Lipschitz maps.   □

6. Dynamical properties

For the sake of convenience, we extend the notation ρ_v introduced in (9) and (10) by defining the maps

$$\rho_v^n : \{0, 1\}^\infty \to \{0, 1\}^\infty$$

for every finite binary word v and every $n \geqslant 1$ as follows.

We start by setting

$$\rho_v^1(w) = \rho_v(w), \tag{46}$$

with $w \in \{0, 1\}^\infty$, and then for $n > 1$ we inductively define

$$\rho_v^n = \rho_{\rho^{n-1}(v)} \circ \rho_v^{n-1}. \tag{47}$$

We emphasise that, in general ρ_v^n does not coincide with $(\rho_v)^n$, the nth iteration of the map ρ_v, when $v \neq \epsilon$ (for $v = \epsilon$ we have $\rho_\epsilon = \rho$, and hence $\rho_\epsilon^n = \rho^n = (\rho_\epsilon)^n$ for every $n \geqslant 1$). Indeed, by expanding the inductive definition we obtain

$$\rho_v^n = \rho_{\rho^{n-1}(v)} \circ \rho_{\rho^{n-2}(v)} \circ \cdots \circ \rho_{\rho(v)} \circ \rho_v, \tag{48}$$

which is a composition of n maps, each equal to ρ or $\widetilde{\rho}$. Namely, the ith map in the composition is ρ or $\widetilde{\rho}$ depending on the parity of the length $|\rho^{n-i}(v)|$ of the image of v under the $(n-i)$th power of ρ. Nevertheless, a straightforward induction on $n \geqslant 1$ based on (10) gives the identity

$$\rho^n(vw) = \rho^n(v)\rho_v^n(w), \tag{49}$$

where only the first two occurrences of n denote iterations of maps.

As an consequence of the surjectivity of ρ, and hence of $\widetilde{\rho}$, also all the maps ρ_v^n are surjective. In fact, as it happens for ρ and $\widetilde{\rho}$, the inverse image $(\rho_v^n)^{-1}(w)$ is countably infinite if w is a finite non-empty word, while it is uncountable if w is an infinite word, that is $w \in \mathcal{B}$.

On the other hand, contrary to what happens for ρ and $\widetilde{\rho}$, the fact that $w \in \mathcal{C}$ is not enough to guarantee that $\rho_v^n(w) \in \mathcal{B}$ when $n > 1$. For example

$$\rho_{110}^2((01)^\infty) = (\rho_{\rho(110)} \circ \rho_{110})((01)^\infty) = (\rho_{01} \circ \rho_{110})((01)^\infty)$$
$$= \rho(\widetilde{\rho}((01)^\infty)) = \rho(0^\infty) = \epsilon.$$

The next two lemmas constitute fundamental technical tools used for establishing the dynamical properties of the map R.

Lemma 6.1. *For any sequence $(w_n)_{n\geqslant 0}$ of non-empty finite binary words and any sequence $(k_n)_{n\geqslant 0}$ of natural numbers with $k_0 = 0$ and $k_n \geqslant k_{n-1} + n_\epsilon(w_{n-1})$ for every $n \geqslant 1$, there is an uncountable set $B \subset [w_0] \cap \mathcal{C} \subset \mathcal{B}$ such that $\rho^{k_n}(b) \in [w_n]$ for every $b \in B$ and every $n \geqslant 1$.*

Proof. For any sequences $(w_n)_{n\geqslant 0}$ and $(k_n)_{n\geqslant 0}$ as above, we construct the generic element of B as an infinite concatenation $b = v_0 v_1 \cdots \in \mathcal{C}$, where the v_n's are non-empty finite binary words inductively defined as follows. We start with $v_0 = w_0$. Then, given $v_0, v_1, \cdots, v_{n-1}$ with $n > 0$, we put $u_n = v_0 v_1 \cdots v_{n-1}$ and we choose v_n to be any of the infinitely many (non-empty) words in $(\rho_{u_n}^{k_n})^{-1}(w_n)$ ending with 1. We observe that, $\rho^{k_n}(u_n) = \epsilon$ for every $n \geqslant 1$, whatever the choice of the v_n's. In fact, $\rho^{k_1}(u_1) = \rho^{k_1}(v_0) = \epsilon$ since $k_1 \geqslant n_\epsilon(v_0)$, while for $n > 1$, taking into account that $k_n - k_{n-1} \geqslant n_\epsilon(w_{n-1})$, we have by induction

$$\rho^{k_n}(u_n) = \rho^{k_n - k_{n-1}}(\rho^{k_{n-1}}(u_{n-1}v_{n-1}))$$
$$= \rho^{k_n - k_{n-1}}(\rho^{k_{n-1}}(u_{n-1})\rho_{u_{n-1}}^{k_{n-1}}(v_{n-1}))$$
$$= \rho^{k_n - k_{n-1}}(w_{n-1})) = \epsilon.$$

The condition that each v_n ends with 1 prevents different choices for the v_n's to produce the same final word b, as it can be easily deduced from the fact that $|v_n|_1 = |w_n|$ does not depend on the specific choice of v_n. In this way, as the result of infinitely many infinite choices, one for each $n > 0$, we get an uncountable set B of words $b = v_0 v_1 \cdots \in [w_0] \cap \mathcal{C}$ such that

$$\rho^{k_n}(b) = \rho^{k_n}(u_n v_n v_{n+1} \cdots) = \rho^{k_n}(u_n)\rho_{u_n}^{k_n}(v_n)\rho_{u_n v_n}^{k_n}(v_{n+1} \cdots)$$
$$= w_n \rho_{u_n v_n}^{k_n}(v_{n+1} \cdots) \in [w_n]$$

for every $n \geqslant 1$, where the second equality can be obtained by two applications of (49), while the third one immediately follows from $\rho^{k_n}(u_n) = \epsilon$ and $\rho_{u_n}^{k_n}(v_n) = w_n$. $\square$

Lemma 6.2.   *For any sequence $(w_n)_{n \geqslant 0}$ of non-empty finite binary words there is an uncountable set $B \subset [w_0] \cap \mathcal{C} \subset \mathcal{B}$ with the property that for every $b \in B$ and every $n, m \geqslant 0$, there exists $k \geqslant m$ such that $\rho^k(b) \in [w_n] \subset \mathcal{B}$.*

Proof.   Possibly by replacing $(w_n)_{n \geqslant 0}$ with the concatenation of all its finite initial subsequences, we can assume that each word w_n appears infinitely many times in the sequence. Under this assumptions, it is enough to prove the existence of an uncountable subset $B \subset [w_0] \cap \mathcal{C}$ with the weaker property that for every $b \in B$ and $n \geqslant 1$ there is $k \geqslant 1$ such that $\rho^k(b) \in [w_n]$. The existence of such a set B is guaranteed by the previous lemma. $\square$

We now proceed discussing some asymptotic properties of the orbits of the map R. First of all, we show that the set of rationals $\mathcal{Q} = \mathbb{Q} \cap [0, 1]$ is R-invariant and establish the asymptotic behaviour of the restriction of $R|_{\mathcal{Q}} : \mathcal{Q} \to \mathcal{Q}$.

Proposition 6.3.   *The set $\mathcal{Q} \subset [0, 1]$ is R-invariant. Moreover, $R|_{\mathcal{Q}}$ admits only two periodic orbits, namely $C_0 = \{0, 2/3\}$ and $C_1 = \{1, 1/3\}$, and $\mathcal{Q}$ decomposes as the disjoint union of two dense subsets*

$$\mathcal{Q} = \mathcal{Q}_0 \cup \mathcal{Q}_1, \tag{50}$$

where $\mathcal{Q}_i$ consists of all the rationals in $\mathcal{Q}$ whose forward orbit contains C_i, for $i = 1, 2$.

Proof.   The fact that $C_0 = \{0, 2/3\}$ or $C_1 = \{1, 1/3\}$ are two-cycles can be trivially verified. Then, the R-invariance of $\mathcal{Q}$ follows once we prove that $R(x) \in \mathcal{Q}$ for every $x \in \mathcal{Q} \setminus \{0, 1\}$.

Given any $x \in \mathcal{Q} \setminus \{0, 1\}$, we can write $\beta(x) = wp^{\infty} \in \mathcal{C}$ with w and p finite binary words of minimal length such that $|p|_1 > 0$. Then, by applying ρ we obtain

$$\rho(\beta(x)) = \rho(wp^{\infty}) = \rho(w)\rho_w(p)\rho_{wp}(p)\rho_{wp^2}(p) \cdots$$

If $|p|$ is even, $|wp^k|$ has the same parity of $|w|$, and hence $\rho_{wp^k} = \rho_w$ for every $k \geqslant 1$. If instead $|p|$ is odd, $|wp^k|$ has the same parity of $|w| + k$, and so ρ_{wp^k} coincides with ρ_w for k even while it coincides with the complementary map $\widetilde{\rho}_w$ for k odd. Thus, we can rewrite $\rho(\beta(x))$ as

$$\rho(\beta(x)) = \begin{cases} \rho(w)(\rho_w(p))^{\infty} & \text{if } |p| \text{ is even} \\ \rho(w)(\rho_w(p)\widetilde{\rho}_w(p))^{\infty} & \text{if } |p| \text{ is odd} \end{cases}.$$

In both cases $\rho(\beta(x))$ is periodic, which implies that $R(x) = \xi(\rho(\beta(x)) \in \mathcal{Q}$. This concludes the proof of the R-invariance of $\mathcal{Q}$.

Now, we pass to prove that the forward orbit of any $x \in \mathcal{Q} \setminus \{0, 1\}$ contains one of C_0 or C_1. From the last formula we derive that $\rho(\beta(x)) \in \mathcal{C}$, that is $\rho(\beta(x)) = \beta(R(x))$, except when $|p|$ is even and $|\rho_w(p)|_1 = 0$, in which case $\rho(\beta(x)) = \beta'(R(x)) = \rho(w)0^{\infty}$. Here, we have two possibilities, either $|\rho(w)|_1 = 0$ and then $R(x) = 0$ or else $|\rho(w)|_1 > 0$ and then $\beta(R(x)) = v1^{\infty}$ with $n_\epsilon(v) \leqslant n_\epsilon(\rho(w))$ (in fact $v = \rho(w)_{|k-1}0$, where k the position of the last 1 in $\rho(w)$). By iterating the process, we can conclude that $R^n(x)$ admits a purely periodic binary expansion for every $n \geqslant n_\epsilon(w)$.

On the other hand, for $x' \in \mathcal{Q} \setminus \{0, 1\}$ such that $\beta(x') = q^{\infty}$ with q a finite binary word of minimal length such that $0 < |q|_1 < |q|$, we have

$$\rho(\beta(x')) = \begin{cases} (\rho(q))^{\infty} & \text{if } |q| \text{ is even} \\ (\rho(q)\widetilde{\rho}(q))^{\infty} & \text{if } |q| \text{ is odd} \end{cases}.$$

The minimality of q, implies that $|\rho(q)| = |q|_1 < |q|$ if $|q|$ is even, while $|\rho(q)\widetilde{\rho}(q)| = 2|\rho(q)| = 2|q|_1 < 2|q|$ if $|q|$ is odd. In this last case, $|\rho(q)\widetilde{\rho}(q)|$ is even, $|\rho(q)\widetilde{\rho}(q)|_1 = |\rho(q)\widetilde{\rho}(q)|/2$, and

$\rho(\beta(x')) = \beta(R(x'))$. Therefore, either $R(x') = \xi(\rho(\beta(x')))$ or $R^2(x') = \xi(\rho^2(\beta(x')))$ admits a purely periodic binary expansion whose period length is strictly less than $|q|$. By iteration, we eventually get $R^n(x') = 0$ or $R^n(x') = 1$ for a sufficiently large n.

At this point, we are left to show that $\mathcal{Q}_0$ and $\mathcal{Q}_1$ are dense, or equivalently that they meet all the dyadic intervals $\xi([w]) \subset [0,1]$ with w any finite binary word.

First, let us argue for Q_0. We put $n = n_\epsilon(w)$, and define a sequence of finite binary words $p_0, p_1, \cdots, p_n$ by induction on i decreasing from n to 0, as follows. We start with $p_n = 0$, and given p_i with $0 < i \leqslant n$, we let p_{i-1} be any even length element of $(\rho_{\rho^{i-1}(w)})^{-1}(p_i)$. The existence of such an element is guaranteed by the surjectivity of $\rho_{\rho^{i-1}(w)}$ and the possibility of appending a 0, if needed to make the length even, without changing the image under $\rho_{\rho^{i-1}(w)}$.

Then, we put $x = \xi(wp_0^\infty) \in \xi([w])$. By induction on i, we get

$$R^i(x) = \xi(\rho^i(w)p_i^\infty)$$

for every $i = 0, \cdots, n$, as follows. The base of the induction is the trivial case of $i = 0$, while the inductive step is given by

$$
\begin{aligned}
R^i(x) &= R(\xi(\rho^{i-1}(w)p_{i-1}^\infty)) && \text{(the inductive hypothesis)} \\
&= \xi(\rho(\rho^{i-1}(w)p_{i-1}^\infty)) && \text{(since } \rho^{i-1}(w)p_{i-1}^\infty \in \mathcal{C}) \\
&= \xi(\rho^i(w)\rho_{\rho^{i-1}(w)}(p_{i-1}^\infty)) && \text{(thanks to equation (47))} \\
&= \xi(\rho^i(w)(\rho_{\rho^{i-1}(w)}p_{i-1})^\infty) && \text{(since } p_{i-1} \text{ has even length)} \\
&= \xi(\rho^i(w)p_i^\infty) && \text{(by definition of } p_{i-1}).
\end{aligned}
$$

In particular, we have $R^n(x) = \xi(\rho^n(w)p_n^\infty) = \xi(0^\infty) = 0$, as desired.

This concludes the proof of the density of $\mathcal{Q}_0$. The same argument, but starting from $p_n = 1$ instead of $p_n = 0$, proves that $\mathcal{Q}_1$ is dense. $\qquad\square$

We remark that C_0 and C_1 are the only periodic R-orbits in $\mathcal{Q}$ and that $\mathcal{Q}$ does not contain any dense R-orbit. Thus, in the following, when considering any other periodic R-orbit or any dense R-orbit, we can always assume that they are disjoint from $\mathcal{Q}$, avoiding in this way the technicalities due to the double binary expansion of the dyadic rationals.

Proposition 6.4. *The set of periodic points of R is dense in* $[0,1]$ *and contains uncountably many n-periodic points for any given minimal period $n \geqslant 1$.*

Proof. In the light of the above observation, apart from the two-cycles C_0 and C_1, all the other n-periodic R-orbits are contained in $[0,1] \backslash \mathcal{Q}$ and then they bijectively correspond to those of ρ through the unique binary expansion $\beta(x)$ of every $x \in [0,1] \backslash \mathcal{Q}$.

Therefore, the first part of the statement follows once we prove that for any non-empty finite binary word w there are uncountably many infinite binary ρ-periodic words b in $[w] \cap \mathcal{C}$, whose period is the vanishing order $n_\epsilon(w)$. In fact, it is enough to observe that the corresponding points $\xi(b)$ in the generic dyadic interval $\xi([w]) \subset [0,1]$ are R-periodic with the same period.

Fixed any non-empty finite binary word w, we put $n = n_\epsilon(w) \geqslant 1$ and consider all the infinite concatenations

$$b = w_0 w_1 w_2 \cdots w_k \cdots \in \mathcal{C},$$

where $w_0 = w$ and for every integer $k \geqslant 1$ we let w_k be any of the infinitely many (non-empty) words in $(\rho^n_{w_0 w_1 \cdots w_{k-1}})^{-1}(w_{k-1})$ which ends with 1. As in the proof of lemma 6.1, this last condition guarantees that there are no duplicates in the resulting words b, which are thus uncountably

many. Then, we have

$$
\begin{aligned}
\rho^n(b) &= \rho^n(w_0 w_1 w_2 \cdots w_k \cdots) \\
&= \rho^n(w_0)\rho^n_{w_0}(w_1)\rho^n_{w_0 w_1}(w_2)\cdots\rho^n_{w_0 w_1\cdots w_{k-1}}(w_k)\cdots \\
&= \epsilon w_0 w_1 \cdots w_{k-1} \cdots = b,
\end{aligned}
$$

and hence b has period n.

For the second part we need the additional fact that for every $n \geqslant 1$ is the vanishing order of all the non-empty finite binary words w in $\rho^{-n}(\epsilon)$, for example $w = \sigma^{n-1}(0)$, and that such a word w can be chosen so that no $\rho^k(w)$ is a prefix of w for $k < n$, for example $w = 00\sigma^{n-1}(0)$. This last property immediately implies that $\rho^k(b) \neq b$ for every b as above and $k = 1, \ldots, n - 1$, that is n is the minimal period of b. $\square$

In particular, proposition 6.4 tells us that the set Fix R of fixed points of R is uncountable, meaning that the graph of R crosses the diagonal uncountably many times. In fact, by taking $w = 0^\ell$ in the previous proof, the resulting word $b = w_0 w_1 w_2 \cdots \in \mathcal{C}$ belongs to Fix ρ, and hence $\xi(b) \in$ Fix R. In the special case when $w = w_0 = 0$ and each w_k is chosen to have minimal length among the words with the prescribed properties, that is either $w_k = \sigma(w_{k-1})$ or $w_k = \sigma(\widetilde{w}_{k-1})$ according to the parity of $|w_0 w_1 \cdots w_{k-1}|$, we get as b the word

$$
\begin{aligned}
b^0 &= 00101\prod_{h=0}^{\infty}(1^{3\cdot 2^h}(01)^{3\cdot 2^h}) \\
&= 00101111010101111111010101010101\cdots. \tag{51}
\end{aligned}
$$

The corresponding real number

$$
x^0 = \xi(b^0) = 0.00101111010101111111010101010101\cdots \tag{52}
$$

turns out to be the largest element in Fix R, as we will see in a while.

Now, we want to make more explicit the procedure given in the proof of proposition 6.4 in the specific case of fixed points. To any sequence $a = (a_1, a_2, \cdots) \in \mathcal{A} = \mathbb{N}^\omega$, we associates a (distinct) fixed point $\langle a \rangle^{\mathrm{F}} \in Fix\,\rho \subset \mathcal{C}$ and a corresponding (distinct) fixed point $x^a = \xi(\langle a \rangle^{\mathrm{F}}) = 0.b^a \in$ Fix $R \subset [0, 1]$, with b^0 and x^0 associated to the null sequence.

First, we observe that a non-empty finite binary word w can be completed to a fixed point $wb \in$ Fix ρ for a suitable $b \in \mathcal{B}$ if and only if $\rho(w)$ is a prefix of it. Moreover, if this is the case then there is a unique word $b \in \mathcal{S}$ such that $b^w = wb \in$ Fix ρ. The word b^w can be constructed, by generalising the construction of b^0 outlined above. Since $\rho(w)$ is a prefix of w, we can write $w = \rho(w)v_0$ for a certain word v_0, which can be easily seen to be non-empty. Then, we define b^w as the infinite concatenation

$$
b^w = wv_1 v_2 \cdots \in Fix\,\rho,
$$

where the words v_k for $k \geqslant 1$ are inductively defined by starting from v_0 and putting $v_k = \sigma(v_{k-1})$ or $v_k = \sigma(\widetilde{v}_{k-1})$ if $|wv_1 \cdots v_{k-1}|$ is even or odd, respectively (compare with the construction of b^0 above). Then $b = v_1 v_2 \cdots \in \mathcal{S}$, since each v_k contains no pair of consecutive 0's and ends with 1. Moreover, these properties of v_k make it uniquely determined by v_{k-1} for every $k \geqslant 1$, which implies the uniqueness of b in $\mathcal{S}$ by induction.

Now, given any $a = (a_1, a_2, \cdots) \in \mathcal{A}$, we put $\langle a \rangle^{\mathrm{F}}_0 = b^0$ and define the sequence $(\langle a \rangle^{\mathrm{F}}_n)_{n \geqslant 0} \subset$ Fix ρ as follows. For every $n \geqslant 1$, let k_n be the position of the nth occurrence of

1 in $\langle a \rangle_{n-1}^{\mathrm{F}}$, and w_n^a be the word obtained from $\langle a \rangle_{n-1}^{\mathrm{F}}|k_n$ by inserting the word 0^{2a_n} before the last digit 1. The word w_n^a share with $\langle a \rangle_{n-1}^{\mathrm{F}}|k_n$ the property that its image $\rho(w_n^a) = \rho(\langle a \rangle_{n-1}^{\mathrm{F}}|k_n)$ is a prefix of it, so we can set $\langle a \rangle_n^{\mathrm{F}} = b^{w_n^a}$. Notice that $|w_n^a| = k_n + 2a_n \geqslant k_n > k_{n-1} + 2a_{n-1} = |w_{n-1}^a|$ for every $n > 1$. Since the sequence $(k_n)_{n \geqslant 1}$ is increasing and $\langle a \rangle_n^{\mathrm{F}}|k_n - 1 = \langle a \rangle_{n-1|k_{n-1}}^{\mathrm{F}}$ for every $n \geqslant 1$, there exists the limit

$$\langle a \rangle^{\mathrm{F}} = \lim_{n \to \infty} \langle a \rangle_n^{\mathrm{F}}, \tag{53}$$

which is a fixed point for R, admitting by construction all the prefixes w_n^a for $n \geqslant 1$.

In order to characterise the structure of Fix R, we need to prove that the previous construction produces all the points in Fix ρ. Actually, we prove something more in the following lemma.

Lemma 6.5. *The map* $\phi : \mathcal{A} \to \mathrm{Fix}\,\rho$ *given by* $a \mapsto \phi(a) = \langle a \rangle^{\mathrm{F}}$ *is a Lipschitz homeomorphism.*

Proof. For any $b \in \mathrm{Fix}\,\rho$, we consider the sequence $a^b = (a_1^b, a_2^b \cdots) \in \mathcal{A}$, where a_n^b is the number of (disjoint) pairs of 0^s immediately preceding the nth occurrence of 1 in b. Then, the map $\psi : \mathrm{Fix}\,\rho \to \mathcal{A}$ given by $b \mapsto a^b$ is the inverse of ϕ, as straightforward verification shows.

Now, for any $a, a' \in \mathcal{A}$ we have that $d(a, a') = 1/2^{n+1}$ if and only if $a_{|n} = a'_{|n}$ and $a_{n+1} \neq a'_{n+1}$, or equivalently, with the above notations, $w_n^a = w_n^{a'}$ and $w_{n+1}^a \neq w_{n+1}^{a'}$. Therefore, $\langle a \rangle^{\mathrm{F}}$ and $\langle a' \rangle^{\mathrm{F}}$ share the same prefix of length $|w_n^a| = k_n + 2a_n \geqslant k_n \geqslant n$, while they have different prefixes of length $|w_{n+1}^a| = k_{n+1} + 2a_{n+1}$. So, we have $1/2^{k_{n+1}+2a_{n+1}} \leqslant d(\phi(a)\phi(a')) \leqslant 1/2^{n+1}$, which means that ϕ is a Lipschitz map (second inequality) and that ψ is continuous at $\phi(a)$ (first inequality). $\qquad\square$

Proposition 6.6. *The fixed point set* Fix R *is a uncountable subset of* $[0,1] \backslash \mathcal{D}$ *and its closure* Cl(Fix R) *is a null measure Cantor set in* $[0,1]$ *with a countable remainder* Cl(*Fix R*)*Fix R* $\subset \mathcal{D}$.

Proof. The inclusion *Fix R* $\subset [0,1] \backslash \mathcal{D}$ follows by proposition 6.3. As a consequence, the maps ξ and β restrict to a bijective maps Fix $\rho \leftrightarrow$ Fix R, and hence Fix R is uncountable.

We now prove that Cl(Fix R) $\subset$ Fix $R \cup \mathcal{D}$, which trivially gives Cl(Fix R)*Fix R* $\subset \mathcal{D}$. The argument is similar to the first part of the proof of proposition 5.3. Let $x \in$ Cl(Fix R) and $(x_n)_{n \geqslant 1} \subset$ *Fix R* be a sequence such that $\lim_{n \to \infty} x_n = x$. Then, by lemma 6.5 there is a sequence $(a_n)_{n \geqslant 1} \subset \mathcal{A}$ such that $x_n = \xi(\phi(a_n))$ for every $n \geqslant 1$. If $x \in \mathcal{D}$ we are done, otherwise the continuity of β at x implies that $\lim_{n \to \infty} \phi(a_n) = \beta(x) \in \mathcal{C}$. If the sequence of the ith components $((a_n)_i)_{n \geqslant 1}$ is bounded for every $i \geqslant 1$, then by compactness we can replace the sequence $(a_n)_{n \geqslant 1}$ by a subsequence converging to $a \in \mathcal{A}$, and by the continuity of ϕ we have $\beta(x) = \phi(a)$, that is $x = \xi(\phi(a)) \in$ Fix R. If instead some sequence $((a_n)_i)_{n \geqslant 1}$ is unbounded, let m be minimum index $i \geqslant 1$ for which this happens. In this case, we replace the sequence $(a_n)_{n \geqslant 1}$ by a suitable subsequence such that the sequence of truncations $((a_n)_{|m-1})_{n \geqslant 1}$ is constant, and hence $\phi(a_n)_{|k_{m-1}}$, where k_{m-1} is the index of the $(m-1)$th occurrence of 1, coincides with the same finite word d for every $n \geqslant 1$. Then, we get $\beta(x) = \lim_{n \to \infty} \phi(a_n) = \mathrm{d}0^{\infty}$, which is absurd since $\beta(x) \in \mathcal{C}$. This concludes the proof of the inclusion Cl(*Fix R*) $\subset$ *Fix R* $\cup \mathcal{D}$.

Concerning the null measure property, let us consider the map

$$\phi \circ \phi_{1^{\infty}}^{-1} : \langle\!\langle 1^{\infty} \rangle\!\rangle \to \textit{Fix } \rho,$$

where $\phi_{1^{\infty}} : \mathcal{A} \to \langle\!\langle 1^{\infty} \rangle\!\rangle$ and $\phi : \mathcal{A} \to \textit{Fix } \rho$ are the homeomorphisms defined in the proof of lemma 5.4 and in lemma 6.5, respectively. This map sends $\langle a \rangle^{1^{\infty}} \in \langle\!\langle 1^{\infty} \rangle\!\rangle$ to $\langle a \rangle^{\mathrm{F}} \in$ Fix ρ for every $a \in \mathcal{A}$, and it can be shown to be a one-Lipschitz map as follows. Given any $a, a' \in \mathcal{A}$

such that $a_{|n} = a'_{|n}$ and $a_{n+1} \neq a'_{n+1}$ we put $h(a, a') = a_1 + \cdots + a_n + \min\{a_{n+1}, a'_{n+1}\}$, as in the proof of lemma 5.4. Then, the maximum index i for which $\langle a \rangle^{1^\infty}_{|i} = \langle a \rangle^{1^\infty}_{|i}$ is given by $2h(a, a') + n$, while the maximum index i for which $\langle a \rangle^{\mathrm{F}}_{|i} = \langle a' \rangle^{\mathrm{F}}_{|i}$ is given by $2h(a, a') + k_n$, where k_n is the index of the nth occurrence of 1 in the word $\langle a \rangle^{\mathrm{F}}_{n-1}$ defined in the construction of $\langle a \rangle^{\mathrm{F}}$ just before lemma 6.5. Since $k_n \geqslant n$, we can conclude that $\phi \circ \phi^{-1}_{1^\infty}$ is a one-Lipschitz map. From this, arguing as in the proof of lemma 5.5, we can derive the Lipschitz condition for the map

$$\xi \circ \phi \circ \phi^{-1}_{1^\infty} \circ \beta : \xi(\langle\!\langle 1^\infty \rangle\!\rangle) \to Fix\, R.$$

Therefore, thanks to proposition 5.8 and its proof, we have $\dim_H Fix\, R \leqslant \dim_H \xi(\langle\!\langle 1^\infty \rangle\!\rangle) = \log_2 \varphi < 1$, which implies that $Fix\, R$ has a null Lebesgue measure. Since the remainder $\mathrm{Cl}(Fix\, R) \backslash Fix\, R$ is countable, also $\mathrm{Cl}(Fix\, R)$ has a null Lebesgue measure.

As a consequence, $\mathrm{Cl}(Fix\, R)$ is topologically 0-dimensional. On the other hand, it is a perfect set, because for every $a \in \mathcal{A}$, we have $\xi(\langle a \rangle^{\mathrm{F}}) = \lim_{n \to \infty} \xi(\langle a_n \rangle^{\mathrm{F}})$ with $a_n \in \mathcal{A}$ defined by $(a_n)_i = a_i + \delta_{i,n}$, for any $i, n \geqslant 1$. So, $\mathrm{Cl}(Fix\, R)$ is a Cantor set (see [22]). $\qquad\square$

We observe that the proof of proposition 6.6 is essentially based on the parametrisation of $Fix\, \rho$ given by lemma 6.5, which is nothing else than an explicit reformulation in the case of fixed points of the general construction provided in the proof of proposition 6.4 of periodic points of any period. So, we expect that similar results hold for periodic points of any given period.

The idea is that, in order to generate all the periodic orbits of a given period $n > 1$, one could start from a certain finite set of 'simplest' ones, and than progressively enlarge this set by iterating in a suitable order the following procedure. Take any such n-orbit $(b_1, b_2, \cdots, b_n)$, insert some pairs of consecutive 0's in any position of any b_i, and then reconstruct all the orbit accordingly.

For example, starting from the two-cycle $(1^\infty, (01)^\infty)$ of ρ, corresponding to the two-cycle C_1 of R (see proposition 6.3), one can obtain in this way the Thue–Morse two-cycle $(u, \widetilde{u})$ of ρ described in the next remark, and hence a corresponding two-cycle for R.

Remark 6.7. Consider the well known Thue–Morse substitution rule on $\mathcal{B}$

$$\widehat{\tau} : \begin{cases} 0 \mapsto 01 \\ 1 \mapsto 10 \end{cases},$$

and the digit-wise generated map $\widehat{\tau} : \mathcal{B} \to \mathcal{B}$. The map $\widehat{\tau}$ has two fixed points, the Thue–Morse sequence starting with 0

$$u = 0110100110010110100101100110100 \cdots$$

and its complementary sequence $\widetilde{u}$, the Thue–Morse sequence starting with 1, which was used in [40] to found the combinatorics of words. Recalling the definition of $\overline{\tau}$ in (13) and noting that

$$\overline{\tau}^2(0) = \widehat{\tau}^2(0) = 0110 \quad \text{and} \quad \overline{\tau}^2(1) = \widehat{\tau}^2(1) = 1001,$$

we can easily verify that $(u, \widetilde{u})$ is a two-cycle of ρ. Then, letting $t = \xi(u)$ be the Thue–Morse number, we have the corresponding two-cycle $(t, 1 - t)$ of R.

Concerning the periodic points of R having higher order n, in analogy with the description of the maximal fixed point x^0, we limit ourselves to consider the periodic points whose binary

expansion is the 'simplest' infinite binary word among those generated by the construction described in the proof of proposition 6.4. Namely, the binary word obtained by starting from the word $w_0 = 0$ and choosing each w_k to be the shortest word in $(\rho^n_{w_0 w_1 \cdots w_{k-1}})^{-1}(w_{k-1})$, that is the image of w_{k-1} under a suitable composition of σ's and $\widetilde{\sigma}$'s.

The case of n even is not so interesting, since we always obtain 0.1^∞, while the binary expression of the 'simplest' periodic point of odd order $n = 2\ell + 1 \geqslant 3$ is given by

$$x^\ell = 0.(01)^{2^{\ell-1}} \prod_{h=0}^\infty \left(1^{2^{(2\ell+1)h+\ell}} (01)^{2^{(2\ell+1)h+\ell-1}} 1^{(2^\ell-1)2^{(2\ell+1)h+\ell}} \right.$$
$$\left. (01)^{2^{(2\ell+1)h+2\ell}} 1^{2^{(2\ell+1)h+2\ell}} (01)^{(2^\ell-1)2^{(2\ell+1)h+2\ell}} \right). \tag{54}$$

In order to see that this point has order $n = 2\ell + 1$, we write it as $0.w \prod_{h=0}^\infty (u_h v_h)$, with

$$w = (01)^{2^{\ell-1}}, u_h = 1^{2^{(2\ell+1)h+\ell}} (01)^{2^{(2\ell+1)h+\ell-1}} 1^{(2^\ell-1)2^{(2\ell+1)h+\ell}},$$
$$v_h = (01)^{2^{(2\ell+1)h+2\ell}} 1^{2^{(2\ell+1)h+2\ell}} (01)^{(2^\ell-1)2^{(2\ell+1)h+2\ell}}. \tag{55}$$

Then, we first note that the word w vanish to the empty word ϵ in exactly n iterates of ρ. Moreover, by a somewhat tedious but straightforward calculation, repeatedly using the rules

$$\rho : \begin{cases} (01)^k \mapsto 1^k \\ 1^{2k} \mapsto (01)^k, \end{cases}$$

for every $k \geqslant 1$, one checks that

$$\rho^n_w(u_0) = \widetilde{\rho}^2 \circ \rho^{2\ell-1}(u_0) = w,$$

while

$$\rho^n_{wu_0v_0\cdots u_h}(v_h) = \rho^{2\ell+1}(v_h) = u_h \quad \text{and} \quad \rho^n_{wu_0v_0\cdots v_h}(u_{h+1}) = \rho^{2\ell+1}(u_{h+1}) = v_h$$

for every $h \geqslant 0$. Putting all those equation together, one can conclude that $w \prod_{h=0}^\infty (u_h v_h)$ is a fixed point for ρ^n, and hence that $0.w \prod_{h=0}^\infty (u_h v_h)$ is a fixed point for R^n.

The rest of this section is devoted to analyse some chaotic properties of the map R.

Proposition 6.8. *The set of points with a dense R-orbit is uncountable and dense in* $[0, 1]$.

Proof. To get the existence of uncountably many points with a dense R-orbit, it suffices to take the sequence $(w_n)_{n \geqslant 0}$ in lemma 6.2 with the property that it contains all the non-empty finite binary words. The density of the points with a dense R-orbit follows immediately observing that we can take any non-empty finite word as w_0. □

Proposition 6.9. *For every interval $I \subset [0, 1]$, there exists $n \geqslant 1$ such that $R^n(I) = [0, 1]$. In particular, R is topologically (strongly) mixing and has sensitive dependence on initial conditions.*

Proof. Let w be any non-empty finite binary word such that $\xi([w]) \subset I$, and set $n = n_\epsilon(w)$. Pick $y \in [0, 1]$. By the surjectivity of ρ^n_w, there is $v \in \mathcal{C}$ such that $\rho^n_w(v) = \beta(y)$, and hence $\rho^n(wv) = \rho^n(w)\rho^n_w(v) = \beta(y)$ as well. Then, for $x = \xi(wv)$ we have $x \in I$ and $R^n(x) = y$.

This implies sensitivity to initial conditions with sensitivity constant $1/2$. Since every open set is a countable union of open intervals, it also follows that, for every open sets $A, B \subset [0, 1]$, there is $n \geqslant 1$ such that $R^m(A) \cap B \neq \emptyset$ for every $m > n$, that is R is topologically mixing. □

Corollary 6.10. *The map R is chaotic in the sense of Devaney (see e.g.* [12]*).*

Proof. It follows by propositions 6.4, 6.8 and 6.9. ☐

Concerning distributional chaos, a concept introduced within the context of topological dynamical systems in 1994 (see [35]), the map R satisfies the strongest form of distributional chaos, namely the distributional chaos of type 1, DC1-chaos in short [1]. Actually, we prove that the DC1-chaos exhibited by R is uniform in the sense of [28], as specified in the next proposition.

Proposition 6.11. *The map R is uniformly DC1-chaotic, meaning that it admits an uncountable uniformly DC1-scrambled set, that is an uncountable subset $S \subset [0, 1]$ such that for every $\delta > 0$ and some $\overline{\delta} > 0$ the following two conditions hold, where # denotes the cardinality,*

$$\limsup_{n \to \infty} \frac{1}{n} \#\{k = 0, \cdots, n - 1 \quad such\ that\ |R^k(x) - R^k(y)| < \delta\} = 1, \qquad (56)$$

$$\liminf_{n \to \infty} \frac{1}{n} \#\{k = 0, \cdots, n - 1 \quad such\ that\ |R^k(x) - R^k(y)| < \overline{\delta}\} = 0, \qquad (57)$$

for every $x \neq y$ in S.

Proof. We will explicitly construct a set $S = \{x_\alpha\}_{\alpha \in A}$ satisfying the properties (56) and (57). As the set of indices A we take the image $A = \mu(\mathcal{B}) \subset \mathcal{B}$ of the map $\mu : \mathcal{B} \to \mathcal{B}$ defined by

$$\mu(b)_n = \begin{cases} b_k & \text{if } n = p^k \text{ with } p \text{ prime and } k \geqslant 1 \\ 0 & \text{otherwise} \end{cases},$$

for every $b \in \mathcal{B}$. We observe that A is an uncountable set, being μ injective, and that all the elements $\alpha \in A$ share the same nth component $\alpha_n = 0$ for infinitely many $n \geqslant 1$. Furthermore, since each $\alpha = \mu(b)$ contains infinite copies of b as subsequences, any two different elements $\alpha, \alpha' \in A$ cannot be ultimately coinciding, that is for every $m \geqslant 1$ there is $n \geqslant m$ with $\alpha_n \neq \alpha'_n$.

Now, let $(e_n)_{n \geqslant 1}$ be the increasing sequence of natural numbers inductively defined by

$$e_1 = 1 \quad \text{and} \quad e_n = n^2 \sum_{k=1}^{n-1} e_k \quad \text{for } n > 1, \qquad (58)$$

and let $(v_n)_{n \geqslant 1}$ and $(u_n)_{n \geqslant 1}$ be the two sequences of binary words inductively defined by

$$\begin{aligned} v_1 &= 1 \quad \text{and } v_n = \sigma^{2e_n}(v_{n-1}) \quad \text{for } n > 1, \\ u_1 &= 001 \text{ and } u_n = 00\sigma^{2e_n}(u_{n-1}) \text{ for } n > 1. \end{aligned} \qquad (59)$$

Clearly, we have $n_\epsilon(v_n) = n_\epsilon(u_n) = \sum_{k=1}^n 2e_k$ for every $n \geqslant 1$. Moreover, the equality $\sigma^2(1^k) = 1^{2k}$ implies that

$$v_n = 1^{2\sum_{k=2}^n e_k}, \qquad (60)$$

for every $n > 1$.

To any index $\alpha \in A$ we associate the sequence of non-empty finite binary words $(w_n^\alpha)_{n \geqslant 0}$, where w_0^α can be arbitrarily chosen, while for $n \geqslant 1$

$$
w_n^\alpha = \begin{cases} v_n \text{ if } \alpha_n = 0 \\ u_n \text{ if } \alpha_n = 1 \end{cases}.
$$

We set $k_n = n_\epsilon(w_1^\alpha) + \cdots + n_\epsilon(w_{n-1}^\alpha)$ and observe that k_n is the same for all $\alpha \in A$.

Then, we apply lemma 6.1 the sequences $(w_n^\alpha)_{n \geqslant 0}$ and $(k_n)_{n \geqslant 1}$ and pick a single binary word $b_\alpha \in [w_0^\alpha] \cap \mathcal{C}$, out of the uncountable resulting set B, such that $\rho^{k_n}(b_\alpha) \in [w_n^\alpha]$ for every $n \geqslant 1$. Finally, we put $S = \{x_\alpha = \xi(b_\alpha)\}_{\alpha \in A}$. Since all the words b_α and $\rho^k(b_\alpha)$ belong to $\mathcal{C}$, by their construction in the proof of lemma 6.1, we have that all the points x_α are distinct and that $R^k(x_\alpha) = \xi(\rho^k(b_\alpha))$ for every $k \geqslant 1$ and $\alpha \in A$.

To see that S is uniformly DC1-scrambled, take any $\alpha \neq \alpha'$ in A. For any $\delta > 0$, let $c(\delta) \geqslant 1$ be such that

$$
1/2^{2^{c(\delta)/2}} < \delta,
$$

in such a way that the interval $\xi([\rho^h(w_n^\alpha)])$ has width

$$
\lambda(\xi([\rho^h(w_n^\alpha)])) < \delta \quad \text{for every} \quad h = 1, \cdots, n_\epsilon(w_n^\alpha) - c(\delta) = \sum_{k=1}^n 2e_k - c(\delta).
$$

(61)

By construction, there is a sequence $(n_i)_{i \geqslant 1}$ such that $\alpha_{n_i} = \alpha'_{n_i}$ for every $i \geqslant 1$. By (61), this implies that

$$
\frac{1}{k_{n_i+1}} \#\{k = 1, \ldots, k_{n_i+1} \quad \text{such that } |R^k(x_\alpha) - R^k(x_{\alpha'})| < \delta\}
$$

$$
\geqslant \frac{\sum_{k=1}^{n_i} 2e_k - c(\delta)}{k_{n_i+1}}.
$$

(62)

Thanks to equation (58), we have

$$
\sum_{k=1}^{n_i} 2e_k = \sum_{k=1}^{n_i-1} 2e_k + 2e_{n_i} = 2(n_i^2 + 1)\frac{e_{n_i}}{n_i^2},
$$

(63)

while

$$
\begin{aligned}
k_{n_i+1} &= \sum_{k=1}^{n_i} n_\epsilon(w_k^\alpha) = \sum_{k=1}^{n_i}\sum_{h=1}^{k} 2e_h = \sum_{k=1}^{n_i-1}\sum_{h=1}^{k} 2e_h + \sum_{h=1}^{n_i} 2e_h \\
&\leqslant 2(n_i - 1)\sum_{k=1}^{n_i-1} e_k + \sum_{k=1}^{n_i} 2e_k \\
&= 2(n_i - 1)\frac{e_{n_i}}{n_i^2} + \sum_{k=1}^{n_i} 2e_k \\
&= 2(n_i - 1)\frac{e_{n_i}}{n_i^2} + 2(n_i^2 + 1)\frac{e_{n_i}}{n_i^2}.
\end{aligned}
$$

(64)

Therefore, the right-hand side of (62) is bounded below by

$$\frac{n_i^2 + 1}{n_i^2 + n_i} - \frac{c(\delta)}{k_{n_i+1}},$$

which tends to 1 as n_i diverges. This completes the proof of property (56).

To prove (57), we consider a different sequence $(n_i)_{i \geqslant 1}$ such that $\alpha_{n_i} \neq \alpha'_{n_i}$ for every $i \geqslant 1$, whose existence is once again guaranteed by construction. Then, (59) and (60) imply that $|R^{k_{n_i}}(x_\alpha) - R^{k_{n_i}}(x_{\alpha'})| > 1/8$. So, putting $\overline{\delta} = 1/8$ and arguing as above, we get

$$\frac{1}{k_{n_i+1}} \#\{k = 1, \ldots, k_{n_i+1} \quad \text{such that } |R^k(x_\alpha) - R^k(x_{\alpha'})| < \overline{\delta}\}$$

$$\leqslant \frac{k_{n_i+1} - \sum_{k=1}^{n_i} 2e_k}{k_{n_i+1}}.$$

Recalling the equations (63) and (64), we see that the last quantity vanishes as n_i diverges, which gives property (57).   □

We emphasise that in the above proof the family of starting words $(w_0^\alpha)_{\alpha \in A}$ is completely arbitrary. This means that the DC1-scrambled sets for R form a dense subset in the space of all subsets of $[0, 1]$ with the Hausdorff pseudo-distance. In particular, the DC1-scrambled set S can be chosen to be a dense subset of any interval $I \subset [0, 1]$. Moreover, by suitably modifying the sequences $(w_n^\alpha)_{n \geqslant 1}$ in the proof, it could be shown that actually R is transitively DC1-chaotic, that is all the points of the uncountable DC1-scrambled set S have dense orbits.

Remark 6.12. Proposition 6.11 is interesting in view of the results of [37, 38], in which the author proves that there is a residual subset in the space of bounded functions of Baire class 1, and of Baire class 2 as well, whose elements are neither Devaney nor Li-Yorke chaotic (see [19, 25]). It follows that the map R, although generated by a very simple erasing substitution rule, is 'much more chaotic' than the topologically generic map in its Baire class and also in the higher one.

In the final part of this section we address topological entropy.

We recall that for a map $f : X \to X$ with X a metric space, the topological entropy can be written as [3, 13]

$$h(f) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log r(n, \varepsilon)}{n}, \tag{65}$$

where $r(n, \varepsilon)$ is the maximum cardinality of an ε-separated set (that is a set whose points are at least ε-apart from each other) in the metric

$$d_n(x, y) = \max_{0 \leqslant i \leqslant n} |f^i(x) - f^i(y)|. \tag{66}$$

Although this definition is expressed in metric terms, it turns out to depend only on the underlying topology—for a modern general reference on topological entropy see [14].

For continuous interval maps, positive topological entropy is known to be equivalent distributional chaos, and in particular to DC1-chaos. In our more general context, the positivity of the entropy of the map R should be proved directly. In fact, in the next proposition we prove that $h(R) = \infty$. This can be thought as a consequence of the fact that the substitution ρ is 'efficient enough' in erasing the prefixes, so that all the points 'forget' quickly where they came from. Namely, proposition 2.1 tells us that the vanishing order of any finite binary word w satisfies the inequality

$$n_\epsilon(w) \leqslant 2 \lfloor \log_2 |w| \rfloor + 2. \tag{67}$$

Proposition 6.13. *The map R has infinite topological entropy.*

Proof. For every $k, n \geqslant 1$ and every sequence of words $w_0, w_1, \ldots, w_n \in \{0,1\}^k$, we apply lemma 6.1 to get a single binary word $b_{w_0,w_1,\ldots,w_n} \in [w_0 0] \cap \mathcal{C}$ such that for every $i = 0, 1, \ldots, n$

$$\rho^{i\ell}(b_{w_0,w_1,\ldots,w_n}) \in [w_i 0],$$

where $\ell = 2\lfloor \log_2 |w| \rfloor + 2 \geqslant n_\epsilon(w_i)$ according to proposition 2.1. Then, we put

$$S_{k,n} = \{x_{w_0,w_1,\ldots,w_n} = \xi(b_{w_0,w_1,\ldots,w_n}) \mid w_0, w_1, \ldots, w_n \in \{0,1\}^k\} \subset [0,1]. \tag{68}$$

Since all the words $b_{w_0,w_1,\ldots,w_n}$ and $\rho^k(b_{w_0,w_1,\ldots,w_n})$ belong to $\mathcal{C}$, by their construction in the proof of lemma 6.1, we have that all the points $x_{w_0,w_1,\ldots,w_n}$ are distinct and that $R^k(x_{w_0,w_1,\ldots,w_n}) = \xi(\rho^k(b_{w_0,w_1,\cdots,w_n}))$ for every $k \geqslant 1$ and $w_0, w_1, \ldots, w_n \in \{0,1\}^k$.

We note that, for every $k, n \geqslant 1$ the set $S_{k,n}$ has cardinality $\#S_{k,n} = 2^{(n+1)k}$ and it is ε-separated in the metric d_n if $\varepsilon < 1/2^{k+1}$. Indeed, given any two different points $x = x_{w_0,w_1,\ldots,w_n}$ and $y = x_{w_0',w_1',\ldots,w_n'}$ in $S_{k,n}$, we have $w_i \neq w_i'$ for some $i = 0, \ldots, n$, and hence $|R^i(x) - R^i(y)| \geqslant \varepsilon$, as it can be easily deduced from $x \in \xi([w_i 0])$ and $y \in \xi([w_i' 0])$.

Then, based on equation (65) and putting $k = \lceil -\log_2 2\varepsilon \rceil - 1$ we get

$$h(R) \geqslant \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log 2^{(n+1)k}}{n} = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{(n+1) \log 2^{(\lceil -\log_2 2\varepsilon \rceil - 1)}}{n} = \infty.$$

$\square$

Recently some attention has been devoted to the study of finer properties of topological entropy, in particular to the characterisation of the points around which the entropy concentrates. More precisely, one can introduce the notion of relative entropy $h(f, K)$ for a map $f : X \to X$, with X a metric space, and a subset $K \subset X$, as the limit (65) where $r(n, \epsilon)$ is defined by constraining the ε-separated sets to be contained in K. Then, following [42], if f has positive entropy, $x \in X$ is called an entropy point (resp. a full entropy point) if $h(f, K) > 0$ (resp. $h(f, K) = h(f)$) for every closed neighborhood of x in X.

Remark 6.14. It is known that, in case of continuous maps with positive entropy on a compact metric space, minimality implies that every point is a full entropy point [42]. The system $([0, 1], R)$ is not minimal (there are periodic points of every period), nevertheless every point is a full entropy point, as an immediate consequence of proposition 6.9.

7. Arithmetic properties

We recall that a Liouville number is a real number x such that, for every n, there exists a sequence of rationals p_i/q_i ($p_i, q_i \in \mathbb{Z}, q_i \neq 0$) verifying

$$\left| x - \frac{p_i}{q_i} \right| < \frac{1}{q_i^n}. \tag{69}$$

Proposition 7.1. *The map R admits uncountably many fixed points that are Liouville numbers, hence transcendental real numbers.*

Proof. For the existence of uncountably many Liouville fixed points, it suffices to observe that, by construction, the fixed point $0.\langle a \rangle^{\mathrm{F}}$ is a Liouville number as soon as the sequence $a \in A$ increases fast enough (see definition (53) and proposition 6.6). By a classical result [24], all the fixed points which are Liouville numbers are transcendental. □

A similar result could be established analogously for periodic points of any given period. However, this simple argument does not apply to the largest fixed point, for which we have to provide a different proof of transcendence.

Proposition 7.2. *The largest fixed point x^0 of the map R is transcendental.*

Proof. Let $p(n, x^0)$ be the cardinality of the set of words of length n appearing in the binary expansion of x^0. Then we have

$$\limsup_{n \to \infty} \frac{p(n, x^0)}{n} < 8. \tag{70}$$

Indeed, given any $n > 3$, let k be the least positive integer such that $k \geqslant \log_2(n/3)$, hence $3 \cdot 2^k \geqslant n$ and $3 \cdot 2^{k-1} < n$. Then, we split the binary expansion b^0 of x^0 as the product $b^0 = b_0' b_0''$, where $b_0' = 00101 \prod_{h=0}^{k-1} 1^{3 \cdot 2^h} (01)^{3 \cdot 2^h}$ and $b_0'' = \prod_{h=k}^{\infty} 1^{3 \cdot 2^h} (01)^{3 \cdot 2^h}$. The length of b_0' is

$$5 + 9 \sum_{h=0}^{k-1} 2^h = 5 + 9(2^k - 1) = 9 \cdot 2^k - 4 < 6n - 4 < 6n. \tag{71}$$

Therefore, there are less than $6n$ different subwords of b^0 having length n and starting in b_0'. We are left to consider the different subwords of length n that are contained in b_0''. Due to the inequality $3 \cdot 2^k \geqslant n$, each of them consists of a (possibly empty) sequence of 1's followed/preceded by a (possibly empty or truncated on the right/left) sequence of 01's. So, there are at most $2n + 3$ such subwords. Therefore, $p(n, x^0) < 8n + 3$ for every $n \geqslant 1$, which gives (70). Thus, by theorem 3.1 in [6], x^0 is transcendental. □

Along the same lines we can prove the following.

Proposition 7.3. *The 'simplest' periodic point x^ℓ of period $2\ell + 1$ defined in (54) is transcendental for each $\ell \geqslant 1$.*

Proof. We know that the general form for the binary expansion of x^ℓ is $b^\ell = w\prod_{h=0}^{\infty}u_h v_h$ where w, u_h and v_h are the words in (55). Proceeding as above, let k be the least positive integer such that $2^{(2\ell+1)k+\ell} \geqslant n$ and $2^{(2\ell+1)(k-1)+\ell} < n$. We then split the binary expansion of b^ℓ as $b'_\ell b''_\ell$, where $b'_\ell = w\prod_{h=0}^{k-1}u_h v_h$ and $b''_\ell = \prod_{h=k}^{\infty}u_h v_h$. Now, one shows that the length of b'_ℓ is bounded above by $12 \cdot 2^{2\ell+1}n$, which also gives a bound for the number of different subwords having length n and starting in b'_ℓ. Moreover, by the same argument used above yields that b''_ℓ contributes with at most $2n + 3$ to $p(n, b^\ell)$. $\qquad\square$

8. Some open questions

Many natural questions concerning the map R do not have yet an answer. We limit ourselves to list a few of them.

Concerning basic properties of the map, we propose the following.

Question 8.1. What is the Hausdorff dimension of the graph G of R?

We note that the self-affine structure of G discussed in section 4 seems not to fit into any of the known criteria for the computation of the Hausdorff dimension.

From the ergodic point of view, the main question is probably the next one.

Question 8.2. Does R admit an invariant measure which is not purely atomic?

We recall that the singularity of R prevents the existence of absolutely continuous (with respect to Lebesgue) invariant measures. We add finally that it can be proved that the *even-order* iterates of the map preserve the Lebesgue measure (work in progress).

Although the main interest in studying the dynamics of a map such as R comes from a natural generalisation of topological dynamics questions to a wider class of functions, like Baire class 1 functions, it could be of interest also in an ergodic context to model specific 'non-equilibrium' situations in which events that are impossible at one time become possible at another time.

Concerning points with a dense orbit, heuristics indicate that there is plenty of them, as we have uncountability for two different reasons: in the procedure described in the proof of lemma 6.2 there are countably many steps each involving countably many choices; in proposition 6.8 there are uncountably many possible choices for the sequence $(w_n)_{n\geqslant1}$.

Therefore, it seems reasonable to ask the following.

Question 8.3. Is the set of points with a dense R-orbit of second category (or even co-meagre)? Does it have positive (or even full) Lebesgue measure?

About the arithmetic properties of R, in the light of the results of section 7, and of the fact that the Thue–Morse number is known to be transcendental, one could ask the question below.

Question 8.4. Is every fixed point of R transcendental? Is every periodic point of R not in C_0 or C_1 transcendental?

Finally, as said in the introduction, the map R is the model-case of a more general class of interval maps generated by erasing block substitutions, meaning (in the binary case) substitution rules $s : \{0, 1\}^k \to \{0, 1\}^*$ such that at least one word w of length k is mapped to ϵ. Looking at these more general maps poses of course new problems.

Question 8.5.   What of the properties we established for the map R are true for interval maps generated by which classes of (binary) erasing block substitutions? What new phenomena arise in this extended context?

Let us list some speculation about this last question.

(a) While the map ρ has an 'almost-inverse' map σ (up to insertions of the word 00), this is not true for a general erasing substitution rule s as above. Indeed, in the s-preimages of a given infinite word $v \in \{0, 1\}^\omega$, there could be in principle infinitely many words $\prod_{i=1}^\infty w_i$ with $w_i \in \{0, 1\}^k$, such that $s(w_i) \neq \epsilon$ for every $i \geqslant 1$. This could make the topological structure of the fibers much more complex than in our present case.

(b) Even assuming that there is only one word $w_\epsilon \in \{0, 1\}^k$ mapped by s to the empty word, if $w_\epsilon \neq 0^k$ we have to consider separately three objects, that is the set $\mathcal{D}$ of dyadic rationals in $[0, 1]$ and the two sets

$$S_1 = \{x \in [0, 1] \mid x = 0.v, \qquad v = s(w0^\infty), \qquad w \in \{0, 1\}^*\},$$
$$S_2 = \{x \in [0, 1] \mid x = 0.ww_\epsilon^\infty, \qquad w \in \{0, 1\}^{nk}\}.$$

It is not difficult to see that these sets coincide in the case when $w_\epsilon = 0^k$, but this is not true in general. Moreover, if $s^{-1}(\epsilon) \subset \{0, 1\}^k$ has more than one element, the interval map $f_s : [0, 1] \to [0, 1]$ generated by the symbolic action of s is in general discontinuous also at some (actually uncountably many) irrationals.

(c) For suitable erasing substitution rules s, the map f_s can have stronger regularity properties than R. For instance, s can be chosen so that f_s is a Darboux function.

A first study of interval maps generated by general erasing block substitutions is done in [11], but several problems are still open, in particular concerning point a) above.

Acknowledgments

List of symbols	Definition	Eq. number
$\mathbb{N}$	Natural numbers (including 0)	
$\mathbb{D} \subset \mathbb{Q}$	Dyadic rationals and all rationals, respectively	
φ	The golden ratio $(1 + \sqrt{5})/2$	
$\mathcal{Q} = \mathbb{Q} \cap [0, 1]$	Rationals in $[0, 1]$	
$\mathcal{D} = \mathbb{D} \cap [0, 1]$	Dyadic rationals in $[0, 1]$	
$\mathcal{N} \subset [0, 1]$	Normal binary numbers in $[0, 1]$	
$\mathcal{A} = \mathbb{N}^\omega$	Infinite sequences of natural numbers	
$\mathcal{B} = \{0, 1\}^\omega$	Infinite binary sequences/words	
$\mathcal{C} \subset \mathcal{B}$	Infinite binary sequences/words which are not eventually 0	
$\mathcal{C}' \subset \mathcal{B}$	Infinite binary sequences/words which are not eventually 1	
$\mathcal{S} \subset \mathcal{C}$	Infinite binary sequences/words with no pair of consecutive 0's	
$\eta : \mathcal{A} \to \mathcal{C}$	Map $(a_1, a_2, \cdots) \mapsto 0^{a_1} 1 0^{a_2} 1 \cdots$	(1)
$\xi : \mathcal{B} \to [0, 1]$	Map $(b_1, b_2, \cdots) \mapsto x = 0.b_1 b_2 \cdots$	(3)
$\beta : (0, 1] \to \mathcal{C}$	Map $x \mapsto (b_1, b_2, \cdots)$, the unique sequence in $\mathcal{C}$ such that $x = 0.b_1 b_2 \cdots$	(4)
$\beta' : [0, 1) \to \mathcal{C}'$	Map $x \mapsto (b_1, b_2, \cdots)$, the unique sequence in $\mathcal{C}'$ such that $x = 0.b_1 b_2 \cdots$	(4)
$|s|, |w|$	Length of a sequence s or word w, respectively	
$|s|_1, |w|_1$	Number of 1's in a binary sequence or word, respectively	
$\widetilde{s}, \widetilde{w}$	Binary sequence/word obtained by complementing s/w, respectively	
$s_{|n}, w_{|n}$	n-truncation of a sequence s or a word w, respectively	(5)
$[s], [w]$	Set of all infinite sequences/words having prefix s/w, respectively	(6)
$\{0, 1\}^*$	Possibly empty finite binary sequences/words	
$\{0, 1\}^\infty$	Possibly empty finite or infinite binary sequences/words	
$\rho : \{0, 1\}^\infty \to \{0, 1\}^\infty$	The main substitution map	(7)
$\tau : \{0, 1\}^\infty \to \{0, 1\}^\infty$	An equivalent two-block substitution map (defined on even/infinite length binary sequences/words)	(11)
$\overline{\tau} : \{0, 1\}^\infty \to \{0, 1\}^\infty$	The substitution map inverse to τ	(13)
$\sigma : \{0, 1\}^\infty \to \{0, 1\}^\infty$	The 'simplest' substitution map inverse to ρ	(14)
$R : [0, 1] \to [0, 1]$	The real map corresponding to ρ	(16)
$S : [0, 1] \to [0, 1]$	The section of R corresponding to σ	(22)
ϵ	The empty sequence/word	
n_ϵ	Vanishing order function of binary sequences/words under the action or ρ	(12)
$\widetilde{\rho}$	The substitution map complementary to ρ	(8)
ρ_k	ρ or $\widetilde{\rho}$, depending on k being an even or odd natural, respectively	(9)
ρ_v	Auxiliary substitution map, coinciding with $\rho_{|v|}$, such that $\rho(vw) = \rho(v)\rho_v(w)$	(10)
ρ_v^n	Auxiliary substitution map such that $\rho^n(vw) = \rho^n(v)\rho_v^n(w)$	(48)

(continued on next page)

$\langle a \rangle^b$	Binary sequence/word obtained from b by inserting 0^{2a_k} before each kth 1	(2)
$\langle\!\langle b \rangle\!\rangle$	The auxiliary set $\{\langle a \rangle^b \mid a \in \mathcal{A}\}$ used to construct the fibers of ρ and R	(15, 21)
$\langle a \rangle^{\mathrm{F}}$	The element of Fix ρ generated from the 'simples' one b^0 according to $a \in \mathcal{A}$	(53)
b^0	The 'simplest' fixed point $\langle 0^\infty \rangle^{\mathrm{F}}$ of the map ρ	(51)
x^0	The largest fixed point $\xi(b^0)$ of the map R	(52)
x^ℓ	The 'simplest' periodic point of the map R having odd order $2\ell + 1$	(54)
C_0, C_1	The two rational R-cycles $\{0, 2/3\}$ and $\{1, 1/3\}$ respectively	
Q_0, Q_1	The subsets of $\mathcal{Q}$ attracted by C_0 and C_1 respectively	(50)

References

[1] Balibrea F, Smítal J and Štefánková M 2005 The three versions of distributional chaos *Chaos Solitons Fractals* **23** 1581–3

[2] Émile Borel M 1909 Les probabilités dénombrables et leurs applications arithmétiques *Rend. Circ. Matem. Palermo* **27** 247–71

[3] Bowen R 1971 Entropy for group endomorphisms and homogeneous spaces *Trans. Am. Math. Soc.* **153** 401–14

[4] Bressoud D M 2008 *A Radical Approach to Lebesgue's Theory of Integration* (Cambridge: Cambridge University Press)

[5] Bugeaud Y 2012 *Distribution Modulo One and Diophantine Approximation* (*Cambridge Tracts in Mathematics* vol 193) (Cambridge: Cambridge University Press)

[6] Bugeaud Y 2015 Expansions of algebraic numbers *Four Faces of Number Theory* (*EMS Ser. Lect. Math.*) (Zürich: Eur. Math. Soc.) pp 31–75

[7] Cassels J W S 1952 On a paper of Niven and Zuckerman *Pac. J. Math.* **2** 555–7

[8] Čiklová M 2005 Dynamical systems generated by functions with connected G_δ graphs *Real Anal. Exchange* **30** 617–737

[9] Čiklová M 2007 Minimal and ω-minimal sets of functions with connected G_δ graphs *Real Anal. Exchange* **32** 397–408

[10] Császár Á 1952 Sur la propriété de Darboux *Comptes Rendus du Premier Congrès des Mathématiciens Hongrois, 27 Août-2 Septembre 1950* (Budapest: Akadémiai Kiadó) pp 551–60

[11] Della Corte A and Farotti M 2022 Dynamics of interval maps generated by erasing substitutions *Journal of Differential Equations* **323** 86–112

[12] Devaney R 2022 *An Introduction to Chaotic Dynamical Systems* (*Studies in Nonlinearity*) (London: CRC Press)

[13] Dinaburg E I 1970 A correlation between topological entropy and metric entropy *Dokl. Akad. Nauk SSSR* **190** 19–22

[14] Downarowicz T 2011 *Entropy in Dynamical Systems* (*New Mathematical Monographs* vol 18) (Cambridge: Cambridge University Press)

[15] Durand F and Rigo M 2009 Syndeticity and independent substitutions *Adv. Appl. Math.* **42** 1–22

[16] Fernau H 1994 Infinite iterated function systems *Math. Nachr.* **170** 79–91

[17] Fogg N P 2002 *Substitutions in Dynamics, Arithmetics and Combinatorics* (*Lecture Notes in Mathematics* vol 1794) (Berlin: Springer)

[18] Hanson B, Pierce P and Steele T H 2019 Dynamics of typical Baire-1 functions on a compact n-manifold *Aequat. Math.* **93** 1111–25

[19] Li T-Y and Yorke J A 1975 Period three implies chaos *Am. Math. Mon.* **82** 985–92

[20] Kahng B 2009 On Devaney's definition of chaos for discontinuous dynamical systems *AMATH'09: Proc. 15th American Conf. Applied Mathematics* (World Scientific and Engineering Academy and Society) pp 89–94

[21] Korczak-Kubiak E, Loranty A and Pawlak R J 2015 On the topological entropy of discontinuous functions. Strong entropy points and Zahorski classes *Monograph on the Occasion of 100th Birthday Anniversary of Zygmunt Zahorski* (Gliwice: Wydawnictwo Politechniki Śląskiej) pp 109–24

[22] Kuratowski K 1966 *Topology* vol 1 and 2 (New York: Academic)

[23] Lebesgue M H 1905 Sur les fonctions représentables analytiquement *J. Math. Appl.* **1** 139–216

[24] Liouville J 1851 Sur des classes très-étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques *J. Math. Appl.* **1** 133–142

[25] Mai J-H 2004 Devaney's chaos implies existence of s-scrambled sets *Proc. Am. Math. Soc.* **132** 2761–7

[26] Moran M 1996 Hausdorff measure of infinitely generated self-similar sets *Monatsh. Math.* **122** 387–99

[27] Niven I and Zuckerman H 1951 On the definition of normal numbers *Pacific J. Math.* **1** 103–9

[28] Oprocha P 2009 Distributional chaos revisited *Trans. Am. Math. Soc.* **361** 4901–25

[29] Pawlak H and Pawlak R J 2009 Dynamics of Darboux functions *Tatra Mt. Math. Publ.* **42** 51–60

[30] Pawlak R J 2009 On the entropy of Darboux functions *Colloq. Math.* **116** 227–41

[31] Pillai S S 1940 On normal numbers *Proc. Indian Acad. Sci.* **12** 179–84

[32] Purves R 1966 Bimeasurable functions *Fund. Math.* **58** 149–57

[33] Queffélec M 2010 *Substitution Dynamical Systems—Spectral Analysis* (*Lecture Notes in Mathematics* vol 1294) (Berlin: Springer)

[34] Schneider J C 2012 On the ambiguity of erasing morphisms *Dissertation* Logos Verlag Berlin GmbH

[35] Schweizer B and Smítal J 1994 Measures of chaos and a spectral decomposition of dynamical systems on the interval *Trans. Am. Math. Soc.* **344** 737–54

[36] Steele T H 2018 The space of ω-limit sets for Baire-1 functions on the interval *Topol. Appl.* **248** 59–63

[37] Steele T H 2019 Dynamics of Baire-1 functions on the interval *Eur. J. Math.* **5** 138–49

[38] Steele T H 2019 Dynamics of Baire-2 functions on the interval *Topol. Appl.* **265** 106821

[39] Szuca P 2007 The covering principle for Darboux Baire 1 functions *Fund. Math.* **193** 133–40

[40] Thue A 1906 *Über Unendliche Zeichenreihen* Mat. Nat. Kl. 7 (*Norske Vid. Selsk. Skr. I* vol 7) (Kristiania (Oslo): Dybwad in Komm) pp 1–22

[41] Van Rooij A C and Schikhof W H 1982 *A Second Course on Real Functions* (Cambridge: Cambridge University Press)

[42] Ye X and Zhang G 2007 Entropy points and applications *Trans. Am. Math. Soc.* **359** 6167–86