On the Hausdorff measure of regular \( \omega \)-languages in Cantor space

Ludwig Staiger

Institut für Informatik, Martin-Luther-Universität Halle-Wittenberg, Germany


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

The first of these conditions 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.

Keywords: Hausdorff measure, \( \omega \)-language, Muller automata, decomposition, set of locally positive measure

Regular \( \omega \)-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 \( \omega \)-languages have been proved to be computable (cf. [Ban89] [MW88] [Edg08] or [MS94] [Sta98a]). The computation of the Hausdorff measure of a regular \( \omega \)-language uses several properties which do not hold for larger classes of \( \omega \)-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 \( \omega \)-languages closed in Cantor space, and the measure-category-theorem of [Sta98b] shows that this connection can be extended to arbitrary regular \( \omega \)-languages.

Our investigations relate the Hausdorff measure of a subset of the Cantor space to the Hausdorff measure of its closure. The result in Section 4.1 shows that under a certain homogeneity condition the measure of a regular \( \omega \)-language coincides with the measure of its closure. The proof uses the decomposition theorem of [Sta98a] which is based on McNaughton’s theorem [McN69] and extends in some sense earlier decompositions of [Am83], [SW74] and [Wag79]. In our paper the decomposition is directed to a partition of the set of final sets \( T \) of a Muller automaton \( A \) accepting a given \( \omega \)-language

*The results of this paper were presented at the Dagstuhl-Seminar “Topological and Game-Theoretic Aspects of Infinite Computations”, Schloß Dagstuhl, 29.06. - 04.07.2008. A preliminary version appeared in [Sta14].

(i) See also a more recent algebraic decomposition in [PP04] Chapter II, Theorem 9.3.

1365–8050 © 2015 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France
F = Lω(A) into the ones contributing to the Hausdorff measure and the ones not contributing. Crucial for this partition is the fact that only those final sets in T maximal w.r.t. set inclusion can contribute to the Hausdorff measure of the accepted ω-language.

Another result (in Section 4.2) is a sufficient condition under which infinite intersections of regular ω-languages 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] (see also [VY06, Section 4.4]). The extension to sets of infinite measure requires the closer inspection of regular ω-languages closed in Cantor space as given in Section 3.3.

The paper is organised as follows. After introducing some notation in Section 2 several 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 N = {0, 1, 2, ...} we denote the set of natural numbers. Its elements will be usually denoted by letters i, . . . , n. Let X be an alphabet of cardinality |X| = r ≥ 2. Then X∗ is the set of finite words on X, including the empty word e, and Xω is the set of infinite strings (ω-words) over X. Subsets of X∗ will be referred to as languages and subsets of Xω as ω-languages.

For w ∈ X∗ and η ∈ X∗ ∪ Xω let w · η be their concatenation. This concatenation product extends in an obvious way to subsets W ⊆ X∗ and B ⊆ X∗ ∪ Xω. For a language W let W∗ := ∪i∈N W i, and Wω := {w1 · · · wi ∈ W \ {e}} be the set of infinite strings formed by concatenating non-empty words in W. Furthermore, |w| is the length of the word w ∈ X∗ and pref(B) is the set of all finite prefixes of strings in B ⊆ X∗ ∪ Xω. We shall abbreviate w ∈ pref(η) (η ∈ X∗ ∪ Xω) by w ⊆ η.

As usual, we consider Xω as a topological space (Cantor space). The closure (smallest closed set containing F) of a subset F ⊆ Xω, C(F), is described as C(F) := {ξ : pref({ξ}) ⊆ pref(F)}. The open sets in Cantor space are the ω-languages of the form W · Xω.

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 Xω. In the setting of languages and ω-languages this can be read as follows (see [Sta93, Sta98a]). For F ⊆ Xω, r = |X| ≥ 2 and 0 ≤ α ≤ 1 the equation

\[ \mathbb{L}_α(F) := \lim_{l \to \infty} \inf \left\{ \sum_{w \in W} r^{-\alpha |w|} : F \subseteq W \cdot X^\omega \land \forall w (w \in W \to |w| \geq l) \right\} \]

(1)
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 $F \subseteq \{ \xi : \xi \in X^\omega \land \text{pref}(\xi) \subseteq V \}$ and $\sum_{w \in V} r^{-\alpha |w|} < \infty$ then $\mathbb{L}_\alpha(F) = 0$.

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

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

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

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

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

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 $\omega$-language $F$ has locally positive $\alpha$-dimensional measure provided $\mathbb{L}_\alpha(F \cap w \cdot X^\omega) > 0$ for all $w \in \text{pref}(F)$. 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 $\alpha$-dimensional measure. If $F' \subseteq F$ and $\mathbb{L}_\alpha(F') = \mathbb{L}_\alpha(F)$ then $\text{pref}(F') = \text{pref}(F)$ and, consequently, $\mathcal{C}(F') = \mathcal{C}(F)$.

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]) $\mathbb{L}_\alpha(E) = \sum_{w \in X^*} \mathbb{L}_\alpha(E \cap w \cdot X^\omega)$ 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, $\text{pref}(F') \subseteq \text{pref}(F)$. Let now $w \in \text{pref}(F)$. 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 $w \in \text{pref}(F')$. \hfill \Box

We add a further relation of the Hausdorff dimension and the measure $\mathbb{L}_\alpha$ for $\omega$-languages 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 $W^\omega$ has locally positive $\alpha$-dimensional measure.

Proof: The first part is Proposition 6.6 of [Sta93]. Let $w \in \text{pref}(W^\omega)$. 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 $0 < \mathbb{L}_\alpha(wv \cdot W^\omega) \leq \mathbb{L}_\alpha(W^\omega \cap w \cdot X^\omega)$. \hfill \Box
3 Decomposition of Regular $\omega$-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 such that $W = \{ w : \delta(s_0; w) \in S' \}$ for some fixed set $S' \subseteq S$.

An $\omega$-language $F \subseteq X^\omega$ is called regular provided there are a finite (deterministic) automaton $A = (X; S; s_0; \delta)$ and a table $T \subseteq \{ S' : S' \subseteq S \}$ such that for $\xi \in X^\omega$ it holds $\xi \in F$ if and only if $\text{Inf}(A; \xi) \in T$ 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 $\xi$. 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 $\delta(s', v) = s$.

Such sets were referred to as essential sets [Wag79] or loops [Sta97, Section 5.1].

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 $T$ to be contained in the set of loops $\{ \text{Inf}(A; \xi) : \xi \in X^\omega \}$.

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

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. $\{ \delta(s, v') : v' \sqsubseteq v'' \} \subseteq S'$ 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. $\delta(s_0; w) = s$ and
2. if $w = w' \cdot x$ for some $x \in X$ then $\delta(s_0; w') \notin S'$.

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 constructible 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 $\xi = w \cdot v_1 \cdot v_2 \cdots$ 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 $s \notin S'$ in the sequence $\langle \delta(s_0; u) \rangle_{u \in \xi}$. Thus we obtain the following (see [Sta98a, Lemma 3]).

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

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

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

(ii) We use the same symbol $\delta$ to denote the usual extension of the function $\delta$ to $S \times X^*$. 


Hausdorff measure of regular \( \omega \)-languages

As an immediate consequence of the Decomposition Lemma we obtain that every regular \( \omega \)-language has the form \( \bigcup_{i=1}^n W_i \cdot V_i^\omega \) where \( W_i, V_i \) are regular languages (see \[Buč62\], \[PP04\], \[Sta97\] 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 \( \omega \)-languages. Note, however, that the representation of Eq. (3) is much finer, since it splits a regular \( \omega \)-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 \( \omega \)-languages.

Since regular \( \omega \)-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 \( \omega \)-languages. Thus we have the following (cf. \[Pal86\]).

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

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 \( \omega \)-language and \( \alpha = \dim F \) then \( \mathbb{L}_\alpha(F) > 0 \).

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

From Eq. (3), Proposition 5 and Proposition 13 we obtain a formula for the Hausdorff measure \( \mathbb{L}_\alpha(F) \) of \( F \):

\[
\mathbb{L}_\alpha(F) = \sum_{S' \in \mathcal{T}} \sum_{s \in S'} \left( \sum_{\omega \in W(s, S')} r^{-\alpha \cdot |\omega|} \right) \cdot \mathbb{L}_\alpha(V_{\omega(s,S')}^\omega). \tag{4}
\]

The following lemma shows that several of the sets \( w \cdot V_{\omega(s,S')}^\omega \) 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 \( V_{\omega(s,S')}^\omega \) be a non-empty regular \( \omega \)-language and \( \omega = \dim F \). Then \( S'' \subseteq S' \) implies \( V_{\omega(s,S')}^\omega \subseteq C(V_{\omega(s,S')}^\omega) \).

Moreover, we have \( \mathbb{L}_\alpha(V_{\omega(s,S')}^\omega) = 0 \) for \( \alpha = \dim V_{\omega(s,S')}^\omega \).

**Proof:** To prove the first assertion it suffices to show \( \text{pref}(V_{\omega(s,S')}^\omega) \subseteq \text{pref}(V_{\omega(s,S')}^\omega) \).

Let \( A_s := (X; S; s; \delta) \). Then \( \zeta \in V_{\omega(s,S')}^\omega \) if and only if \( \text{Inf}(A_s, s; \zeta) = S' \) and \( \{s(s, u) : u \sqsubseteq \zeta\} \subseteq S' \). Consequently, for \( u \in V_{\omega(s,S')}^\omega \) and \( \tau \in V_{\omega(s,S')}^\omega \) we have \( \text{Inf}(A_s, v; \tau) = S' \) whence \( V_{\omega(s,S')}^\omega \cdot V_{\omega(s,S')}^\omega = V_{\omega(s,S')}^\omega \) and thus \( \text{pref}(V_{\omega(s,S')}^\omega) \subseteq \text{pref}(V_{\omega(s,S')}^\omega) \).

As \( V_{\omega(s,S')}^\omega \) and \( V_{\omega(s,S')}^\omega \) are disjoint subsets of \( C(V_{\omega(s,S')}^\omega) \) the second assertion follows from the first one and Proposition 4.

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

If \( \alpha = \dim F \) and we choose among the maximal sets \( S' \subseteq \mathcal{T} \) those for which \( \mathbb{L}_\alpha(w \cdot V_{\omega(s,S')}^\omega) > 0 \) we eliminate all sets \( w \cdot V_{\omega(s,S')}^\omega \) with \( \mathbb{L}_\alpha(w \cdot V_{\omega(s,S')}^\omega) = 0 \) in Eq. (3) and obtain the following.
Theorem 9 Let \( A = (X; S; s_0; \delta) \) be a finite automaton, \( \mathcal{T} \subseteq \{ \inf(A; \xi) : \xi \in X^\omega \} \) be a table, \( F = \{ \xi : \inf(A; \xi) \in \mathcal{T} \} \) and \( \alpha = \dim F \).

If \( S := \{ S' : S' \in \mathcal{T} \land \exists s \in S' \land \bot\alpha(V^\omega_{s, S'}) > 0 \} \) the \( \omega \)-language \( F' = \{ \xi : \inf(A; \xi) \in S \} \) satisfies \( F' \subseteq F \) and \( \mathbb{L}_\alpha(F') = \mathbb{L}_\alpha(F) \).

Moreover, the \( \omega \)-language \( F' \) has a decomposition
\[
F' = \bigcup_{S' \in S} \bigcup_{s \in S'} \bigcup_{w \in W(e, S')} w \cdot V^\omega_{s, S'},
\]
where \( \mathbb{L}_\alpha(w \cdot V^\omega_{s, S'}) > 0 \) for all sets \( w \cdot V^\omega_{s, S'} \).

3.3 The case of closed \( \omega \)-languages

In [SW74, Wag79] it was observed that the tables \( \mathcal{T} \) of finite automata \( A = (X; S; s_0; \delta) \) accepting regular \( \omega \)-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 \( \mathcal{T} \subseteq \{ \inf(A; \xi) : \xi \in X^\omega \} \) be a table such that the \( \omega \)-language \( F = \{ \xi : \inf(A; \xi) \in \mathcal{T} \} \) is closed. Then \( \mathcal{T} \) satisfies the following properties.

1. If \( S' \in \mathcal{T}, S'' \in \{ \inf(A; \xi) : \xi \in X^\omega \} \) and \( S' \cap S'' \neq \emptyset \) then \( S' \cup S'' \in \mathcal{T} \).

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

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

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

As we shall see in the following theorem the strongly connected components \( S' \subseteq S_+ \) the terminal ones play a special role. A similar observation was made in [MS94 Section 3] in connection with the calculation of the Hausdorff measure of closed regular \( \omega \)-languages. Here we call a strongly connected component \( S' \in \mathcal{T} \) terminal in \( \mathcal{T} \) 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 \( \mathcal{T} \subseteq \{ \inf(A; \xi) : \xi \in X^\omega \} \) be a table such that the \( \omega \)-language \( F = \{ \xi : \inf(A; \xi) \in \mathcal{T} \} \) is closed. Let \( \tilde{\mathcal{T}} \subseteq \mathcal{T} \) be the set of all strongly connected components terminal in \( \mathcal{T} \) and \( F' = \{ \xi : \inf(A; \xi) \in \tilde{\mathcal{T}} \} \).

Then \( F' \subseteq \bigcup_{S' \in \tilde{\mathcal{T}}} \bigcup_{s \in S'} W(e, S') \cdot C(V^\omega_{s, S'}) \subseteq F' = C(F') \).

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

Proof: Obviously, \( F' \subseteq F \). First we show that \( F \subseteq C(F') \). Let \( \xi \in F \). Then \( \xi \in \inf(A; \xi) \in \mathcal{T} \). Choose some \( s'' \in \inf(A; \xi) \in \tilde{\mathcal{T}} \). Since the automaton is finite, there are a terminal strongly connected component \( S' \in \tilde{\mathcal{T}} \), a state \( s' \in S' \) and a \( v \in X^* \) such that \( \delta(s'', v) = s' \). Consider the set \( \text{pref}(\xi; s'') := \{ u : u \subseteq \xi \land \delta(s_0, u) = s'' \} \). Then \( \text{pref}(\xi; s'') = \text{pref}(\xi) \).
Let \( w \in X^* \setminus \{e\} \) satisfy \( \delta(s', w) = s' \) and \( \{\delta(s', w') : w' \sqsubseteq w\} = S' \). Then \( \text{Inf}(A; w w^\omega) = S' \) for every \( u \in \text{pref}(\xi; s''\cdot) \), that is, \( \text{pref}(\xi) \subseteq \text{pref}(F') \).

Next observe that in view of Lemma 4 we have \( F' = \bigcup_{S' \in \mathcal{F}} \bigcup_{s \in S'} W(s; S') \cdot V^\omega_{(s; S')} \). 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 \( w \cdot V^\omega_{(s; S')} \subseteq F \cap w \cdot X^\omega \subseteq w \cdot C(V^\omega_{(s; S')}) \). If \( \xi \in w \cdot V^\omega_{(s; S')} \) then \( \text{Inf}(A; \xi) = S' \) 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, \( \text{Inf}(A; \xi) \subseteq S \) implies \( \{\delta(s_0, w) : w \cdot u \cap \xi \} \subseteq S' \). As \( S' \) is a strongly connected component, for every \( u, \omega \cdot u \cap \xi \), there is a \( u' \) such that \( \delta(s_0, w \cdot u \cdot u') = \delta(s_0, w) = s \) and \( \{\delta(s_0, w \cdot u'') : u'' \subseteq w \cdot u'\} = S' \). This shows \( \text{pref}(\xi) \subseteq \text{pref}(w \cdot V^\omega_{(s; S')}) \), that is, \( \xi \in C(w \cdot V^\omega_{(s; S')}) = w \cdot C(V^\omega_{(s; S')}) \).

4 Results on Hausdorff Measure

4.1 Sets of locally positive measure

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

If, moreover, \( \alpha = \dim F = \dim C(F) \) then \( C(F) \) has locally positive \( \alpha \)-dimensional measure.

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

From Theorem 5 we know that \( F \) contains a regular \( \omega \)-language \( F' = \bigcup_{i=1}^n W_i \cdot V^\omega_i \) with \( \mathcal{L}_\alpha(F') = \mathcal{L}_\alpha(F) \) where the sets \( W_i, V_i \) are regular, the \( V_i \) are prefix-free with \( \mathcal{L}_\alpha(V^\omega_i) > 0 \), and the sets \( W_i \cdot V^\omega_i \) are mutually disjoint.

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

If \( \text{pref}(\xi) \subseteq \text{pref}(W_i \cdot V^\omega_i) \) then there is a \( u \in W_i \) such that \( \text{pref}(\xi) \setminus \text{pref}(u) \subseteq w \cdot \text{pref}(V^\omega_i) \). This shows \( \mathcal{C}(F') = \bigcup_{i=1}^n W_i \cdot C(V^\omega_i) \cup \bigcup_{i=1}^n \{\xi : \text{pref}(\xi) \subseteq \text{pref}(W_i)\} \).

Since \( \mathcal{L}_\alpha(w \cdot C(V^\omega_i)) = \mathcal{L}_\alpha(w \cdot V^\omega_i) \) the assumption \( \mathcal{L}_\alpha(C(F)) > \mathcal{L}_\alpha(F) \) implies \( \mathcal{L}_\alpha(\{\xi : \text{pref}(\xi) \subseteq \text{pref}(W_i)\}) > 0 \) for some \( i \in \{1, \ldots, n\} \) which in view of Proposition 2 yields \( \sum_{u \in \text{pref}(W_i)} r^{-\alpha |w|} = \infty \).

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

The additional assertion follows from \( \mathcal{L}(F \cap w \cdot X^\omega) \leq \mathcal{L}(C(F \cap w \cdot X^\omega)) \) for \( w \in \text{pref}(F) \).

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

**Example 1** Let \( X = \{0, 1\} \), \( F_1 = \{0, 1\}^\omega \cdot 0^\omega \) and \( F_2 := 0^\omega \cup 1 \cdot \{0, 1\}^\omega \). Then \( \mathcal{C}(F_1) = \{0, 1\}^\omega \) and \( \mathcal{C}(F_2) = 0^\omega \cup 1 \cdot \{0, 1\}^\omega \). Then \( \dim F_1 < \dim \mathcal{C}(F_1) = 1 \) for \( i = 1, 2 \). \( \mathcal{C}(F_1) \) has locally positive \( 1 \)-dimensional measure whereas, since \( \mathcal{L}_1(F_2 \cap 0 \cdot \{0, 1\}^\omega) = 0 \), \( \mathcal{C}(F_2) \) has not.
As an immediate consequence we obtain the following.

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

In case \( \mathbb{L}_\alpha(F) = 1 \) the measure of the difference \( \mathbb{L}_\alpha(C(F) \setminus F) \) 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 \( \mathcal{C}(F_1) = 0^\omega \cup 0^* \cdot 1^\omega \) and \( \mathcal{C}(F_2) = 0^* \cdot 1^\omega \cup 0^* \cdot 1^* \cdot 0^\omega \), and, consequently, \( \mathbb{L}_0(C(F_1) \setminus F_1) = \mathbb{L}_0(0^\omega) = 1 \) and \( \mathbb{L}_0(C(F_2) \setminus F_2) = \mathbb{L}_0(0^* \cdot 1^\omega) = \infty \).

In Theorem \[12\] the hypothesis that \( F \) has locally positive \( \alpha \)-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, \( \mathcal{C}(F) = (0 \cdot \{0, 1\})^\omega \cup (1 \cdot \{0, 1\})^\omega \) and one easily calculates \( \dim F = \dim \mathcal{C}(F) = \frac{1}{2} \), \( \mathbb{L}_{1/2}(F) = \frac{1}{\sqrt{2}} > 0 \) and \( \mathbb{L}_{1/2}(\mathcal{C}(F)) = 2 \cdot \mathbb{L}_{1/2}(F) = \sqrt{2} \).

From Proposition \[3\] and Theorem \[12\] we obtain the following relationship for the Hausdorff measure of regular \( \omega \)-power languages.

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

\[
\mathbb{L}_\alpha(C(W^\omega)) = \mathbb{L}_\alpha(W^\omega).
\]

Corollary \[14\] and, consequently, Theorem \[12\] are not valid for non-regular \( \omega \)-languages. In [Sta05 Section 3.5] examples of prefix-free non-regular languages \( W \) fulfilling various relationships between \( \mathbb{L}_\alpha(W^\omega) \) and \( \mathbb{L}^\alpha(C(W^\omega)) \) 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 \( \omega \)-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 \( T \subseteq \{\text{Inf}(A; \xi) : \xi \in X^\omega\} \) be a table such that the \( \omega \)-language \( F' = \{\xi : \text{Inf}(A; \xi) \in T\} \) is closed. Let \( \hat{T} \subseteq T \) be the set of all strongly connected components terminal in \( T \) and \( F = \{\xi : \text{Inf}(A; \xi) \in \hat{T}\} \).

If \( \dim F' = \alpha \) and \( F' \) has locally positive \( \alpha \)-dimensional measure then \( F' = \mathcal{C}(F) \) 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 \( \alpha \)-dimensional measure, that is, \( \mathbb{L}_\alpha(F \cap w \cdot X^\omega) > 0 \) for all \( w \in \text{pref}(F') = \text{pref}(F) \).

Let \( w \in \text{pref}(F) \) and consider the identity \( F = \bigcup_{s' \in \hat{T}} \bigcup_{s \in S} W_{(s; s')} \cdot V_{(s; s')}^\omega \) derived in the proof of Theorem \[11\].

Then there are an \( s' \in \hat{T} \) and an \( s \in S \) such that \( w \in \text{pref}(W_{(s; s')}) \) or \( w \in W_{(s; S')} \cdot \text{pref}(V_{(s; s')}^\omega) \).

In both cases \( v \cdot u \in W_{(s; S')} \cdot V_{(s; s')}^\omega \) for some \( u \in X^\omega \), in particular \( s = \delta(s_0, v \cdot u) \).

By Theorem \[11\] we have \( F' \cap v \cdot u \cdot X^\omega = v \cdot u \cdot \mathcal{C}(V_{(s; s')}^\omega) \) and, since \( v \cdot u \cdot V_{(s; s')}^\omega \subseteq F \), with Proposition \[8\] we obtain \( 0 < \mathbb{L}_{\alpha}(F' \cap v \cdot u \cdot X^\omega) = \mathbb{L}_{\alpha}(v \cdot u \cdot V_{(s; s')}^\omega) \). Thus Theorem \[12\] proves the assertion. \( \square \)
4.2 The measure of sets residual in its closure

This last part shows that an \( \omega \)-language having a regular closure and which is topologically large in its closure 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 \( G_\delta \)-set. Moreover, we call a set \( F \) nowhere dense in \( E \) provided \( C(E \setminus C(F)) = C(E) \), that is, if \( C(F) \) 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, \( G_\delta \)-sets \( E \) in Cantor space are residual in \( C(E) \).

The following lemma shows a connection between Hausdorff dimension and relative density of regular \( \omega \)-languages.

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

Then every regular \( \omega \)-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 \( \omega \)-language which is a countable intersection of regular \( \omega \)-languages, residual in \( C(E) \) and let \( C(E) \) be regular. If \( \alpha = \dim C(E), \mathbb{L}_\alpha(C(E)) < \infty \), and \( C(E) \) has locally positive \( \alpha \)-dimensional measure then \( \mathbb{L}_\alpha(C(E)) = \mathbb{L}_\alpha(E) \).

Proof: Observe that \( C(E) \) is a regular \( \omega \)-language. Since \( E \) is supposed to be residual in \( C(E) \), \( C(E) \setminus E \) is of first Baire category in \( C(E) \), and, since \( E \) is a countable intersection of regular \( \omega \)-languages, say \( E = \bigcap_{i \in \mathbb{N}} F_i \), the difference \( C(E) \setminus E = \bigcup_{i \in \mathbb{N}} (C(E) \setminus F_i) \) is a countable union of regular \( \omega \)-languages \( C(E) \setminus F_i \), each of which is of first Baire category in \( C(E) \). Then according to Lemma 16 \( \mathbb{L}_\alpha(C(E) \setminus F_i) = 0 \), and the assertion follows.

Using Proposition 15 we can drop the requirement \( \mathbb{L}_\alpha(C(E)) < \infty \) in Theorem 17.

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

Proof: The case \( \mathbb{L}_\alpha(C(E)) < \infty \) is proved in Theorem 17. Let \( \mathbb{L}_\alpha(C(E)) = \infty \).

Proposition 15 shows that \( \mathbb{L}_\alpha(E') = \infty \) for the set \( E' = \bigcup_{S' \subseteq \hat{S}} \bigcup_{s \in S'} W(s;S') \cdot V_{(s;S')}^\omega \) derived via Theorem 11 from a finite automaton accepting \( C(E) \). Then, similarly as in the proof of Theorem 12 one finds a set \( \hat{W}_{(s;S')} \) such that \( \sum_{w \in \hat{W}_{(s;S')}} r^{-\alpha |w|} = \infty \) and \( \mathbb{L}_\alpha(V_{(s;S')}^\omega) > 0 \).

Again, using Theorem 11 one has \( W(s;S') \cdot C(V_{(s;S')}^\omega) \subseteq C(E) \), and, moreover, \( w \cdot C(V_{(s;S')}^\omega) = C(E) \cap w \cdot X^\omega \) for \( w \in W(s;S') \).

Now, Propositions 3 and 7 show \( \mathbb{L}_\alpha(C(E) \cap w \cdot X^\omega) < \infty \) and we can apply Theorem 17. This yields \( \mathbb{L}_\alpha(C(E) \cap w \cdot X^\omega) = \mathbb{L}_\alpha(E \cap w \cdot X^\omega) = r^{-\alpha |w|} \cdot \mathbb{L}_\alpha(V_{(s;S')}^\omega) \). Summing over \( w \in W(s;S') \) and taking into account that \( \mathbb{L}_\alpha(V_{(s;S')}^\omega) > 0 \) we obtain \( \mathbb{L}_\alpha(E) \geq \sum_{w \in W(s;S')} r^{-\alpha |w|} \cdot \mathbb{L}_\alpha(V_{(s;S')}^\omega) = \infty \).

The Theorems 17 and 18 can be applied also to non-regular \( \omega \)-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 $\omega$-words which contain every word as an infix. Those $\omega$-words are referred to as disjunctive [JST83] or rich [Sta98b]. $E$ is a $G_\delta$-set in Cantor space, hence residual in $C(E) = X^\omega$.

We have $\dim C(E) = 1$ and obtain $L_1(C(E)) = L_1(E) = 1$. \hfill \Box$

The condition that $E$ be residual in $C(E)$ 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 $\omega$-languages $\{0, 1\}^* \cdot \{0^n 1, 0\}^\omega$.

Then $C(E) = \{0, 1\}^\omega$ and $\alpha = \dim C(E) = 1$, $L_1(C(E) \cap w \cdot \{0, 1\}^\omega) > 0$ for all $w \in \{0, 1\}^*$ but, as $E$ is countable, $\dim E = 0$ and hence $L_1(E) = 0$. \hfill \Box$

Acknowledgements

The author would like to thank Klaus W. Wagner and the two referees for several helpful comments.

References


Hausdorff measure of regular $\omega$-languages


