CDMTCS Research Report Series

Shift-Invariant Topologies for the Cantor Space X^ω

Stefan Hoffmann
Universität Trier

Sibylle Schwarz
HTWK Leipzig

Ludwig Staiger
Martin-Luther-Universität
Halle-Wittenberg

CDMTCS-501
September 2016

Centre for Discrete Mathematics and
Theoretical Computer Science

Shift-Invariant Topologies for the Cantor Space X^ω

Stefan Hoffmann*
Universität Trier, Fachbereich IV – Informatik
D-54286 Trier, Germany

Sibylle Schwarz†
Hochschule für Technik, Wirtschaft und Kultur Leipzig
Fakultät Informatik, Mathematik und Naturwissenschaften
D-04251 Leipzig, Germany

and

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

Abstract

The space of one-sided infinite words plays a crucial rôle in several parts of Theoretical Computer Science. Usually, it is convenient to regard this space as a metric space, the CANTOR space. It turned out that for several purposes topologies other than the one of the CANTOR space are useful, e.g. for studying fragments of first-order logic over infinite words or for a topological characterisation of random infinite words.

It is shown that these topologies refine the topology of the CANTOR space. Moreover, from common features of these topologies we extract properties which characterise a large class of topologies. It turns out that,

*email: `hoffmanns@uni-trier.de`
†email: `sibylle.schwarz@htwk-leipzig.de`
‡email: `ludwig.staiger@informatik.uni-halle.de`

for this general class of topologies, the corresponding closure and interior operators respect the shift operations and also, to some extent, the definability of sets of infinite words by finite automata.

Keywords: CANTOR space, shift-invariance, finite automata, subword metrics, ω-languages

Contents

The space of one-sided infinite words plays a crucial rôle in several parts of Theoretical Computer Science (see [PP04, TB73] and the surveys [HR86, Sta97, Tho90, Tho97]). Several properties known from automata theory or its applications in specification and verification can be characterised using topological properties of the accepted languages of infinite words (see e.g. [Sch04] or the recent survey [DMW15]). Usually, it is convenient to regard this space as a topological space provided with the CANTOR topology.

It turned out that for several purposes other topologies on the space of infinite words are also useful [Red86, Sta87], e.g. for investigations in first-order logic [DK11], to characterise the set of random infinite words [CMS03] or the set of disjunctive infinite words [Sta05] and to describe the convergence behaviour of not necessarily hyperbolic iterative function systems [FS01, Sta03].

Most of these papers use topologies on the space of infinite words which are refinements of the CANTOR topology showing a kind of shift invariance. The aim of this paper is to give a unified treatment of those topologies and to investigate their relations to the CANTOR topology.

Special attention is paid to subsets of the space of infinite words definable by finite automata. It turns out that several of the refinements of the CANTOR topology under consideration behave well with respect to finite automata, that is, the corresponding closure and interior operators preserve at least one of the classes of finite-state or regular ω-languages.

The paper is organised as follows. After some necessary notation in Section 1 we introduce the concept of general topological spaces and we derive a sufficient condition for their metrisability. Then in Section 2 we consider the CANTOR space provided with its classical metric topology and we briefly introduce ω-languages (sets of infinite words) definable by finite automata and their properties w.r.t. the CANTOR topology. Though the paper deals with automaton-definable ω-languages we will refer to automata explicitly mainly in Section 2.2. As described in [DMW15] automata-theoretic properties used here are in fact topological properties.

The third section is devoted to general properties of shift-invariant topologies on the CANTOR space. Then in the fourth part we investigate four particular shift-invariant topologies generated by bases consisting of automaton-definable ω-languages. Two subword metrics and their topologies are the subject of Section 5. Here also automaton-definable ω-languages play a major rôle. All seven topologies are compared with each other in the sixth section, and the final Section 5.4 draws a connection to the subword complexity of infinite words.

The paper combines, in a self-contained manner, results of the conference papers [SS10] and [HS15].

1 Notation and Preliminaries

We introduce the notation used throughout the paper. By $\mathbb{N} = \{0,1,2,\ldots\}$ we denote the set of natural numbers. Let X be a finite alphabet of cardinality $|X| \geq 2$. By X^* we denote the set (monoid) of words on X, including the *empty word* e, and X^ω is the set of infinite sequences (ω-words) over X. For $w \in X^*$ and $\eta \in X^* \cup X^\omega$ let $w \cdot \eta$ be their *concatenation*. This concatenation product extends in an obvious way to subsets $W \subseteq X^*$ and $P \subseteq X^* \cup X^\omega$. For a language W let $W^* := \bigcup_{i\in\mathbb{N}} W^i$ be the *submonoid* of X^* generated by W, and by $W^\omega := \{w_1 \cdots w_i \cdots : w_j \in W \setminus \{e\}\}$ we denote the set of infinite strings formed by concatenating words in W. Furthermore $|w|$ is the *length* of the word $w \in X^*$ and $\mathbf{pref}(P)$ ($\mathbf{infix}(P)$) is the set of all finite prefixes (infixes) of strings in $P \subseteq X^* \cup X^\omega$. We shall abbreviate $w \in \mathbf{pref}(\eta)$ ($\eta \in X^* \cup X^\omega$) by $w \sqsubseteq \eta$. If $\xi \in X^\omega$ by $\mathbf{infix}^\infty(\xi) \subseteq \mathbf{infix}(\xi)$ we denote the set of infixes occurring infinitely often in ξ.

Further we denote by $P/w := \{\eta : w \cdot \eta \in P\}$ the *left derivative* or *state* of the set $P \subseteq X^* \cup X^\omega$ generated by the word w. We refer to P as *finite-state* provided the set of states $\{P/w : w \in X^*\}$ is finite. It is well-known that a language $W \subseteq X^*$ is finite state if and only if it is accepted by a finite automaton, that is, it is a regular language.[1]

Regular ω-languages, that is, ω-languages accepted by finite automata, are the finite unions of sets of the form $W \cdot V^\omega$, where W and V are regular languages (cf. e.g. [Sta97]). Every regular ω-language is finite-state, but, as it was observed in [Tra62], not every finite-state ω-language is regular (cf. also [Sta83]).

1.1 Topological spaces in general

A topological space is a pair $(\mathscr{X}, \mathscr{O})$ where $\mathscr{X}$ is a non-empty set and $\mathscr{O} \subseteq 2^{\mathscr{X}}$ is a family of subsets of $\mathscr{X}$ containing $\mathscr{X}$ itself and being closed under arbitrary union and under finite intersection. The family $\mathscr{O}$ is usually called the family of *open* subsets of the space $\mathscr{X}$. Their complements are referred to as *closed* sets of the space $\mathscr{X}$.

As usually, a set $\mathbb{B} \subseteq \mathscr{O}$ is a *base* for a topology $(\mathscr{X}, \mathscr{O})$ provided every set $M \in \mathscr{O}$ is the (possibly empty) union of sets from $\mathbb{B}$. Thus it does no harm if one considers bases containing $\emptyset$. It is well-known that a family of subsets $\mathbb{B}$ of a set $\mathscr{X}$ which is closed under finite intersection generates a topology on $\mathscr{X}$ in this way.

KURATOWSKI observed that topological spaces can be likewise defined using closure or interior operators. A topological *interior* operator $\mathscr{J}$ is a mapping

[1] Observe that the relation $\sim_P$ defined by $w \sim_P v$ iff $P/w = P/v$ is the NERODE right congruence of P.

$\mathcal{J} : 2^{\mathcal{X}} \to 2^{\mathcal{X}}$ satisfying the following relations.

$$\begin{array}{rcl} \mathcal{J}\mathcal{X} &=& \mathcal{X} \\ M &\supseteq& \mathcal{J}M \;=\; \mathcal{J}\mathcal{J}M \text{ , and} \\ \mathcal{J}(M_1 \cap M_2) &=& \mathcal{J}M_1 \cap \mathcal{J}M_2 \end{array} \tag{1}$$

It assigns to a subset $M \subseteq \mathcal{X}$ the largest open set contained in M. The interior operator $\mathcal{J}$ can be described by a base $\mathbb{B}$ as follows.

$$\mathcal{J}(M) := \bigcup\{B : B \subseteq M \wedge B \in \mathbb{B}\} \tag{2}$$

Using the complementary (duality) relation between open and closed sets one defines the *closure* (smallest closed set containing M) as follows.

$$\mathcal{C}M := \mathcal{X} \smallsetminus \mathcal{J}(\mathcal{X} \smallsetminus M) \tag{3}$$

Then the following holds.

$$\begin{array}{rcl} \mathcal{C}\emptyset &=& \emptyset \\ M &\subseteq& \mathcal{C}M \;=\; \mathcal{C}\mathcal{C}M \text{ , and} \\ \mathcal{C}(M_1 \cup M_2) &=& \mathcal{C}M_1 \cup \mathcal{C}M_2 \end{array} \tag{4}$$

For topologies $\mathcal{T}_1 = (\mathcal{X}, \mathcal{O}_1)$ and $\mathcal{T}_2 = (\mathcal{X}, \mathcal{O}_2)$ we say that $\mathcal{T}_1$ is *finer* than $\mathcal{T}_2$ provided $\mathcal{O}_1 \supseteq \mathcal{O}_2$, which is equivalent to $\forall M(M \subseteq \mathcal{X} \to \mathcal{J}_1(M) \supseteq \mathcal{J}_2(M))$ or, equivalently, $\forall M(M \subseteq \mathcal{X} \to \mathcal{C}_1(M) \subseteq \mathcal{C}_2(M))$. In order to prove that a topology $\mathcal{T}_1$ is finer than $\mathcal{T}_2$ it suffices to show $\mathbb{B}_2 \subseteq \mathcal{O}_1$ for some base $\mathbb{B}_2$ of $\mathcal{T}_2$.

An element $x \in \mathcal{X}$ is called an *isolated point* if the singleton $\{x\}$ is an open subset of $\mathcal{X}$. We denote by $\mathbb{I}_{\mathcal{T}}$ the set of isolated points of $\mathcal{T} = (\mathcal{X}, \mathcal{O})$. Every $M \subseteq \mathbb{I}_{\mathcal{T}}$ is an open subset of $\mathcal{X}$. Moreover, the following is easily seen.

Property 1 *Let $\mathcal{T} = (\mathcal{X}, \mathcal{O})$ be a topological space, $\mathbb{I}_{\mathcal{T}}$ its set of isolated points and let $M \subseteq \mathbb{I}_{\mathcal{T}}$ be a closed set. Then every $M' \subseteq M$ is simultaneously open and closed.*

Moreover, x is an isolated point of $\mathcal{T}$ if and only if $\{x\}$ is in some (every) base of $\mathcal{T}$.

Proof. The first assertion follows because $\mathcal{X} \smallsetminus M' = (\mathcal{X} \smallsetminus M) \cup (M \smallsetminus M')$ is a union of two open sets, and the second because $\{x\}$ has to be a union of base sets. ❑

1.2 Metrisability

Usually, a topology $\mathscr{T} = (\mathscr{X}, \mathscr{O})$ is given by the set of open sets $\mathscr{O}$ or a base $\mathbb{B}$. If the topological space $\mathscr{T}$, however, satisfies certain separation properties it might be possible to introduce a metric. A *metric* on $\mathscr{X}$ is a function $\rho : \mathscr{X} \times \mathscr{X} \to [0,\infty)$ which satisfies the following conditions.

$$\begin{aligned} \rho(x,x) &= 0 \text{ if and only if } x = y, \\ \rho(x,y) &= \rho(y,x), \text{ and} \\ \rho(x,y) &\le \rho(x,y) + \rho(y,z) \end{aligned}$$

A metric ρ on $\mathscr{X}$ induces a topology $\mathscr{T}_\rho$ in the following way: Take the set of open balls $\mathbb{B}_\rho = \{K(x,\varepsilon) : x \in \mathscr{X} \wedge \varepsilon > 0\}$ where $K(x,\varepsilon) := \{y : \rho(x,y) < \varepsilon\}$ as a base for the topology.

If ρ satisfies the stronger ultra-metric inequality $\rho(x,y) \le \max\{\rho(x,y), \rho(y,z)\}$ then all open balls $K(x,\varepsilon)$ are also closed.

A topological space $\mathscr{T} = (\mathscr{X}, \mathscr{O})$ is referred to as *metrisable* if there is a metric ρ on $\mathscr{X}$ such that $\mathbb{B}_\rho$ is a base for the topology $\mathscr{T}$. It is well-known that not every topological space is metrisable [Eng77, Kur66]. In Chapters 4 and 5 of [Eng77] several metrisation theorems are given. Here we rely on special properties of the topologies introduced below.

To show the metrisability of the spaces considered below we have the following idea from profinite methods in mind (e.g. [Pin09]): two elements $x_1, x_2 \in \mathscr{X}$ are close if a base set of large weight is required to separate them. Here a *weight* of base sets $B \in \mathbb{B}$ is a function $\nu : \mathbb{B} \to \mathbb{N}$ such that the pre-image $\nu^{-1}(i)$ is finite for every $i \in \mathbb{N}$.

Theorem 2 *Let $(\mathscr{X}, \mathscr{O})$ be an infinite topological space with a countable base $\mathbb{B}$ satisfying the following conditions.*

1. *Every set $B \in \mathbb{B}$ is also closed, and*
2. *for every two points $x, y \in \mathscr{X}$, $x \neq y$, there are disjoint base sets $B_x, B_y \in \mathbb{B}$ such that $x \in B_x$ and $y \in B_y$.*

Then $(\mathscr{X}, \mathscr{O})$ is metrisable.

For the sake of completeness we add a proof.

Proof. Let $\mathbb{B} = \{B_i : i \in \mathbb{N}\}$, and assign to every base set B_i a weight $\nu(i) \in \mathbb{N}$ such that every $\nu^{-1}(k) = \{B_i : \nu(i) = k\}$ is finite. Then define

$$\rho(x,y) := \begin{cases} 0, & \text{if } x = y \text{, and} \\ \sup\{2^{-\nu(i)} : |\{x,y\} \cap B_i| = 1\}, & \text{otherwise.} \end{cases}$$

Observe the following two properties of the metric ρ.

1. $\rho(x,y) \leq \max\{\rho(x,z), \rho(y,z)\}$
2. $\rho(x,y) < \varepsilon$ if and only if $\forall j\,(\varepsilon \leq 2^{-\nu(j)} \to (x \in B_j \leftrightarrow y \in B_j))$

For open balls $K(x,\varepsilon)$ of radius $\varepsilon > 0$ around $x \in \mathscr{X}$, this implies

$$\text{on the one hand} \quad K(x,\varepsilon) = \bigcap_{\varepsilon \leq 2^{-\nu(i)}, x \in B_i} B_i \cap \bigcap_{\varepsilon \leq 2^{-\nu(i)}, x \notin B_i} (\mathscr{X} \setminus B_i),$$

$$\text{and on the other hand} \quad B_i = \bigcup_{\varepsilon \leq 2^{-\nu(i)}, x \in B_i} K(x,\varepsilon)\,.$$

The proof follows from these identities and the facts that the sets $\mathscr{X} \setminus B_i$ are also open and that in view of $\nu(i) \leq -\log_2 \varepsilon$ the intersections are finite. ❑

As usual, in a topological space, we denote the classes of countable unions of closed sets as $\mathbf{F}_\sigma$ and of countable intersections of open sets as $\mathbf{G}_\delta$.

Lemma 3 *Let $(\mathscr{X},\mathscr{O})$ be a topological space. Then the classes of $\mathbf{F}_\sigma$-sets and of $\mathbf{G}_\delta$-sets are closed under finite union and intersection, and the class of sets being simultaneously of type $\mathbf{F}_\sigma$ and $\mathbf{G}_\delta$ is a Boolean algebra.*

In metric spaces the following holds (cf. [Eng77, Kur66]).

Lemma 4 *If $(\mathscr{X},\mathscr{O})$ is a metrisable topological space then every closed subset is a $\mathbf{G}_\delta$-set, and, hence, every open subset is an $\mathbf{F}_\sigma$-set.*

2 The CANTOR topology

In this section we list some properties of the CANTOR topology on X^ω and of regular ω-languages (see [PP04, Sta97, Tho90, TB73]).

2.1 Basic properties

We consider the space of infinite words (ω-words) X^ω as a metric space with metric ρ defined as follows[2]

$$\rho(\xi,\eta) := \sup\{r^{1-|w|} : w \in \mathbf{pref}(\xi) \,\Delta\, \mathbf{pref}(\eta)\}, \tag{5}$$

that is, the distance $\rho(\xi,\eta)$ is specified by the shortest non-common prefix of ξ and η. Here $r > 1$ is a real number[3], Δ denotes the symmetric difference of sets and we set $\sup \emptyset := 0$, that is, $\rho(\xi,\eta) = 0$ if and only if $\xi = \eta$..

[2] Observe that $e \notin \mathbf{pref}(\xi) \,\Delta\, \mathbf{pref}(\eta)$ and Eq. (5) imply $\rho(\xi,\eta) = \inf\{r^{-|w|} : w \sqsubset \xi \wedge w \sqsubset \eta\}$.

[3] It is convenient to choose $r = |X|$. Then every ball of radius r^{-n} is partitioned into exactly r balls of radius $r^{-(n+1)}$

Since $\mathbf{pref}(\xi)\,\Delta\,\mathbf{pref}(\eta) \subseteq (\mathbf{pref}(\xi)\,\Delta\,\mathbf{pref}(\zeta)) \cup (\mathbf{pref}(\zeta)\,\Delta\,\mathbf{pref}(\eta))$, the metric ρ satisfies the ultra-metric inequality

$$\rho(\xi,\eta) \leq \max\{\rho(\xi,\zeta),\rho(\zeta,\eta)\}\,.$$

As it was explained above, the space (X^ω,ρ) can be also considered as a topological space with base $\mathbb{B}_C := \{w\cdot X^\omega : w \in X^*\}$. Here the non-empty base sets $w\cdot X^\omega$ are open (and closed) balls with radius $r^{-|w|}$.

The following is well-known.

Property 5 *The following holds for the* CANTOR *topology.*

1. *Open sets in (X^ω,ρ) are of the form $W\cdot X^\omega$ where $W\subseteq X^*$.*
2. *A subset $E\subseteq X^\omega$ is open and closed (clopen) if and only if $E = W\cdot X^\omega$ where $W\subseteq X^*$ is finite.*
3. *A subset $F\subseteq X^\omega$ is closed if and only if $F = \{\xi : \mathbf{pref}(\xi)\subseteq \mathbf{pref}(F)\}$.*
4. $$\begin{aligned}\mathcal{C}_C(F) &:= \{\xi : \xi\in X^\omega \wedge \mathbf{pref}(\xi)\subseteq\mathbf{pref}(F)\}\\ &= \bigcap\{W\cdot X^\omega : W\subseteq X^* \wedge W \text{ is finite} \wedge F\subseteq W\cdot X^\omega\}\end{aligned}$$
   is the closure of F.

Moreover, the space (X^ω,ρ) is a compact space, that is, for every family of open sets $(E_i)_{i\in I}$ such that $\bigcup_{i\in I} E_i = X^\omega$ there is a finite sub-family $(E_i)_{i\in I'}$ satisfying $\bigcup_{i\in I'} E_i = X^\omega$. This property is in some sense characteristic for the CANTOR topology on X^ω. In particular, no topology refining the CANTOR topology is compact.

Lemma 6 *Let $(X^\omega,\mathcal{O})$ be a topology with $\{W\cdot X^\omega : W\subseteq X^*\}\subset\mathcal{O}$. Then the space $(X^\omega,\mathcal{O})$ is not compact.*

The proof uses Corollary 3.1.14 in [Eng77]. We give an example illustrating Lemma 6 in the CANTOR space.

Example 1 Let $X = \{0,1\}$ and $(X^\omega,\mathcal{O})$ be a topology with $\{W\cdot X^\omega : W\subseteq X^*\}\subseteq\mathcal{O}$. Let further $F\notin\{W\cdot X^\omega : W\subseteq X^*\}$ be an open and closed subset of X^ω with $0^\omega\in F$. Then F and the sets $0^n1\cdot X^\omega\setminus F$ are an infinite partition of X^ω into open sets. ❏

2.2 Regular ω-languages

In this part we mention some facts on regular ω-languages known from the literature, e.g. [HR86, PP04, Sta97, Tho90, TB73].

The first one shows the importance of ultimately periodic ω-words. Denote by $\mathsf{Ult} := \{w \cdot v^\omega : w, v \in X^* \wedge v \neq e\}$ the set of ultimately periodic ω-words.

Theorem 7 (Büchi [Büc62]) *The class of regular ω-languages is a Boolean algebra and, if $F \subseteq X^\omega$ is regular, $w \in X^*$, and $W \subseteq X^*$ is a regular language, then $W \cdot F$ and F/w are regular.*

Every non-empty regular ω-language contains an ultimately periodic ω-word, and regular ω-languages $E, F \subseteq X^\omega$ coincide if and only if $E \cap \mathsf{Ult} = F \cap \mathsf{Ult}$.

As a consequence we obtain that Ult itself is an example of a finite-state ω-language not being regular.

Example 2 Ult is finite-state, in particular, $\mathsf{Ult}/w = \mathsf{Ult}$, but in view of $X^\omega \cap \mathsf{Ult} = \mathsf{Ult}$ and $X^\omega \neq \mathsf{Ult}$ not regular.

The class of finite-state ω-languages has similar closure properties (see [Sta83]).

Lemma 8 *The class of finite-state ω-languages is a Boolean algebra and, if $F \subseteq X^\omega$ is finite-state, $w \in X^*$, and $W \subseteq X^*$ is a regular language, then $W \cdot F$ and F/w are also finite-state.*

For regular ω-languages we have the following topological characterisations analogous to Property 5.

Property 9 *Let $F \subseteq X^\omega$ be regular. Then in the* CANTOR *topology the following hold true.*

1. *F is open if and only if $F = W \cdot X^\omega$ where $W \subseteq X^*$ is a regular language.*
2. *$F \subseteq X^\omega$ is closed if and only if $F = \{\xi : \mathbf{pref}(\xi) \subseteq \mathbf{pref}(F)\}$ and $\mathbf{pref}(F)$ is regular.*
3. *If F is an $\mathbf{F}_\sigma$-set then F is the (countable) union of closed regular ω-languages.*
4. *F is the (countable) union of regular ω-languages in the class $\mathbf{G}_\delta$.*

The class of finite-state ω-languages has the following additional closure properties.

Property 10 *Let $E \subseteq X^\omega$ be finite-state.*

1. **pref**(E) *is a regular language.*

2. $\mathscr{C}_C(E)$ *and* $\mathscr{J}_C(E)$ *are regular* ω*-languages.*

And, finally, we mention a topological sufficient condition when finite-state ω-languages are regular.

Theorem 11 ([Sta83]) *Every finite-state* ω*-language in the class* $\mathbf{F}_\sigma \cap \mathbf{G}_\delta$ *is a Boolean combination of open regular* ω*-languages, thus, in particular, a regular* ω*-language.*

We conclude this part with examples of regular ω-languages having certain topological properties.

Example 3 Let $\emptyset \subset A \subset X$ and $w \in X^* \setminus \{e\}$. The ω-languages A^ω and also $w \cdot A^\omega$ are regular and closed but not open. Thus the complementary ω-languages $X^\omega \setminus A^\omega = X^* \cdot (X \setminus A) \cdot X^\omega$ and $X^\omega \setminus w \cdot A^\omega = \bigcup_{u \in X^{|w|} \setminus \{w\}} u \cdot X^\omega \cup w \cdot X^* (X \setminus A) \cdot X^\omega$ are regular and open but not closed.

Then, for $w \not\sqsubseteq v$ and $v \not\sqsubseteq w$, the ω-languages $w \cdot A^\omega \cup v \cdot X^* \cdot (X \setminus A) \cdot X^\omega$ are Boolean combinations of open regular ω-languages and neither open nor closed. ❏

Example 4 As a consequence of the BAIRE category theorem (see [Kur66, Chapter 3.9.3] or [Kur66, Chapter 3, §34.V]) we know that a set $F \subseteq X^\omega$ which is dense in itself, that is, $\xi \in \mathscr{C}_C(F \setminus \{\xi\})$ for $\xi \in F$, and countable is an $\mathbf{F}_\sigma$-set but not a $\mathbf{G}_\delta$-set in the CANTOR topology. Hence, for $A \subseteq X, |A| \geq 2$ and $w \in A^*$, the ω-languages $A^* \cdot w^\omega$ are regular ones in the class $\mathbf{F}_\sigma \setminus \mathbf{G}_\delta$.

In fact, using the results of [SW74, Wag79] and Theorem 12 of the next part one can show that $A^* \cdot w^\omega \in \mathbf{F}_\sigma \setminus \mathbf{G}_\delta$ by automata-theoretic means. ❏

2.3 Acceptance by finite automata

Next we give a connection between accepting devices and topology. This makes it possible to prove the metrisability of the topologies in Section 4 via Theorem 2.

The typical accepting devices for regular ω-languages are finite automata. A finite (deterministic) automaton over the alphabet X is a quadruple $\mathscr{A} = (X; S; s_0; \delta)$ where S is a finite non-empty set, $s_0 \in S$ and $\delta : \subseteq S \times X \to S$ is a partial function.

From McNaughton's paper [McN66] we know that an ω-language $F \subseteq X^\omega$ is regular provided there are a finite automaton $\mathscr{A}$ and a table $\mathscr{S} \subseteq \{S' : S' \subseteq S\}$ such that, for $\xi \in X^\omega$, $\xi \in F$ holds if and only if $\mathrm{Inf}(\mathscr{A};\xi) \in \mathscr{S}$ where $\mathrm{Inf}(\mathscr{A};\xi)$ is the set of all states $s \in S$ through which the automaton $\mathscr{A}$ runs infinitely often when reading the input ξ.

A simpler acceptance mode is obtained when we use deterministic BÜCHI automata, that is, $\xi \in F \subseteq X^\omega$ if and only if there are a finite automaton $\mathscr{A}$ and a subset $S' \subseteq S$ such that $\mathrm{Inf}(\mathscr{A};\xi) \cap S' \neq \emptyset$. Then we have the following.

Theorem 12 (Landweber [Lan69]) *An ω-language F is accepted by a deterministic* BÜCHI *automaton if and only if F is regular and a $\mathbf{G}_\delta$-set.*

It follows from Example 4 that not every regular ω-language is accepted by a deterministic BÜCHI automaton.

A particular case of BÜCHI acceptance is the following. Observe that we consider automata where δ is not fully defined.

Proposition 13 *An ω-language F is accepted by a (deterministic)* BÜCHI *automaton $\mathscr{A} = (X; S; s_0; \delta)$ with $S' = S$ if and only if F is regular and closed.*

3 Topologies Refining the CANTOR Topology: Shift-invariant Topologies

In this section we consider topologies on X^ω which are invariant under left and right shifts. To this end we define the following.

Definition 1 We will refer to a family $\mathbb{M} \subseteq 2^{X^\omega}$ as *shift-invariant* provided

$$\forall F \forall w \forall v (F \in \mathbb{M} \wedge w \in X^* \wedge v \in \mathbf{pref}(F) \to w \cdot F, F/v \in \mathbb{M})\,. \tag{6}$$

Theorem 7 and Lemma 8 show, in particular, that the classes of regular or finite-state ω-languages are shift-invariant.

We will call a topology $\mathscr{T} = (X^\omega, \mathscr{O})$ *shift-invariant* provided its set of open sets $\mathscr{O}$ is shift-invariant. It is easy to see that a topology $\mathscr{T}$ is shift-invariant if it has a shift-invariant base $\mathbb{B}_{\mathscr{T}}$ and that the base $\mathbb{B}_C$ of the CANTOR-topology is shift-invariant. Since $X^\omega \in \mathscr{O}$ for every topology $\mathscr{T} = (X^\omega, \mathscr{O})$, every shift-invariant topology on X^ω refines the CANTOR-topology.

Shift-invariant topologies on the space of finite words were investigated in [Pro80].

Next we are going to describe the interior and closure operator of shift-invariant topologies $\mathscr{T}$ on X^ω. To this end we call a set of open sets $M \subseteq 2^{X^\omega}$ a *shift generator* of $\mathscr{T}$ provided $\{w \cdot E : w \in X^* \wedge E \in M\}$ is a base of $\mathscr{T}$. In particular, if a base $\mathbb{B}$ of $\mathscr{T}$ itself is shift-invariant, $\mathbb{B}$ is a shift generator of $\mathscr{T}$. For the CANTOR topology, for instance, $M = \{X^\omega\}$ is a minimal shift generator of $\mathscr{T}_C$.

Now, the interior operator of a shift-invariant topology can be described using the following construction. Let $E, F \subseteq X^{\omega}$. We set

$$L(F;E) := \{w : w \in X^* \wedge F \supseteq w \cdot E\}. \tag{7}$$

Lemma 14 *Let M be a shift generator of the topology $\mathcal{T}$ on X^{ω}. If $\mathcal{J}$ is the corresponding interior operator then*

$$\mathcal{J}(F) = \bigcup_{E \in M} L(F;E) \cdot E$$

for every $F \subseteq X^{\omega}$.

Proof. Since $\mathcal{J}(F)$ is open and M is a shift generator of $\mathcal{T}$ there are a family of sets $E_j \in M$ and a family of words $w_j \in X^*$ such that $\mathcal{J}(F) = \bigcup_{j \in J} w_j \cdot E_j$. Thus $F \supseteq w_j \cdot E_j$ for $j \in J$, that is, $w_j \in L(F;E_j)$. Now, the assertion follows with $\bigcup_{j \in J} w_j \cdot E_j = \bigcup_{j \in J} L(F;E_j) \cdot E_j$. ❑

It should be mentioned that the languages $L(F;E)$ have a simple structure, if only F has a simple structure.

Lemma 15 *If $F \subseteq X^{\omega}$ is finite-state then $L(F;E)$ is a regular language.*

Proof. It suffices to prove the identity

$$L(F/v;E) = L(F;E)/v. \tag{8}$$

Indeed, we have $w \in L(F/v;E)$ if and only if $F \supseteq (v \cdot w) \cdot E$ which, in turn, is equivalent to $v \cdot w \in L(F;E)$, that is, $w \in L(F;E)/v$. ❑

The subsequent lemma shows that for shift-invariant topologies on X^{ω} the closure and the interior operators are stable with respect to the derivative.

Lemma 16 *If $\mathcal{T}$ is a shift-invariant topology on X^{ω} then $\mathcal{J}(F)/v = \mathcal{J}(F/v)$ and $\mathcal{C}(F)/v = \mathcal{C}(F/v)$ for all $F \subseteq X^{\omega}$ and $v \in X^*$.*

Proof. Let M be a shift generator for $\mathcal{T}$. Then, in view of Lemmas 14 and 15 and Eq. (8) we have

$$\begin{aligned}
\mathcal{J}(F)/v &= \Big(\bigcup_{E \in M} L(F;E) \cdot E\Big)/v \\
&= \bigcup_{E \in M} (L(F;E)/v) \cdot E \cup \bigcup_{E \in M} \bigcup_{\substack{v' \cdot v'' = v \\ v' \in L(F;E)}} E/v'' \text{, and} \\
\mathcal{J}(F/v) &= \bigcup_{E \in M} (L(F;E)/v) \cdot E.
\end{aligned}$$

Thus it remains to show that $E'/v'' \subseteq \mathcal{J}(F/v)$ whenever $E' \in M$ and $v = v' \cdot v''$ with $v' \in L(F;E')$.

From $v = v' \cdot v''$ and $v' \in L(F;E')$ we have $v \cdot (E'/v'') \subseteq v' \cdot E' \subseteq F$ whence $E'/v'' \subseteq F/v$. The assertion $E'/v'' \subseteq \mathcal{J}(F/v)$ follows since E'/v'' is open.

The proof for $\mathcal{C}$ follows from the identity $X^\omega \setminus E/w = (X^\omega \setminus E)/w$ and Eq. (3). ❑

As a consequence of Lemma 16 we obtain

Corollary 17 *If a topology* $\mathcal{T}$ *on* X^ω *is shift-invariant then* $\mathcal{J}(v \cdot F) = v \cdot \mathcal{J}(F)$ *and* $\mathcal{C}(v \cdot F) = v \cdot \mathcal{C}(F)$ *for all* $F \subseteq X^\omega$ *and* $v \in X^*$.

Proof. First recall that the topology $\mathcal{T}$ refines the CANTOR topology on X^ω, hence every set $v \cdot X^\omega$ is closed and open in $\mathcal{T}$. Consequently, $v \cdot F \subseteq v \cdot X^\omega$ implies $\mathcal{C}(v \cdot F) \subseteq v \cdot X^\omega$.

Now according to Lemma 16 the identities $\mathcal{C}(F) = \mathcal{C}((v \cdot F)/v) = \mathcal{C}(v \cdot F)/v$ hold. This yields $v \cdot \mathcal{C}(F) = \mathcal{C}(v \cdot F) \cap v \cdot X^\omega$ and the assertion follows with $\mathcal{C}(v \cdot F) \subseteq v \cdot X^\omega$. The proof for $\mathcal{J}$ is similar. ❑

Property 10.2 showed that interior and closure of finite-state ω-languages in the CANTOR topology are regular ω-languages. Here we investigate whether this property holds for all shift-invariant topologies on X^ω. First we derive the following consequences of the Lemmas 14, 15 and 16.

Corollary 18 *Let a topology* $\mathcal{T}$ *on* X^ω *be shift-invariant and let* $F \subseteq X^\omega$ *be a finite-state* ω*-language.*

1. *Then* $\mathcal{J}(F)$ *and* $\mathcal{C}(F)$ *are finite-state* ω*-languages.*
2. *If moreover, there is a finite shift generator* M *of* $\mathcal{T}$ *consisting solely of regular* ω*-languages then* $\mathcal{J}(F)$ *and* $\mathcal{C}(F)$ *are even regular* ω*-languages.*

Proof. The first assertion follows from Lemma 16.

For proving the assertion on the regularity of the ω-languages $\mathcal{J}(F)$ and $\mathcal{C}(F)$ we observe that the strong assumption on M and Lemmas 14 and 15 yield $\mathcal{J}(F) = \bigcup_{E \in M_{\mathbb{B}}} L(F;E) \cdot E$ where the union is finite and $L(F;E) \subseteq X^*$ and $E \subseteq X^\omega$ are regular. Thus $\mathcal{J}(F)$ is also regular. The assertion for $\mathcal{C}(F)$ now follows from Eq. (3). ❑

The following necessary and sufficient condition for the regularity of the interior and closure is easily seen.

Property 19 *Let* $\mathcal{T}$ *be a shift-invariant topology on* X^ω*. Then, for all finite-state* $F \subseteq X^\omega$*, the* ω*-languages* $\mathcal{J}_{\mathcal{T}}(F)$ *and* $\mathcal{C}_{\mathcal{T}}(F)$ *are regular if and only if every open and finite-state* $E \subseteq X^\omega$ *is a regular* ω*-language.*

4 Topologies Related to Finite Automata

In this section we consider shift-invariant topologies refining the CANTOR topology which are closely related to finite automata. Common to these topologies is the fact that they are defined by shift-invariant bases consisting of regular ω-languages. The study of those topologies was inspired from the alphabetic topologies introduced by DIEKERT and KUFLEITNER [DK11], $\mathcal{T}_\alpha$ and $\mathcal{T}_s$ in the sequel, which are useful for investigations in restricted first-order theories for infinite words.

As a first topology we consider the topology, called BÜCHI toplogy $\mathcal{T}_B$, having all regular ω-languages as open (and closed) sets. This is the coarsest topology having all regular ω-languages as open sets.

Then we turn to the automatic topology $\mathcal{T}_A$ having all regular ω-languages which are closed in the CANTOR topology as open sets. This topology is remarkable because here by Property 9 and Theorem 12 all ω-languages accepted by deterministic BÜCHI-automata are closed.

The subsequent two topologies are the alphabetic topologies $\mathcal{T}_\alpha$ and $\mathcal{T}_s$ mentioned above.

Each of the four topologies considered has an infinite set of isolated points and, in view of Lemma 6, none of them is a compact topology on X^ω.

Before proceeding to the study of the particular topologies we mention some general properties of topologies having bases consisting solely of regular ω-languages.

4.1 Isolated points

Our first property concerns isolated points in the topologies.

As was mentioned in Theorem 7, for ω-languages definable by finite automata ultimately periodic ω-words play a special rôle. The following property applies to topologies related to finite automata.

Property 20 *If $\mathcal{T} = (X^\omega, \mathcal{O})$ has a base $\mathbb{B}$ such that every $F \in \mathbb{B}$ contains an ultimately periodic ω-word then*

1. *every non-empty open subset of $\mathcal{T}$ contains an ultimately periodic ω-word,*
2. *the set of isolated points $\mathbb{I}_{\mathcal{T}}$ is a subset of* Ult, *and*
3. $\mathcal{J}(X^\omega \setminus \mathsf{Ult}) = \emptyset$ *and* $\mathcal{C}(\mathsf{Ult}) = X^\omega$.
4. *If, moreover, $\mathcal{T} = (X^\omega, \mathcal{O})$ has a basis consisting solely of regular ω-languages then every finite or countable open set is a subset of* Ult.

Proof. 1. Every non-empty open set contains a base set.

2. If $\xi \in X^\omega$ is an isolated point then $\{\xi\}$ is open, and Item 1 implies $\xi \in \mathsf{Ult}$.

3. The first identity is obvious and the second follows by the complementation property Eq. (3).

4. From Corollary 4.1 in [Sta97] we know that every finite or countable regular ω-language is a subset of Ult. ❑

It is, however, not true that Property 20 applies to all shift-invariant topologies on X^ω. For example it does not apply to the topologies investigated in [CMS03, Sta05]. In [CMS03] the set of isolated points is disjoint to Ult, and the other topology has all $\xi \in X^\omega$ with $\mathbf{infix}(\xi) \neq X^*$ as isolated points (see [Sta05, Theorem 14]).

4.2 Metrisability

As we mentioned in Section 1, a topology on a set might be given in several ways: by specifying all open sets or a particular base or, if possible, by a suitable metric. In the latter case, the topology is referred to as metrisable.

For our topologies on X^ω the following lemma applies.

Lemma 21 *Let* $\mathcal{T} = (X^\omega, \mathcal{O})$ *be a shift-invariant topology on* X^ω *having a base* $\mathbb{B}$ *consisting solely of regular* ω*-languages. If every base set* $F \in \mathbb{B}$ *is also closed in* $\mathcal{T}$ *then* $\mathcal{T}$ *is metrisable.*

Proof. By our assumptions the Condition 1 of Theorem 2 is satisfied. Observe that, since $\mathcal{T}$ is shift-invariant, it refines the CANTOR topology on X^ω. Thus Condition 2 holds, too. Then we may take

$$\nu(B) := \min\{|S| : (X, S, s_0, \delta) \text{ is an automaton which accepts } B\}$$

as a weight for $B \in \mathbb{B}$ and the assertion follows from Theorem 2. ❑

Remark. In the proof of Lemma 21, in general, it is not possible to take the number of left derivatives $\nu(B) = |\{B/w : w \in X^*\}|$ as a weight for B. Then infinitely many regular ω-languages may have the same weight, for example the ω-languages $X^* \cdot w^\omega$ all have weight 1, and the construction in the proof of Theorem 2 fails. ❑

4.3 The BÜCHI topology

As the first topology we introduce a topology which has all ω-languages definable by finite automata as open sets.

Definition 2 The *BÜCHI topology* $\mathcal{T}_B$ is defined by the base

$$\mathbb{B}_B := \{F : F \subseteq X^\omega \wedge F \text{ is a regular } \omega\text{-language}\}.$$

Since the class of regular ω-languages is closed under complement, every set in the base $\mathbb{B}_B$ is simultaneously open and closed. As every set of the form $w \cdot X^\omega$ is a regular ω-language the BÜCHI topology refines the CANTOR topology. Moreover, the singletons $\{w \cdot v^\omega\}$ are regular. Thus, in view of Property 20 the set of isolated points of the BÜCHI topology is $\mathbb{I}_B = \mathsf{Ult}$.

As a consequence of Theorem 2 and Lemma 21 we obtain.

Corollary 22 *The* BÜCHI *topology on* X^ω *is metrisable.*

In BÜCHI topology, trivially, interior and closure of regular ω-languages are again regular. Unlike the CANTOR topology the interior and closure of finite-state ω-languages need not be regular in $\mathcal{T}_B$: the finite-state non-regular ω-language Ult is open and its complement $X^\omega \setminus \mathsf{Ult}$ is closed.

In view of Property 9.4 one can choose a smaller base for $\mathcal{T}_B$.

$$\mathbb{B}'_B := \{F : F \subseteq X^\omega \wedge F \text{ is a regular } \omega\text{-language in } \mathbf{G}_\delta\}$$

4.4 The automatic topology

In this section we introduce a topology where not all regular ω-languages are simultaneously open and closed; the topology arises from the CANTOR topology by adding all closed (in the CANTOR topology) regular ω-languages to the base. By Property 9 and Theorem 12 this will result in having all deterministic BÜCHI-acceptable ω-languages, that is, all regular ω-languages in the BOREL class $\mathbf{G}_\delta$, as closed ones.

4.4.1 Definition and general properties

Definition 3 The *automatic topology* $\mathcal{T}_A$ on X^ω is defined by the base

$$\mathbb{B}_A := \{F : F \subseteq X^\omega \wedge F \text{ is a regular and } F \text{ is closed in the CANTOR topology}\}.$$

The sets (open balls) $w \cdot X^\omega$ are regular and closed in the CANTOR topology. Thus the base $\mathbb{B}_A$ contains $\mathbb{B}_C$, and the automatic topology refines the CANTOR topology. Since $\mathbb{B}_A \subseteq \mathbb{B}_B$, the BÜCHI topology is finer than the automatic topology.

Property 23

1. *If $F \subseteq X^\omega$ is open (closed) in the* CANTOR *topology $\mathcal{T}_C$ then F is open (closed) in $\mathcal{T}_A$, and if $E \subseteq X^\omega$ is open (closed) in $\mathcal{T}_A$ then E is open (closed) in $\mathcal{T}_B$.*

2. *Every non-empty set open in $\mathcal{T}_A$ contains an ultimately periodic ω-word.*

3. *The set* Ult *of ultimately periodic* ω*-words is the set* $\mathbb{I}_A$ *of all isolated points in* $\mathcal{T}_A$.

Proof. 1. and 2. were explained above.

3. $\mathbb{I}_A \subseteq$ Ult follows from Property 20.2. Conversely, every ω-language $\{w \cdot v^\omega\} = w \cdot \{v\}^\omega$ is regular and closed in the CANTOR topology, thus also open in $\mathcal{T}_A$. ❑

Furthermore, the family of regular ω-languages and the family of ω-languages closed in the CANTOR topology are shift-invariant. This shows that $\mathbb{B}_A$ and the topology $\mathcal{T}_A$ are shift-invariant. Following Property 23.1, every $F \in \mathbb{B}_A$ is also closed in $\mathcal{T}_A$. Thus, as in the case of the BÜCHI topology, we can prove the following.

Corollary 24 *The automatic topology on* X^ω *is metrisable.*

The following theorem characterises the closure operator for the automatic topology via regular ω-languages open in the CANTOR topology in a way similar to Property 5.4.

Theorem 25

$$\mathcal{C}_A(F) = \bigcap\{W \cdot X^\omega : F \subseteq W \cdot X^\omega \wedge W \subseteq X^* \textit{ is regular}\}$$

Proof. If $W \subseteq X^*$ is a regular language, then $W \cdot X^\omega$ is a regular ω-language open in the CANTOR topology, and consequently $X^\omega \setminus W \cdot X^\omega \in \mathbb{B}_A$. Hence $W \cdot X^\omega$ is closed in $\mathcal{T}_A$. Thus the inclusion "$\subseteq$" follows.

Let, conversely, $\xi \notin \mathcal{C}_A(F)$. Then there is a set $F' \in \mathbb{B}_A$ such that $\xi \in F'$ and $F \cap F' = \emptyset$. F' is a regular ω-language closed in $\mathcal{T}_C$. Thus $X^\omega \setminus F' = W' \cdot X^\omega \supseteq F$ for some regular language $W' \subseteq X^*$. Consequently, $\xi \notin W' \cdot X^\omega \supseteq \bigcap\{W \cdot X^\omega : W \subseteq X^* \wedge F \subseteq W \cdot X^\omega \wedge W$ is regular$\}$. ❑

As an immediate consequence of the choice of $\mathbb{B}_A$ we obtain the following.

Corollary 26 *Every set open in* $\mathcal{T}_A$ *is an* $\mathbf{F}_\sigma$*-set in the* CANTOR *topology, and every set closed in* $\mathcal{T}_A$ *is a* $\mathbf{G}_\delta$*-set in the* CANTOR *topology.*

The converse of Corollary 26 is not true in general.

Example 5 Let $\eta \notin$ Ult and consider the countable ω-language $F := \{0^n \cdot 1 \cdot \eta : n \in \mathbb{N}\}$.

Then, in the CANTOR topology, $F = (\{0\}^\omega \cup F) \cap 0^* \cdot 1 \cdot \{0,1\}^\omega \subseteq \{0,1\}^\omega$ is the intersection of a closed set with an open set, hence, both F and $X^\omega \setminus F$ are simultaneously $\mathbf{F}_\sigma$-sets and a $\mathbf{G}_\delta$-sets.

As F does not contain any ultimately periodic ω-word, it cannot be open in $\mathcal{T}_A$. Thus $X^\omega \setminus F$ is not closed in $\mathcal{T}_A$.

Consequently, $0 \cdot F \cup 1 \cdot (X^\omega \setminus F)$ is neither open nor closed in $\mathcal{T}_A$ but being simultaneously an $\mathbf{F}_\sigma$-set and a $\mathbf{G}_\delta$-set in the CANTOR topology. ❑

For regular ω-languages, however, we have the following characterisation of ω-languages closed or open, respectively, in $\mathcal{T}_A$ via topological properties in the CANTOR topology. The second item, however, shows a difference to the CANTOR topology.

Proposition 27 *1. Let $F \subseteq X^\omega$ be a regular ω-language. Then F is an $\mathbf{F}_\sigma$-set in the* CANTOR *topology if and only if F is open in $\mathcal{T}_A$, and F is a $\mathbf{G}_\delta$-set in the* CANTOR *topology if and only if F is closed in $\mathcal{T}_A$.*

2. There are clopen sets in $\mathcal{T}_A$ which are not regular.

Proof. 1. In the CANTOR topology, every regular ω-language F being an $\mathbf{F}_\sigma$-set is a countable union of closed regular ω-languages (see [SW74] or the surveys [HR86, Sta97]).

2. The ω-language $F_\square := \bigcup_{n\in\mathbb{N}} 0^{n^2} \cdot 1 \cdot X^\omega$ and its complement $X^\omega \setminus F_\square = \{0^\omega\} \cup \bigcup_{n \text{ is not a square}} 0^n \cdot 1 \cdot X^\omega$ partition the whole space $X^\omega = \{0,1\}^\omega$ into two non-regular ω-languages open in $\mathcal{T}_A$. ❑

4.4.2 Non-preservation of regularity by $\mathcal{J}_A$ and $\mathcal{C}_A$

From Corollary 18.1 we know that finite-state ω-languages are preserved by the interior $\mathcal{J}_A$ and closure $\mathcal{C}_A$.

The same examples, Ult and $X^\omega \setminus \mathsf{Ult}$, as for the BÜCHI topology show that the interior or the closure of finite-state ω-languages need not be regular. A still more striking difference to the CANTOR topology (see Property 10.1) is the fact that the closure (and, by complementation, also the interior) of a regular ω-language need not be regular again.

Example 6 We use the fact (Theorem 7) that two regular ω-languages E, F coincide if only $E \cap \mathsf{Ult} = F \cap \mathsf{Ult}$. Then, if $F \subseteq X^\omega$ is regular and $\mathcal{C}_A(F) \cap \mathsf{Ult} = F \cap \mathsf{Ult}$ we have either $\mathcal{C}_A(F) = F$ or $\mathcal{C}_A(F)$ is not regular.

Let $X = \{0,1\}$ and consider $F = \{0,1\}^* \cdot 0^\omega \subseteq \mathsf{Ult}$. According to Example 4 $\{0,1\}^* \cdot 0^\omega$ is not a $\mathbf{G}_\delta$-set in the CANTOR-topology, hence not closed in $\mathcal{T}_A$, whence $\mathcal{C}_A(\{0,1\}^* \cdot 0^\omega) \neq \{0,1\}^* \cdot 0^\omega$.

Utilising Theorem 25 we get $\mathcal{C}_A(\{0,1\}^* \cdot 0^\omega) \subseteq \bigcap_{k\in\mathbb{N}} \{0,1\}^* \cdot 0^k \cdot \{0,1\}^\omega$. Consequently, $\mathcal{C}_A(\{0,1\}^* \cdot 0^\omega) \cap \mathsf{Ult} = \{0,1\}^* \cdot 0^\omega = \{0,1\}^* \cdot 0^\omega \cap \mathsf{Ult}$, and from the above consideration we obtain that $\mathcal{C}_A(\{0,1\}^* \cdot 0^\omega)$ cannot be regular. ❑

4.5 The alphabetic topology

This topology was introduced by DIEKERT and KUFLEITNER in [DK11]. It is defined by the following base.

Definition 4 The *alphabetic topology* is defined by the base

$$\mathbb{B}_\alpha := \{w \cdot A^\omega : w \in X^* \land A \subseteq X\}.$$

Then, obviously, $\mathbb{B}_C \subseteq \mathbb{B}_\alpha$ and Example 3 shows $\mathbb{B}_\alpha \subseteq \mathbb{B}_A$. Similar to Property 23, the following holds for $\mathcal{T}_\alpha$.

Property 28 *1. If $F \subseteq X^\omega$ is open (closed) in the* CANTOR *topology $\mathcal{T}_C$ then F is open (closed) in $\mathcal{T}_\alpha$, and if $E \subseteq X^\omega$ is open (closed) in $\mathcal{T}_\alpha$ then E is open (closed) in $\mathcal{T}_A$.*

2. Every non-empty set open in $\mathcal{T}_\alpha$ contains an ultimately periodic ω-word of the form $w \cdot a^\omega$, $w \in X^*$ and $a \in X$.

3. The set $X^* \cdot \{a^\omega : a \in X\}$ is the set $\mathbb{I}_\alpha$ of all isolated points in $\mathcal{T}_\alpha$.

Proof. 1. follows from the inclusion relations of the bases explained above, and 2. immediately from Definition 4.

3. The proof follows from the fact that ξ is an isolated point of a topology $\mathcal{T}$ if and only if $\{\xi\}$ is in some base $\mathbb{B}$ of $\mathcal{T}$. ❏

All base sets are regular and closed in the CANTOR topology $\mathcal{T}_C$, so they are also closed in $\mathcal{T}_\alpha$ and, since $\mathbb{B}_\alpha$ is shift-invariant, from Lemma 21 we obtain the following.

Corollary 29 *The alphabetic topology on X^ω is metrisable.*

The base $\mathbb{B}_\alpha$ has the finite shift generator $M_\alpha = \{A^\omega : A \subseteq X\}$consisting of regular ω-languages. Thus we have the following consequence of Corollary 18.2.

Lemma 30 *Let $E \subseteq X^\omega$ be a finite-state ω-language. Then $\mathcal{I}_\alpha(E)$ and $\mathcal{C}_\alpha(E)$ are regular ω-languages.*

4.6 The strict alphabetic topology

For the next definition we fix the following notation (cf. [DK11]). For $A \subseteq X$ the ω-language A^{im} is the set of all ω-words $\xi \in X^\omega$ where exactly the letters in A occur infinitely often, that is, $A^{\text{im}} = X^* \cdot (\bigcap_{a \in A}(A^* \cdot a)^\omega)$. In particular, $A^{\text{im}} = X^* \cdot A^{\text{im}}$.

Definition 5 The *strict alphabetic topology* is defined by the base
$$\mathbb{B}_s := \{w \cdot (A^\omega \cap A^{\mathrm{im}}) : w \in X^* \wedge A \subseteq X\}.$$

The following elementary relations hold for the strict alphabetic topology $\mathcal{T}_s$.

Property 31 *1. If $F \subseteq X^\omega$ is open (closed) in $\mathcal{T}_\alpha$ then F is open (closed) in $\mathcal{T}_s$, and if $E \subseteq X^\omega$ is open (closed) in $\mathcal{T}_s$ then E is open (closed) in $\mathcal{T}_B$.*

2. Every non-empty set open in $\mathcal{T}_s$ contains an ultimately periodic ω-word.

3. The set $X^* \cdot \{a^\omega : a \in X\}$ is the set II_s of all isolated points in $\mathcal{T}_\alpha$.

Proof. 1. Since for $A \subseteq X$ we have $w \cdot A^\omega = \bigcup_{B \subseteq A} w \cdot A^* \cdot (B^\omega \cap B^{\mathrm{im}})$, every $F \in \mathbb{B}_\alpha$ is open in $\mathcal{T}_s$. The assertions for $\mathcal{T}_\alpha$ then follow. The assertions for the BÜCHI topology are obvious.

2. Let $v \in A^*$ contain every letter of A. Then $w \cdot v^\omega \in w \cdot (A^\omega \cap A^{\mathrm{im}})$. Thus every base set contains an ultimately periodic ω-word.

3. According to Property 1, for an isolated point ξ the singleton $\{\xi\}$ is in $\mathbb{B}_s$. A base set $w \cdot (A^\omega \cap A^{\mathrm{im}})$ of $\mathcal{T}_s$ is a singleton if and only if $|A| = 1$. Thus the assertion follows. ❑

However, the automatic topology is not finer than the strict alphabetic topology. To this end, in view of Example 4 and Proposition 27 it suffices to show that in $\mathcal{T}_s$ there are closed sets of the form $X^* \cdot a^\omega, a \in X$.

Corollary 32 *There are closed sets in the strict alphabetic topology $\mathcal{T}_s$ on X^ω which are not $\mathbf{G}_\delta$-sets in* CANTOR *topology.*

Proof. Let $a \in X$. Then every set $X^* \cdot (B^\omega \cap B^{\mathrm{im}})$ is open in $\mathcal{T}_s$ and $X^* \cdot a^\omega = X^\omega \setminus \bigcup_{B \neq \{a\}} X^* \cdot (B^\omega \cap B^{\mathrm{im}})$. ❑

Next we show that the strict alphabetic topology is also metrisable. As it is clearly shift-invariant, it suffices to show that all $F \in \mathbb{B}_s$ are also closed. To this end consider the following identity.

Proposition 33 *Let $A \subseteq X$.*
Then $$A^\omega \cap A^{\mathrm{im}} = \Big(X^\omega \setminus \bigcup_{b \notin A} X^* \cdot b \cdot X^\omega\Big) \setminus \bigcup_{B \subset A} X^* \cdot (B^\omega \cap B^{\mathrm{im}}).$$

Proof. The relations $A^\omega = X^\omega \setminus \bigcup_{b \notin A} X^* \cdot b \cdot X^\omega$ and $A^{\mathrm{im}} \supseteq A^\omega \setminus \bigcup_{B \subset A} X^* \cdot (B^\omega \cap B^{\mathrm{im}})$ prove the asserted identity. ❑

The metrisability of $\mathcal{T}_s$ now follows from Lemma 21.

Corollary 34 *The strict alphabetic topology on X^ω is metrisable.*

Since $M_s := \{A^\omega \cap A^{\mathrm{im}} : A \subseteq X\}$ is a finite shift generator consisting of regular ω-languages for the base $\mathbb{B}_s$, we obtain via Corollary 18.2 the analogue to Lemma 30.

Corollary 35 *If* $F \subseteq X^\omega$ *is finite-state then* $\mathcal{J}_s(F)$ *and* $\mathcal{C}_s(F)$ *are regular* ω*-languages.*

5 Topologies Defined by Subword Metrics

It was shown (see [Sta93, Section 5] and [Sta12]) that regular ω-languages are closely related to the (asymptotic) subword complexity of infinite words. Therefore, as another two refinements of the CANTOR topology we introduce two topologies which are defined via metrics on X^ω. These metrics are derived from the metric of the CANTOR space by taking into account not only the common prefixes but also the subwords occurring or occurring infinitely often in the ω-words.

Definition 6 (Subword metrics)

$$\rho_I(\xi,\eta) := \sup\{r^{1-|w|} : w \in (\mathbf{pref}(\xi)\,\Delta\,\mathbf{pref}(\eta)) \cup (\mathbf{infix}(\xi)\,\Delta\,\mathbf{infix}(\eta))\}$$
$$\rho_\infty(\xi,\eta) := \sup\{r^{1-|w|} : w \in (\mathbf{pref}(\xi)\,\Delta\,\mathbf{pref}(\eta)) \cup (\mathbf{infix}^\infty(\xi)\,\Delta\,\mathbf{infix}^\infty(\eta))\}$$

These metrics respect the length of the shortest non-common prefix of ξ and η as well as the length of the shortest non-common subword (non-common subword occurring infinitely often). Thus

$$\rho_I(\xi,\eta) \geq \rho(\xi,\eta) \text{ and } \rho_\infty(\xi,\eta) \geq \rho(\xi,\eta), \tag{9}$$
$$\rho_I(\xi,\eta) = \max\big\{\rho(\xi,\eta), \sup\{r^{1-|w|} : w \in \mathbf{infix}(\xi)\,\Delta\,\mathbf{infix}(\eta)\}\big\}, \text{ and} \tag{10}$$
$$\rho_\infty(\xi,\eta) = \max\big\{\rho(\xi,\eta), \sup\{r^{1-|w|} : w \in \mathbf{infix}^\infty(\xi)\,\Delta\,\mathbf{infix}^\infty(\eta)\}\big\}. \tag{11}$$

Similar to the case of ρ one can verify that ρ_I and ρ_∞ satisfy the ultra-metric inequality. Therefore, balls in the metric spaces (X^ω, ρ_I) or (X^ω, ρ_∞) are simultaneously open and closed. Moreover, Eq. (9) shows that both topologies refine the CANTOR topology of X^ω.

5.1 Shift-invariance

In this part we show that the topologies $\mathcal{T}_\infty$ and $\mathcal{T}_I$ induced by the metrics ρ_∞ and ρ_I, respectively, are shift-invariant. First we derive some simple properties of the metrics.

Lemma 36 *Let $u \in X^*$ and $v, w \in X^m$. Then*

$$\begin{aligned} \rho_\infty(u\cdot\xi, u\cdot\eta) &\le \rho_\infty(\xi,\eta), & (12)\\ \rho_\infty(\xi,\eta) &\le r^m\cdot\rho_\infty(w\cdot\xi, v\cdot\eta), & (13)\\ \rho_I(u\cdot\xi, u\cdot\eta) &\le \rho_I(\xi,\eta), \textit{ and} & (14)\\ \rho_I(\xi,\eta) &\le r^m\cdot\rho_I(w\cdot\xi, v\cdot\eta). & (15)\end{aligned}$$

Proof. If $\xi = \eta$, all inequalities are trivially satisfied. So, in the following, we may assume $\xi \ne \eta$.

As $\mathbf{infix}^\infty(\xi) = \mathbf{infix}^\infty(u\cdot\xi)$, Eqs. (12) and (13) follow from Eq. (11) and the respective properties of the metric ρ of the Cantor topology $\rho(u\cdot\xi, u\cdot\eta) \le \rho(\xi,\eta)$ and $\rho(w\cdot\xi, v\cdot\eta) \ge \rho(w\cdot\xi, w\cdot\eta) = r^{-|w|}\cdot\rho(\xi,\eta)$.

Let $\rho_I(\xi,\eta) = r^{-n}$, that is, $\mathbf{infix}(\xi)\cap X^n = \mathbf{infix}(\eta)\cap X^n$ and $w \sqsubset \xi$ and $w \sqsubset \eta$ for some $w \in X^n$. Then, obviously, $v \sqsubset u\cdot\xi$ and $v \sqsubset u\cdot\eta$ for some $v \in X^n$. Moreover, $\mathbf{infix}(u\cdot\xi)\cap X^n = (\mathbf{infix}(u\cdot w)\cap X^n)\cup(\mathbf{infix}(\xi)\cap X^n) = \mathbf{infix}(u\cdot\eta)\cap X^n$. This proves Eq. (14).

Finally, we have to prove Eq. (15). If $w \ne v$ then $\rho(w\cdot\xi, v\cdot\eta) \ge r^{-m}$, and Eq. (15) is obvious. Let $w = v$ and $\rho_I(\xi,\eta) = r^{-n}$ for some $n \in \mathbb{N}$. We have to show that $\rho_I(w\cdot\xi, w\cdot\eta) \ge r^{-(n+m)}$.

If $\rho(\xi,\eta) = r^{-n}$ then $\rho(w\cdot\xi, w\cdot\eta) = r^{-(n+m)}$ and Eq. (10) proves $\rho_I(w\cdot\xi, w\cdot\eta) \ge r^{-(n+m)}$.

If $\rho(\xi,\eta) < r^{-n}$ in view of $\rho_I(\xi,\eta) = r^{-n}$ we have $u \in \mathbf{infix}(\xi)\,\Delta\,\mathbf{infix}(\eta)$ for some $u \in X^{n+1}$. Now, it suffices to show $(\mathbf{infix}(w\xi)\,\Delta\,\mathbf{infix}(w\eta))\cap X^{n+m+1} \ne \emptyset$. Assume $u \in \mathbf{infix}(\xi)$ and $v'u \notin \mathbf{infix}(w\xi)\,\Delta\,\mathbf{infix}(w\eta)$ for all $v' \in X^m$. Then $u \in \mathbf{infix}(\xi)$ implies $v'u \in \mathbf{infix}(w\xi)\cap\mathbf{infix}(w\eta)$ for some $v' \in X^m$. Since $|w| = |v'| = m$, we have $u \in \mathbf{infix}(\eta)$, a contradiction. ❏

As a consequence we obtain our result.

Corollary 37 *The topologies $\mathcal{T}_I$ and $\mathcal{T}_\infty$ are shift invariant.*

Proof. We use the fact that, in view of Lemma 36, the mappings Φ_u and Φ_m defined by $\Phi_u(\xi) := u\cdot\xi$ and $\Phi_m(w\cdot\xi) := \xi$ for $w \in X^m$ are continuous w.r.t. the metrics ρ_I and ρ_∞, respectively.

Thus, if $F \subseteq X^\omega$ is open in $\mathcal{T}_I$ or $\mathcal{T}_\infty$ then $\Phi_u^{-1}(F) = F/u$ and, for $m = |w|$, also $w\cdot F = \Phi_m^{-1}(F)\cap w\cdot X^\omega$ are open sets. ❏

5.2 Balls in $\mathcal{T}_I$ and $\mathcal{T}_\infty$

Denote by $K_I(\xi, r^{-n})$ and $K_\infty(\xi, r^{-n})$ the open balls of radius r^{-n} around ξ in the spaces (X^ω, ρ_I) and (X^ω, ρ_∞), respectively. Since ρ_I and ρ_∞ satisfy the ultra-metric inequality, they are also closed balls of radius $r^{-(n+1)}$. For $w \sqsubset \xi$ with

$|w| = n+1$ and $W := X^{n+1} \cap \mathbf{infix}(\xi)$, $V := X^{n+1} \cap \mathbf{infix}^\infty(\xi)$, $\overline{W} := X^{n+1} \setminus \mathbf{infix}(\xi)$ and $\overline{V} := X^{n+1} \setminus \mathbf{infix}^\infty(\xi)$ we obtain the following description of balls via regular ω-languages.

$$K_I(\xi, r^{-n}) = w \cdot X^\omega \cap \bigcap_{u \in W} X^* \cdot u \cdot X^\omega \setminus \bigcup_{u \in \overline{W}} X^* \cdot u \cdot X^\omega, \text{ and} \quad (16)$$

$$K_\infty(\xi, r^{-n}) = w \cdot X^\omega \cap X^* \cdot ([\bigcap_{u \in V} (X^* \cdot u)^\omega] \setminus \bigcup_{u \in \overline{V}} X^* \cdot u \cdot X^\omega). \quad (17)$$

An immediate consequence of the representations in Eqs. (16) and (17) is the following.

Lemma 38 *1. Every ball $K_I(\xi, r^{-n})$ is a Boolean combination of regular ω-languages open in the* CANTOR *topology and, therefore, simultaneously open and closed in the automatic topology $\mathcal{T}_A$.*

2. Every ball $K_\infty(\xi, r^{-n})$ is a regular ω-language and, therefore, simultaneously open and closed in the BÜCHI *topology $\mathcal{T}_B$.*

Proof. 1. From Lemma 4 we know that open sets in a metric space are simultaneously $\mathbf{F}_\sigma$- and $\mathbf{G}_\delta$-sets. Then, according to Lemma 3, the set $K_I(\xi, r^{-n})$ is simultaneously an $\mathbf{F}_\sigma$- and $\mathbf{G}_\delta$-set in the CANTOR topology. Now the assertion follows from Proposition 27.

2. This is is obvious. ❑

Using the MORSE-HEDLUND Theorem (or the proof of Theorem 1.3.13 of [Lot02]) one obtains special representations of small balls containing ultimately periodic ω-words. To this end we derive the following lemma.

Lemma 39 *Let $w, u \in X^*$, $u \neq e$ and $\xi \in X^\omega$. Then $w \cdot u \sqsubset \xi$ and $\mathbf{infix}(\xi) \cap X^{|w \cdot u|} = \mathbf{infix}(w \cdot u^\omega) \cap X^{|w \cdot u|}$ imply $\xi = w \cdot u^\omega$.*

Proof. First observe that $|\mathbf{infix}(w \cdot u^\omega) \cap X^{|w \cdot u|}| = |\mathbf{infix}(w \cdot u^\omega) \cap X^{|w \cdot u|+1}|$. Thus, for every $v \in \mathbf{infix}(w \cdot u^\omega) \cap X^{|w \cdot u|}$, there is a unique $v' \in \mathbf{infix}(w \cdot u^\omega) \cap X^{|w \cdot u|}$ such that $v \sqsubset a \cdot v'$ for some $a \in X$. Consequently, the ω-word $\xi \in X^\omega$ with $w \cdot u \sqsubset \xi$ and $\mathbf{infix}(\xi) \cap X^{|w \cdot u|} = \mathbf{infix}(w \cdot u^\omega) \cap X^{|w \cdot u|}$ is uniquely specified. ❑

Lemma 40 *Let $w \cdot u^\omega \in X^\omega$ where $|w| \leq |u|$ and let $m \geq |w| + |u|$ and $n > |u|$. Then*

$$K_I(w \cdot u^\omega, r^{-m}) = \{w \cdot u^\omega\}, \text{ and} \quad (18)$$

$$K_\infty(w \cdot u^\omega, r^{-n}) = w' \cdot X^* \cdot u^\omega \text{ where } w' \sqsubset w \cdot u \text{ and } |w'| = n. \quad (19)$$

Proof. Every $\xi \in K_I(w \cdot u^\omega, r^{-m})$ satisfies $\mathbf{infix}(\xi) \cap X^m = \mathbf{infix}(w \cdot u^\omega) \cap X^m$ and $w \cdot u \sqsubset \xi$, and the assertion of Eq. (18) follows from Lemma 39.

If $\xi \in K_\infty(w \cdot u^\omega, r^{-n})$ then there is a tail ξ' of ξ such that $u \sqsubset \xi'$ and $\mathbf{infix}^\infty(\xi) \cap X^n = \mathbf{infix}(\xi') \cap X^n = \mathbf{infix}(u^\omega) \cap X^n$ whence, again by Lemma 39, $\xi' = u^\omega$. ❑

As a corollary we obtain the following.

Corollary 41 1. *The set of isolated points of the space* (X^ω, ρ_I) *is* $\mathbb{I}_I = \mathsf{Ult}$.

2. *The space* (X^ω, ρ_∞) *has no isolated points and all sets of the form* $X^* \cdot u^\omega$ *are simultaneously closed and open.*

3. *In the space* (X^ω, ρ_∞) *there are open sets which are not* $\mathbf{F}_\sigma$*-sets in the* CANTOR *topology.*

Proof. Since every non-empty open subset of (X^ω, ρ_I) and also of (X^ω, ρ_∞) contains an ultimately periodic ω-word, every isolated point has to be ultimately periodic. Now Eq. (18) shows that every $w \cdot u^\omega$ is an isolated point in (X^ω, ρ_I), and Eq. (19) proves that (X^ω, ρ_∞) has no isolated points. The remaining part of Item 2 follows from Eq. (19), $X^* \cdot u^\omega = \bigcup_{w \in X^n} w \cdot X^* \cdot u^\omega$ and that the balls are closed and open.

Finally, it is known from Example 4 that $X^\omega \setminus X^* \cdot u^\omega$ is not an $\mathbf{F}_\sigma$-set in the CANTOR topology. ❑

5.3 Non-preservation of regular ω-languages

In this section we investigate whether regular ω-languages are preserved by $\mathscr{I}_I$, $\mathscr{C}_I$, $\mathscr{I}_\infty$ and $\mathscr{C}_\infty$.

As an immediate consequence of Corollary 18 and the fact that the topologies $\mathscr{T}_I$ and $\mathscr{T}_\infty$ are shift-invariant we obtain that the closure and interior of a finite-state $F \subseteq X^\omega$ are also finite-state.

Similar to the case of the automatic topology $\mathscr{T}_A$ (see Example 6) we obtain that the closure (and the interior) of regular ω-languages in the spaces (X^ω, ρ_I) and (X^ω, ρ_∞) need not be regular again. To this end we use the same argument as in Example 6: We present regular ω-languages F such that their respective closures satisfy $\mathscr{C}_I(F) \cap \mathsf{Ult} = F \cap \mathsf{Ult}$ and $\mathscr{C}_I(F) \neq F$, or $\mathscr{C}_\infty(F) \cap \mathsf{Ult} = F \cap \mathsf{Ult}$ and $\mathscr{C}_\infty(F) \neq F$.

Example 7 Let $X = \{0,1\}$ and $F = \{0,1\}^* \cdot 0^\omega$.

First we show that $\mathscr{C}_I(\{0,1\}^* \cdot 0^\omega) \cap \mathsf{Ult} = \{0,1\}^* \cdot 0^\omega$. Let $w \cdot u^\omega \notin \{0,1\}^* \cdot 0^\omega$. Then $u \notin \{0\}^*$ and $0^{|w \cdot u|} \notin \mathbf{infix}(w \cdot u^\omega)$. Now Eq. (18) yields $K(w \cdot u^\omega, r^{-|w \cdot u|}) \cap$

$X^* \cdot 0^{|w\cdot u|} \cdot X^\omega = \emptyset$. Thus $\rho_I(w\cdot u^\omega, v\cdot 0^\omega) \geq r^{-|w\cdot u|}$ for all $v \in X^*$ whence $w\cdot u^\omega \notin \mathcal{C}_I(\{0,1\}^* \cdot 0^\omega)$.

In order to prove $\mathcal{C}_I(\{0,1\}^* \cdot 0^\omega) \supset \{0,1\}^* \cdot 0^\omega$ we observe that $\mathcal{C}_I(\{0,1\}^* \cdot 0^\omega)$ contains every ζ with $\mathbf{infix}(\zeta) = \{0,1\}^*$. Indeed, let $\mathbf{infix}(\zeta) = \{0,1\}^*$ and define for $w_n \sqsubset \zeta$, $|w_n| \geq n$, the ω-word $\xi_n := w_n \cdot (\prod_{v\in X^n} v) \cdot 0^\omega$. Then $\rho_I(\xi_n,\zeta) \leq r^{-n}$, and, therefore, $\lim_{n\to\infty} \xi_n = \zeta \in \mathcal{C}_I(\{0,1\}^* \cdot 0^\omega)$.

Thus the ω-language $\mathcal{C}_I(\{0,1\}^* \cdot 0^\omega)$ is not regular. ❏

Since $\{0,1\}^* \cdot 0^\omega$ is closed in $\mathcal{T}_\infty$, we cannot use this ω-language in the case of $\mathcal{T}_\infty$.

Example 8 Let $X = \{0,1\}$ and $F := \{0,1\}^* \cdot ((00)^*1)^\omega$.

If $w\cdot u^\omega \in \mathcal{C}_\infty(F)$, then there is a $\xi \in F$ such that $\rho_\infty(w\cdot u^\omega, \xi) < r^{-|wu|}$. According to Lemma 40 we have $\xi \in w \cdot X^* \cdot u^\omega$. Thus $u^\omega = u' \cdot \eta$ where $\eta \in ((00)^*1)^\omega$ whence $w\cdot u^\omega = w\cdot u' \cdot \eta \in F$.

Finally consider the ω-words $\xi_i := \prod_{j=0}^{2i} 10^j \cdot (1\cdot 0^{2i})^\omega \in F$ and $\zeta = \prod_{j=0}^{\infty} 10^j = 110100\cdots$. Since ζ has infinitely many infixes 10^j1 where j is odd, $\zeta \notin F$.

For the ω-words ξ_i and ζ we have $\mathbf{pref}(\xi_i) \cap X^n = \mathbf{pref}(\zeta) \cap X^n$ and $\mathbf{infix}^\infty(\xi_i) \cap X^n = \mathbf{infix}^\infty(\zeta) \cap X^n = \{0^n\} \cup \{0^j \cdot 1 \cdot 0^{n-j-1} : 0 \leq j < n\}$ when $n \leq 2i$. This implies $\rho_\infty(\xi_i,\zeta) \leq r^{-2i}$, that is, $\lim_{i\to\infty} \xi_i = \zeta \in \mathcal{C}_\infty(F)$.

Thus the ω-language $\mathcal{C}_\infty(\{0,1\}^* \cdot ((00)^*1)^\omega)$ is not regular. ❏

5.4 Subword Complexity

Above we mentioned that regular ω-languages are closely related to the (asymptotic) subword complexity of infinite words. In this part we show that the level sets $F_\gamma^{(\tau)}$ of the asymptotic subword complexity (see [Sta93, Sta12]) are open in the topologies defined by the subword metrics ρ_I and ρ_∞.

First we introduce the concept of asymptotic subword complexity.

Definition 7 (Asymptotic subword complexity)

$$\tau(\xi) := \lim_{n\to\infty} \frac{\log_{|X|} |\mathbf{infix}(\xi) \cap X^n|}{n}$$

Using the inequality $|\mathbf{infix}(\xi) \cap X^{n+m}| \leq |\mathbf{infix}(\xi) \cap X^n| \cdot |\mathbf{infix}(\xi) \cap X^m|$ it is easy to see that the limit in Definition 7 exists and

$$\tau(\xi) = \inf\left\{ \frac{\log_{|X|} |\mathbf{infix}(\xi) \cap X^n|}{n} : n \in \mathbb{N} \wedge n \geq 1 \right\}. \tag{20}$$

Eq. (5.2) of [Sta93] shows that in Definition 7 and Eq. (20) one can replace the term $\mathbf{infix}(\xi)$ by $\mathbf{infix}^\infty(\xi)$.

Corollary 42

$$\tau(\xi) = \inf\left\{\frac{\log_{|X|}|\mathbf{infix}^\infty(\xi)\cap X^n|}{n} : n \in \mathbb{N} \wedge n \geq 1\right\}$$

Let, for $0 < \gamma \leq 1$, $F_\gamma^{(\tau)} := \{\xi : \xi \in X^\omega \wedge \tau(\xi) < \gamma\}$ be the *lower level sets* of the asymptotic subword complexity. For $\gamma = 0$ we set $F_0^{(\tau)} := \mathsf{Ult}$ (instead of $F_0^{(\tau)} = \emptyset$). We want to show that these sets are open in (X^ω, ρ_I) and (X^ω, ρ_∞). As a preparatory result we derive the subsequent Lemma 43.

Let $E_n(\xi) := \{\eta : \mathbf{infix}(\eta) \cap X^n \subseteq \mathbf{infix}(\xi)\}$ and $E'_n(\xi) := \{\eta : \mathbf{infix}^\infty(\eta) \cap X^n \subseteq \mathbf{infix}^\infty(\xi)\}$ be the sets of ω-words having only infixes or infixes occurring infinitely often of length n of ξ, respectively. These sets can be equivalently described as

$$\begin{aligned} E_n(\xi) &= X^\omega \setminus X^* \cdot (X^n \setminus \mathbf{infix}(\xi)) \cdot X^\omega \text{ and} \\ E'_n(\xi) &= X^* \cdot (X^\omega \setminus X^* \cdot (X^n \setminus \mathbf{infix}^\infty(\xi)) \cdot X^\omega), \text{ respectively} \end{aligned}$$

which resembles in some sense the characterisation of open balls in Eqs. (16) and (17). In fact, it appears that the sets $E_n(\xi)$ and $E'_n(\xi)$ are open in the respective spaces (X^ω, ρ_I) and (X^ω, ρ_∞).

Lemma 43 *Let $\xi \in X^\omega$. Then $\xi \in E_n(\xi) \cap E'_n(\xi)$, the set $E_n(\xi)$ is open in (X^ω, ρ_I) and the set $E'_n(\xi)$ is open in (X^ω, ρ_∞).*

Proof. Clearly, $\xi \in E_n(\xi) \cap E'_n(\xi)$.

For a proof that $E_n(\xi)$ is open in (X^ω, ρ_I) we show that $\eta \in E_n(\xi)$ implies that the ball $K_I(\eta, r^{-n})$ is contained in $E_n(\xi)$. Let $\eta \in E_n(\xi)$ and $\zeta \in K_I(\eta, r^{-n})$. Then, $\rho_I(\eta,\zeta) < r^{-n}$, that is, in particular, $\mathbf{infix}(\eta) \cap X^n = \mathbf{infix}(\zeta) \cap X^n$, whence $\zeta \in E_n(\xi)$

The proof for $E'_n(\xi)$ is similar. ❑

This much preparation enables us to show that the level sets are open sets.

Theorem 44 *Let $0 \leq \gamma \leq 1$. Then the sets $F_\gamma^{(\tau)}$ are open in (X^ω, ρ_I) and (X^ω, ρ_∞).*

Proof. For $\gamma = 0$ we have $F_\gamma^{(\tau)} = \mathsf{Ult}$ which, according to Corollary 41, is open in (X^ω, ρ_I) as well as in (X^ω, ρ_∞).

Let $\gamma > 0$ and $\tau(\xi) < \gamma$. We show that then $E_n(\xi) \subseteq F_\gamma^{(\tau)}$ and $E'_n(\xi) \subseteq F_\gamma^{(\tau)}$ for some $n \in \mathbb{N}$. Together with Lemma 43 this shows that $F_\gamma^{(\tau)}$ contains, with every ξ, open sets containing this ξ.

If $\tau(\xi) < \gamma$ then in view of Eq. (20) we have $\frac{1}{n} \cdot \log_{|X|}|\mathbf{infix}(\xi) \cap X^n| < \gamma$ for some $n \in \mathbb{N}$. Then for every $\eta \in E_n(\xi)$ it holds $\tau(\eta) \leq \frac{1}{n} \cdot \log_{|X|}|\mathbf{infix}(\xi) \cap X^n| < \gamma$ and, consequently, $E_n(\xi) \subseteq F_\gamma^{(\tau)}$.

The proof for (X^ω, ρ_∞) is similar using Corollary 42. ❑

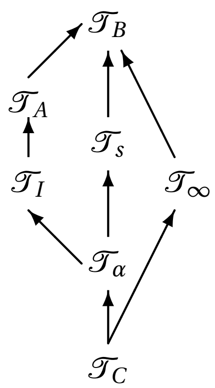
The proof shows also that $\xi \in F_\gamma^{(\tau)}$ implies that $X^\omega \smallsetminus X^* \cdot (X^n \smallsetminus \mathbf{infix}(\xi)) \cdot X^\omega \subseteq F_\gamma^{(\tau)}$ for some $n > 0$. Thus $F_\gamma^{(\tau)}$ is a countable union of regular ω-languages closed in the CANTOR topology, hence an $\mathbf{F}_\sigma$-set in the CANTOR topology. The sets $F_\gamma^{(\tau)}$ are finite-state[4] non-regular ω-languages because their complement $X^\omega \smallsetminus F_\gamma^{(\tau)}$ is non-empty and does not contain any ultimately periodic ω-word. Thus, in view of Theorem 11, they are not $\mathbf{G}_\delta$-sets in CANTOR-space and they are examples of sets open in (X^ω, ρ_I) and (X^ω, ρ_∞) which are non-regular $\mathbf{F}_\sigma$-sets in CANTOR-space.

Finally we show that the level sets $F_\gamma^{(\tau)}$ for $0 \le \gamma \le 1$ are not open in the strict alphabetic topology $\mathcal{T}_s$ and hence also not in $\mathcal{T}_\alpha$.

Example 9 Let $X = \{a_1, \ldots, a_k\}$. Then $\xi = (a_1 \cdots a_k)^\omega \in \mathsf{Ult} \subseteq F_\gamma^{(\tau)}$ for all γ. Since ξ contains every letter infinitely often, every base set $E \in \mathbb{B}_s$ with $\xi \in E$ is of the form $w \cdot X^{\mathrm{im}}$. Thus it contains all $\eta \in w \cdot X^\omega$ with $\mathbf{infix}(\eta) = X^*$, that is $\tau(\eta) = 1$. Consequently, $\eta \notin F_1^{(\tau)} \supseteq F_\gamma^{(\tau)}$. ❑

6 The hierarchy of topologies

Finally we show that the following inclusions hold for the topologies considered so far. All inclusions are proper and other ones than the ones indicated do not exist.



First, the obvious inclusions of the bases $\mathbb{B}_B \supseteq \mathbb{B}_A \supseteq \mathbb{B}_\alpha \supseteq \mathbb{B}_C$ and $\mathbb{B}_B \supseteq \mathbb{B}_s$ imply the inclusions for the corresponding topologies (see also Properties 23, 28 and 31). $\mathcal{T}_s \supseteq \mathcal{T}_\alpha$ follows from Property 31.

[4]In particular, they satisfy $F_\gamma^{(\tau)}/w = F_\gamma^{(\tau)}$ for all $w \in X^*$.

Next we verify how the topologies $\mathcal{T}_I$ and $\mathcal{T}_\infty$ fit into the diagram. The inclusions $\mathcal{T}_C \subseteq \mathcal{T}_\infty \subseteq \mathcal{T}_B$ follow from the fact that $\mathcal{T}_\infty$ is shift-invariant (Corollary 37) and that every ball in $\mathcal{T}_\infty$ is a regular ω-language (Lemma 38.2). Similarly, Lemma 38.1 proves the inclusion $\mathcal{T}_I \subseteq \mathcal{T}_A$.

In order to show $\mathcal{T}_\alpha \subseteq \mathcal{T}_I$ fix, for $\emptyset \neq B \subseteq X$, an ω-word $\xi_B \in B^\omega \cap B^{\mathrm{im}}$, that is, ξ_B contains only letters in B and every letter of B infinitely often. Then $A^\omega = \bigcup_{B\subseteq A}\bigcup_{a\in B} K_I(a\cdot\xi_B, 1)$, as a union of open balls, is open in $\mathcal{T}_I$, and Corollary 37 shows that $\mathbb{B}_\alpha$ consists of sets open in $\mathcal{T}_I$.

The following relations show the properness of the inclusions.

$$\begin{aligned}
\mathcal{T}_\alpha &\not\subseteq \mathcal{T}_\infty, && (21)\\
\mathcal{T}_\infty &\not\subseteq \mathcal{T}_s, && (22)\\
\mathcal{T}_\infty &\not\subseteq \mathcal{T}_A, && (23)\\
\mathcal{T}_I &\not\subseteq \mathcal{T}_s, && (24)\\
\mathcal{T}_s &\not\subseteq \mathcal{T}_A, \text{ and} && (25)\\
\mathcal{T}_A &\not\subseteq \mathcal{T}_I. && (26)
\end{aligned}$$

Since the sets of isolated points satisfy $\mathbb{I}_I \supset \mathbb{I}_s = \mathbb{I}_\alpha \supset \mathbb{I}_\infty$ we obtain the inequalities Eqs. (21) and (24).

The topologies $\mathcal{T}_s$ and $\mathcal{T}_\infty$ have open sets which are not $\mathbf{F}_\sigma$-sets in the CANTOR topology (see Corollary 32 and Corollary 41.3. Hence Eqs. (23) and (25) follow from Corollary 26.

Finally, the remaining inequalities Eqs. (22) and (26) are proved by the following two examples.

Example 10 Let $0,1 \in X$. Corollary 41.2 shows that the countable ω-language $F = X^*\cdot(01)^\omega$ is open in $\mathcal{T}_\infty$. Since for $|A| \geq 2$ the base sets $w\cdot(A^\omega \cap A^{\mathrm{im}}) \in \mathbb{B}_s$ are uncountable, $0 \neq 1$ implies that F does not contain a non-empty $E \in \mathbb{B}_s$.

Hence F cannot be open in $\mathcal{T}_s$. ❑

Example 11 ([Hof14]) The regular ω-language $F = \{1,00\}^\omega \subseteq \{0,1\}^\omega$ is closed in the CANTOR topology, hence open in $\mathcal{T}_A$. Assume F to be open in $\mathcal{T}_I$. We have $\eta = \prod_{i\in\mathbb{N}} 10^{2i} \in F$ and, therefore, $K_I(\eta, r^{-n}) \subseteq F$ for some $n \in \mathbb{N}$.

Consider $\xi = \prod_{i=0}^{n} 10^{2i}\cdot\prod_{i=2n+1}^{\infty} 10^i \notin F$. Then we have $\prod_{i=0}^{n} 10^{2i} \sqsubset \eta$ and $\prod_{i=0}^{n} 10^{2i} \sqsubset \xi$. Consequently,

$$\begin{aligned}
\mathbf{infix}(\xi)\cap\{0,1\}^{2n} &= \Big(\mathbf{infix}(\textstyle\prod_{i=0}^{n} 10^{2i}) \cup 0^*\cdot 1\cdot 0^* \cup 0^*\Big)\cap\{0,1\}^{2n}\\
&= \mathbf{infix}(\eta)\cap\{0,1\}^{2n}.
\end{aligned}$$

It follows $\rho_I(\xi,\eta) \leq r^{-2n}$, that is, $\xi \in K_I(\eta, r^{-n}) \subseteq \{1,00\}^\omega$, a contradiction. ❑

References

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

[CMS03] Cristian S. Calude, Solomon Marcus, and Ludwig Staiger. A topological characterization of random sequences. *Inform. Process. Lett.*, 88(5):245–250, 2003.

[DK11] Volker Diekert and Manfred Kufleitner. Fragments of first-order logic over infinite words. *Theory Comput. Syst.*, 48(3):486–516, 2011.

[DMW15] Volker Diekert, Anca Muscholl, and Igor Walukiewicz. A note on monitors and büchi automata. In Martin Leucker, Camilo Rueda, and Frank D. Valencia, editors, *Theoretical Aspects of Computing - ICTAC 2015 - 12th International Colloquium Cali, Colombia, October 29-31, 2015, Proceedings*, volume 9399 of *Lecture Notes in Computer Science*, pages 39–57. Springer, 2015.

[Eng77] Ryszard Engelking. *General topology*. PWN—Polish Scientific Publishers, Warsaw, 1977. Translated from the Polish by the author, Monografie Matematyczne, Tom 60. [Mathematical Monographs, Vol. 60].

[FS01] Henning Fernau and Ludwig Staiger. Iterated function systems and control languages. *Inform. and Comput.*, 168(2):125–143, 2001.

[Hof14] Stefan Hoffmann. Metriken zur Verfeinerung des Cantor-Raumes auf X^ω. Diploma thesis, Martin-Luther-Universität Halle-Wittenberg, Institut für Informatik, Halle, 2014.

[HR86] Hendrik Jan Hoogeboom and Grzegorz Rozenberg. Infinitary languages: basic theory and applications to concurrent systems. In J. W. de Bakker, W. P. de Roever, and G. Rozenberg, editors, *Current trends in concurrency (Noordwijkerhout, 1985)*, volume 224 of *Lecture Notes in Comput. Sci.*, pages 266–342. Springer, Berlin, 1986. Overviews and tutorials.

[HS15] Stefan Hoffmann and Ludwig Staiger. Subword metrics for infinite words. In Frank Drewes, editor, *Implementation and Application of*

Automata - 20th International Conference, CIAA 2015, Umeå, Sweden, August 18-21, 2015, Proceedings, volume 9223 of *Lecture Notes in Computer Science*, pages 165–175. Springer, 2015.

[Kur66] Kazimierz Kuratowski. *Topology. Vol. I.* Academic Press, New York-London; Państwowe Wydawnictwo Naukowe, Warsaw, 1966.

[Lan69] Lawrence H. Landweber. Decision problems for ω-automata. *Math. Systems Theory*, 3:376–384, 1969.

[Lot02] M. Lothaire. *Algebraic combinatorics on words*, volume 90 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 2002.

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

[Pin09] Jean-Éric Pin. Profinite methods in automata theory. In Susanne Albers and Jean-Yves Marion, editors, *STACS*, volume 3 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 31–50, Dagstuhl, Germany, 2009. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany.

[PP04] Dominique Perrin and Jean-Éric Pin. *Infinite words. Automata, semigroups, logic and games.* Amsterdam: Elsevier/Academic Press, 2004.

[Pro80] Helmut Prodinger. Topologies on free monoids induced by closure operators of a special type. *RAIRO Inform. Théor. Appl.*, 14(2):225–237, 1980.

[Red86] Roman R. Redziejowski. Infinite-word languages and continuous mappings. *Theoret. Comput. Sci.*, 43(1):59–79, 1986.

[RS97] Grzegorz Rozenberg and Arto Salomaa, editors. *Handbook of Formal Languages*, volume 3. Springer-Verlag, Berlin, 1997.

[Sch04] Klaus Schneider. *Verification of Reactive Systems - Formal Methods and Algorithms.* Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2004.

[SS10] Sibylle Schwarz and Ludwig Staiger. Topologies refining the Cantor topology on X^ω. In Cristian S. Calude and Vladimiro Sassone, editors, *IFIP TCS*, volume 323 of *IFIP*, pages 271–285. Springer, 2010.

[Sta83] Ludwig Staiger. Finite-state ω-languages. *J. Comput. System Sci.*, 27(3):434–448, 1983.

[Sta87] Ludwig Staiger. Sequential mappings of ω-languages. *RAIRO Inform. Théor. Appl.*, 21(2):147–173, 1987.

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

[Sta97] Ludwig Staiger. ω-languages. In Rozenberg and Salomaa [RS97], pages 339–387.

[Sta03] Ludwig Staiger. Weighted finite automata and metrics in Cantor space. *J. Autom. Lang. Comb.*, 8(2):353–360, 2003. Selected papers of the workshop "Weighted automata: theory and applications" (Dresden, 2002).

[Sta05] Ludwig Staiger. Topologies for the set of disjunctive ω-words. *Acta Cybern.*, 17(1):43–51, 2005.

[Sta12] Ludwig Staiger. Asymptotic subword complexity. In Henning Bordihn, Martin Kutrib, and Bianca Truthe, editors, *Languages Alive*, volume 7300 of *Lecture Notes in Computer Science*, pages 236–245. Springer, Heidelberg, 2012.

[SW74] Ludwig Staiger and Klaus Wagner. Automatentheoretische und automatenfreie Charakterisierungen topologischer Klassen regulärer Folgenmengen. *Elektron. Informationsverarb. Kybernetik*, 10(7):379–392, 1974.

[TB73] Boris A. Trakhtenbrot and Yan M. Barzdiń. *Finite automata, Behavior and synthesis.* North-Holland, Amsterdam, 1973.

[Tho90] Wolfgang Thomas. Automata on infinite objects. In Jan van Leeuwen, editor, *Handbook of theoretical computer science, Vol. B*, pages 133–191. Elsevier, Amsterdam, 1990.

[Tho97] Wolfgang Thomas. Languages, automata, and logic. In Rozenberg and Salomaa [RS97], pages 389–455.

[Tra62] Boris A. Trahtenbrot. Finite automata and the logic of one-place predicates. *Sibirsk. Mat. Ž.*, 3:103–131, 1962.

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