Topological Entropy of Formal Languages

Friedrich Martin Schneider, Daniel Borchmann

July 14, 2015

In this work we shall consider a notion of complexity of formal languages L that is inspired by the concept of *entropy* from dynamical systems. More precisely, we shall define the *topological entropy* of L to be the exponential growth-rate of the restrictions of the Nerode congruence relation of L to words of length at most n. We shall show that the topological entropy of regular languages is always zero, but that there are also non-regular languages with vanishing entropy, for example Dyck-languages. Furthermore, we shall establish a way of how to compute the entropy of a formal language that is given by a topological automaton accepting it. Finally, we shall point out that the topological entropy of a formal language can be seen as the entropic dimension of a suitable pre-compact pseudo-metric space.

1 Introduction

A variety of notions has been developed to assess different aspects of complexity of formal languages. Most of these notions have been devised with an understanding of complexity in mind that comes with classical complexity theory, and thus these notions are formulated as decision problems. Examples for this are the word problem and the equivalence problem for formal languages, and the complexity of the formal languages is measured by the complexity class for which these problems are complete. Other notions quantify complexity by other means. Examples are the *state complexity* [18] of a regular language, which gives the complexity of the language as the number of states in its minimal automaton, or the *syntactic complexity* of a regular language, which instead considers the size of corresponding syntactic semigroup [10].

The core idea of the present article is to expand the methods of measuring a formal language's complexity by a topological approach in terms of *topological entropy*, which proved tremendously useful to dynamical systems. Topological entropy was introduced by Adler et al. [1] for single homeomorphisms (or continuous transformations) on a compact Hausdorff space. The literature provides several essentially different extensions of this concept for continuous group and semigroup actions. Among others, there is an approach towards topological entropy for continuous actions of finitely generated (pseudo-)groups due to Ghys et al. [13] (see also [3, 5, 7, 16]), which has also been investigated for continuous semigroup actions in [4, 6, 14].

By a dynamical system we mean a continuous semigroup action on a compact Hausdorff topological space. Topological entropy measures the ability of an observer to distinguish between points of the dynamical system just by recognizing transitions *at equal time intervals*, i.e., with respect to a fixed generating system of transformations, starting from the initial state. Since the above notion of dynamical system may very well be regarded as the topological counterpart of a finite automaton, it seems natural to utilize the dynamical approach for applications to automata theory.

To view a formal language L over an alphabet Σ as some kind of dynamical system we take inspiration from the characterization of regular languages as languages whose Myhill-Nerode congruence relation $\Theta(L)$ has finite index. Recall that for $u, v \in \Sigma$ we have

$$(u,v) \in \Theta(L) \iff \forall w \in \Sigma^*\colon (uw \in L \iff vw \in L).$$

The relation $\Theta(L)$ can be seen as some way of measuring the complexity of L: if L is regular, the number of equivalence classes is finite and equals the number of states in the minimal automaton of L. This is the idea behind the notion of state complexity.

However, if L is not regular, $\Theta(L)$ does not have finite index, and asking for the state complexity of L is not a reasonable undertaking. The notion of *topological entropy* that we shall introduce in this paper tries to overcome this issue in the following way. Instead of considering the relation $\Theta(L)$ alone, we shall also take into account how $\Theta(L)$ arises as a limit of a sequences of certain equivalence relations $\Theta(\Sigma^{(n)}, L)$. The behavior of this sequence $(\Theta(\Sigma^{(n)}, L) \mid n \in \mathbb{N})$ then gives rise to our notion of topological complexity of L.

It is the purpose of the this paper to present a first investigation of the notion of topological entropy of formal languages. We shall show that all regular languages have vanishing topological entropy, but that there are also non-regular languages whose topological entropy is zero, most notably Dyck languages [2]. We shall also give examples of context-free languages with non-vanishing entropy. Furthermore, we shall discuss how the topological entropy of a formal language can be computed if the language itself is represented by means of a topological automaton [15].

This paper is structured as follows. We shall first introduce and investigate the topological entropy of formal languages in Section 2. In Section 3 we shall have a closer look into the connection of topological entropy of formal languages and *topological complexity* of semigroup actions on compact Hausdorff spaces. In particular we shall show how the topological entropy of a language can be obtained if the language is given by a topological automaton accepting it. Finally, Section 4 shall show that the topological entropy coincides with the entropic dimension of a suitable precompact pseudo-ultrametric space.

2 Topological entropy of formal languages

In this section we shall introduce our new notion of topological entropy of formal languages. As already sketched in the introduction, this notion is inspired by the characterization of regular languages as those languages whose Nerode congruence has finite index. To make our argumentation easier to follow, we shall thus first recall this famous result.

Thereafter, we shall introduce and investigate our notion of topological entropy. In particular, we shall show that all regular languages have entropy zero, but that there are also non-regular languages with vanishing entropy. This latter discussion shall be embedded in a more general observation about languages defined via groups with sub-exponential growth. Finally, we shall discuss some more examples of languages and determine their entropy.

Let us recall some basic notation. Let Θ be an equivalence relation on a set Y. For $y \in Y$ we put $[y]_\Theta := \{ x \in Y \mid (x,y) \in \Theta \}$. Then $Y/\Theta := \{ [y]_\Theta \mid y \in Y \}$. Furthermore, the *index* of Θ on Y is defined as $\operatorname{ind}(\Theta) := |Y/\Theta|$. For a mapping $f\colon X \to Y$ we set $f^{-1}(\Theta) := \{ (s,t) \in X \times X \mid (f(x), f(y)) \in \Theta \}$. Clearly, $f^{-1}(\Theta)$ then constitutes an equivalence relation on X.

Now let Σ be an alphabet, i.e., a finite and non-empty set. The *Nerode congruence* of a language $L \subseteq \Sigma^*$ is the equivalence relation

$$\Theta(L) := \{(u,v) \in \Sigma^* \times \Sigma^* \mid \forall w \in \Sigma^*\colon uw \in L \Leftrightarrow vw \in L\}.$$

Recall that L is regular if and only if it is accepted by an automaton. The following characterization of regular languages in terms of the Nerode congruence relation is well-known.

Theorem 2.1 (Myhill-Nerode) *Let Σ be a finite alphabet. A language $L \subseteq \Sigma^*$ is regular if and only if $\Theta(L)$ has finite index.*

For regular languages L the number of equivalence classes of the Nerode congruence relation can thus be considered as a measure of complexity of the L. However, if L is not regular this measure is not available anymore. We shall remedy this by not considering the *number* of equivalence classes of $\Theta(L)$, but by considering the *growth* of the number of equivalence classes of a particular approximation of $\Theta(L)$. Based on this growth we introduce our notion of topological entropy of L.

Definition 2.2 *Let Σ be an alphabet. For $F \subseteq \Sigma^*$ finite and $L \subseteq \Sigma^*$ define*

$$\Theta(F,L) := \{(u,v) \in \Sigma^* \times \Sigma^* \mid \forall w \in F\colon uw \in L \Leftrightarrow vw \in L\}. \qquad \triangle$$

Now, the equivalence relations $\Theta(F,L)$ constitute an approximation of $\Theta(L)$ in the sense that

$$\Theta(L) = \bigcap \{ \Theta(F,L) \mid F \subseteq \Sigma^* \text{ finite} \}. \tag{1}$$

Furthermore, it can be seen quite easily that $\Theta(F,L)$ always has finite index. The mapping $\Phi_{F,L}\colon \Sigma^*/\Theta(F,L) \to \{0,1\}^F$ given by

$$\Phi_{F,L}([u]_{\Theta(F,L)})(w) := \begin{cases} 1 & \text{if } uw \in L, \\ 0 & \text{otherwise} \end{cases}$$

is a well-defined injection. Because of this we have $\operatorname{ind}\Theta(F,L) \leq 2^{|F|}$, and thus $\Theta(F,L)$ has finite index. Thus the following definition is reasonable.

Definition 2.3 *Let Σ be an alphabet, and denote with $\mathcal{F}(\Sigma^*)$ the set of finite subsets of Σ^*. Define*

$$\gamma\colon \mathcal{F}(\Sigma^*) \times \mathcal{P}(\Sigma^*) \to \mathbb{N},\ (F, L) \mapsto \operatorname{ind}\Theta(F, L).$$

Given $L \subseteq \Sigma^*$, we call

$$\gamma_L\colon \mathcal{F}(\Sigma^*) \to \mathbb{N},\ F \mapsto \gamma(F, L)$$

the topological complexity function *of L. The* topological entropy *of a language $L \subseteq \Sigma^*$ is defined to be*

$$h(L) := \limsup_{n\to\infty} \frac{\log_2 \gamma_L(\Sigma^{(n)})}{n},$$

where $\Sigma^{(n)}$ is the set of all words over Σ of length at most n. $\triangle$

Next we want to collect some obvious but useful properties of topological complexity functions.

Proposition 2.4 *Let Σ be a finite alphabet, let $E, F \subseteq \Sigma^*$ be finite, and let $L, L_0, L_1 \subseteq \Sigma^*$. Then*

(a) $\gamma(F, \emptyset) = \gamma(F, \Sigma^*) = 1$, *and thus* $h(\emptyset) = h(\Sigma^*) = 0$.
(b) $\gamma(E \cup F, L) \le \gamma(E, L) \cdot \gamma(F, L)$. *If* $E \subseteq F$, *then* $\gamma(E, L) \le \gamma(F, L)$.
(c) $\gamma(F, L) = \gamma(F, \Sigma^* \setminus L)$, *and hence* $h(L) = h(\Sigma^* \setminus L)$.
(d) $\gamma(F, L_0 \cup L_1) \le \gamma(F, L_0) \cdot \gamma(F, L_1)$, *and thus* $h(L_0 \cup L_1) \le h(L_0) + h(L_1)$.
(e) $\gamma(F, L_0 \cap L_1) \le \gamma(F, L_0) \cdot \gamma(F, L_1)$, *and thus* $h(L_0 \cap L_1) \le h(L_0) + h(L_1)$.

Proof Clearly $\Theta(F, \emptyset) = \Theta(F, \Sigma^*) = \Sigma^* \times \Sigma^*$ and hence $\gamma(F, \emptyset) = \gamma(F, \Sigma^*) = 1$. Secondly, because of $\Theta(E \cup F, L) = \Theta(E, L) \cap \Theta(F, L)$ we obtain $\gamma(E \cup F, L) \le \gamma(E, L) \cdot \gamma(F, L)$. In particular, if $E \subseteq F$, then $\Theta(F, L) \subseteq \Theta(E, L)$ and hence $\gamma(E, L) \le \gamma(F, L)$. Moreover, $\Theta(F, L) = \Theta(F, \Sigma^* \setminus L)$ and therefore $\gamma(F, L) = \gamma(F, \Sigma^* \setminus L)$. Finally, it is easy to check that $\Theta(F, L_0) \cap \Theta(F, L_1) \subseteq \Theta(F, L_0 \cup L_1)$. Consequently, $\gamma(F, L_0 \cup L_1) \le \gamma(F, L_0) \cdot \gamma(F, L_1)$.

Utilizing the previous observations, we can furthermore conclude that

$$\begin{aligned}
\Theta(F, L_0) \cap \Theta(F, L_1) &= \Theta(F, \Sigma^* \setminus L_0) \cap \Theta(F, \Sigma^* \setminus L_1) \\
&\subseteq \Theta(F, (\Sigma^* \setminus L_0) \cup (\Sigma^* \setminus L_1)) \\
&= \Theta(F, \Sigma^* \setminus (L_0 \cap L_1)) \\
&= \Theta(F, L_0 \cap L_1)
\end{aligned}$$

and therefore $\gamma(F, L_0 \cap L_1) \le \gamma(F, L_0) \cdot \gamma(F, L_1)$. $\square$

As a consequence of this proposition we immediately obtain the result that the class of languages with zero entropy and the class of languages with finite entropy are closed under Boolean operations. Since it can be seen easily that finite languages have zero entropy we immediately obtain that all regular languages must have zero entropy as well.

The following result gives a precise formulation of this fact, and provides an alternative proof for the claim.

Theorem 2.5 *Let Σ be an alphabet and $L \subseteq \Sigma^*$. The following are equivalent:*

(a) L is regular,
(b) γ_L is bounded, and
(c) there exists some finite subset $F \subseteq \Sigma^*$ such that $\Theta(F, L) = \Theta(L)$.

Proof $(a) \Longrightarrow (b)$. Due to 2.1, $\Theta(L)$ has finite index. Note that $\Theta(L) \subseteq \Theta(F, L)$ and hence $\gamma_L(F) \leq \operatorname{ind} \Theta(L)$ for all $F \subseteq \Sigma^*$ finite. Thus, γ_L is bounded.

$(b) \Longrightarrow (c)$. Suppose that γ_L is bounded. Then there exists some finite $F_0 \subseteq \Sigma^*$ such that $\gamma_L(F_0) = \sup\{\, \gamma_L(F) \mid F \subseteq \Sigma^* \text{ finite} \,\}$. We shall show that $\Theta(F_0, L) = \Theta(L)$. Of course, $\Theta(L) \subseteq \Theta(F_0, L)$. Let $(u, v) \in (\Sigma^* \times \Sigma^*) \setminus \Theta(L)$. By (1) there exists some finite $F_1 \subseteq \Sigma^*$ such that $(u, v) \notin \Theta(F_1, L)$. Obviously, $F_0 \cup F_1 \subseteq \Sigma^*$ is finite and $\Theta(F_0 \cup F_1, L) \subseteq \Theta(F_0, L)$. By assumption, $\gamma_L(F_0 \cup F_1) \leq \gamma_L(F_0)$. Consequently, $\Theta(F_0 \cup F_1, L) = \Theta(F_0, L)$ and therefore $(u, v) \in (\Sigma^* \times \Sigma^*) \setminus \Theta(F_1, L) \subseteq (\Sigma^* \times \Sigma^*) \setminus \Theta(F_0 \cup F_1, L) = (\Sigma^* \times \Sigma^*) \setminus \Theta(F_0, L)$. This substantiates that $\Theta(F_0, L) = \Theta(L)$.

$(c) \Longrightarrow (a)$. By assumption $\Theta(L) = \Theta(F, L)$, and since $\Theta(F, L)$ has finite index, $\Theta(L)$ has finite index as well. Hence, L is regular due to 2.1. □

Corollary 2.6 *Let Σ be an alphabet. If $L \subseteq \Sigma^*$ is regular, then $h(L) = 0$.*

The converse of this corollary does not hold, i.e., there are non-regular languages with vanishing topological entropy. To see this we shall show that *Dyck languages* always have zero entropy (c.f. 2.12). We shall put the corresponding argumentation in a more general framework, by estimating the entropy of languages defined by groups. For this purpose, we recall the concept of *growth* in groups. Consider a finitely generated group G. Let S be a finite symmetric generating subset of G containing the neutral element. The *exponential growth rate* of G with respect to S is defined to be

$$\operatorname{egr}(G, S) := \limsup_{n \to \infty} \frac{\log_2 |S^n|}{n}.$$

Note that this quantity is finite as $|S^n| \leq |S|^n$ for every $n \in \mathbb{N}$. Furthermore,

$$\operatorname{egr}(G, S) := \lim_{n \to \infty} \frac{\log_2 |S^n|}{n}$$

due to a well-known result by Fekete [12]. Of course, the precise value of the exponential growth rate depends upon the particular choice of a generating set.

However, if T is another finite symmetric generating subset of G containing the neutral element, then

$$\frac{1}{k} \cdot \operatorname{egr}(G, T) \leq \operatorname{egr}(G, S) \leq l \cdot \operatorname{egr}(G, T)$$

where $k := \inf\{m \in \mathbb{N} \setminus \{0\} \mid T \subseteq S^m\}$ and $l := \inf\{m \in \mathbb{N} \setminus \{0\} \mid S \subseteq T^m\}$. This justifies the following definition: G is said to have *sub-exponential growth* if $\operatorname{egr}(G, S) = 0$ for some (and thus any) symmetric generating set S of G containing the neutral element. The class of finitely generated groups with sub-exponential growth encompasses all finitely generated abelian groups. In fact, if G is abelian, then

$$S^n \subseteq \left\{ \prod_{s \in S} s^{\alpha(s)} \;\middle|\; \alpha \colon S \to \{\, 0, \dots, n \,\} \right\}$$

and thus $|S^n| \leq (n+1)^{|S|}$ for all $n \in \mathbb{N}$. Now let us return to formal languages.

Theorem 2.7 *Let Σ be an alphabet. Let G be a group, $\varphi\colon \Sigma^* \to G$ a homomorphism, $H \subseteq G$, and $E \subseteq G$ finite. Define*

$$\begin{aligned} P_\varphi(H) &:= \{w \in \Sigma^* \mid \forall u \text{ prefix of } w\colon \varphi(u) \in H\}, \\ L_\varphi(H, E) &:= P_\varphi(H) \cap \varphi^{-1}(E). \end{aligned}$$

Then $\gamma(F, L_\varphi(H, E)) \leq |E| \cdot |\varphi(F)| + 1$ for all finite $F \subseteq \Sigma^*$. In particular,

$$h(L_\varphi(H, E)) \leq \limsup_{n\to\infty} \frac{\log_2 |\varphi(\Sigma^{(n)})|}{n} \leq \log_2 |\Sigma|.$$

Furthermore, if S is a finite symmetric generating subset of G containing the neutral element and $k := \inf\{m \in \mathbb{N} \setminus \{0\} \mid \varphi(\Sigma) \subseteq S^m\}$, then

$$h(L_\varphi(H, E)) \leq k \cdot \operatorname{egr}(G, S).$$

Proof We abbreviate $P := P_\varphi(H)$ and $L := L_\varphi(H, E)$. Consider a finite subset $F \subseteq \Sigma^*$. Then $Q := E\varphi(F)^{-1}$ is a finite subset of G. Fix any object $\infty \notin Q$ and define $Q_\infty := Q \cup \{\infty\}$. Let us consider the map $\psi\colon \Sigma^* \to Q_\infty$ given by

$$\psi(u) := \begin{cases} \varphi(u) & \text{if } u \in P \cap \varphi^{-1}(Q), \\ \infty & \text{otherwise} \end{cases} \qquad (u \in \Sigma^*).$$

We show $\ker\psi \subseteq \Theta(F, L)$. To this end, let $(u, v) \in \ker\psi$. We proceed by case analysis.

First case: $\psi(u) = \psi(v) \neq \infty$. Now, $u, v \in P \cap \varphi^{-1}(Q)$ and $\varphi(u) = \psi(u) = \psi(v) = \varphi(v)$. Let $w \in F$ and suppose that $uw \in L$. We show $vw \in L$. We observe that

$$\varphi(vw) = \varphi(v)\varphi(w) = \varphi(u)\varphi(w) = \varphi(uw) \in E,$$

i.e., $vw \in \varphi^{-1}(E)$. In order to prove that $vw \in P$, let x be a prefix of vw. If x is a prefix of v, then $\varphi(x) \in H$ as $v \in P$. Otherwise, there exists a prefix y of w such that $x = vy$, and so we conclude that $\varphi(x) = \varphi(vy) = \varphi(v)\varphi(y) = \varphi(u)\varphi(y) = \varphi(uy) \in H$, because $uw \in P$ and uy is a prefix of uw. Hence, $vw \in L$. On account of symmetry, it follows that $(u, v) \in \Theta(F, L)$.

Second case: $\psi(u) = \psi(v) = \infty$. Let $x \in \{u, v\}$. If $x \notin \varphi^{-1}(Q)$, then we conclude that $\varphi(xw) = \varphi(x)\varphi(w) \notin E$ and thus $xw \notin L$ for any $w \in F$. If $x \notin P$, then $xw \notin P$ and hence $xw \notin L$ for any $w \in F$. This proves that $\{uw, vw\} \cap L = \emptyset$ for all $w \in F$. Consequently, $(u, v) \in \Theta(F, L)$.

This substantiates that $\ker\psi \subseteq \Theta(F, L)$. Therefore

$$\gamma(F, L) = \operatorname{ind}\Theta(F, L) \leq \operatorname{ind}(\ker\psi) \leq |Q_\infty| \leq |Q| + 1 \leq |E| \cdot |\varphi(F)| + 1.$$

In particular, it follows that

$$\begin{aligned} h(L) = \limsup_{n\to\infty} \frac{\log_2 \gamma_L(\Sigma^{(n)})}{n} &\leq \limsup_{n\to\infty} \frac{\log_2(|E| \cdot |\varphi(\Sigma^{(n)})| + 1)}{n} \\ = \limsup_{n\to\infty} \frac{\log_2 |\varphi(\Sigma^{(n)})|}{n} &\leq \limsup_{n\to\infty} \frac{\log_2(n \cdot |\Sigma|^n)}{n} = \log_2 |\Sigma|. \end{aligned}$$

Finally, suppose S to be a finite symmetric generating subset of G containing the neutral element. Since Σ is finite, $M := \{m \in \mathbb{N} \setminus \{0\} \mid \varphi(\Sigma) \subseteq S^m\}$ is not empty. Let $k := \inf M$. Our considerations above now readily imply that

$$h(L_\varphi(S,E)) \leq \limsup_{n\to\infty} \frac{\log_2 |\varphi(\Sigma^{(n)})|}{n} \leq k \cdot \limsup_{n\to\infty} \frac{\log_2 |S^n|}{n} = k \cdot \mathrm{egr}(G,S). \qquad \square$$

For groups whose growth is sub-exponential the previous theorem yields that the corresponding languages $L_\varphi(S,E)$ have zero entropy.

Corollary 2.8 *Let Σ be an alphabet, let G be a group with sub-exponential growth, and $\varphi\colon \Sigma^* \to G$ a homomorphism. Then for each $S \subseteq G$ and finite $E \subseteq G$, it is true that $h(L_\varphi(S,E)) = 0$.*

We immediately obtain the following statement.

Corollary 2.9 *Let Σ be an alphabet, let G be a finitely generated abelian group, and $\varphi\colon \Sigma^* \to G$ a homomorphism. Then for each $S \subseteq G$ and finite $E \subseteq G$, it is true that $h(L_\varphi(S,E)) = 0$.*

The following corollaries are immediate consequences of Theorem 2.7 for $S = G$.

Corollary 2.10 *Let Σ be a finite alphabet and $L \subseteq \Sigma^*$. Let G be a group, $\varphi\colon \Sigma^* \to G$ a homomorphism and $E \subseteq G$ finite such that $L = \varphi^{-1}(E)$. Then $\gamma(F,L) \leq |E| \cdot |\varphi(F)| + 1$ for all finite $F \subseteq \Sigma^*$. In particular,*

$$h(L) \leq \limsup_{n\to\infty} \frac{\log_2 |\varphi(\Sigma^{(n)})|}{n} \leq \log_2 |\Sigma|.$$

Corollary 2.11 *Let Σ be a finite alphabet, $L \subseteq \Sigma^*$. Let G be an abelian group, $\varphi\colon \Sigma^* \to G$ a homomorphism and $E \subseteq G$ finite such that $L = \varphi^{-1}(E)$. Then $h(L) = 0$.*

With the previous results in place, we are now able to argue that *Dyck languages* have finite entropy. Recall that the *Dyck language with k sorts of parentheses* consists of all balanced strings over $\{\, (_1,)_1, \ldots, (_k,)_k \,\}$. Alternatively, we can view the Dyck language with k sorts of parentheses as the set of all strings that can be reduced to the empty word by successively eliminating matching pairs of parentheses.

We can formalize this as follows. Let $\Sigma, \overline{\Sigma}$ be two alphabets, $\Delta := \Sigma \cup \overline{\Sigma}$, and let $\kappa\colon \Sigma \to \overline{\Sigma}$ be a bijection. Consider the the free group $F(\Sigma)$ with generator set Σ, and denote with $\varphi\colon \Delta^* \to \Sigma$ the unique homomorphism satisfying $\varphi(a) = a$ and $\varphi(\kappa(a)) = a^{-1}$ for all $a \in \Sigma$. Define

$$D(\kappa) := \{w \in \Delta^* \mid \varphi(w) = e \wedge (\forall u \text{ prefix of } w : |w|_a \geq |w|_{\kappa(a)})\}.$$

If $\Sigma = \{\, (_1, \ldots, (_k \,\}$, $\overline{\Sigma} = \{\,)_1, \ldots,)_k \,\}$, and $\kappa((_i) =)_i$, then the set $D(\kappa)$ coincides with the Dyck language with k sorts of parentheses.

Theorem 2.12 *Let $\kappa\colon \Sigma \to \overline{\Sigma}$ be a bijection between finite sets. Then*

$$\log_2|\Sigma| \leq h(D(\kappa)) \leq \operatorname{egr}(F(\Sigma), S)$$

for $S := \Sigma \cup \Sigma^{-1} \cup \{e\}$, where e denotes the neutral element of $F(\Sigma)$.

Proof We first show $\operatorname{ind}\Theta(\Sigma^{(n)}, L) \geq |\Sigma^n|$, since this implies $\log_2|\Sigma| \leq h(D(\kappa))$. For this let $u, v \in \Sigma^n$, $u \neq v$. Define $\kappa(u) := \kappa(u_{|u|}) \dots \kappa(u_1)$, where $u = u_1 \dots u_{|u|}$. Then $u \cdot \kappa(u) \in L$, but $v \cdot \kappa(u) \notin L$. Thus $(u, v) \notin \Theta(\Sigma^{(n)}, L)$ and therefore $\operatorname{ind}\Theta(\Sigma^{(n)}, L) \geq |\Sigma^n|$ as required.

For the second inequality let us consider the unique homomorphism $\psi\colon F(\Sigma) \to \mathbb{Z}^\Sigma$ satisfying

$$\psi(b)(a) := \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{otherwise} \end{cases} \qquad (a, b \in \Sigma).$$

We observe $D(\kappa) = L_\varphi(\psi^{-1}(\mathbb{N}^\Sigma), \{e\})$, where the mapping φ is as above. Hence, we have $h(D(\kappa)) \leq \operatorname{egr}(F(\Sigma), S)$ by 2.9. $\square$

The reason that Dyck languages with more than one type of parentheses have non-zero positive entropy is cause by the requirement that in a word w the different types of parentheses need to mutually balanced, i.e., $\varphi(w) = e$. In other words, if we replace this requirement by the weaker condition that each opening parenthesis has to be closed eventually, then we obtain a class of languages with zero entropy.

Theorem 2.13 *Let $\kappa\colon \Sigma \to \overline{\Sigma}$ be a bijection between finite sets, let $\Delta := \Sigma \cup \overline{\Sigma}$, and consider the language*

$$D'(\kappa) := \{w \in \Delta^* \mid \forall a \in \Sigma\colon (|w|_a = |w|_{\kappa(a)}) \wedge (\forall u \text{ prefix of } w : |w|_a \geq |w|_{\kappa(a)})\}.$$

Then $h(D'(\kappa)) = 0$.

Proof Let us consider the homomorphism $\varphi\colon \Delta^* \to \mathbb{Z}^\Sigma$ given by

$$\varphi(w)(a) := |w|_a - |w|_{\kappa(a)} \qquad (w \in \Delta^*,\ a \in \Sigma).$$

We observe that $D(\kappa) = L_\varphi(\mathbb{N}^\Sigma, \{0\})$, wherefore $h(D(\kappa)) = 0$ by 2.9. $\square$

Note that for $|\Sigma| = 1$ we have $D(\kappa) = D'(\kappa) = 0$, and thus we obtain an example of a language with zero entropy that is not regular. Other non-regular languages with vanishing entropy are discussed in the following examples.

Example 2.14 Let Σ be an alphabet.

(a) Let $m \in \mathbb{N}$ and $a, b \in \Sigma$, $a \neq b$. Then $L := \{w \in \Sigma^* \mid |w|_a = |w|_b + m\}$ is not regular. However, $h(L) = 0$ by Corollary 2.11. To see this, note that the mapping $\varphi\colon \Sigma^* \to \mathbb{Z}$, $w \mapsto |w|_a - |w|_b$ constitutes a homomorphism where $L = \varphi^{-1}(\{m\})$.

(b) Suppose that $\Sigma = \{ a, b, c \}$. Then $L := \{ a^n b^n c^n \mid n \in \mathbb{N} \}$ is not context-free, but $h(L) = 0$. To see this we show that $\Theta = \Theta(\Sigma^{(n)}, L)$ has the equivalence classes

$$\begin{aligned}
&[c^k]_\Theta, \quad k \leq n/2 \\
&[b^\ell c^k]_\Theta, \quad 1 \leq \ell \leq k,\, 2k - \ell \leq n \\
&[a^\ell b^k c^k]_\Theta, \quad 1 \leq \ell \leq k,\, k - \ell \leq n \\
&[b]_\Theta.
\end{aligned} \tag{2}$$

From this it follows that $\operatorname{ind} \Theta(\Sigma^{(n)}, L) \in \mathcal{O}(n^2)$, and thus $h(L) = 0$.
To see that the sets in (2) are indeed all equivalence classes of $\Theta(\Sigma^{(n)}, L)$, let $w \in \Sigma^*$ such that w is not an element of the first three types of classes in (2). We need to show that then $w \in [b]_{\Theta(\Sigma^{(n)}, L)}$. We do this by showing that there is no $u \in \Sigma^{(n)}$ such that $uw \in L$.
Assume by contradiction that such a word u exists. Then u must be of one of the following forms

$$\begin{aligned}
u &= a^k b^k c^\ell, \quad 2k + \ell \leq n,\, 0 \leq \ell \leq k, \\
u &= a^k b^\ell, \quad k + \ell \leq n,\, 0 \leq \ell < k, \\
u &= a^\ell, \quad 0 \leq \ell < n
\end{aligned}$$

If $u = a^k b^k c^\ell$, $2k + \ell \leq n$, $0 \leq \ell \leq k$, then $w = c^{k-\ell}$, $k - \ell \leq n/2$, and therefore $w \in [c^{k-\ell}]_{\Theta(\Sigma^{(n)}, L)}$, a contradiction. If $u = a^k b^\ell$, $k + \ell \leq n, 0 \leq \ell < k$, then $w = b^{k-\ell} c^k$, and $k - \ell > 0$, $2k - (k - \ell) \leq n$, thus $w \in [b^{k-\ell} c^k]_{\Theta(\Sigma^{(n)}, L)}$, again a contradiction. If $u = a^\ell$, then $w = a^{k-\ell} b^k c^k$, and $k - (k - \ell) \leq n$, so $w \in [a^\ell b^k c^k]_{\Theta(\Sigma^{(n)}, L)}$, another contradiction.
Thus, our assumption that u exists is false. The same is true for the word b, and thus $w \in [b]_{\Theta(\Sigma^{(n)}, L)}$, as required. ◇

The following example shows that there are natural examples of "simple" languages whose entropy is not zero, but still finite.

Example 2.15 Suppose $|\Sigma| \geq 2$. Then the *palindrome language*

$$L := \{ w w^R \mid w \in \Sigma^* \}$$

is not regular, but context-free, and $h(L) \in (0, \infty)$.

To see $h(L) > 0$, observe that for each $n \in \mathbb{N}$ and all $u, v \in \Sigma^n$, if $(u, v) \in \Theta(\Sigma^{(n)}, L)$, then $u = v$. This is because of $v v^R \in L$, we also have $u v^R \in L$, and hence $u = v$. Thus

$$[u]_{\Theta(\Sigma^{(n)}, L)} \neq [v]_{\Theta(\Sigma^{(n)}, L)} \quad (u \neq v)$$

Thus $\operatorname{ind} \Theta(\Sigma^{(n)}, L) \geq |\Sigma^n| = |\Sigma|^n$, and we obtain

$$h(L) = \limsup_{n \to \infty} \frac{\log_2 |\Sigma|^n}{n} = \log_2 |\Sigma| > 0.$$

To see $h(L) < \infty$ we shall consider the relation Θ^* defined by

$$(u,v) \in \Theta^* \iff (u,v) \in \Theta(\Sigma^{(n)}, L) \text{ and } (|u| \leq n \iff |v| \leq n).$$

Then $\operatorname{ind}\Theta(\Sigma^{(n)}, L) \leq \operatorname{ind}\Theta^*$. We shall show

$$\limsup_{n\to\infty} \frac{\log_2(\operatorname{ind}\Theta^*)}{n} < \infty.$$

There are at most $|\Sigma^{(n)}|$ many equivalence classes $[u]_{\Theta^*}$ for $u \in \Sigma^*$, $|u| < n$. To count the number of equivalence classes for $|u| \geq n$ we define

$$\ell_n(u) := \{\, a_1 \dots a_i \mid 1 \leq i \leq n, a_1, \dots, a_i \in \Sigma, u = a_1 \dots a_i u', u' \in L \,\}.$$

Then for $u, v \in \Sigma^* \setminus \Sigma^{(n)}$ we have

$$(u,v) \in \Theta^* \iff (u,v) \in \Theta(\Sigma^{(n)}, L) \iff \ell_n(u) = \ell_n(v). \tag{3}$$

The first equivalence is clear. To see the second equivalence let $(u,v) \in \Theta(\Sigma^{(n)}, L)$, and let $a_1 \dots a_i \in \ell_n(u)$. By definition of $\ell_n(u)$ it is then true that

$$u(a_1 \dots a_i)^R \in L.$$

Because $(u,v) \in \Theta(\Sigma^{(n)}, L)$ we therefore obtain

$$v(a_1 \dots a_i)^R \in L,$$

i.e., v is of the form $v = a_1 \dots a_i v'$ for some $v' \in L$. This yields $a_1 \dots a_i \in \ell_n(v)$. By symmetry we obtain $\ell_n(u) = \ell_n(v)$ as required.

Conversely, assume $\ell_n(u) = \ell_n(v)$, and let $w \in \Sigma^{(n)}$ be such that $uw \in L$. Because $|u| \geq n$, there exists $u' \in L$ with $uw = w^R u' w$. Then $w^R \in \ell_n(u) = \ell_n(v)$, and therefore $v = w^R v'$ for some $v' \in L$. But then $vw \in L$. By symmetry $vw \in L \implies uw \in L$ for each $w \in \Sigma^{(n)}$, and therefore $(u,v) \in \Theta(\Sigma^{(n)}, L)$, as required.

Using the characterization from Equation (3) we have

$$\left| {}^{\Sigma^* \setminus \Sigma^{(n)}} \big/ {}_{\Theta^*} \right| = \left| \{\, \ell_n(u) \mid u \in \Sigma^* \setminus \Sigma^{(n)} \,\} \right|$$

Now every set $\ell_n(u)$ with $u = u_1 \dots u_k$, $k \geq n$, can be represented by the prefix $u_1 \dots u_n$ of u of length n together with a tuple $t \in \{\, 0, 1 \,\}^n$ defined by

$$t_i = 1 \iff u_1 \dots u_i \in \ell_n(u).$$

Therefore,

$$\left| {}^{\Sigma^* \setminus \Sigma^{(n)}} \big/ {}_{\Theta^*} \right| = \left| \{\, \ell_n(u) \mid u \in \Sigma^* \setminus \Sigma^{(n)} \,\} \right| \leq |\Sigma|^n \cdot 2^n.$$

This yields

$$\operatorname{ind}\Theta^* = \left|{}^{\Sigma^{(n)}}\!\!\diagup_{\Theta^*}\right| + \left|{}^{\Sigma^* \setminus \Sigma^{(n)}}\!\!\diagup_{\Theta^*}\right| \leq |\Sigma^{(n)}| + |\Sigma|^n \cdot 2^n,$$

and thus

$$\limsup_{n\to\infty} \frac{\log_2(\operatorname{ind}\Theta^*)}{n} = \log_2(2|\Sigma|) < \infty. \qquad \diamond$$

It is not hard to see that the entropy of a formal language can very well be infinite. This is illustrated by the following example

Example 2.16 Let $|\Sigma| \geq 2$, and choose mappings $\varphi_n \colon \Sigma^{2^n} \to \mathfrak{P}(\Sigma^n)$ for each $n \in \mathbb{N}$ such that $|\operatorname{im}(\varphi_n)| = |\Sigma|^{2^n} = 2^{2^n}$. Then define a language $L \subseteq \Sigma^*$ by

$$L \cap \Sigma^m := \begin{cases} \{\, uv \mid u \in \Sigma^{2^n}, v \in \varphi_n(u) \,\} & \text{if } m = 2^n + n \text{ for some } n \in \mathbb{N}, \\ \emptyset & \text{otherwise.} \end{cases}$$

Then $2^{2^n} \leq \gamma_L(n)$, i.e.,

$$2^{2^n} \leq \operatorname{ind}\Theta(\Sigma^n, L). \tag{4}$$

To see this we shall show that each word $\varphi_n(u)$ defines its own equivalence class, i.e., for words $u_0, u_1 \in \Sigma^{2^n}$ with $\varphi_n(u_0) \neq \varphi_n(u_1)$ we have $(u_0, u_1) \notin \Theta(\Sigma^n, L)$. This is because if $\varphi_n(u_0) \neq \varphi_n(u_1)$ we can assume without loss of generality that there exists some word $v \in \varphi_n(u_0) \setminus \varphi_n(u_1)$. By definition of L we then have $u_0 v \in L$, but since $|u_1 v| = 2^n + n$ and $v \notin \varphi_n(u_1)$ we also get $u_1 v \notin L$. Thus $(u_0, u_1) \notin \Theta(\Sigma^n, L)$.

But then (4) implies

$$\limsup_{n\to\infty} \frac{\log_2 \gamma_L(n)}{n} \geq \limsup_{n\to\infty} \frac{\log_2 2^{2^n}}{n} = \infty,$$

and thus $h(L) = \infty$. $\diamond$

3 Entropy of semigroup actions and topological automata

Regular languages are exactly those accepted by finite automata. For some non-regular languages, there exist similar characterizations in terms of finite state machines, e.g., push-down automata, linearly bounded Turing machines, and Turing machines. However, there is a different approach in terms of *topological automata* [15], which in contrast to finite automata have an infinite state set which itself is equipped with a compact Hausdorff topology. In this case, *every* language L is accepted by a topological automaton, and one can ask whether the topological entropy of L can be expressed in terms of a topological automaton accepting it. In this section we shall show that this question has a positive answer.

The concept of topological automata arises from the observation that the transition function of a finite automata is a *monoid action* on the set Σ^* of all words over Σ. Recall that an *action* of a monoid S on a set X is a mapping $\alpha \colon X \times S \to X$ such that

$$\begin{aligned} \alpha(x, e_S) &= x, \\ \alpha(x, st) &= \alpha(\alpha(x,s), t) \end{aligned}$$

for all $x \in X, s, t \in S$.

Recall that a *(deterministic) automaton* over an alphabet Σ is a tuple $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ consisting of a finite set Q of *states*, a *transition function* $\delta \colon Q \times \Sigma \to Q$, a set $F \subseteq Q$ of *final* states, and an *initial state* $q_0 \in Q$. The transition function is usually extended to the set of all words over Σ by virtue of

$$\begin{aligned} \delta^*(q, \varepsilon) &:= q, \\ \delta^*(q, wa) &:= \delta(\delta(q, w), a) \end{aligned}$$

for $q \in Q$, $a \in \Sigma$, and $w \in \Sigma^*$. It is not difficult to see that in this case δ is a monoid action of Σ^* on Q. The *language accepted* by $\mathcal{A}$ is then

$$\mathcal{L}(\mathcal{A}) := \{w \in \Sigma^* \mid \delta^*(q_0, w) \in F\}.$$

We can extend the notion of deterministic finite automata to an infinite state set as follows [15].

Definition 3.1 *A* topological automaton *over an alphabet Σ is a tuple $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ consisting of*

- *a compact Hausdorff space X, called the set of* states *of $\mathcal{A}$*
- *a continuous action α of Σ^* on X, called the* transition function *of $\mathcal{A}$,*
- *a point $x_0 \in X$, called the* initial state *of $\mathcal{A}$, and*
- *a clopen subset $F \subseteq X$, called the set of* final states *of $\mathcal{A}$.*

We say that $\mathcal{A}$ is trim *if $\alpha(x_0, \Sigma^*)$ is dense in X. The* language recognized by *$\mathcal{A}$ is defined as*

$$\mathcal{L}(\mathcal{A}) := \{w \in \Sigma^* \mid \alpha(x_0, w) \in F\}.$$

Let $\mathcal{B} = (Y, \Sigma, \beta, y_0, G)$ be another topological automaton. We shall say that $\mathcal{A}$ and $\mathcal{B}$ are isomorphic*, and write $\mathcal{A} \cong \mathcal{B}$, if there exists a homeomorphism $\varphi \colon X \to Y$ such that*

$$\varphi(\alpha(x, \sigma)) = \beta(\varphi(x), \sigma)$$

for all $x \in X, \sigma \in \Sigma$, $\varphi(x_0) = y_0$, and $\varphi(F) = G$. $\triangle$

Evidently, isomorphic automata accept the same language.

Observe that every automaton accepting L can be turned into an automaton that is trim: if $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ is a topological automaton accepting L, then replacing X with $\overline{\alpha(x_0, \Sigma^*)}$ and F with $F \cap \overline{\alpha(x_0, \Sigma^*)}$ always yields a trim automaton accepting the same language L.

As already stated, and in contrast to regular languages, every formal language $L \subseteq \Sigma^*$ is accepted by a topological automaton, cf. [15, Proposition 2.1].

Proposition 3.2 *Let $L \subseteq \Sigma^*$ and χ_L the characteristic function of L. Equip $X := \{0, 1\}^{\Sigma^*}$ with the product topology, and define the mapping $\delta \colon X \times \Sigma^* \to X$ by*

$$\delta(f, u)(v) := f(uv).$$

Then L is accepted by the topological automaton $(X, \Sigma, \delta, \chi_L, T)$ for $T := \{ f \in X \mid f(\varepsilon) = 1 \}$.

With the notation of 3.2, we define the *minimal automaton of L* to be

$$\mathcal{A}_L = \big(\overline{\chi_L(\Sigma^*)}, \Sigma, \delta, \chi_L, T_L\big),$$

where $\overline{\chi_L(\Sigma^*)}$ is the closure of $\chi_L(\Sigma^*)$ in $\{0,1\}^{\Sigma^*}$, and $T_L = T \cap \overline{\chi_L(\Sigma^*)}$. Clearly, $\mathcal{A}_L$ is trim. Indeed we have the following fact, cf. [15, Theorem 2.2].

Proposition 3.3 *Let $L \subseteq \Sigma^*$, and let $\mathcal{A} = (X, \Sigma, x_0, \delta, F)$ be a topological automaton accepting L. Then $\mathcal{A} \cong \mathcal{A}_L$ if and only if for every automaton $\mathcal{B} = (Y, \Sigma, y_0, \lambda, G)$ accepting L there exists a uniquely determined surjective continuous function*

$$\varphi\colon Y \to X$$

satisfying $\varphi(\lambda(y,\sigma)) = \delta(\varphi(y),\sigma)$, $\varphi(y_0) = x_0$, and $F = \varphi^{-1}(G)$.

Since $\mathcal{A}_L \cong \mathcal{A}_L$, this proposition immediately yields that the minimal automaton is indeed minimal in the above sense. Moreover, in the case that L is regular, $\mathcal{A}_L$ is finite and is the usual minimal automaton of regular languages.

Example 3.4 Let Σ be a finite alphabet and let $a, b \in \Sigma$, $a \neq b$. We consider the *Alexandroff compactification* $\mathbb{Z}_\infty$ of the discrete space of integers $\mathbb{Z}$, that is the set $\mathbb{Z}_\infty = \mathbb{Z} \cup \{\infty\}$ equipped with the topology

$$\{M \subseteq \mathbb{Z} \cup \{\infty\} \mid \infty \in M \Longrightarrow \mathbb{Z} \setminus M \text{ is finite}\}.$$

We define an action α of Σ^* on $\mathbb{Z}_\infty$ by setting $\alpha(m,a) = m+1$, $\alpha(m,b) = m-1$ and $\alpha(m,c) = m$ for all $m \in \mathbb{Z}_\infty$ and $c \in \Sigma \setminus \{a,b\}$. Then α constitutes a continuous action of Σ^* on $\mathbb{Z}_\infty$, and for each $n \in \mathbb{N}$ the topological automaton $\mathcal{A} = (\mathbb{Z}_\infty, \Sigma, \alpha, 0, \{n\})$ accepts the language $L = \big\{w \in \Sigma^* \;\big|\; |w|_a = |w|_b + n\big\}$. $\diamond$

We now shall express the topological complexity of the language L accepted by a topological automaton $\mathcal{A} = (Q, \Sigma, \alpha, x_0, F)$ by the *topological entropy* of the continuous action α of Σ^* on Q [1, 8, 17]. To this end, we shall first fix some useful notation and recall some important definitions about continuous actions on compact Hausdorff spaces.

Let X again be a compact Hausdorff space. We shall denote by $\mathcal{C}(X)$ the set of all finite open covers of X. If $f\colon X \to X$ is continuous and $\mathcal{U} \in \mathcal{C}(X)$, then $f^{-1}(\mathcal{U}) := \{f^{-1}(U) \mid U \in \mathcal{U}\}$ is a finite open cover of X as well. Given $\mathcal{U}, \mathcal{V} \in \mathcal{C}(X)$, we say that $\mathcal{V}$ *refines* $\mathcal{U}$ and write $\mathcal{U} \preceq \mathcal{V}$ if

$$\forall V \in \mathcal{V}\, \exists U \in \mathcal{U}\colon\; V \subseteq U,$$

and we say that $\mathcal{U}$ and $\mathcal{V}$ are *refinement-equivalent* and write $\mathcal{U} \equiv \mathcal{V}$ if $\mathcal{U} \preceq \mathcal{V}$ and $\mathcal{V} \preceq \mathcal{U}$. Furthermore, if $(\mathcal{U}_i \mid i \in I)$ is a finite family of finite open covers of X, then

$$\bigvee_{i\in I} \mathcal{U}_i := \{\, \bigcap_{i\in I} U_i \mid (U_i)_{i\in I} \in \prod_{i\in I} \mathcal{U}_i \,\}.$$

is a finite open cover of X as well. For $\mathcal{U} \in \mathcal{C}(X)$ let

$$N(\mathcal{U}) := \inf\{\, |\mathcal{V}| \mid \mathcal{V} \subseteq \mathcal{U}, X = \bigcup \mathcal{V} \,\}.$$

In preparation for some later considerations, let us recall the following basic observations.

Remark 3.5 ([1]) Let X be a compact Hausdorff space, $\mathcal{U}, \mathcal{V} \in \mathcal{C}(X)$, I be a finite set, $(\mathcal{U}_i)_{i\in I}, (\mathcal{V}_i)_{i\in I} \in \mathcal{C}(X)^I$, and $f\colon X \to X$ be a continuous map. Then the following statements hold:

(1) $\mathcal{U} \preceq \mathcal{V} \implies N(\mathcal{U}) \leq N(\mathcal{V})$,
(2) $\mathcal{U} \preceq \mathcal{V} \implies f^{-1}(\mathcal{U}) \preceq f^{-1}(\mathcal{V})$,
(3) $(\forall i \in I\colon \mathcal{U}_i \preceq \mathcal{V}_i) \implies \bigvee_{i\in I} \mathcal{U}_i \preceq \bigvee_{i\in I} \mathcal{V}_i$. ◇

Now we come to dynamical systems, i.e., continuous semigroup actions. Let S be a semigroup and consider a continuous action $\alpha\colon X \times S \to X$ of S on X, that is, α is supposed to be an action of S on X where $\alpha_s\colon X \to X$, $\alpha_s(x) = \alpha(x,s)$ is continuous for every $s \in S$. For $\mathcal{U} \in \mathcal{C}(X)$ we write

$$s^{-1}(\mathcal{U}) := \alpha_s^{-1}(\mathcal{U}).$$

For every finite $F \subseteq S$ and $\mathcal{U} \in \mathcal{C}(X)$ let

$$(F : \mathcal{U})_\alpha := N\big(\bigvee_{s\in F} s^{-1}(\mathcal{U})\big).$$

Assume F to be a finite generating subset of S. If $\mathcal{U}$ is a finite open cover of X, then we define

$$\eta(\alpha, F, \mathcal{U}) := \limsup_{n\to\infty} \frac{\log_2 (F^n : \mathcal{U})_\alpha}{n}.$$

Furthermore, the *topological entropy of α with respect to F* is defined to be the quantity

$$\eta(\alpha, F) := \sup\{\eta(\alpha, F, \mathcal{U}) \mid \mathcal{U} \in \mathcal{C}(X)\}.$$

Of course, the precise value of this quantity depends on the choice of a finite generating system. However, we observe the following fact.

Proposition 3.6 *Let S be a semigroup and let α be a continuous action of S on some compact Hausdorff space X. Suppose $E, F \subseteq S$ to be finite subsets generating S. Then*

$$\frac{1}{m} \cdot \eta(\alpha, F) \leq \eta(\alpha, E) \leq n \cdot \eta(\alpha, F),$$

where $m := \inf\{k \in \mathbb{N} \mid F \subseteq E^k\}$ and $n := \inf\{k \in \mathbb{N} \mid E \subseteq F^k\}$.

Proof Let $\mathcal{U} \in \mathcal{C}(X)$. Evidently, $(E^k : \mathcal{U}) \leq (F^{kn} : \mathcal{U})$ for all $k \in \mathbb{N}$, whence

$$\begin{aligned}
\eta(\alpha, E, \mathcal{U}) &= \limsup_{k\to\infty} \frac{\log_2 (E^k : \mathcal{U})_\alpha}{k} \leq \limsup_{k\to\infty} \frac{\log_2 (F^{kn} : \mathcal{U})_\alpha}{k} \\
&= n \limsup_{k\to\infty} \frac{\log_2 (F^{kn} : \mathcal{U})_\alpha}{kn} \leq n \limsup_{k\to\infty} \frac{\log_2 (F^k : \mathcal{U})_\alpha}{k} = n \cdot \eta(\alpha, F, \mathcal{U})
\end{aligned}$$

Thus, $\eta(\alpha, E, \mathcal{U}) \leq n \cdot \eta(\alpha, F, \mathcal{U})$. This shows that $\eta(\alpha, E) \leq n\eta(\alpha, F)$. Due to symmetry, it follows that $\eta(\alpha, F) \leq m \cdot \eta(\alpha, E)$ as well. $\square$

We shall now show that the entropy of a formal language is bounded from above by the entropy of any topological automaton accepting it. In the case that the automaton is trim, these two notions even coincide.

Theorem 3.7 *Suppose $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ to be a topological automaton. Consider $S := \Sigma \cup \{\varepsilon\}$ and $\mathcal{U} := \{F, X \setminus F\}$. Then $h(L(\mathcal{A})) \leq \eta(\alpha, S, \mathcal{U})$. If $\mathcal{A}$ is trim, then $h(L(\mathcal{A})) = \eta(\alpha, S, \mathcal{U})$.*

We prove this theorem with the following three auxiliary statements.

Lemma 3.8 *Let $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ be a topological automaton. Let $\Phi\colon \Sigma^* \to X$, $w \mapsto \alpha(x, w)$ and $\mathcal{U} := \{F, X \setminus F\}$. Consider a finite subset $E \subseteq \Sigma^*$ as well as the equivalence relation*

$$\Lambda_E := \{(x, y) \mid \forall w \in E\colon\ \alpha(x, w) \in F \Longleftrightarrow \alpha(y, w) \in F\}.$$

Then the following statements hold:

(1) $X/\Lambda_E = \left(\bigvee_{w \in E} w^{-1}(\mathcal{U})\right) \setminus \{\emptyset\}$.
(2) $\Theta(E, L(\mathcal{A})) = (\Phi \times \Phi)^{-1}(\Lambda_E)$.
(3) If $\mathcal{A}$ is trim, then $\Phi(\Sigma^*) \cap V \neq \emptyset$ for every $V \in X/\Lambda_E$.

Proof (1): We observe that $\mathcal{V} := \left(\bigvee_{w \in E} w^{-1}(\mathcal{U})\right) \setminus \{\emptyset\}$ constitutes a finite partition of X into clopen subsets. For any $V \in \mathcal{V}$ and $x \in V$, we observe that

$$\begin{aligned}
[x]_{\Lambda_E} &= \{y \in Y \mid \forall w \in E\colon \alpha(x, w) \in F \Leftrightarrow \alpha(y, w) \in F\} \\
&= \{y \in Y \mid \forall w \in E\colon x \in w^{-1}(F) \Leftrightarrow y \in w^{-1}(F)\} \\
&= \{y \in Y \mid \forall w \in E\, \forall U \in \mathcal{U} : x \in w^{-1}(U) \Leftrightarrow y \in w^{-1}(U)\} \\
&= \{y \in Y \mid \forall W \in \mathcal{V}\colon x \in W \Leftrightarrow y \in W\} \\
&= \{y \in Y \mid y \in V\} \\
&= V
\end{aligned}$$

We conclude that $X/\Lambda_E = \mathcal{V}$.

(2): Let $L := L(\mathcal{A})$. For any two words $u, v \in \Sigma^*$, it follows that

$$\begin{aligned}
(u, v) \in \Theta(E, L) &\Longleftrightarrow \forall w \in E\colon uw \in L \Leftrightarrow vw \in L \\
&\Longleftrightarrow \forall w \in E\colon \alpha(x_0, uw) \in F \Leftrightarrow \alpha(x_0, vw) \in F \\
&\Longleftrightarrow \forall w \in E\colon \alpha(\alpha(x_0, u), w) \in F \Leftrightarrow \alpha(\alpha(x_0, v), w) \in F \\
&\Longleftrightarrow \forall w \in E\colon \alpha(\Phi(u), w) \in F \Leftrightarrow \alpha(\Phi(v), w) \in F \\
&\Longleftrightarrow (\Phi(u), \Phi(v)) \in \Lambda_E.
\end{aligned}$$

That is, $\Theta(E, L) = (\Phi \times \Phi)^{-1}(\Lambda_E)$.

(3): By (1), the set X/Λ_E is a collection of open, non-empty subsets of X. If $\mathcal{A}$ is trim, then $\Phi(\Sigma^*)$ is dense in X, and thus $\Phi(\Sigma^*) \cap V \neq \emptyset$ for every $V \in X/\Lambda_E$. $\square$

Lemma 3.9 *Let $f\colon X \to Y$ be a mapping and let Θ be an equivalence relation on Y. Then* $\operatorname{ind}(f^{-1}(\Theta)) \leq \operatorname{ind}(\Theta)$. *If $f(X) \cap V \neq \emptyset$ for all $V \in Y/\Theta$, then* $\operatorname{ind}(f^{-1}(\Theta)) = \operatorname{ind}(\Theta)$.

Proof The mapping

$$\Psi\colon X/f^{-1}(\Theta) \to Y/\Theta\colon [x]_{f^{-1}(\Theta)} \mapsto [f(x)]_\Theta$$

is well-defined and injective, since

$$\begin{aligned} [x_1]_{f^{-1}(\Theta)} = [x_2]_{f^{-1}(\Theta)} &\iff (x_1, x_2) \in f^{-1}(\Theta) \\ &\iff (f(x_1), f(x_2)) \in \Theta \\ &\iff [f(x_1)]_\Theta = [f(x_2)]_\Theta. \end{aligned}$$

Therefore $\operatorname{ind}(f^{-1}(\Theta)) \leq \operatorname{ind}(\Theta)$.

If $f(X) \cap V \neq \emptyset$ for each $V \in Y/\Theta$, then Ψ is also surjective: for $V \in Y/\Theta$ let $x \in X$ be such that $f(x) \in V$. Then $V = [f(x)]_\Theta = \Psi([x]_{f^{-1}(\Theta)})$. It follows that $\operatorname{ind}(f^{-1}(\Theta)) = \operatorname{ind}(\Theta)$. □

Proposition 3.10 *Let $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ be a topological automaton and let $\mathcal{U} := \{F,\ X \setminus F\}$. Consider a finite subset $E \subseteq \Sigma^*$. Then $\gamma_{L(\mathcal{A})}(E) \leq (E : \mathcal{U})_\alpha$. Furthermore, if $\mathcal{A}$ is trim, then $\gamma_{L(\mathcal{A})}(E) = (E : \mathcal{U})_\alpha$.*

Proof Let $L := L(\mathcal{A})$ and $\mathcal{V} := \bigvee_{w \in E}(w^{-1}(\mathcal{U}))$. Since $\mathcal{V} \setminus \{\emptyset\}$ constitutes a finite partition of X into clopen subsets, $\mathcal{V} \setminus \{\emptyset\}$ does not admit any proper subcover. Consequently, $N(\mathcal{V}) = |\mathcal{V} \setminus \{\emptyset\}|$. Applying 3.8 and 3.9, we conclude

$$\gamma_L(E) = |\Sigma^*/\Theta(E, L)| \overset{3.8(2),3.9}{\leq} |X/\Lambda_E| \overset{3.8(1)}{=} |\mathcal{V} \setminus \{\emptyset\}| = N(\mathcal{V}) = (E : \mathcal{U})_\alpha.$$

Finally, if $\mathcal{A}$ is trim, then 3.8 (3) and 3.9 assert $|\Sigma^*/\Theta(E, L)| = |X/\Lambda_E|$ and therefore $\gamma_L(E) = (E : \mathcal{U})_\alpha$. □

The particular choice of the cover $\mathcal{U} = \{F, X \setminus F\}$ seems arbitrary, but this is not the case. Indeed, if the automaton $\mathcal{A} = (Q, \Sigma, \alpha, x_0, F)$ is minimal, then the entropy $\eta(\alpha, \Sigma \cup \{\varepsilon\})$ of the automaton equals $\eta(\alpha, \Sigma \cup \{\varepsilon\}, \mathcal{U})$. We shall show this fact in 3.14. As a preparation, we shall first investigate three auxiliary statements.

Lemma 3.11 *Let $f\colon \mathbb{N} \to \mathbb{N}$ be monotone. Then*

$$\limsup_{n\to\infty} \frac{\log_2(f(n))}{n} = \limsup_{n\to\infty} \frac{\log_2(f(n+k))}{n}$$

for every $k \in \mathbb{N}$.

Proof The interesting inequality we have to show is $\geq$. For this we assume without loss of generality that

$$s := \limsup_{n\to\infty} \frac{\log_2(f(n))}{n} < \infty.$$

Let $\varepsilon > 0$. We shall show

$$\limsup_{n\to\infty} \frac{\log_2(f(n+k))}{n} \leq s + \varepsilon.$$

For this we observe that there exists $m \in \mathbb{N}$ such that for all $n \geq m$ we have

$$\log_2(f(n)) \leq n \cdot \Big(s + \frac{\varepsilon}{2}\Big).$$

This implies that

$$\begin{aligned} \frac{\log_2(f(n+k))}{n} &\leq \frac{(n+k)\cdot(s+\varepsilon/2)}{n} \\ &= \Big(s+\frac{\varepsilon}{2}\Big) + \Big(s+\frac{\varepsilon}{2}\Big)\cdot\frac{k}{n} \end{aligned}$$

and for sufficiently large n we obtain

$$\begin{aligned} &\leq s + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= s + \varepsilon, \end{aligned}$$

and the claim follows. □

Lemma 3.12 *Let X be a set, let S be a semigroup, and let $\alpha\colon S\times X \to X$ be an action of S on X. Let $\mathcal{U}$ be a finite cover of X and let $M, N \subseteq S$ be finite. Then*

$$\bigvee_{s\in MN} s^{-1}(\mathcal{U}) \equiv \bigvee_{s\in N} s^{-1}\big(\bigvee_{t\in M} t^{-1}(\mathcal{U})\big).$$

In particular, the complexities of those two covers coincide.

Proof Without loss of generality we may assume that $\mathcal{U}$ is closed under intersection: in fact, $\mathcal{U}$ is refinement-equivalent to the finite cover $\tilde{\mathcal{U}} := \{\bigcap\mathcal{V} \mid \mathcal{V}\subseteq\mathcal{U}\}$. Hence, if the desired statement was true for $\tilde{\mathcal{U}}$, then this would imply

$$\bigvee_{s\in MN} s^{-1}(\mathcal{U}) \equiv \bigvee_{s\in MN} s^{-1}(\tilde{\mathcal{U}}) \equiv \bigvee_{s\in N} s^{-1}\big(\bigvee_{t\in M} t^{-1}(\tilde{\mathcal{U}})\big) \equiv \bigvee_{s\in N} s^{-1}\big(\bigvee_{t\in M} t^{-1}(\mathcal{U})\big)$$

due to the statements (2) and (3) of Remark 3.5.

Henceforth, assume that $\mathcal{U}$ is closed under intersection. We shall show an even stronger claim, namely

$$\bigvee_{s\in MN} s^{-1}(\mathcal{U}) = \bigvee_{s\in N} s^{-1}\big(\bigvee_{t\in M} t^{-1}(\mathcal{U})\big). \tag{5}$$

To ease readability, let us denote the left-hand side by $\mathcal{L}$, and the right-hand side by $\mathcal{R}$.

Let $Y \in \mathcal{L}$. Then

$$Y = \bigcap_{s\in MN} s^{-1}(U_s)$$

for some $(U_s \mid s \in MN) \in \prod_{s \in MN} s^{-1}(\mathcal{U})$. For each $s \in MN$ we can choose $\tau_s \in M, \sigma_s \in N$ such that $s = \tau_s \sigma_s$. Then

$$\begin{aligned}
Y &= \bigcap_{s \in MN} s^{-1}(U_s) \\
&= \bigcap_{s \in MN} (\tau_s \sigma_s)^{-1}(U_s) \\
&= \bigcap_{s \in MN} \sigma_s^{-1}\big(\tau_s^{-1}(U_{\tau_s \sigma_s})\big) \\
&= \bigcap_{\sigma \in N} \bigcap_{\tau \in M} \sigma^{-1}\big(\tau^{-1}(U_{\tau\sigma})\big) \\
&= \bigcap_{\sigma \in N} \sigma^{-1}\big(\bigcap_{\tau \in M} \tau^{-1}(U_{\tau\sigma})\big) \in \mathcal{R}
\end{aligned}$$

Conversely, let $Y \in \mathcal{R}$. Then

$$Y = \bigcap_{\sigma \in N} \sigma^{-1}\big(\bigcap_{\tau \in M} \tau^{-1}(U_{\sigma,\tau})\big)$$

for some $(U_{\sigma,\tau} \mid \sigma \in N, \tau \in M) \in \prod_{(\sigma,\tau) \in M \times N} \mathcal{U}^{M \times N}$. Then

$$\begin{aligned}
Y &= \bigcap_{\sigma \in N} \bigcap_{\tau \in M} \sigma^{-1}\big(\tau^{-1}(U_{\sigma,\tau})\big) \\
&= \bigcap_{\sigma \in N} \bigcap_{\tau \in M} (\tau\sigma)^{-1}(U_{\sigma,\tau}) \\
&= \bigcap_{s \in MN} s^{-1}\Big(\bigcap\big\{U_{\sigma,\tau} \mid \sigma \in N, \tau \in M, s = \tau\sigma\big\}\Big)
\end{aligned}$$

Define

$$U_s := \big\{U_{\sigma,\tau} \mid \sigma \in N, \tau \in M, s = \tau\sigma\big\}.$$

Then $U_s \in \mathcal{U}$ for each $s \in MN$, as $\mathcal{U}$ is closed under intersections. But then

$$Y = \bigcap_{s \in MN} s^{-1}(U_s) \in \mathcal{L}$$

as required.

Finally, Equation 5 and Remark 3.5 (1) yield

$$N\big(\bigvee_{s \in MN} s^{-1}(\mathcal{U})\big) = N\Big(\bigvee_{s \in N} s^{-1}\big(\bigvee_{t \in M} t^{-1}(\mathcal{U})\big)\Big),$$

as it has been claimed. □

Lemma 3.13 *Let $L \subseteq \Sigma^*$ and let $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ be the minimal automaton of L. Consider $S := \Sigma \cup \{\varepsilon\}$ and $\mathcal{U} := \{F,\ X \setminus F\}$. If $\mathcal{V}$ is a finite open cover of X, then there exists some $n \in \mathbb{N}$ such that $\mathcal{V} \preceq \bigvee_{s \in S^n} s^{-1}(\mathcal{U})$.*

Proof For $n \in \mathbb{N}$, let us consider the equivalence relation

$$\Lambda_n := \Lambda_{\Sigma^{(n)}} = \{ (x,y) \in X \times X \mid \forall w \in \Sigma^{(n)} \colon \alpha(w,x) \in F \iff \alpha(w,y) \in F \}$$

(cf. Lemma 3.8). We are going to show that

$$\mathcal{W} := \{[x]_{\Lambda_n} \mid n \in \mathbb{N},\, x \in X,\, \exists V \in \mathcal{V} \colon\ [x]_{\Lambda_n} \subseteq V\}$$

is an open cover of X. By Lemma 3.8 (1), it follows that $\mathcal{W}$ is a collection of open subsets of X. Thus, we only need to argue that $X = \bigcup \mathcal{W}$. To this end, let $x \in X$. Since $\mathcal{V}$ is a cover of X, there exists some $V \in \mathcal{V}$ with $x \in V$. As V is open in X with respect to the subspace topology inherited from $\{0,1\}^{\Sigma^*}$, we find a finite set $E \subseteq \Sigma^*$ such that $W := \{y \in X \mid \forall w \in E \colon\ x(w) = y(w)\} \subseteq V$. Let $n \in \mathbb{N}$ where $E \subseteq S^n$. We observe that

$$\begin{aligned}
[x]_{\Lambda_n} &= \{y \in X \mid \forall w \in S^n \colon\ \alpha(x,w) \in F \iff \alpha(y,w) \in F\} \\
&= \{y \in X \mid \forall w \in S^n \colon\ \alpha(x,w)(\varepsilon) = 1 \iff \alpha(y,w)(\varepsilon) = 1\} \\
&= \{y \in X \mid \forall w \in S^n \colon\ x(w) = 1 \iff y(w) = 1\} \\
&= \{y \in X \mid \forall w \in S^n \colon\ x(w) = y(w)\} \\
&\subseteq W \subseteq V.
\end{aligned}$$

Accordingly, $[x]_{\Lambda_n} \in \mathcal{W}$ and hence $x \in \bigcup \mathcal{W}$. This proves the claim. Now, since X is compact, there exists a finite subset $\mathcal{W}_0$ where $X = \bigcup \mathcal{W}_0$. Due to finiteness of $\mathcal{W}_0$, there is some $n \in \mathbb{N}$ such that $\mathcal{W}_0 \preceq X/\Lambda_n$. We conclude that

$$\mathcal{V} \preceq \mathcal{W} \preceq \mathcal{W}_0 \preceq X/\Lambda_n \overset{3.8(1)}{=} \Big(\bigvee\nolimits_{s \in S^n} s^{-1}(\mathcal{U})\Big) \setminus \{\emptyset\} \equiv \Big(\bigvee\nolimits_{s \in S^n} s^{-1}(\mathcal{U})\Big),$$

which completes the proof. □

Theorem 3.14 *Let $L \subseteq \Sigma^*$ and let $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ be the minimal automaton of L. Consider $S := \Sigma \cup \{\varepsilon\}$ and $\mathcal{U} := \{F,\ X \setminus F\}$. Then $h(L) = \eta(\alpha, S, \mathcal{U}) = \eta(\alpha, S)$.*

Proof Define $\mathcal{U} := \{ F,\ X \setminus F \}$. Since $\mathcal{A}$ is trim, we know that $h(L) = \eta(\alpha, S, \mathcal{U})$ by Theorem 3.7 and hence $h(L) \leq \eta(\alpha, S)$. To show the converse inequality, let $\mathcal{V}$ be a finite open cover of X. We show that $\eta(\alpha, S, \mathcal{V}) \leq \eta(\alpha, S, \mathcal{U})$. According to 3.13, there exists some $m \in \mathbb{N}$ such that $\mathcal{V}$ is refined by $\bigvee_{s \in S^m} s^{-1}(\mathcal{U})$. Then

$$N\Big(\bigvee\nolimits_{s \in S^n} s^{-1}(\mathcal{V})\Big) \leq N\Big(\bigvee\nolimits_{s \in S^n} s^{-1}\big(\bigvee\nolimits_{t \in S^m} t^{-1}(\mathcal{U})\big)\Big) = N\Big(\bigvee\nolimits_{s \in S^{m+n}} s^{-1}(\mathcal{U})\Big)$$

by 3.12. Now we obtain

$$\begin{aligned}
\eta(\alpha, S, \mathcal{V}) = \limsup_{n \to \infty} \frac{\log_2 (S^n : \mathcal{V})_\alpha}{n} &\overset{3.5(1)}{\leq} \limsup_{n \to \infty} \frac{\log_2 (S^{n+m} : \mathcal{U})_\alpha}{n} \\
&\overset{3.11}{=} \limsup_{n \to \infty} \frac{\log_2 (S^n : \mathcal{U})_\alpha}{n} = \eta(\alpha, S, \mathcal{U}).
\end{aligned}$$

Therefore, $\eta(\alpha, S) \leq \eta(\alpha, S, \mathcal{U})$ and hence $\eta(\alpha, S) = \eta(\alpha, S, \mathcal{U}) = h(L)$ by 3.7. □

4 Topological entropy and entropic dimension

Another interesting characterization of the entropy of formal languages is in terms of the entropic of a suitable precompact pseudo-ultrametric space. For this recall that a pseudo-metric space (X, d) is called *precompact* if for each $r \in (0, \infty)$ there exists some finite set $F \subseteq X$ such that

$$X = \bigcup \{ \mathrm{B}_d(x, r) \mid x \in F \}.$$

If (X, d) is a precompact pseudo-metric space, then define

$$\gamma_{(X,d)}(r) := \inf \{ |F| \bigm| F \subseteq X \text{ finite}, X = \bigcup \{ \mathrm{B}_d(x, r) \mid x \in F \} \}.$$

Then the *entropic dimension* $\dim(X, d)$ of the precompact pseudo-metric space (X, d) is defined as [11]

$$\dim(X, d) := \limsup_{r \to 0+} \frac{\log_2(\gamma_{(X,d)}(r))}{\log_2(1/r)}.$$

To now obtain a precompact pseudo-metric space (X, d) whose entropic dimension is the same as the topological entropy of a given language L, we shall first start with a general observation. Let X be a non-empty set and let $\Theta = (\Theta_n \mid n \in \mathbb{N})$ be a descending sequence of equivalence relations on X. Define $d_\Theta \colon X \times X \to [0, \infty)$ as

$$d_\Theta(x, y) := 2^{-\inf\{ n \in \mathbb{N} \mid (x,y) \notin \Theta_n \}} \quad (x, y \in X).$$

It is easy to see that $d_\Theta(x, x) = 0$ and $d_\Theta(x, y) = d_\Theta(y, x)$ is true for all $x, y \in X$. Moreover, as

$$\begin{aligned} \{ n \in \mathbb{N} \mid (x, z) \notin \Theta_n \} &\subseteq \{ n \in \mathbb{N} \mid (x, y) \notin \Theta_n \vee (y, z) \notin \Theta_n \} \\ &= \{ n \in \mathbb{N} \mid (x, y) \notin \Theta_n \} \cup \{ n \in \mathbb{N} \mid (y, z) \notin \Theta_n \}, \end{aligned}$$

we also have $d_\Theta(x, z) \leq \max\{ d_\Theta(x, y), d_\Theta(y, z) \}$ for all $x, y, z \in X$. Because of this (X, d_Θ) is a pseudo-ultrametric space.

Proposition 4.1 *Let X be a non-empty set and let $\Theta = (\Theta_n \mid n \in \mathbb{N})$ be a descending sequence of equivalence relations on X such that each Θ_n has finite index in X. Then (X, d_Θ) is precompact and it is true that*

$$\dim(X, d_\Theta) = \limsup_{n \to \infty} \frac{\log_2 |X/\Theta_n|}{n}.$$

Proof We first observe that for all $x, y \in X$ and $n \in \mathbb{N}$

$$d_\Theta(x, y) < 2^{-n} \iff n < \inf\{ m \in \mathbb{N} \mid (x, y) \notin \Theta_m \} \iff (x, y) \in \Theta_n.$$

Therefore, $X/\Theta_n = \{ \mathrm{B}_{d_\Theta}(x, 2^{-n}) \mid x \in X \}$. Since X/Θ_n is finite, (X, d_Θ) is precompact and

$$\gamma_{(X,d_\Theta)}(2^{-n}) = |X/\Theta_n|.$$

Consequently,

$$
\begin{aligned}
\dim(X, d_\Theta) &= \limsup_{r \to 0+} \frac{\log_2(\gamma_{(X,d_\Theta)}(r))}{\log_2(1/r)} \\
&= \limsup_{n \to \infty} \frac{\log_2(\gamma_{(X,d_\Theta)}(2^{-n}))}{n} \\
&= \limsup_{n \to \infty} \frac{\log_2 |X/\Theta_n|}{n}
\end{aligned}
$$

as required. □

A straightforward application of this lemma is the following theorem.

Corollary 4.2 *Let Σ be an alphabet and let $L \subseteq \Sigma^*$. Then with $\Theta := (\Theta(\Sigma^{(n)}, L) \mid n \in \mathbb{N})$ it is true that*

$$\dim(\Sigma^*, d_\Theta) = h(L).$$

In the case that the language L is represented by a topological automaton we obtain the following result.

Theorem 4.3 *Let $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ be a topological automaton. Let $\Lambda = (\Lambda_n \mid n \in \mathbb{N})$ where*

$$\Lambda_n := \Lambda_{\Sigma^{(n)}} = \{ (x, y) \in X \times X \mid \forall w \in \Sigma^{(n)} \colon \alpha(w, x) \in F \iff \alpha(w, y) \in F \}$$

whenever $n \in \mathbb{N}$ (cf. 3.8). Then $h(L(\mathcal{A})) \leq \dim(X, d_\Lambda)$. Furthermore, if $\mathcal{A}$ is trim, then $h(L(\mathcal{A})) = \dim(X, d_\Lambda)$.

Proof Let $L := L(\mathcal{A})$ and $n \in \mathbb{N}$. We observe that $\gamma_L(\Sigma^{(n)}) = |\Sigma^*/\Theta(\Sigma^{(n)}, L)| \leq X/\Lambda_n$ by 3.8 (2). Moreover, if $\mathcal{A}$ is trim, then $\gamma_L(\Sigma^{(n)}) = X/\Lambda_n$ due to 3.8 (3). Hence, 4.1 yields the desired statements. □

The pseudo-metric considered in the theorem above does not necessarily generate the topology of the respective automaton. In fact, this happens to be true if and only if the automaton is minimal, i.e., isomorphic to the minimal automaton of the accepted language. Furthermore, this case can be characterized in terms of a separation property: a topological automaton is minimal if and only if the induced pseudo-metric is a metric.

Proposition 4.4 *Let $\mathcal{A} = (X, \Sigma, \alpha, x_0, F)$ be a topological automaton and $L := L(\mathcal{A})$. Then the topology generated by d_Λ is contained in the topology of X. Furthermore, the following statements are equivalent:*

(1) $\mathcal{A} \cong \mathcal{A}(L)$.
(2) d_Λ is a metric.
(3) d_Λ generates the topology of X.

Proof By 3.8 (1), the subset $B_{d_\Lambda}(x,\varepsilon) = [x]_{\Lambda_{-\lceil \log_2 \varepsilon \rceil}}$ is open in X for all $x \in X$ and $\varepsilon \in (0,\infty)$. Hence, the topology generated by d_Λ is contained in the original topology of X. Now let us prove the claimed equivalences:

(2)$\Longrightarrow$(3): Suppose that d_Λ is a metric. Then the topology generated by d_Λ is a Hausdorff topology. Since this topology is contained in the compact topology of X, both topologies coincide due to a basic result from set-theoretic topology (see [9, §9.4, Corollary 3]).

(3)$\Longrightarrow$(1): Assume that d_Λ generates the topology of X. This clearly implies d_Λ to be a metric. Consider the unique surjective continuous homomorphism $\varphi\colon \mathcal{A} \to \mathcal{A}(L)$. We are going to show that φ is injective. To this end, let $x,y \in X$ such that $\varphi(x) = \varphi(y)$. We argue that $d_\Lambda(x,y) = 0$. Let $n \in \mathbb{N}$. For every $w \in \Sigma^{(n)}$, we observe that

$$\begin{aligned}\alpha(x,w) \in F &\iff \varphi(\alpha(x,w)) \in T_L \iff \delta(\varphi(x),w) \in T_L \\ &\iff \delta(\varphi(y),w) \in T_L \iff \varphi(\alpha(y,w)) \in T_L \iff \alpha(y,w) \in F.\end{aligned}$$

Thus, $(x,y) \in \Lambda_n$. It follows that $(x,y) \in \bigcap_{n\in\mathbb{N}} \Lambda_n$ and hence $d_\Lambda(x,y) = 0$. Since d_Λ is a metric, we conclude that $x = y$. Accordingly, φ is a bijective continuous map between compact Hausdorff spaces and therefore a homeomorphism. This again is due to an elementary result from set-theoretic topology (see [9, §9.4, Corollary 2]).

(1)$\Longrightarrow$(2): Suppose $\varphi\colon \mathcal{A} \to \mathcal{A}(L)$ to be the necessarily unique isomorphism. Concerning any two points $x,y \in X$, we observe that

$$\begin{aligned}(x,y) \in \Lambda_n(\mathcal{A}) &\iff \forall w \in \Sigma^{(n)}\colon\ \alpha(x,w) \in F \Leftrightarrow \alpha(y,w) \in F \\ &\iff \forall w \in \Sigma^{(n)}\colon\ \varphi(\alpha(x,w)) \in T_L \Leftrightarrow \varphi(\alpha(y,w)) \in T_L \\ &\iff \forall w \in \Sigma^{(n)}\colon\ \delta(\varphi(x),w) \in T_L \Leftrightarrow \delta(\varphi(y),w) \in T_L \\ &\iff (\varphi(x),\varphi(y)) \in \Lambda_n(\mathcal{A}(L))\end{aligned}$$

for every $n \in \mathbb{N}$. Hence, $d_{\Lambda(\mathcal{A})}(x,y) = d_{\Lambda(\mathcal{A}(L))}(\varphi(x),\varphi(y))$ for all $x,y \in X$. Accordingly, it suffices to show that $d_{\Lambda(\mathcal{A}(L))}$ is a metric. To this end, let $f,g \in \overline{\chi_L(\Sigma^*)}$ such that $d_{\Lambda(\mathcal{A}(L))}(f,g) = 0$. We argue that $f = g$. Let $w \in \Sigma^*$. Then there exists $n \in \mathbb{N}$ where $w \in \Sigma^{(n)}$. Since $d_{\Lambda(\mathcal{A}(L))}(f,g) = 0$, we conclude that $(f,g) \in \Lambda_n(\mathcal{A}(L))$ and thus

$$\begin{aligned}f(w) = 1 &\iff \delta(f,w)(\varepsilon) = 1 \iff \delta(f,w) \in T_L \\ &\iff \delta(g,w) \in T_L \iff \delta(g,w)(\varepsilon) = 1 \iff g(w) = 1.\end{aligned}$$

Therefore, $f(w) = g(w)$. It follows that $f = g$. This shows that $d_{\Lambda(\mathcal{A}(L))}$ is a metric and hence completes the proof. □

References

[1] Roy L. Adler, Alan G. Konheim, and Michael H. McAndrew. "Topological Entropy". In: *Transactions of the American Mathematical Society* (2 1965).

[2] Jean-Michel Autebert, Jean Berstel, and Luc Boasson. "Context-free Languages and Pushdown Automata". In: *Handbook of Formal Languages*. Ed. by Grzegorz Rozenberg and Arto Salomaa. Vol. 1. New York, NY, USA: Springer-Verlag, 1997, pp. 111–174.

[3] Andrzej Biś. "An analogue of the variational principle for group and pseudogroup actions". In: *Université de Grenoble. Annales de l'Institut Fourier* 63.3 (2013), pp. 839–863.

[4] Andrzej Biś. "Entropies of a semigroup of maps". In: *Discrete and Continuous Dynamical Systems. Series A* 11.2-3 (2004), pp. 639–648.

[5] Andrzej Biś. "Partial variational principle for finitely generated groups of polynomial growth and some foliated spaces". In: *Colloquium Mathematicum* 110.2 (2008), pp. 431–449.

[6] Andrzej Biś and Mariusz Urbański. "Some remarks on topological entropy of a semigroup of continuous maps". In: *Cubo. A Mathematical Journal* 8.2 (2006), pp. 63–71.

[7] Andrzej Biś and Paweł G. Walczak. "Entropy of distal groups, pseudogroups, foliations and laminations". In: *Annales Polonici Mathematici* 100.1 (2011), pp. 45–54.

[8] François Blanchard, Bernard Host, and Alejandro Maass. "Topological complexity". In: *Ergodic Theory and Dynamical Systems* (3 2000), pp. 641–662.

[9] Nicolas Bourbaki. *Elements of mathematics. General topology. Part 1.* Addison-Wesley Publishing Co., 1966, pp. vii+437.

[10] Janusz A. Brzozowski and Yuli Ye. "Syntactic Complexity of Ideal and Closed Languages". In: *Developments in Language Theory - 15th International Conference, DLT 2011, Milan, Italy, July 19-22, 2011. Proceedings*. Ed. by Giancarlo Mauri and Alberto Leporati. Vol. 6795. Lecture Notes in Computer Science. Springer, 2011, pp. 117–128.

[11] Joel M. Cohen. "Cogrowth and amenability of discrete groups". In: *Journal of Functional Analysis* 48 (3 1982), pp. 301–309.

[12] Michael Fekete. "Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten". In: *Mathematische Zeitschrift* 17.1 (1923), pp. 228–249.

[13] Étienne Ghys, Rémi Langevin, and Paweł G. Walczak. "Entropie géométrique des feuilletages". In: *Acta Mathematica* 160.1-2 (1988), pp. 105–142.

[14] Fagner B. Rodrigues and Paulo Varandas. "Specification properties and thermodynamical properties of semigroup actions". In: *ArXiv e-prints* (2015). arXiv:`1502.01163 [math.DS]`.

[15] Benjamin Steinberg. "Topological dynamics and recognition of languages". In: *CoRR* abs/1306.1468 (2013). URL: `http://arxiv.org/abs/1306.1468`.

[16] Paweł G. Walczak. "Expansion growth, entropy and invariant measures of distal groups and pseudogroups of homeo- and diffeomorphisms". In: *Discrete and Continuous Dynamical Systems. Series A* 33.10 (2013), pp. 4731–4742.

[17] Xinhua Yan and Lianfa He. "Topological Complexity of Semigroup Actions". In: *Journal of Korean Mathematical Society* (1 2008), pp. 221–228.

[18] Sheng Yu. "State Complexity of Regular Languages". In: *Journal of Automata, Languages and Combinatorics* 6 (2000), pp. 221–234.