CDMTCS
Research
Report
Series
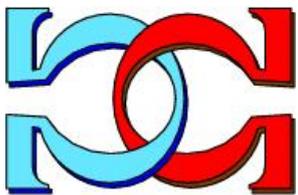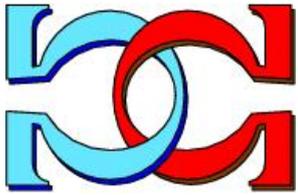
On the Hausdorff measure of regular ω-languages in Cantor space
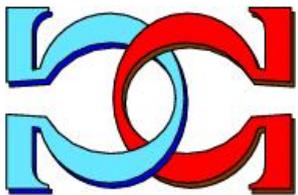
Ludwig Staiger
Martin-Luther-Universität
Halle-Wittenberg
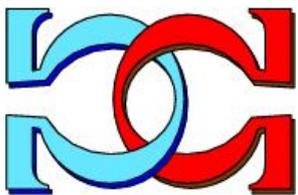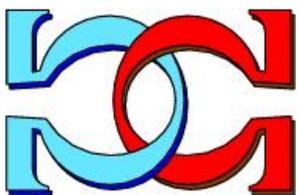
Centre for Discrete Mathematics and
Theoretical Computer Science
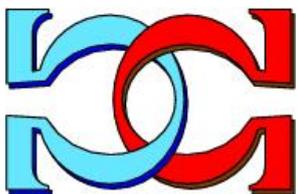
On the Hausdorff measure of regular ω-languages in Cantor space[*]

Ludwig Staiger[†]
Martin-Luther-Universität Halle-Wittenberg
Institut für Informatik
von-Seckendorff-Platz 1, D–06099 Halle (Saale), Germany

Abstract

This paper deals with the calculation of the Hausdorff measure of regular ω-languages, that is, subsets of the Cantor space definable by finite automata. Using methods for decomposing regular ω-languages into disjoint unions of parts of simple structure we derive two sufficient conditions under which ω-languages with a closure definable by a finite automaton have the same Hausdorff measure as this closure.

The first of these condition is related to the homogeneity of the local behaviour of the Hausdorff dimension of the underlying set, and the other with a certain topological density of the set in its closure.

[†]email: `ludwig.staiger@informatik.uni-halle.de`

Contents

Regular ω-languages are not only famous because they are definable by finite automata but also because they are the ones definable in Büchi's [Büc62] restricted monadic second order arithmetic (cf. the survey [Tho90] or [PP04]).

Hausdorff dimension and Hausdorff measure for regular ω-languages have been proved to be computable (cf. [Ban89, MW88, Edg08] or [MS94, Sta98a]). The computation of the Hausdorff measure of a regular ω-language uses several properties which do not hold for larger classes of ω-languages (cf. [Sta93, MS94, Sta98b]). These properties show that subsets of the Cantor space definable by finite automata really deserve the name "regular".

For instance, Theorem 21 of [MS94] shows a strong connection of Hausdorff dimension and topological density for regular ω-languages closed in Cantor space, and the measure-category-theorem of [Sta98b] shows that this connection can be extended to arbitrary regular ω-languages.

Our investigations relate the Hausdorff measure of a subset of the Cantor space to the Hausdorff measure of its closure. A first result shows that under a certain homogeneity condition the measure of a regular ω-language coincides with the measure of its closure. The proof uses the decomposition theorem of [Sta98a] which is based on McNaughton's theorem [McN66] and extends in some sense earlier decompositions of [Arn83, SW74, Wag79].

A second result is a sufficient condition under which sets topologically large relative to its closure have the same Hausdorff measure as their closure. Here, for the case of finite measure, we rely on the measure-category theorem derived in [Sta98b, Theorem 4]. The extension to sets of infinite measure requires a closer inspection of the decomposition of regular ω-languages closed in Cantor space.

The paper is organised as follows. After introducing some notation in Section 2 some properties of Hausdorff measure and dimension are listed. Then the third section deals with decompositions of regular ω-languages derived from the accepting automata. This concerns the general decomposition as in [Sta98a], a new decomposition according to non-null Hausdorff measure and the decomposition of closed sets mentioned above. Then in Section 4 we derive the results on the coincidence of the Hausdorff measures of ω-languages of a certain shape and their closures.

1   Notation

In this section we introduce the notation used throughout the paper. By $\mathbb{IN} = \{0, 1, 2, \ldots\}$ we denote the set of natural numbers. Its elements will be usually denoted by letters $i, \ldots, n$. Let X be an alphabet of cardinality $|X| = r \geq 2$.

Then X^* is the set of finite words on X, including the *empty word e*, and X^ω is the set of infinite strings (ω-words) over X. Subsets of X^* will be referred to as *languages* and subsets of X^ω as *ω-languages*.

For $w \in X^*$ and $\eta \in X^* \cup X^\omega$ let $w \cdot \eta$ be their *concatenation*. This concatenation product extends in an obvious way to subsets $W \subseteq X^*$ and $B \subseteq X^* \cup X^\omega$. For a language W let $W^* := \bigcup_{i \in \mathbb{N}} W^i$, and $W^\omega := \{w_1 \cdots w_i \cdots : w_i \in W \setminus \{e\}\}$ be the set of infinite strings formed by concatenating non-empty words in W. Furthermore, $|w|$ is the *length* of the word $w \in X^*$ and $\mathbf{pref}(B)$ is the set of all finite prefixes of strings in $B \subseteq X^* \cup X^\omega$. We shall abbreviate $w \in \mathbf{pref}(\{\eta\})$ $(\eta \in X^* \cup X^\omega)$ by $w \sqsubseteq \eta$.

As usual, we consider X^ω as a topological space (Cantor space). The *closure* (smallest closed set containing F) of a subset $F \subseteq X^\omega, \mathscr{C}(F)$, is described as $\mathscr{C}(F) := \{\xi : \mathbf{pref}(\{\xi\}) \subseteq \mathbf{pref}(F)\}$. The *open sets* in Cantor space are the ω-languages of the form $W \cdot X^\omega$.

We assume the reader to be familiar with the basic facts of the theory of regular languages and finite automata. We postpone the definition of regularity for ω-languages to Section 3. For more details on ω-languages and regular ω-languages see the book [PP04] or the survey papers [Sta97, Tho90].

2 Hausdorff Dimension and Hausdorff Measure

First, we shall describe briefly the basic formulae needed for the definition of Hausdorff dimension and Hausdorff measure. For more background and motivation see Section 1 of [MS94].

We recall the definition of the Hausdorff measure and Hausdorff dimension (see [Edg08, Fal90]) of a subset of X^ω. In the setting of languages and ω-languages this can be read as follows (see [Sta93, Sta98a]). For $F \subseteq X^\omega$, $r = |X| \geq 2$ and $0 \leq \alpha \leq 1$ the equation

$$\mathbb{L}_\alpha(F) := \lim_{l \to \infty} \inf\Big\{ \sum_{w \in W} r^{-\alpha \cdot |w|} : F \subseteq W \cdot X^\omega \wedge \forall w(w \in W \to |w| \geq l) \Big\} \quad (1)$$

defines the α-dimensional metric outer measure on X^ω. The measure $\mathbb{L}_\alpha$ satisfies the following properties (see [Edg08, Fal90, MS94]).

Proposition 1 *Let $F \subseteq X^\omega$, $V \subseteq X^*$ and $\alpha \in [0,1]$.*

1. If $\mathbb{L}_\alpha(F) < \infty$ then $\mathbb{L}_{\alpha+\varepsilon}(F) = 0$ for all $\varepsilon > 0$.

2. If $F \subseteq \{\xi : \xi \in X^\omega \wedge \mathbf{pref}(\xi) \subseteq V\}$ and $\sum_{v \in V} r^{-\alpha \cdot |v|} < \infty$ then $\mathbb{L}_\alpha(F) = 0$.

3. It holds the scaling property $\mathbb{L}_\alpha(w \cdot F) = r^{-\alpha \cdot |w|} \cdot \mathbb{L}_\alpha(F)$.

4

4. *If V is prefix-free then $\mathbb{L}_\alpha(F \cap V \cdot X^\omega) = \sum_{w \in V} \mathbb{L}_\alpha(F \cap w \cdot X^\omega)$.*

Then the *Hausdorff dimension* of F is defined as

$$\dim F := \sup\{\alpha : \alpha = 0 \vee \mathbb{L}_\alpha(F) = \infty\} = \inf\{\alpha : \mathbb{L}_\alpha(F) = 0\}.$$

It should be mentioned that dim is countably stable and invariant under scaling, that is, for $F_i \subseteq X^\omega$ we have

$$\dim \bigcup_{i \in \mathbb{N}} F_i = \sup\{\dim F_i : i \in \mathbb{N}\} \quad \text{and} \quad \dim w \cdot F_0 = \dim F_0. \qquad (2)$$

In particular, every at most countable subset $E \subseteq X^\omega$ has Hausdorff dimension $\dim E = 0$, and the measure $\mathbb{L}_0$ is the counting measure, that is, $\mathbb{L}_0(E) = |E|$ if E is finite and $\mathbb{L}_0(E) = \infty$, otherwise.

Hausdorff dimension and measure need not be distributed uniformly on a set. In order to describe a certain homogeneity we use the following concept (cf. [MS94, Section 4]). We say that an ω-language F has *locally positive α-dimensional measure* provided $\mathbb{L}_\alpha(F \cap w \cdot X^\omega) > 0$ for all $w \in \mathbf{pref}(F)$. Then the following technical result is true.

Proposition 2 *Let $F \subseteq X^\omega$ have $\dim F = \alpha$, $\mathbb{L}_\alpha(F) < \infty$ and locally positive α-dimensional measure. If $F' \subseteq F$ and $\mathbb{L}_\alpha(F') = \mathbb{L}_\alpha(F)$ then $\mathbf{pref}(F') = \mathbf{pref}(F)$ and, consequently, $\mathscr{C}(F') = \mathscr{C}(F)$.*

Proof. First observe that $F' \subseteq F$ implies $\mathbb{L}_\alpha(F' \cap w \cdot X^\omega) \leq \mathbb{L}_\alpha(F \cap w \cdot X^\omega)$ for all $w \in X^*$. Then the general identity (see Proposition 1.4) $\mathbb{L}_\alpha(E) = \sum_{w \in X^n} \mathbb{L}_\alpha(E \cap w \cdot X^\omega)$ and the hypothesis $\mathbb{L}_\alpha(F') = \mathbb{L}_\alpha(F) < \infty$ imply $\mathbb{L}_\alpha(F' \cap w \cdot X^\omega) = \mathbb{L}_\alpha(F \cap w \cdot X^\omega)$.

Obviously, $\mathbf{pref}(F') \subseteq \mathbf{pref}(F)$. Let now $w \in \mathbf{pref}(F)$. Then $\mathbb{L}_\alpha(F \cap w \cdot X^\omega) > 0$ which in view of $\mathbb{L}_\alpha(F' \cap w \cdot X^\omega) = \mathbb{L}_\alpha(F \cap w \cdot X^\omega)$ implies $w \in \mathbf{pref}(F')$. ❑

We add a further relation of the Hausdorff dimension and the measure $\mathbb{L}_\alpha$ for ω-languages (see [Sta93, MS94]) of a special shape.

Proposition 3 *Let $\alpha = \dim W^\omega$. Then $\mathbb{L}_\alpha(W^\omega) \leq 1$, and if, moreover, $\mathbb{L}_\alpha(W^\omega) > 0$ then W^ω has locally positive α-dimensional measure.*

Proof. The first part is Proposition 6.6 of [Sta93]. Let $w \in \mathbf{pref}(W^\omega)$. Then there is a $v \in X^*$ such that $wv \in W^*$, and, consequently $wv \cdot W^\omega \subseteq W^\omega \cap w \cdot X^\omega$. Now Proposition 1.3 yields $0 < \mathbb{L}_\alpha(wv \cdot W^\omega) \leq \mathbb{L}_\alpha(W^\omega \cap w \cdot X^\omega)$. ❑

3 Decomposition of Regular ω-languages

As usual we call a language $W \subseteq X^*$ *regular* if there is a finite (deterministic) automaton $A = (X; S; s_0; \delta)$, where S is the finite set of states, $s_0 \in S$ is the initial state and $\delta : S \times X \to S$ is the transition function[1], such that $W = \{w : \delta(s_0; w) \in S'\}$ for some fixed set $S' \subseteq S$.

An ω-language $F \subseteq X^\omega$ is called *regular* provided there are a finite (deterministic) automaton $A = (X; S; s_0; \delta)$ and a table $\mathcal{T} \subseteq \{S' : S' \subseteq S\}$ such that for $\xi \in X^\omega$ it holds $\xi \in F$ if and only if $\text{Inf}(A; \xi) \in \mathcal{T}$ where $\text{Inf}(A; \xi)$ is the set of all states $s \in S$ through which the automaton A runs infinitely often when reading the input ξ. Observe that $S' = \text{Inf}(A; \xi)$ holds for a subset $S' \subseteq S$ if and only if

1. there is a word $u \in X^*$ such that $\delta(s_0; u) \in S'$, and

2. for all $s, s' \in S'$ there are non-empty words $w, v \in X^*$ such that $\delta(s, w) = s'$ and $\delta(s', v) = s$.

Thus, to ease our notation, unless stated otherwise in the sequel we will assume all automata to be initially connected, that is, $S = \{\delta(s_0; w) : w \in X^*\}$ and the tables $\mathcal{T}$ to be contained in $\{\text{Inf}(A; \xi) : \xi \in X^\omega\}$.

The ω-language $F = \{\xi : \text{Inf}(A; \xi) \in \mathcal{T}\}$ is the disjoint union of all sets $F_{S'} = \{\xi : \text{Inf}(A; \xi) = S'\}$ where $S' \in \mathcal{T}$.

We are going to split F into smaller mutually disjoint parts. Let $A = (X; S; s_0; \delta)$ be fixed. We refer to a word $v \in X^*$ as $(s; S')$-*loop completing* if and only if

1. v is not the empty word,

2. $\delta(s, v) = s$ and $\{\delta(s, v') : v' \sqsubseteq v\} = S'$, and

3. $\{\delta(s, v') : v' \sqsubseteq v''\} \subset S'$ for all proper prefixes $v'' \sqsubset v$ with $\delta(s, v'') = s$.

and we call a word $w \in X^*$ $(s; S')$-*loop entering* provided

1. $\delta(s_0; v) = s$ and

2. if $w = w' \cdot x$ for some $x \in X$ then $\delta(s_0; w') \notin S'$.

---

[1] We use the same symbol δ to denote the usual extension of the function δ to $S \times X^*$.

3.1 The general case

Denote by $V_{(s;S')}$ the set of all $(s;S')$-loop completing words and by $W_{(s;S')}$ the set of all $(s;S')$-loop entering words. Both languages are regular and constructable from the finite automaton $A = (X;S;s_0;\delta)$. Moreover, $V_{(s;S')}$ is prefix-free, whereas $W_{(s;S')}$ need not be so. Nevertheless, every $\xi \in F_{S'}$ has a unique representation $\xi = w \cdot v_1 \cdots v_i \cdots$ where $w \in W_{(s;S')}$ and $v_i \in V_{(s;S')}$. Here the state $s \in S'$ is uniquely determined as the state succeeding the last state $\hat{s} \notin S'$ in the sequence $(\delta(s_0;u))_{u \sqsubset \xi}$. Thus we obtain the following (see [Sta98b]).

Lemma 4 (Decomposition Lemma) *Let $A = (X;S;s_0;\delta)$ be a finite automaton, $\mathscr{T} \subseteq \{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$ be a table and let $F = \{\xi : \mathrm{Inf}(A;\xi) \in \mathscr{T}\}$. Then*

$$F = \bigcup_{S' \in \mathscr{T}} \bigcup_{s \in S'} \bigcup_{w \in W_{(s;S')}} w \cdot V_{(s;S')}^\omega, \tag{3}$$

and the sets $w \cdot V_{(s;S')}^\omega$ are pairwise disjoint.

As an immediate consequence of the Decomposition Lemma we obtain that every regular ω-language has the form $\bigcup_{i=1}^n W_i \cdot V_i^\omega$ where W_i, V_i are regular languages (see [Büc62, PP04, Sta97] or [Tho90]). The converse is also true, that is, if $W \subseteq X^*$ and $F, E \subseteq X^\omega$ are regular then also $W^\omega, W \cdot E$ and $E \cup F$ are regular ω-languages. Note, however, that the representation of Eq. (3) is much finer, since it splits a regular ω-language $F = \bigcup_{i=1}^n W_i \cdot V_i^\omega$ into mutually disjoint parts $w \cdot V_i^\omega, w \in W_i, i \in \{1,\ldots,n\}$, where, additionally, the languages V_i are prefix-free.

3.2 Decomposition according to Hausdorff measure

Next we are going to construct, depending on the automaton $A = (X;S;s_0;\delta)$, a subset F' of the set F in Eq. (3) on which the Hausdorff measure $\mathbb{L}_\alpha$ is concentrated. To this end we need some properties of the measure $\mathbb{L}_\alpha$ for regular ω-languages.

Since regular ω-languages are Borel sets in Cantor space (cf. [Sta97, Tho90]), $\mathbb{L}_\alpha$ is not only an outer measure but a measure on the class of regular ω-languages. Thus we have the following (cf. [Fal86]).

Proposition 5 *If $(F_i)_{i \in \mathbb{N}}$ is a family of mutually disjoint regular ω-languages then $\mathbb{L}_\alpha(\bigcup_{i \in \mathbb{N}} F_i) = \sum_{i \in \mathbb{N}} \mathbb{L}_\alpha(F_i)$.*

Moreover, the following are shown in [Sta93] and [MS94].

Proposition 6 ([Sta93, Theorem 4.7]) *If $F \subseteq X^\omega$ is a non-empty regular ω-language and $\alpha = \dim F$ then $\mathbb{L}_\alpha(F) > 0$.*

7

Proposition 7 ([Sta93, Theorem 4.6],[MS94, Theorem 6]) *Let $V \subseteq X^*$ be regular and prefix-free. Then $\mathbb{L}_\alpha(V^\omega) = \mathbb{L}_\alpha(\mathscr{C}(V^\omega))$.*

From Eq. (3), Proposition 5 and Proposition 1.3 we obtain a formula for the Hausdorff measure $\mathbb{L}_\alpha(F)$ of F:

$$\mathbb{L}_\alpha(F) = \sum_{S' \in \mathscr{T}} \sum_{s \in S'} \left(\sum_{w \in W_{(s;S')}} r^{-\alpha \cdot |w|} \right) \cdot \mathbb{L}_\alpha(V^\omega_{(s;S')}). \tag{4}$$

The following lemma shows that several of the sets $w \cdot V^\omega_{(s;S')}$ do not contribute to the measure $\mathbb{L}_\alpha(F)$ of F.

Proposition 8 *Let $A = (X;S;s_0;\delta)$ be a finite automaton and $V_{(s;S')} \neq \emptyset$. Then $S'' \subset S'$ implies $V^\omega_{(s;S'')} \subseteq \mathscr{C}(V^\omega_{(s;S')})$.*
  Moreover, we have $\mathbb{L}_\alpha(V^\omega_{(s;S'')}) = 0$ for $\alpha = \dim V^\omega_{(s;S')}$.

Proof. To prove the first assertion it suffices to show $\mathbf{pref}(V^\omega_{(s;S'')}) \subseteq \mathbf{pref}(V^\omega_{(s;S')})$.
  Let $A_s := (X;S;s;\delta)$. Then $\zeta \in V^\omega_{(s;S')}$ if and only if $\mathrm{Inf}(A_s;\zeta) = S'$ and $\{\delta(s,u) : u \sqsubset \xi\} \subseteq S'$. Consequently, for $v \in V^*_{(s;S'')}$ and $\xi \in V^\omega_{(s;S')}$ we have $\mathrm{Inf}(A_s;v \cdot \xi) = S'$ whence $V^*_{(s;S'')} \cdot V^\omega_{(s;S')} = V^\omega_{(s;S')}$ and thus $\mathbf{pref}(V^\omega_{(s;S'')}) \subseteq \mathbf{pref}(V^\omega_{(s;S')})$.
  As $V^\omega_{(s;S'')}$ and $V^\omega_{(s;S')}$ are disjoint subsets of $\mathscr{C}(V^\omega_{(s;S')})$ the second assertion follows from the first one and Proposition 7. ❏

Proposition 8 shows that for an ω-language F accepted by an automaton $A = (X;S;s_0;\delta)$ and a table $\mathscr{T} \subseteq \{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$ the measure $\mathbb{L}_\alpha(F)$ for $\alpha = \dim F$ is concentrated only on subsets $w \cdot V^\omega_{(s,S')}$ for which S' is maximal w.r.t. set inclusion in $\mathscr{T}$.

If $\alpha = \dim F$ and we choose among the maximal sets $S' \in \mathscr{T}$ those for which $\mathbb{L}_\alpha(w \cdot V^\omega_{(s;S')}) > 0$ we eliminate all sets $w \cdot V^\omega_{(s;S')}$ with $\mathbb{L}_\alpha(w \cdot V^\omega_{(s;S')}) = 0$ in Eq. (3) and obtain the following.

Theorem 9 *Let $A = (X;S;s_0;\delta)$ be a finite automaton, $\mathscr{T} \subseteq \{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$ be a table, $F = \{\xi : \mathrm{Inf}(A;\xi) \in \mathscr{T}\}$ and $\alpha = \dim F$.*

If $\mathscr{S} := \{S' : S' \in \mathscr{T} \wedge \exists s(s \in S' \wedge \mathbb{L}_\alpha(V^\omega_{(s,S')}) > 0)\}$ the ω-language $F' = \{\xi : \mathrm{Inf}(A;\xi) \in \mathscr{S}\}$ satisfies $F' \subseteq F$ and $\mathbb{L}_\alpha(F) = \mathbb{L}_\alpha(F')$.

Moreover, the ω-language F' has a decomposition

$$F' = \bigcup_{S' \in \mathscr{S}} \bigcup_{s \in S'} \bigcup_{w \in W_{(s;S')}} w \cdot V^\omega_{(s;S')}, \tag{5}$$

where $\mathbb{L}_\alpha(w \cdot V^\omega_{(s;S')}) > 0$ for all sets $w \cdot V^\omega_{(s;S')}$.

8

3.3 The case of closed ω-languages

In [SW74, Wag79] it was observed that the tables $\mathscr{T}$ of finite automata $A = (X;S;s_0;\delta)$ accepting regular ω-languages closed in Cantor topology have the following simple structure.

Lemma 10 *Let $A = (X;S;s_0;\delta)$ be an initially connected finite automaton and let $\mathscr{T} \subseteq \{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$ be a table such that the ω-language $F = \{\xi : \mathrm{Inf}(A;\xi) \in \mathscr{T}\}$ is closed. Then $\mathscr{T}$ satisfies the following properties.*

1. *If $S' \in \mathscr{T}$, $S'' \in \{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$ and $S' \cap S'' \neq \emptyset$ then $S' \cup S'' \in \mathscr{T}$.*

2. *If $S' \in \mathscr{T}$, $S'' \in \{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$ and $\delta(s'',v) \in S'$ for some $s'' \in S''$ and $v \in X^*$ then $S'' \in \mathscr{T}$.*

Informally speaking, Condition 1 of Lemma 10 shows that the table $\mathscr{T}$ is fully determined by the automaton A and its strongly connected components (SCCs), that is, subsets $S' \in \mathscr{T}$ satisfying the condition $\forall s \forall s'(s,s' \in S' \rightarrow \exists w \exists v(w \neq e \neq v \wedge \delta(s,w) = s' \wedge \delta(s',v) = s))$. In connection with Proposition 8 one observes that strongly connected components are maximal sets in $\{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$.

Condition 2 implies that we can partition the set of states into an accepting part $S_+ := \{s : \exists S' \exists v(S' \in \mathscr{T} \wedge v \in X^* \wedge \delta(s,v) \in S')\}$ and a rejecting part $S_- := S \setminus S_+$ such that $F = \{\xi : \mathrm{Inf}(A;\xi) \subseteq S_+\}$.

As we shall see in the following theorem among the strongly connected components $S' \subseteq S_+$ the terminal ones play a special rôle. A similar observation was made in [MS94, Section 3] in connection with the calculation of the Hausdorff measure of closed regular ω-languages. Here we call a strongly connected component $S' \in \mathscr{T}$ *terminal in* $\mathscr{T}$ provided $\delta(s,v) \in S'$ or $\delta(s,v) \in S_-$ for $s \in S'$ and arbitrary $v \in X^*$.

Theorem 11 *Let $A = (X;S;s_0;\delta)$ be an initially connected finite automaton and let $\mathscr{T} \subseteq \{\mathrm{Inf}(A;\xi) : \xi \in X^\omega\}$ be a table such that the ω-language $F = \{\xi : \mathrm{Inf}(A;\xi) \in \mathscr{T}\}$ is closed. Let $\widehat{\mathscr{T}} \subseteq \mathscr{T}$ be the set of all strongly connected components terminal in $\mathscr{T}$ and $F' = \{\xi : \mathrm{Inf}(A;\xi) \in \widehat{\mathscr{T}}\}$.*

Then $F' \subseteq \bigcup_{S' \in \widehat{\mathscr{T}}} \bigcup_{s \in S'} W_{(s;S')} \cdot \mathscr{C}(V^\omega_{(s;S')}) \subseteq F = \mathscr{C}(F')$.

Moreover, if $s = \delta(s_0,w) \in S'$, for some $S' \in \widehat{\mathscr{T}}$, then $w \cdot X^\omega \cap F = w \cdot \mathscr{C}(V^\omega_{(s;S')})$.

Proof. Obviously, $F' \subseteq F$. First we show that $F \subseteq \mathscr{C}(F')$. Let $\xi \in F$. Then $S'' = \mathrm{Inf}(A;\xi) \in \mathscr{T}$. Choose some $s'' \in S''$. Since the automaton is finite, there are a terminal strongly connected component $S' \in \widehat{\mathscr{T}}$, a state $s' \in S'$ and a $v \in X^*$ such that $\delta(s'',v) = s'$. Consider the set $\mathrm{pref}(\xi;s'') := \{u : u \sqsubseteq \xi \wedge \delta(s_0,u) = s''\}$. Then $\mathbf{pref}(\mathrm{pref}(\xi;s'')) = \mathbf{pref}(\xi)$.

Let $w \in X^* \setminus \{e\}$ satisfy $\delta(s', w) = s'$ and $\{\delta(s', w') : w' \sqsubseteq w\} = S'$. Then $\mathrm{Inf}(A; uvw^\omega) = S'$ for every $u \in \mathrm{pref}(\xi; s'')$, that is, $\mathbf{pref}(\xi) \subseteq \mathbf{pref}(F')$.

Next observe that in view of Lemma 4 we have $F' = \bigcup_{S' \in \widehat{\mathcal{T}}} \bigcup_{s \in S'} W_{(s;S')} \cdot V_{(s;S')}^\omega$. Then the inclusion relations follow from $F' \subseteq F$ and the fact that F is closed.

For the proof of second assertion it suffices to show $w \cdot V_{(s;S')}^\omega \subseteq F \cap w \cdot X^\omega \subseteq w \cdot \mathcal{C}(V_{(s;S')}^\omega)$. If $\xi \in w \cdot V_{(s;S')}^\omega$ then $\mathrm{Inf}(A; \xi) = S'$ whence $\xi \in F \cap w \cdot X^\omega$.

Let $\xi \in F \cap w \cdot X^\omega$. Since $s = \delta(s_0, w) \in S'$ and S' is a terminal strongly connected component, $\mathrm{Inf}(A; \xi) \subseteq S_+$ implies $\{\delta(s_0, w \cdot u) : w \cdot u \sqsubset \xi\} \subseteq S'$. As S' is a strongly connected component, for every $u, w \cdot u \sqsubset \xi$, there is a u' such that $\delta(s_0, w \cdot u \cdot u') = \delta(s_0, w) = s$ and $\{\delta(s_0, w \cdot u'') : u'' \sqsubseteq u \cdot u'\} = S'$. This shows $\mathbf{pref}(\xi) \in \mathbf{pref}(w \cdot V_{(s;S')}^\omega)$, that is, $\xi \in \mathcal{C}(w \cdot V_{(s;S')}^\omega) = w \cdot \mathcal{C}(V_{(s;S')}^\omega)$. $\qquad \square$

4 Results on Hausdorff Measure

4.1 Sets of locally positive measure

Theorem 12 *If $F \subseteq X^\omega$ is a regular ω-language, $\alpha = \dim F$, F has locally positive α-dimensional measure then $\mathbb{L}_\alpha(\mathcal{C}(F)) = \mathbb{L}_\alpha(F)$.*

If, moreover, $\alpha = \dim F = \dim \mathcal{C}(F)$ then $\mathcal{C}(F)$ has locally positive α-dimensional measure.

Proof. It suffices to show that $\mathbb{L}_\alpha(\mathcal{C}(F)) > \mathbb{L}_\alpha(F)$ implies $\mathbb{L}_\alpha(F) = \infty$.

From Theorem 9 we know that F contains a regular ω-language $F' = \bigcup_{i=1}^n W_i \cdot V_i^\omega$ with $\mathbb{L}_\alpha(F) = \mathbb{L}_\alpha(F')$ where the sets W_i, V_i are regular, the V_i are prefix-free with $\mathbb{L}_\alpha(V_i) > 0$, and the sets $w \cdot V_i^\omega, w \in W_i$, $i \in \{1, \dots, n\}$, are mutually disjoint.

Assume $\infty \geq \mathbb{L}_\alpha(\mathcal{C}(F)) > \mathbb{L}_\alpha(F)$. Since F has locally positive α-dimensional measure, by Proposition 2 $\mathbf{pref}(F') = \mathbf{pref}(F)$ whence $\mathcal{C}(F') = \mathcal{C}(F)$.

If $\mathbf{pref}(\xi) \subseteq \mathbf{pref}(W_i \cdot V_i^\omega)$ then there is a $w \in W_i$ such that $\mathbf{pref}(\xi) \setminus \mathbf{pref}(w) \subseteq w \cdot \mathbf{pref}(V_i^\omega)$ or $\mathbf{pref}(\xi) \subseteq \mathbf{pref}(W_i)$. This shows

$$\mathcal{C}(F') = \bigcup_{i=1}^n W_i \cdot \mathcal{C}(V_i^\omega) \cup \bigcup_{i=1}^n \{\xi : \mathbf{pref}(\xi) \subseteq \mathbf{pref}(W_i)\} \ .$$

Since $\mathbb{L}_\alpha(w \cdot \mathcal{C}(V_i^\omega)) = \mathbb{L}_\alpha(w \cdot V_i^\omega)$ the assumption $\mathbb{L}_\alpha(\mathcal{C}(F)) > \mathbb{L}_\alpha(F)$ implies that $\mathbb{L}_\alpha(\{\xi : \mathbf{pref}(\xi) \subseteq \mathbf{pref}(W_i)\}) > 0$ for some $i \in \{1, \dots, n\}$ which in view of Proposition 1.2 yields $\sum_{w \in \mathbf{pref}(W_i)} r^{-\alpha \cdot |w|} = \infty$.

Since W_i is regular, there is a $k \in \mathbb{N}$ such that for every $v \in \mathbf{pref}(W_i)$ there is a $w \in W_i$ with $v \sqsubseteq w$ and $|w| - |v| \leq k$. Thus $\sum_{v \in \mathbf{pref}(W_i)} r^{-\alpha \cdot |v|} = \infty$ implies $\sum_{w \in W_i} r^{-\alpha \cdot |w|} = \infty$ and we obtain $\mathbb{L}_\alpha(F) \geq \mathbb{L}_\alpha(W_i \cdot V_i^\omega) = (\sum_{w \in W_i} r^{-\alpha \cdot |w|}) \cdot \mathbb{L}_\alpha(V_i^\omega) = \infty$.

The additional assertion follows from $0 < \mathbb{L}_\alpha(F \cap w \cdot X^\omega) \leq \mathbb{L}_\alpha(\mathscr{C}(F) \cap w \cdot X^\omega)$ for $w \in \mathbf{pref}(F)$. ❏

The following example shows that the additional assertion need not be true for $\dim F < \dim \mathscr{C}(F)$.

Example 1 *Let $X = \{0,1\}$, $F_1 = \{0,1\}^* \cdot 0^\omega$ and $F_2 := 0^\omega \cup 1 \cdot \{0,1\}^* \cdot 0^\omega$. Then $\mathscr{C}(F_1) = \{0,1\}^\omega$ and $\mathscr{C}(F_2) = 0^\omega \cup 1 \cdot \{0,1\}^\omega$. Then $0 = \dim F_i < \dim \mathscr{C}(F_i) = 1$ for $i = 1,2$, $\mathscr{C}(F_1)$ has locally positive 1-dimensional measure whereas, since $\mathbb{L}_1(F_2 \cap 0 \cdot \{0,1\}^*) = 0$, $\mathscr{C}(F_2)$ has not.*

As an immediate consequence we obtain the following.

Corollary 13 *If $F \subseteq X^\omega$ is a regular ω-language, $\alpha := \dim F$, F has locally positive α-dimensional measure and $\mathbb{L}_\alpha(\mathscr{C}(F) \setminus F) > 0$ then $\mathbb{L}_\alpha(F) = \infty$.*

In case $\mathbb{L}_\alpha(F) = \infty$ the measure of the difference $\mathbb{L}_\alpha(\mathscr{C}(F) \setminus F)$ may be finite or infinite.

Example 2 *Let $X = \{0,1\}$ and $F_1 := 0^* \cdot 1^\omega$ and $F_2 := 0^* \cdot 1^* \cdot 0^\omega$. Both sets are countable, thus $\dim F_1 = \dim F_2 = 0$. We have $\mathscr{C}(F_1) = 0^\omega \cup 0^* \cdot 1^\omega$ and $\mathscr{C}(F_2) = 0^* \cdot 1^\omega \cup 0^* \cdot 1^* \cdot 0^\omega$, and, consequently, $\mathbb{L}_0(\mathscr{C}(F_1) \setminus F_1) = \mathbb{L}_0(0^\omega) = 1$ and $\mathbb{L}_0(\mathscr{C}(F_2) \setminus F_2) = \mathbb{L}_0(0^* \cdot 1^\omega) = \infty$.* ❏

In Theorem 12 the hypothesis that F has locally positive α-dimensional measure is essential. We give an example.

Example 3 *Let $X = \{0,1\}$ and $F := F_1 \cup F_2$ where $F_1 = (0 \cdot \{0,1\})^\omega$ is a closed set and $F_2 = (1 \cdot \{0,1\})^* \cdot (10)^\omega$.*

Then $\mathbb{L}_\alpha(F \cap 1 \cdot \{0,1\}^\omega) = 0$ for $\alpha > 0$, since F_2 is countable. Moreover, $\mathscr{C}(F) = (0 \cdot \{0,1\})^\omega \cup (1 \cdot \{0,1\})^\omega$ and one easily calculates $\dim F = \dim \mathscr{C}(F) = \frac{1}{2}$, $\mathbb{L}_{1/2}(F) = \frac{1}{\sqrt{2}} > 0$ and $\mathbb{L}_{1/2}(\mathscr{C}(F)) = 2 \cdot \mathbb{L}_{1/2}(F) = \sqrt{2}$. ❏

From Proposition 3 and Theorem 12 we obtain the following relationship for the Hausdorff measure of regular ω-power languages.

Corollary 14 *Let $W \subseteq X^*$. If W^ω is a regular ω-language and $\alpha = \dim W^\omega$ then*

$$\mathbb{L}_\alpha(\mathscr{C}(W^\omega)) = \mathbb{L}_\alpha(W^\omega).$$

Corollary 14 and, consequently, Theorem 12 are not valid for non-regular ω-languages. In [Sta05, Section 3.3] examples of prefix-free non-regular languages W fulfilling various relationships between $\mathbb{L}_\alpha(W^\omega)$ and $\mathbb{L}_\alpha(\mathscr{C}(W^\omega))$ are given.

As a further application of Theorem 12 we derive a result which is in some sense a converse to Proposition 2. It shows that for special closed regular ω-languages F we can find a subset of the form of Eq. (5) having the same measure and the same closure as F. Here we use the approach of Theorem 11.

Proposition 15 *Let $A = (X; S; s_0; \delta)$ be an initially connected finite automaton and let $\mathscr{T} \subseteq \{\mathrm{Inf}(A; \xi) : \xi \in X^\omega\}$ be a table such that the ω-language $F = \{\xi : \mathrm{Inf}(A; \xi) \in \mathscr{T}\}$ is closed. Let $\widehat{\mathscr{T}} \subseteq \mathscr{T}$ be the set of all strongly connected components terminal in $\mathscr{T}$ and $F' = \{\xi : \mathrm{Inf}(A; \xi) \in \widehat{\mathscr{T}}\}$.*

If $\dim F = \alpha$ and F has locally positive α-dimensional measure then $F = \mathscr{C}(F')$ and $\mathbb{L}_\alpha(F) = \mathbb{L}_\alpha(F')$.

Proof. The first assertion was already proved in Theorem 11. Then, in view of Theorem 12, it suffices to show that F' has locally positive α-dimensional measure, that is, $\mathbb{L}_\alpha(F' \cap w \cdot X^\omega) > 0$ for all $w \in \mathbf{pref}(F) = \mathbf{pref}(F')$.

Let $v \in \mathbf{pref}(F')$ and consider the identity $F' = \bigcup_{S' \in \widehat{\mathscr{T}}} \bigcup_{s \in S'} W_{(s;S')} \cdot V_{(s;S')}^\omega$ derived in the proof of Theorem 11.

Then there are an $S' \in \widehat{\mathscr{T}}$ and an $s \in S'$ such that $v \in \mathbf{pref}(W_{(s;S')})$ or $v \in W_{(s;S')} \cdot \mathbf{pref}(V_{(s;S')}^*)$. In both cases $v \cdot u \in W_{(s;S')} \cdot V_{(s;S')}^*$ for some $u \in X^*$, in particular $s = \delta(s_0, v \cdot u)$.

By Theorem 11 we have $F \cap v \cdot u \cdot X^\omega = v \cdot u \cdot \mathscr{C}(V_{(s;S')}^\omega)$ and, since $v \cdot u \cdot V_{(s;S')}^\omega \subseteq F'$, with Proposition 8 we obtain $0 < \mathbb{L}_\alpha(F \cap v \cdot u \cdot X^\omega) = \mathbb{L}_\alpha(v \cdot u \cdot V_{(s;S')}^\omega) \leq \mathbb{L}_\alpha(F' \cap v \cdot X^\omega)$. ❏

4.2   The measure of sets residual in its closure

This last part shows that an ω-language having a regular closure and which is topologically large in its closure and has the same measure as its closure.

Before we proceed to the presentation of the results we have to introduce some necessary prerequisites concerning the topology of the Cantor space.

As usual, a countable intersection of open sets is referred to as a $\mathbf{G}_\delta$-set. Moreover, we call a set F *nowhere dense in E* provided $\mathscr{C}(E \setminus \mathscr{C}(F)) = \mathscr{C}(E)$, that is, if $\mathscr{C}(F)$ does not contain a nonempty subset of the form $E \cap w \cdot X^\omega$, and a subset F is referred to as of *first Baire category in E* if F is a countable union of sets nowhere dense in E. If E is closed and F is of first Baire category in E then $E \setminus F$ is referred to as *residual in E*. In particular, $\mathbf{G}_\delta$-sets E in Cantor space are residual in $\mathscr{C}(E)$.

The following lemma shows a connection between Hausdorff dimension and relative density of regular ω-language.

Lemma 16 ([Sta98b, Theorem 8]) *Let $E \subseteq X^\omega$ be a regular ω-language which is closed in Cantor space, $\alpha = \dim E$ and let E have finite and locally positive α-dimensional measure.*

Then every regular ω-language $F \subseteq E$ is of first Baire category in E if and only if $\mathbb{L}_\alpha(F) = 0$.

This much preparatory apparatus leads to the following result.

Theorem 17 *Let $E \subseteq X^\omega$ be an ω-language which is a countable intersection of regular ω-languages, residual in $\mathscr{C}(E)$ and let $\mathscr{C}(E)$ be regular. If $\alpha = \dim \mathscr{C}(E)$, $\mathbb{L}_\alpha(\mathscr{C}(E)) < \infty$, and $\mathscr{C}(E)$ has locally positive α-dimensional measure then $\mathbb{L}_\alpha(\mathscr{C}(E)) = \mathbb{L}_\alpha(E)$.*

Proof. Observe that $\mathscr{C}(E)$ is a regular ω-language. Since E is supposed to be residual in $\mathscr{C}(E)$, $\mathscr{C}(E) \setminus E$ is of first Baire category in $\mathscr{C}(E)$, and, since E is a countable intersection of regular ω-languages, say $E = \bigcap_{i \in \mathbb{N}} F_i$, the difference $\mathscr{C}(E) \setminus E = \bigcup_{i \in \mathbb{N}} (\mathscr{C}(E) \setminus F_i)$ is a countable union of regular ω-languages $\mathscr{C}(E) \setminus F_i$, each of which is of first Baire category in $\mathscr{C}(E)$. Then according to Lemma 16 $\mathbb{L}_\alpha(\mathscr{C}(E) \setminus F_i) = 0$, and the assertion follows. ❏
Using Proposition 15 we can drop the requirement $\mathbb{L}_\alpha(\mathscr{C}(E)) < \infty$ in Theorem 17.

Theorem 18 *Let $E \subseteq X^\omega$ be an ω-language which is a countable intersection of regular ω-languages, residual in $\mathscr{C}(E)$ and let $\mathscr{C}(E)$ be regular. If $\alpha = \dim \mathscr{C}(E)$ and $\mathscr{C}(E)$ has locally positive α-dimensional measure then $\mathbb{L}_\alpha(\mathscr{C}(E)) = \mathbb{L}_\alpha(E)$.*

Proof. The case $\mathbb{L}_\alpha(\mathscr{C}(E)) < \infty$ is proved in Theorem 17. Let $\mathbb{L}_\alpha(\mathscr{C}(E)) = \infty$.
Proposition 15 shows that $\mathbb{L}_\alpha(E') = \infty$ for the set $E' = \bigcup_{S' \in \widehat{\mathcal{F}}} \bigcup_{s \in S'} W_{(s;S')} \cdot V^\omega_{(s;S')}$ derived via Theorem 11 from a finite automaton accepting $\mathscr{C}(E)$. Then, similarly as in the proof of Theorem 12 one finds a set $W_{(s;S')}$ such that $\sum_{w \in W_{(s;S')}} r^{-\alpha \cdot |w|} = \infty$ and $\mathbb{L}_\alpha(V^\omega_{(s;S')}) > 0$.
Again, using Theorem 11 one has $W_{(s;S')} \cdot \mathscr{C}(V^\omega_{(s;S')}) \subseteq \mathscr{C}(E)$, and, moreover, $w \cdot \mathscr{C}(V^\omega_{(s;S')}) = \mathscr{C}(E) \cap w \cdot X^\omega$ for $w \in W_{(s;S')}$.
Now, Propositions 3 and 7 show $\mathbb{L}_\alpha(\mathscr{C}(E) \cap w \cdot X^\omega) < \infty$ and we can apply Theorem 17. This yields $\mathbb{L}_\alpha(\mathscr{C}(E) \cap w \cdot X^\omega) = \mathbb{L}_\alpha(E \cap w \cdot X^\omega) = r^{-\alpha \cdot |w|} \cdot \mathbb{L}_\alpha(V^\omega_{(s;S')})$. Summing over $w \in W_{(s;S')}$ and taking into account that $\mathbb{L}_\alpha(V^\omega_{(s;S')}) > 0$ we obtain $\mathbb{L}_\alpha(E) \geq \sum_{w \in W_{(s;S')}} r^{-\alpha \cdot |w|} \cdot \mathbb{L}_\alpha(V^\omega_{(s;S')}) = \infty$. ❏
The Theorems 17 and 18 can be applied also to non-regular ω-languages. We give a simple example.

Example 4 *Let $E := \bigcap_{w \in X^*} X^* \cdot w \cdot X^\omega$ be the set of all ω-words which contain every word as an infix. Those ω-words are referred to as disjunctive [JST83] or rich [Sta98b]. E is a $\mathbf{G}_\delta$-set in Cantor space, hence residual in $\mathscr{C}(E) = X^\omega$.*

We have $\dim \mathscr{C}(E) = 1$ *and obtain* $\mathbb{L}_1(\mathscr{C}(E)) = \mathbb{L}_1(E) = 1.$ ❑

The condition that E be residual in $\mathscr{C}(E)$ is really essential as the following example shows.

Example 5 *Let* $X = \{0, 1\}$ *and* $E = \{0, 1\}^* \cdot 0^\omega = \bigcap_{n \in \mathbb{N}} \{0, 1\}^* \cdot \{0^n 1, 0\}^\omega$ *which is an intersection of regular ω-languages* $\{0, 1\}^* \cdot \{0^n 1, 0\}^\omega$.

Then $\mathscr{C}(E) = \{0, 1\}^\omega$ *and* $\alpha = \dim \mathscr{C}(E) = 1$, $\mathbb{L}_1(\mathscr{C}(E) \cap w \cdot \{0, 1\}^\omega) > 0$ *for all* $w \in \{0, 1\}^*$ *but, as E is countable,* $\dim E = 0$ *and hence* $\mathbb{L}_1(E) = 0.$ ❑

Acknowledgement

References

[Arn83] André Arnold. Rational ω-languages are nonambiguous. *Theoret. Comput. Sci.*, 26(1-2):221–223, 1983.

[Ban89] Christoph Bandt. Self-similar sets. III. Constructions with sofic systems. *Monatsh. Math.*, 108(2-3):89–102, 1989.

[Büc62] J. Richard Büchi. On a decision method in restricted second order arithmetic. In *Logic, Methodology and Philosophy of Science (Proc. 1960 Internat. Congr .)*, pages 1–11. Stanford Univ. Press, Stanford, Calif., 1962.

[Edg08] Gerald Edgar. *Measure, Topology, and Fractal Geometry*. Undergraduate Texts in Mathematics. Springer, New York, second edition, 2008.

[Fal86] Kenneth Falconer. *The Geometry of Fractal Sets*, volume 85 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1986.

[Fal90] Kenneth Falconer. *Fractal Geometry*. John Wiley & Sons Ltd., Chichester, 1990.

[JST83] Helmut Jürgensen, Huei Jan Shyr, and Gabriel Thierrin. Disjunctive ω-languages. *Elektron. Informationsverarb. Kybernet.*, 19(6):267–278, 1983.

[McN66] Robert McNaughton. Testing and generating infinite sequences by a finite automaton. *Information and Control*, 9:521–530, 1966.

[MS94]   Wolfgang Merzenich and Ludwig Staiger. Fractals, dimension, and formal languages. *RAIRO Inform. Théor. Appl.*, 28(3-4):361–386, 1994.

[MW88]   R. Daniel Mauldin and Stanley C. Williams. Hausdorff dimension in graph directed constructions. *Trans. Amer. Math. Soc.*, 309(2):811–829, 1988.

[PP04]   Dominique Perrin and Jean-Éric Pin. *Infinite Words*, volume 141 of *Pure and Applied Mathematics*. Elsevier, Amsterdam, 2004.

[Sta93]   Ludwig Staiger. Kolmogorov complexity and Hausdorff dimension. *Inform. and Comput.*, 103(2):159–194, 1993.

[Sta97]   Ludwig Staiger. ω-languages. In Grzegorz Rozenberg and Arto Salomaa, editors, *Handbook of Formal Languages*, volume 3, pages 339–387. Springer-Verlag, Berlin, 1997.

[Sta98a]   Ludwig Staiger. The Hausdorff measure of regular ω-languages is computable. *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS*, (66):178–182, 1998.

[Sta98b]   Ludwig Staiger. Rich ω-words and monadic second-order arithmetic. In Mogens Nielsen and Wolfgang Thomas, editors, *Computer science logic (Aarhus, 1997)*, volume 1414 of *Lecture Notes in Computer Science*, pages 478–490. Springer-Verlag, Berlin, 1998. Selected papers from the 11th International Workshop (CSL '97) held at the 6th Annual Conference of the European Association for Computer Science Logic (EACSL) at the University of Aarhus, Aarhus, August 23–29, 1997.

[Sta05]   Ludwig Staiger. Infinite iterated function systems in Cantor space and the Hausdorff measure of ω-power languages. *Internat. J. Found. Comput. Sci.*, 16(4):787–802, 2005.

[SW74]   Ludwig Staiger and Klaus Wagner. Automatentheoretische und automatenfreie Charakterisierungen topologischer Klassen regulärer Folgenmengen. *Elektron. Informationsverarbeit. Kybernetik*, 10:379–392, 1974.

[Tho90]   Wolfgang Thomas. Automata on infinite objects. In Jan van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 133–192. North Holland, Amsterdam, 1990.

[Wag79]   Klaus Wagner. On ω-regular sets. *Inform. and Control*, 43(2):123–177, 1979.