# On Syntactic Congruences for $\omega$−languages*

**Oded Maler**†
VERIMAG
Miniparc ZIRST
F-38330 Montbonnot
France
`Oded.Maler@imag.fr`

**Ludwig Staiger**‡
Martin-Luther-Universität Halle-Wittenberg
Institut für Informatik
D-06099 Halle (Saale)
Germany
`ludwig.staiger@informatik.uni-halle.de`

Version revised in January, 2008

Recently, *Benjamin Aminof* from Hebrew University, School of Engineering and Computer Science, Jerusalem, discovered an error in Lemma 13 of the version printed in Theoret. Comput. Sci., Vol. **183**, No. 1, pp. 93 – 112. This flaw causes subsequent ones in also in Lemma 14 and Theorem 15 of the paper.
The present version contains correct versions and proofs of the mentioned statements.
We are grateful to *Benjamin* for pointing out the flaw.

*A preliminary version of this paper appeared in: STACS 93 *Proc. $10^{th}$ Annual Symposium on Theoretical Computer Science* (P. Enjalbert, A. Finkel and K.W. Wagner Eds.), Lecture Notes in Computer Science, Vol. 665, pp. 586–594, Springer-Verlag, Berlin 1993

†The results presented in this paper have been obtained while the author was with INRIA/IRISA, Rennes, France.

‡The results presented in this paper have been obtained while the author was with Universität Siegen, Siegen, Germany.

**Abstract**

In this paper we investigate several questions related to syntactic congruences and to minimal automata associated with $\omega$-languages. In particular we investigate relationships between the so-called simple (because it is a simple translation from the usual definition in the case of finitary languages) syntactic congruence and its infinitary refinement (the *iteration* congruence) investigated by Arnold [Ar85]. We show that in both cases not every $\omega$-language having a finite syntactic monoid is regular and we give a characterization of those $\omega$-languages having finite syntactic monoids.

Among the main results we derive a condition which guarantees that the simple syntactic congruence and Arnold's iteration congruence coincide and show that *all* (including infinite-state) $\omega$-languages in the Borel class $F_\sigma \cap G_\delta$ satisfy this condition. We also show that all $\omega$-languages in this class are accepted by their minimal-state automaton — provided they are accepted by any Muller automaton.

Finally we develop an alternative theory of recognizability of $\omega$-languages by families of right-congruence relations, and define a canonical object (much smaller than the iteration monoid) associated with every $\omega$-language. Using this notion of recognizability we give a *necessary and sufficient* condition for a regular $\omega$-language to be accepted by its minimal-state automaton.

# Contents

**1 Introduction** 3

**2 Preliminaries, Congruences and Automata** 4

**3 Some observations on the iteration congruence** 5

**4 The case when $\simeq$ and $\overset{\sim}{=}$ coincide** 7

**5 Acceptance by minimal-state automata** 9

**6 Recognition by right-congruences** 12

**A Proof of Lemma 13** 19

**B Independence examples** 21

# 1 Introduction

The well-known Kleene-Myhill-Nerode theorem for languages states that a language $U \subseteq \Sigma^*$ is regular (rational), iff its syntactic right-congruence $\sim_U$ defined by

$$x \sim_U y \text{ iff } \forall v \in \Sigma^* : xv \in U \iff yv \in U$$

has a finite index. In that case the right-congruence classes correspond to the states of the unique minimal automaton that accepts $U$. An equivalent condition is that the finer two-sided syntactic congruence $\simeq_U$ defined by

$$x \simeq_U y \text{ iff } \forall u \in \Sigma^* : ux \sim_U uy$$

has a finite index. Here the congruence classes correspond to the elements of the transformation monoid associated with the minimal automaton accepting $U$.

As already observed by Trakhtenbrot [Tr62] these same observations are no longer true in the case of $\omega$-languages (cf. also [JT83], [LS77] or [St83]). Here the class of $\omega$-languages having a finite syntactic monoid (so-called finite-state $\omega$-languages) is much larger than the class of $\omega$-languages accepted by finite automata (regular or rational $\omega$-languages) [St83].

Arnold [Ar85] investigated a new concept of syntactic congruence (henceforth called the *iteration* congruence) for $\omega$-languages. As his results show, this concept yields a characterization of regular $\omega$-languages by finite monoids (the iteration monoid), but not in the same simple way as for finitary languages.

As we shall see below, despite the fact that the iteration monoid is indeed more accurate (it is infinite for some $\omega$-languages which are finite-state but not regular), yet there are even non-Borel $\omega$-languages for which the iteration monoid is finite. To this end we shall derive a necessary and sufficient condition for an $\omega$-language for having a finite iteration monoid.

As one of the main results we give a condition on $\omega$-languages that guarantees that the iteration syntactic congruence coincides with the simple one. We show that this condition holds for all (including those which are not finite-state) $\omega$-languages in the Borel-class $F_\sigma \cap G_\delta$. Not only in this sense does the class $F_\sigma \cap G_\delta$ constitute a "well-behaving" fragment of the $\omega$-languages: we show also that such $\omega$-languages once accepted at all by an automaton are accepted by their "minimal-state" automaton, that is, by the automaton isomorphic to their syntactic right-congruence thus extending the result in [St83].

Finally, we introduce an alternative notion of recognizability by a family of *right*-congruence relations, and give a necessary and sufficient condition for a regular $\omega$-language to be acceptable by its minimal-state automaton. This theory complements the existing algebraic theory of recognition by monoids (two-sided congruences).

The rest of the paper is organized as follows: In Section 2 we give the necessary definitions and notations. In Section 3 we investigate the properties of Arnold's iteration congruence. Sections 4 and 5 are devoted to the proofs of two important properties of $F_\sigma \cap G_\delta$ $\omega$-languages: the coincidence of the iteration congruence and the simple congruence, and the acceptability by the minimal-state automaton. In Section 6 (which can be read independently of Sections 3–5) we develop the theory of recognizability by right-congruences, and apply it to derive a necessary and sufficient condition for regular $\omega$-languages to be acceptable by their minimal-state automaton.

## 2 Preliminaries, Congruences and Automata

By $\Sigma^*$ we denote the set (monoid) of finite words on a finite alphabet $\Sigma$, including the *empty word* $e$, let $\Sigma^+$ denote $\Sigma^* - \{e\}$ and $\Sigma^\omega$ the set of infinite words ($\omega$-words). For an $\omega$-word $\alpha = \alpha(1)\alpha(2)\cdots$, we will use $\alpha(i..j)$ to denote the sub-word $\alpha(i)\alpha(i+1)\cdots\alpha(j)$. As usual we will refer to subsets of $\Sigma^*$ as *languages* and to subsets of $\Sigma^\omega$ as *$\omega$-languages*. For $u \in \Sigma^*$ and $\beta \in \Sigma^* \cup \Sigma^\omega$ let $u\beta$ be their concatenation and let $u^\omega$ be the $\omega$-word formed by concatenating the word $u$ infinitely often (provided $u \neq e$) . The concatenation product extends in an obvious way to subsets $U \subseteq \Sigma^*$ and $B \subseteq \Sigma^* \cup \Sigma^\omega$. For a language $U \subseteq \Sigma^*$ let $U^*$ and $U^\omega$ denote, respectively, the set of finite and infinite sequences formed by concatenating words in $U$. By $|u|_a$ we denote the number of occurrences of the letter $a \in \Sigma$ in the word $u \in \Sigma^*$. Finally $u \preceq v$ and $u \prec v$ denote the facts that $u$ is a prefix and a proper prefix of $v$.

An equivalence relation $\simeq$ is a *congruence* on $\Sigma^*$ if $u \simeq v$ implies $xuy \simeq xvy$ for all $u,v,x,y \in \Sigma^*$. We say that $\simeq$ is a *right-congruence* if $u \simeq v$ implies $uy \simeq vy$ for all $u,v,y \in \Sigma^*$. Clearly, every congruence is also a right-congruence. We will denote by $[u] := \{v : v \in \Sigma^* \text{ and } v \simeq u\}$ the equivalence class containing the word $u$, and use $\langle v\rangle$ instead of $[v]$ if the corresponding relation is a right-congruence. We will say that $\simeq$ is *finite* when it has a finite index (or alternatively, the factor-monoid $\Sigma^*/\simeq$ is finite), and that it is *trivial* when $\simeq$ is $\Sigma^* \times \Sigma^*$.

As in [Ar85] we say that a congruence $\simeq$ *covers* an $\omega$-language $E$ provided $E = \bigcup\{[u][v]^\omega : uv^\omega \in E\}$ and we say that an $\omega$-language $E$ is *regular* provided there is a finite congruence $\simeq$ which covers $E$. This is in fact equivalent to the condition that $E = \bigcup_{i=1}^{n} W_i \cdot V_i^\omega$ for some $n \in \mathbb{N}$ and regular languages $W_i, V_i \subseteq X^*$.

The natural (*Cantor*-) topology on the space $\Sigma^\omega$ is defined as follows. A set $E \subseteq \Sigma^\omega$ is *open* iff it is of the form $U\Sigma^\omega$, where $U \subseteq \Sigma^*$ (in other words, $\beta \in E$ iff it has a prefix in $U$). A set is *closed* if its complement is open or equivalently if its elements do not have any prefix in some $U' \subseteq \Sigma^*$. The class $G_\delta$ consists of all countable intersections of open sets. A set is in $F_\sigma$ if its complement is in $G_\delta$. Thus an $F_\sigma$-set can be written as a countable union of closed sets. The rest of the Borel hierarchy is constructed similarly. We note here in passing that every regular $\omega$-language is contained in the Boolean closure of the Borel class $F_\sigma$.

Additional material on $\omega$-languages appears in [Ei74, EH93, HR85, LS77, PP93, St87, Th90, Wa79].

**Definition 1 (Syntactic Congruences)** *Let $E \subseteq \Sigma^\omega$ be an $\omega$-language. We associate with $E$ the following equivalence relations on $\Sigma^*$:*

- *Syntactic right-congruence:*
$$x \sim_E y \text{ iff } \forall\beta \in \Sigma^\omega (x\beta \in E \iff y\beta \in E) \tag{1}$$
- *Simple syntactic congruence:*
$$x \simeq_E y \text{ iff } \forall u \in \Sigma^* (ux \sim_E uy) \tag{2}$$
- *Infinitary syntactic-congruence:*
$$x \approx_E y \text{ iff } \forall u,v \in \Sigma^* (u(xv)^\omega \in E \iff u(yv)^\omega \in E) \tag{3}$$

  (Here we tacitly assume that neither $xv$ nor $yv$ are empty.)

- *Arnold's iteration syntactic-congruence:*

$$x \cong_E y \text{ iff } x \simeq_E y \wedge x \approx_E y \tag{4}$$

By definition $\simeq$ refines $\sim$ and $\cong$ refines both $\simeq$ and $\approx$. In the general case $\simeq$ and $\approx$ are incomparable, since they refer to two different kinds of interchangeability of $x$ and $y$. The following examples give evidence of this fact.

**Example 1** *Let $E_1 := \{a, bb\}^* a^\omega$. Then $a \simeq_{E_1} bb$ but $a \not\approx_{E_1} bb$. Hence the iteration and the simple syntactic congruence associated with $E_1$ are distinct.*

**Example 2** *For $E_2 := abc^\omega$ we have $a \not\simeq_{E_2} b$ but $a \approx_{E_2} b$.* (Nevertheless, since $E_2$ is a closed $\omega-$language as Theorem 10 below shows, $\simeq_{E_2}$ and $\cong_{E_2}$ coincide).

We shall see later that some conditions on $E$ imply that $\simeq$ refines $\approx$. An $\omega$-language $E$ such that $\simeq_E$ (or equivalently, $\sim_E$) is finite is called *finite-state*.
A *deterministic Muller automaton* is a quintuple $\mathcal{A} = (\Sigma, Q, \delta, q_0, \mathcal{T})$ where $\Sigma$ is the input alphabet, $Q$ is the state space, $\delta : Q \times \Sigma \to Q$ is the transition function, $q_0$ the initial state and $\mathcal{T} \subseteq 2^Q$ is a family of accepting subsets (the table). By $Inf(\mathcal{A}, \alpha)$ we denote the subset of $Q$ which is visited infinitely many times while $\mathcal{A}$ is reading $\alpha \in \Sigma^\omega$. The $\omega$-language accepted/recognized by $\mathcal{A}$ is $\{\alpha \in \Sigma^\omega : Inf(\mathcal{A}, \alpha) \in \mathcal{T}\}$. According to the Büchi-McNaughton theorem an $\omega$-language is regular iff it is recognized by some deterministic finite-state Muller automaton.
With every right-congruence relation we can associate an automaton, and in particular with the relation $\sim_E$ for a given $\omega$-language $E$:

**Definition 2 (Minimal-state automaton)** *Let $E$ be an $\omega$-language and let $\sim_E$ be its syntactic right-congruence* (Definition 1). *Its minimal-state automaton is*

$$\mathcal{A}_E := (\Sigma, Q, \delta, q_0)$$

*where $Q := \{\langle u \rangle : u \in \Sigma^*\}$, $q_0 := \langle e \rangle$, and $\delta(\langle u \rangle, a) := \langle ua \rangle$.*

Here, in contrast to the language case, not every (regular) $\omega$-language $E$ can be accepted by its minimal-state automaton $\mathcal{A}_E$. For example, the minimal-state automaton of $\{a, b\}^* a^\omega$ has only one state and does not accept $\{a, b\}^* a^\omega$, whereas there are several non-isomorphic two-state Muller automata accepting $\{a, b\}^* a^\omega$ (cf. [Mu63], [St83], [St87]).

## 3 Some observations on the iteration congruence

In this section we show that despite the fact that $\cong_E$ provides additional information on $E$ which is missing from $\simeq_E$, still it fails to characterize the regular $\omega$-languages in contrast to $\simeq$ for languages.

**Fact 1** *There are $\omega$-languages which are finite-state while their iteration monoid is infinite.*

**Proof**: Let the language $V \subseteq \{a,b\}^*$ be defined by the equation

$$V = a \cup bV^2 \,.$$

Alternatively, $V$ may be defined as the language consisting of those words $v \in \Sigma^*$ satisfying $|v|_a = |v|_b + 1$ and $|u|_a \leq |u|_b$ for every $u \prec v$. Let $E := V^\omega$. Then one easily verifies $E = VE = (a \cup bV^2)E = \{a,b\}E$. Thus $u \simeq_E v$ for every $u, v \in \{a,b\}^*$ and $\simeq_E$ is trivial.
In order to show that $\cong_E$ is infinite we prove that $(b^i a^{i+1})^\omega \in E$ and $(b^j a^{i+1})^\omega \notin E$, that is $b^i \ncong_E b^j$ for $0 < i < j$.
Since $b^i a^{i+1} \in V$, we have $(b^i a^{i+1})^\omega \in V^\omega = E$. On the other hand every word in $V$ contains more occurrences of $a$ than of $b$. Consequently, $j > i$ implies that the $\omega$-word $(b^j a^{i+1})^\omega$ has no prefix in V, whence $(b^j a^{i+1})^\omega \notin V\Sigma^\omega \supseteq E$. ❑

The second observation (as already noted in [Ar85]) is that, in general, the finiteness of $\cong_E$ does not guarantee regularity of $E$:

**Fact 2** *The $\omega$-language $Ult = \{uv^\omega : u \in \Sigma^*, v \in \Sigma^+\}$ of all ultimately periodic $\omega$-words has a trivial syntactic monoid, that is $x \cong_{Ult} y$ for every $x, y \in \Sigma^*$, but is not regular.*

Next we investigate the question which $\omega$-languages have a an iteration congruence of finite index. To this aim we show that with every $\omega$-language $E$ we can associate in a canonical way an $\omega$-language $F_E$ which is covered by $\cong_E$. Define

$$F_E = \bigcup \{[u][v]^\omega : uv^\omega \in E\}$$

where $[\cdot]$ denotes a congruence class of $\cong_E$. The following statement holds true.

**Lemma 3** *$E \cap Ult = F_E \cap Ult$.*

**Proof**: By definition $E \cap Ult \subseteq F_E \cap Ult$. Let $xy^\omega \in F_E$. Then there are $u, v$ such that $uv^\omega \in E$ and $xy^\omega \in [u][v]^\omega$. From this we can obtain words $y_1$ and $y_2$ such that $y = y_1 y_2$, and natural numbers $i, j, m$ and $n$ such that $xy^i y_1 \in [u][v]^m$ and $y_2 y^j y_1 \in [v]^n$. Since $\cong_E$ is a congruence, it follows that $xy^i y_1 \cong_E uv^m$ and $y_2 y^j y_1 \cong_E v^n$ and, because $uv^m (v^n)^\omega = uv^\omega \in E$ by the definition of $\cong_E$, also $xy^i y_1 (y_2 y^j y_1)^\omega = xy^\omega \in E$. ❑

**Theorem 4** *For every $E \subseteq \Sigma^\omega$, the iteration congruence $\cong_E$ is finite iff $E$ is finite-state and there is a regular $\omega$-language $F$ such that $E \cap Ult = F \cap Ult$.*

**Proof**: Let $E$ be finite-state and let the regular $\omega$-language $F$ satisfy $E \cap Ult = F \cap Ult$. It can be easily verified that $x \simeq_E y$ and $x \cong_{F \cap Ult} y$ imply $x \cong_E y$ and thus $\simeq_E \cap \cong_F \ \subseteq \ \cong_E$. But the congruences $\simeq_E$ and $\cong_F$ are both finite and so is $\cong_E$. Conversely, let $\cong_E$ be finite. Then $F_E$ is a regular $\omega$-language satisfying $E \cap Ult = F_E \cap Ult$. ❑

In [St83] it was shown that the cardinality of the set $\{E : \ \simeq_E$ is finite$\}$ is $2^{2^{\aleph_0}}$, in particular, there are already as many subsets of $\Sigma^\omega$ whose simple syntactic monoid is trivial. The following claim shows that the same is true in the case of $\cong_E$:

**Claim 5** *There are $2^{2^{\aleph_0}}$ $\omega$-languages having a trivial iteration congruence.*

**Proof**: Since the set $\{E : \simeq_E$ is trivial$\}$ is closed under the Boolean operations, any $\omega$-language $F$ for which $\simeq_F$ is trivial splits in a unique way into a disjoint union $(F \cap Ult) \cup (F - Ult)$ where for both parts $\simeq$ is trivial. As $Ult$ is countable, there are at most $2^{\aleph_0}$ distinct parts of the form $F \cap Ult$. Consequently, there are $2^{2^{\aleph_0}}$ $\omega$-languages $E \subseteq \Sigma^\omega - Ult$ such that $\simeq_E$ is trivial. But for every such $E$ $\approx_E$ is trivial and hence the iteration congruence of $E$ is trivial; this proves our assertion. ❑

Given that a Borel class in $\Sigma^\omega$ contains only $2^{\aleph_0}$ sets and that there are only countably many Borel classes [Ku66], it follows that there are $\omega$-languages $E$ even beyond the Borel hierarchy for which $\cong_E$ is trivial. This is in sharp contrast with the Myhill-Nerode theorem where the finiteness of the syntactic monoid implies the regularity of the language.

## 4 The case when $\simeq$ and $\cong$ coincide

In Theorem 21 of [St83] it was proved that every finite-state $\omega$-language $E \subseteq \Sigma^\omega$ which is simultaneously in the Borel classes $F_\sigma$ and $G_\delta$ is regular. Our aim is to show that this very condition also guarantees the iteration congruence of $E$ coincides with the simple syntactic congruence of $E$. It is remarkable that this condition holds for all $\omega$-languages in $F_\sigma \cap G_\delta$ not only for those which are finite-state.
First let us mention the following simple properties of the congruences $\simeq_E$ and $\cong_E$:

**Fact 6** *For every $u \in \Sigma^*$, $x, y \in \Sigma^+$:*

1. *If $x \simeq_E y$ then $u\{x,y\}^* x^\omega \cap E \neq \emptyset$ implies $u\{x,y\}^* x^\omega \subseteq E$*
2. *If $x \cong_E y$ then $u\{x,y\}^* x^\omega \cap E \neq \emptyset$ implies $u\{x,y\}^* y^\omega \subseteq E$.*

Now we obtain the following necessary and sufficient condition under which the congruences $\simeq_E$ and $\cong_E$ coincide:

**Lemma 7** *Let $E \subseteq \Sigma^\omega$. Then $\simeq_E = \cong_E$ if and only if the following condition holds*

$$\forall u \in \Sigma^* \, \forall x, y \in \Sigma^+ \big( x \simeq_E y \rightarrow (u\{x,y\}^* x^\omega \subseteq E \rightarrow u\{x,y\}^* y^\omega \cap E \neq \emptyset) \big) .$$

**Proof**: Clearly, the condition is necessary. In order to show its sufficiency we assume $x \simeq_E y$, and we show that then

$$\forall u, v \in \Sigma^* (u(xv)^\omega \in E \rightarrow u(yv)^\omega \in E)$$

that is, the additional condition for $\cong_E$ is satisfied.
If $x \simeq_E y$ and $u(xv)^\omega \in E$ then $xv \simeq_E yv$, and by the above claim we also have $u\{xv, yv\}^*(xv)^\omega \subseteq E$. Now our condition implies $u\{xv, yv\}^*(yv)^\omega \cap E \neq \emptyset$. Again the above claim shows that $u(yv)^\omega \in E$. ❑

As an immediate consequence we obtain the following simple sufficient condition. To express it we define:

**Definition 3** *An $\omega$-language $E$ has the period exchange property (or is period exchanging) provided for all $u \in \Sigma^*,\ x, y \in \Sigma^+$ the inclusion $u\{x, y\}^* x^\omega \subseteq E$ implies that $u\{x, y\}^* y^\omega \cap E \neq \emptyset$.*

**Corollary 8** *If $E$ has the period exchange property then $\simeq_E = \cong_E$.*

In order to prove the announced statement for $\omega$-languages in the Borel-class $F_\sigma \cap G_\delta$ we recall that for every $\omega$-language $E \in G_\delta$ there exists a language $U \in \Sigma^*$ such that for every $\beta \in \Sigma^\omega$, $\beta \in E$ iff $\beta$ has infinitely many prefixes in $U$. Using this we can show the following.

**Lemma 9** *Every $\omega$-language $E$ in the Borel-class $F_\sigma \cap G_\delta$ has the period exchanging property.*

**Proof**: Since both $E$ and its complement are in $G_\delta$, there exist two languages $U$ and $U'$ such that every $\omega$-word in $E$ has infinitely many prefixes in $U$ and every $\omega$-word not in $E$ has infinitely many prefixes in $U'$. Suppose that for some $u, x, y \in \Sigma^*$, $u\{x, y\}^* x^\omega \subseteq E$ and $u\{x, y\}^* y^\omega \subseteq \Sigma^\omega - E$.
Since $ux^\omega \in E$ there is a number $k_1$ such that $ux^{k_1}$ has a prefix in $U$, and since $ux^{k_1}y^\omega \notin E$, the word $ux^{k_1}y^{l_1}$ has a prefix in $U'$ for some $l_1$. Next we consider $ux^{k_1}y^{l_1}x^\omega \in E$: there must be some $k_2$ such that $ux^{k_1}y^{l_1}x^{k_2}$ has at least two prefixes in $U$, etc. Repeating this alternating argument, we construct an infinite sequence $ux^{k_1}y^{l_1} \ldots x^{k_i}y^{l_i} \ldots$ having infinitely many prefixes in $U$ and infinitely many prefixes in $U'$ and thus belonging simultaneously to $E$ and to its complement. ❏

This implies:

**Theorem 10** *For every $\omega$-language $E \in F_\sigma \cap G_\delta$, and every $x, y \in \Sigma^*$ $x \simeq_E y$ iff $x \cong_E y$.*

Note that the converse of Lemma 9 is not true in general: the $\omega$-language *Ult* is period exchanging, but not in $G_\delta$. However, for regular $\omega$-languages the converse is also true, —a similar observation was made in Theorem 6.2 of [Wi93].

**Lemma 11** *Every regular period exchanging $\omega$-language $E$ belongs to the Borel-class $F_\sigma \cap G_\delta$.*

**Proof**: From [SW74] (cf. also [Wa79]) it is known that a regular $\omega$-language $E$ is in $F_\sigma \cap G_\delta$ iff it is accepted by a finite-state Muller automaton $\mathcal{A}$ using a family of accepting subsets $\mathcal{T}$ having the following property: if $T \in \mathcal{T}$, $T = Inf(\mathcal{A}, \zeta)$ for some $\zeta \in \Sigma^\omega$, $T' = Inf(\mathcal{A}, \xi)$ for some $\xi \in \Sigma^\omega$, and $T \cap T' \neq \emptyset$ then $T' \in \mathcal{T}$.
Let $E$ be a regular period exchanging $\omega$-language accepted by a finite Muller automaton $\mathcal{A} = (\Sigma, Q, \delta, q_0, \mathcal{T})$, and let $T = Inf(\mathcal{A}, \zeta) \in \mathcal{T}$ be an accepting subset and let $T'$ be another subset such that $q \in T \cap T'$ for some $q \in Q$ and $Inf(\mathcal{A}, \xi) = T'$ for some $\xi \in \Sigma^\omega$. Among the $\omega$-words whose *Inf* is $T$ there is a word $ux^\omega$ satisfying $\delta(q_0, u) = q$, $\delta(q, x) = q$ and $T = \{\delta(q, x') : x' \preceq x\}$. Similarly there is a word $y$ such that $\delta(q, y) = q$ and $T' = \{\delta(q, y') : y' \preceq y\}$. One can see that for every $\alpha \in u\{x, y\}^* x^\omega$, $Inf(\mathcal{A}, \alpha) = T$ and thus $u\{x, y\}^* x^\omega \subseteq E$ and, since $E$ is period exchanging, we have some $\beta \in u\{x, y\}^* y^\omega \cap E$. But $Inf(\mathcal{A}, \beta) = T'$ and, hence, $T'$ must also be in $\mathcal{T}$. ❏

Although it follows from Claim 5 that $\simeq$ and $\cong$ coincide for some non-Borel sets, in general even for regular $\omega$-languages in the Borel class $F_\sigma$ it happens that $\simeq$ and $\cong$

may not coincide (cf. Example 1). On the other hand the following example shows a regular $\omega$-language in $F_\sigma$ where $\simeq$ and $\cong$ coincide; yet the language is not period exchanging:[1]

**Example 3** *Let $E_3 := \{a,b\}^* a^\omega \cup ca^\omega$. Then $\approx_{E_3}$ has as congruence classes $a^*$ and $\{a,b,c\}^* - a^*$, and the inclusion $\simeq_{E_3} \subseteq \approx_{E_3}$ is easily verified.*
*On the other hand $\{a,b\}^* a^\omega \subseteq E_3$ but $\{a,b\}^* b^\omega \cap E_3 = \emptyset$.*

# 5 Acceptance by minimal-state automata

In this section we will show that $\omega$-languages in $F_\sigma \cap G_\delta$ have another important property, namely they are accepted by their minimal-state automaton. Again, this property is true for arbitrary $\omega$-languages, not necessarily finite-state, provided that they can be accepted at all by a Muller automaton. The last reservation is in order because, as we show below, not every $\omega$-language, even those in $F_\sigma \cap G_\delta$, can be accepted by a Muller automaton.
For a given automaton $\mathcal{A}$ we will denote $\{\beta : Inf(\mathcal{A}, \beta) = \emptyset\}$ by $\mathcal{A}^\emptyset$ and $\{\beta : Inf(\mathcal{A}_E, \beta) = \emptyset\}$ by $E^\emptyset$, where $\mathcal{A}_E$ is the minimal-state automaton of $E$.

**Claim 12** *An $\omega$-language $E$ can be accepted by the Muller automaton $\mathcal{A} = (\Sigma, Q, \delta, q_0, \mathcal{T})$ only if $E^\emptyset \subseteq E$ or $E^\emptyset \cap E = \emptyset$.*

**Proof**: Clearly any Muller automaton can accept $E$ only if $\mathcal{A}^\emptyset \subseteq E$ or $E \cap \mathcal{A}^\emptyset = \emptyset$. Since any automaton $\mathcal{A}$ accepting $E$ refines $\mathcal{A}_E$ we have $E^\emptyset \subseteq \mathcal{A}^\emptyset$ and the result follows. ❏

Claim 12 is irrelevant in the case of a finite-state automaton $\mathcal{A}$, because then $E^\emptyset = \emptyset$. But the following example shows that for an infinite-state automaton $\mathcal{A}$ the set $E^\emptyset$ may indeed be non-empty.

**Example 4** *Let $\xi := aba^2b^2a^3b^3\ldots$ Clearly, $E_4 = \{\xi\}$ is not finite-state, more exactly, we have $u \not\sim_{E_4} v$ whenever $u \prec \xi$ and $u \prec v$, and $u \sim_{E_4} v$ when both $u$ and $v$ are not prefixes of $\xi$. Thus $E_4 = E_4^\emptyset$.*

We continue with an example of a simple $\omega$-language not accepted by a Muller automaton.

**Example 5** *Let $\xi := aba^2b^2a^3b^3\ldots$, let $\eta := bab^2a^2b^3a^3\ldots$ and consider the $\omega$-language $E_5 = \{\xi\} \cup (b\Sigma^\omega - \{\eta\})$. In the same way as above one obtains $\{\xi\} \cup \{\eta\} = E_5^\emptyset$ and Claim 12 shows that $E_5$ cannot be accepted by any Muller automaton.*

Note that, in this case, $E_5$ is the union of the closed set $\{\xi\}$ and the open set $(b\Sigma^\omega - \{\eta\})$, hence in $F_\sigma \cap G_\delta$. Moreover, since similar to Example 4, the $\omega$-languages $(b\Sigma^\omega - \{\eta\})$, $(\Sigma^\omega - \{\eta\})$, and $\{\xi\} \cup b\Sigma^\omega$ are accepted by their corresponding minimal-state automata, Example 5 shows that the class of $\omega$-languages accepted by arbitrary Muller

[1] A first example of this kind was obtained by Th. Wilke (personal communication).

automata is not closed under union and intersection (though it is obviously closed under complementation).
Having demonstrated this phenomenon we will show that $\omega$-languages in $F_\sigma \cap G_\delta$ which are accepted by Muller automata are already accepted by their minimal-state automata. First we mention a property of the right congruence $\sim_E$ for $\omega$-languages $E \in F_\sigma \cap G_\delta$ which follows from results of [St83]. For the sake of completeness we shall give the proof in Appendix A.

**Lemma 13** *If $E \subseteq \Sigma^\omega$ is in $F_\sigma \cap G_\delta$ then*

$$E = (E \cap E^\emptyset) \cup \bigcup_{u \in Pref(E) \wedge E(u) \cap E \neq \emptyset} E(u)$$

*where $Pref(E) = \{u : E \cap u\Sigma^\omega \neq \emptyset\}$ and $E(u) = \{\beta : u \prec \beta \wedge \forall v(v \prec \beta \to \exists x(vx \sim_E u))\}$.*
*Observe that here $vx$ need not be a prefix of $\beta$.*

This is a stronger property than the one given in [DL95] for saturating right congruences of regular $\omega$-languages $E \in F_\sigma \cap G_\delta$. Compare also to the Landweber right congruences for regular $\omega$-languages $E \in G_\delta$ derived in [Le90].
We mention still that, in view of the identity

$$\Sigma^\omega - E(u) = \{u' : |u'| = |u| \wedge u' \neq u\} \cdot \Sigma^\omega \cup \{v : \forall x(vx \not\sim_E u)\} \cdot \Sigma^\omega$$

every set $E(u)$ is closed.
In order to achieve our goal we proceed along the lines of [St83] and show that the connected components of $\mathcal{A}_E$ accept $E \in F_\sigma \cap G_\delta$ provided $E \subseteq E^\emptyset$ or $E \cap E^\emptyset = \emptyset$.
We set $S_u := \{w : \exists x \exists y(wx \sim_E u \wedge uy \sim_E w)\}$. Thus $S_u = S_{u'}$ if and only if $u' \in S_u$, and $\langle S_u \rangle := \{\langle w \rangle : w \in S_u\}$ is the strongly connected component of $\mathcal{A}_E$ containing $\langle u \rangle$. Moreover, $S_u$ is interval closed, that is, $w, w' \in S_u$ and $w \preceq w'$ imply that for all $v$, $w \preceq v \preceq w'$, it holds $v \in S_u$.
First we remove the condition $u \prec \beta$ from $E(u)$. We define

$$E'(u) := \{\beta : \exists w(w \preceq \beta \wedge w \in S_u) \wedge \forall v(v \preceq \beta \to \exists x(vx \sim_E u))\}.$$

$E'(u)$ has the following properties.

**Lemma 14** *1. $E'(u) = \{\beta : Pref(\beta) - S_u$ is finite$\}$*

*2. If $u' \in S_u$ then $E'(u) = E'(u')$.*

*3. $E'(u) = \bigcup_{u' \in S_u} E(u')$*

*4. $E'(u) \cap E \neq \emptyset$ if and only if $E(u) \cap E \neq \emptyset$*

**Proof**: First observe that $w \in S_u$, $w \prec v$ and $v \notin S_u$ imply $S_u \cap v\Sigma^* = \emptyset$.
1. Consider $\beta \in E'(u)$. Then $w \in S_u$ for some $w \prec \beta$. Now $Pref(\beta) = Pref(w) \cup \{v : w \preceq v \prec \beta\}$. It suffices to show $\{v : w \preceq v \prec \beta\} \subseteq S_u$.
Since $w \in S_u$, we have $uy \sim_E w$ for some $y \in \Sigma^*$, and, consequently, $uyy' \sim_E v$ when $wy' = v$. Now, $\exists x(vx \sim_E u))$ is immediate from $v \prec \beta \in E'(u)$.
If $\beta \notin E'(u)$ then either $Pref(\beta) \cap S_u = \emptyset$ or $w \in S_u$ and $v \notin S_u$ for some $w \prec v \prec \beta$. Since $S_u$ is interval closed, it follows $Pref(\beta) \cap S_u \subseteq Pref(v)$.
2. follows from 1. and $S_u = S_{u'}$.

3. In view of 2. the inclusion $\supseteq$ is obvious. Conversely, if $\beta \in E'(u)$ then $u' \prec \beta$ for some $u' \in S_u$ and, consequently, $\beta \in E(u')$.
4. Let $w\alpha \in E'(u) \cap E$ where $w \in S_u$. Then $uy \sim_E w$ for some $y \in \Sigma^*$. Consequently, $uy\alpha \in E$.
In order to show $uy\alpha \in E(u)$ we consider the prefixes $v \prec \alpha$. They satisfy $wvx_v \sim_E u$ for suitable $x_v \in \Sigma^*$. From $uy \sim_E w$ we obtain $uyvx_v \sim_E u$, whence $ux\alpha \in E(u)$. ❑

Now we can prove our result generalising Theorem 24 of [St83].

**Theorem 15** *Let $E \in F_\sigma \cap G_\delta$ such that either $E^\varnothing \subseteq E$ or $E \cap E^\varnothing = \varnothing$. Then $E$ is accepted by its minimal-state automaton $\mathcal{A}_E$.*

**Proof**: We observe that the class of all subsets of $\Sigma^\omega$ acceptable by Muller automata as well as the class $F_\sigma \cap G_\delta$ are closed under complementation. So we may assume without loss of generality that $E^\varnothing \subseteq E$.
From Lemmata 13 and 14 it follows that $E = E^\varnothing \cup \bigcup_{u \in Pref(E) \wedge E(u) \cap E \neq \varnothing} E'(u)$. So for every $u$ we let $T_u = \{\langle v \rangle : \exists x \exists v (vx \sim_E u \wedge uy \sim_E uv)\}$ be the strongly connected component of $\mathcal{A}_E$ which contains $\langle u \rangle$ and we let $\mathcal{T}_u$ be $2^{T_u} - \{\varnothing\}$.
Then for every $\alpha \in \Sigma^\omega$ , $\alpha \in E'(u) - E^\varnothing$ iff $Inf(\mathcal{A}_E, \alpha) \in \mathcal{T}_u$, so by letting $\mathcal{T} = \{\varnothing\} \cup \bigcup_{u \in Pref(E) \wedge E'(u) \cap E \neq \varnothing} \mathcal{T}_u$ we can make $\mathcal{A}_E$ accept $E$. ❑

Since for a finite-state $\omega$-language $E$ the set $E^\varnothing$ is always empty, our theorem yields as an immediate consequence the assertion of Theorems 21 and 24 of [St83].

**Corollary 16** *If $E$ is a finite-state $\omega$-language in $F_\sigma \cap G_\delta$ then $E$ is regular and is accepted by its (finite) minimal-state automaton $\mathcal{A}_E$.*

Note that Example 1 shows that this condition ($E$ being in $F_\sigma \cap G_\delta$) is not a necessary one:

**Example 1 (continued)** *Theorem 10 and Example 1 prove that $E_1 \notin F_\sigma \cap G_\delta$* (In fact $E_1$ is in $F_\sigma$, since it is a countable set, hence $E_1 \notin G_\delta$.)*, but it is easily verified that $\mathcal{A}_{E_1}$ accepts $E_1$* (cf. [St83, Example 1]).[2]

Next we will provide a necessary condition for an $\omega$-language $E$ to be acceptable by its minimal-state automaton $\mathcal{A}_E$. This condition is based on a relation between $\approx_E$ and $\simeq_E$ and is valid for arbitrary (not necessarily regular) $\omega$-languages.
Let us define a congruence relation based on $\sim_E$ which refines $\simeq_E$ by considering two words to be equivalent only if they have the same set of right-factors (modulo $\sim_E$).

**Definition 4 (Factorized congruence)** *The factorization of $\sim_E$ is a congruence $\sim_E^*$ defined as $x \sim_E^* y$* iff $\forall u \in \Sigma^*$

1. *$ux \sim_E uy$ and*
2. *$\forall v (v \preceq x \rightarrow \exists v' (v' \preceq y \wedge uv \sim_E uv'))$ and*

[2] N. Gutleben (personal communication) showed that arbitrary high degrees of Wagner's [Wa79] hierarchy contain regular $\omega$-languages $E$ accepted by $\mathcal{A}_E$.

3. $\forall v'(v' \preceq y \rightarrow \exists v(v \preceq x \wedge uv \sim_E uv'))$

It is more intuitive to see the meaning of this relation in terms of the minimal-state automaton $\mathcal{A}_E$. Here $x \sim_E^* y$ iff from every state $q$ both $x$ and $y$ lead to the same state while visiting the same set of states. (Observe that $x \simeq_E y$ iff from every state $q$ of $\mathcal{A}_E$ both $x$ and $y$ lead to the same state without necessarily visiting the same set of intermediate states). One can see that $u \sim_E v$ and $x \sim_E^* y$ imply that for every $z$, $\mathit{Inf}(\mathcal{A}_E, u(xz)^\omega) = \mathit{Inf}(\mathcal{A}_E, v(yz)^\omega)$. A similar refinement of the right congruence related to a deterministic automaton was introduced in [DL95] as the cycle congruence of an automaton.

**Claim 17** *An $\omega$-language $E$ can be accepted by its minimal-state automaton $\mathcal{A}_E$ using Muller condition only if $x \sim_E^* y$ implies $x \approx_E y$ for all $x, y \in \Sigma^*$.*

**Proof**: Suppose that $x \sim_E^* y$ and $x \not\approx_E y$, that is, for some $x \sim_E^* y$, there exist $u, v$ such that $u(xv)^\omega \in E$ and $u(yv)^\omega \notin E$. But $xv \sim_E^* yv$, hence $\mathit{Inf}(\mathcal{A}_E, u(xv)^\omega) = \mathit{Inf}(\mathcal{A}_E, u(yv)^\omega)$, and $\mathcal{A}_E$ cannot accept $E$. ❑

The condition of the previous claim fails to be sufficient. To this end consider again Example 3.

**Example 3 (continued)** *One verifies that $\mathcal{A}_{E_3}$ cannot accept $E_3 = \{a, b\}^* a^\omega \cup ca^\omega$. But in virtue of $\sim_{E_3}^* \subseteq \simeq_{E_3}$ and $\simeq_{E_3} \subseteq \approx_{E_3}$ we have $\sim_{E_3}^* \subseteq \approx_{E_3}$.*

Intuitively the reason is that $\sim_{E_3}^*$ is too refined: $a \not\sim_{E_3}^* b$ because $ca \not\sim_{E_3} cb$ and yet $\mathit{Inf}(\mathcal{A}_E, aa^\omega) = \mathit{Inf}(\mathcal{A}_E, ab^\omega)$. In the next section we will introduce more suitable definitions for that purpose. Recalling that $\simeq_{E_3}$ and $\cong_{E_3}$ coincide, we can conclude that the questions whether $\mathcal{A}_E$ accepts $E$ and whether $\simeq_E$ and $\cong_E$ coincide, being both related to the study of syntactic congruences, are likewise independent (cf. also Appendix B).

# 6 Recognition by right-congruences

In this section we will develop an alternative theory of recognition of $\omega$-languages by right-congruence relations, as a complement to the recognition by two-sided congruences (monoids) described in [Ar85, Ei74, PP93, Th90]. Using this theory we give a necessary and sufficient condition for a regular $\omega$-language to be accepted by its minimal-state automaton.

**Definition 5 (Family of right-congruences)** *A family of right-congruences* (FORC) *is a pair $\mathcal{R} = (\sim, \{\sim_u\}_{\langle u\rangle \in \Sigma^*/\sim})$ such that:*

1. *$\sim$ is a right-congruence relation.*

2. *$\sim_u$ is a right-congruence relation for every $\langle u\rangle \in \Sigma^*/\sim$.*

3. *For all $u, x, y \in \Sigma^*$, $x \sim_u y$ implies $ux \sim uy$.*

As we can see, a FORC consists of a "leading" relation $\sim$ and a relation associated with each of its classes. We will denote classes of $\sim$ by $\langle u\rangle$ and classes of $\sim_u$ by $\langle v\rangle_u$. A FORC is finite if all the right-congruences are of finite index. As in the case of finite congruences, the following factorization property holds:

**Lemma 18** *Let $\mathcal{R}$ be a finite* FORC*. Then every $\omega$-word $\alpha$ has a factorization $\alpha = uv_1v_2,\ldots$ such that $v_i \sim_u v_{i+1}$ and $uv_i \sim u$ for all $i > 0$.*

**Proof**: (Along the same lines of the proof of Lemma 2.2 in [Th90] for congruences, cf. also Section 2.3 of [PP93]). Let $\alpha = u\beta$ such that $\langle u\rangle$ is a class of $\sim$ that appears infinitely often in $\alpha$ and let $J = \{j_1, j_2, \ldots\}$ be an increasing sequence of indices such that $u\beta(1..j_i) \sim u$ for every $i$. Next we define an equivalence relation on $\mathbb{N}$: $n_1 \sim^\beta n_2$ if for some $m > n_1, n_2$ $\beta(n_1..m) \sim_u \beta(n_2..m)$ (in other words, positions $n_1$ and $n_2$ "merge" after $m$). By the finiteness of $\sim_u$, $\sim_\beta$ is finite too, so we can take an infinite sub-sequence of of indices $K = \{k_1, k_2, \ldots\} \subseteq J$ such that $k_i < k_{i+1}$ and $k_i + 1 \sim^\beta k_{i+1} + 1$, that is, for every $i$ there is some $m_i \geq k_{i+1} + 1$ such that $\beta(k_i + 1..m_i) \sim_u \beta(k_{i+1} + 1..m_i)$. Finally we take a sub-sequence of indices $L = \{l_1, l_2, \ldots\} \subseteq K$ such that for some $v$, $\beta(l_1 + 1..l_i) \in \langle v\rangle_u$ for every $i$, and $\beta(l_i + 1..m) \sim_u \beta(l_{i+1} + 1..m)$ for some $m \leq l_{i+2}$.
Let $v_i := \beta(l_i + 1..l_{i+1})$. Then $v_1 \cdots v_i \sim_u v$.
By definition of $\sim^\beta$ it also implies $\beta(l_i + 1..l_{i+2}) \sim_u \beta(l_{i+1} + 1..l_{i+2})$,that is $v_iv_i + 1 \sim_u v_i + 1$, so, by induction, for every $i \geq 1$, $\beta(l_i + 1..l_{i+1}) = v_i \in \langle v\rangle_u$ and together with $u\beta(1..l_1) \sim u$ we have the desired factorization. ❑

**Definition 6 (Recognition by FORCs)** *An $\omega$-language $E$ is covered by a* FORC *$\mathcal{R} = (\sim, \{\sim_u\}_{\langle u\rangle \in \Sigma^*/\sim})$ if it can be written as a union of sets of the form $\langle u\rangle(\langle v\rangle_u)^\omega$ such that $uv \sim u$. An $\omega$-language $E$ is saturated by $\mathcal{R}$ if for every $u, v$ such that $uv \sim v$, $\langle u\rangle(\langle v\rangle_u)^\omega \cap E \neq \emptyset$ implies $\langle u\rangle(\langle v\rangle_u)^\omega \subseteq E$. An $\omega$-language $E$ is recognized by $\mathcal{R}$ if it is both covered and saturated by it.*

As for congruences, in the special case of finite FORCs, covering and saturation coincide.

**Lemma 19** *A finite* FORC *$\mathcal{R}$ covers an $\omega$-language $E$ if and only if it saturates $E$.*

**Proof**: (See also proof of Lemma 1.1 in [Ar85]). Saturation implies covering by virtue of the Factorization Lemma 18. Now we show that covering implies saturation: Suppose $\langle u\rangle(\langle v\rangle_u)^\omega \cap E \neq \emptyset$. Since $\langle u\rangle(\langle v\rangle_u)^\omega \cap E$ is regular it contains an ultimately-periodic word $xy^\omega$. Since $y$ is finite we have $xy^\omega = z_1z_2^\omega$, where $z_1 = xy^{n_1}y_1$, $z_2 = y_2y^{n_2}y_1$, $y = y_1y_2$, $z_1 \sim uv^{m_1} \sim u$ and $z_2 \sim_u v^{m_2}$. Since $\langle u\rangle(\langle v\rangle_u)^\omega \subseteq \langle uv^{m_1}\rangle(\langle v^{m_2}\rangle_u)^\omega$, by covering we have $\langle u\rangle(\langle v\rangle_u)^\omega \subseteq E$. ❑

Next we will show how every deterministic automaton $\mathcal{A} = (\Sigma, Q, \delta, q_0)$ defines an associated FORC that bears important information about the transition structure of the automaton. For every $q \in Q$ and $u \in \Sigma^*$ we will denote by $Vis(q, u)$ the set of states visited by the automaton while reading $u$ starting at $q$, and let $MSCC(q) := \{q' : \exists x(\delta(q, x) = q') \wedge \exists y(\delta(q', y) = q)\}$ be the maximal strongly-connected component in the transition graph of $\mathcal{A}$ which contains $q$.

**Definition 7 (The FORC of an automaton)** *Let $\mathcal{A} = (\Sigma, Q, \delta, q_0)$ be a deterministic automaton. The* FORC *associated with $\mathcal{A}$ is $\mathcal{R}_{\mathcal{A}} = (\sim, \{\sim_u\}_{\langle u\rangle \in \Sigma^*/\sim})$ defined as:*

1. $x \sim y$ *iff* $\delta(q_0, x) = \delta(q_0, y)$
2. $x \sim_u y$ *iff* $Vis(q, x) \cap MSCC(q') = Vis(q, y) \cap MSCC(q')$ *whenever* $\delta(q_0, u) = q$ *and* $\delta(q, x) = \delta(q, y) = q'$.

In other words $x$ and $y$ are congruent from $q = \delta(q_0, u)$ if they lead to the same state, and if they visit the same set of states which the automaton may still visit in the future. It is easily verified that $\mathcal{R}_{\mathcal{A}}$ is indeed a FORC.

**Claim 20** *Two $\omega$-words have the same Inf in $\mathcal{A}$ if and only if they have equivalent $\mathcal{R}_{\mathcal{A}}$-factorizations into $\langle u\rangle(\langle v\rangle_u)^\omega$ with $uv \sim u$.*

**Proof**: Let $\alpha$ be any $\omega$-word such that $Inf(\mathcal{A}, \alpha) = T = \{q_1, \ldots, q_m\}$ and let $i_1$ be the first occurrence of $q_1$ in the run of the automaton over $\alpha$ after all the states in $Q - T$ have disappeared. For every $k > 1$ let $i_k$ be the first occurrence of $q_1$ such that all the states in $T$ occurred between positions $i_{k-1}$ and $i_k$. By letting $u = \alpha(1..i_1)$ and $v_k = \alpha(i_k + 1..i_{k+1})$ we obtain the desired factorization. Conversely it is immediate to see that such a factorization determines $Inf(\mathcal{A}, \alpha)$. ❑

**Corollary 21** *A Muller automaton $\mathcal{A}$ can accept $E$ if and only if its* FORC *$\mathcal{R}_{\mathcal{A}}$ recognizes $E$.*

**Proof**: If $\mathcal{R}_{\mathcal{A}}$ does not recognize $E$ there must be some $\alpha \in E$ and $\beta \notin E$ having the same *Inf* and $\mathcal{A}$ cannot accept $E$. If $R_{\mathcal{A}}$ recognizes $E$ then for every $T \in 2^Q$ all the words $\alpha \in \Sigma^\omega$ such that $Inf(\mathcal{A}, \alpha) = T$ have an identical factorization, and thus the set $\mathcal{T}$ of accepting subsets can be determined consistently. ❑

**Theorem 22 ("Myhill-Nerode" theorem for $\omega$-languages)** *An $\omega$-language is regular if and only if it is recognized by a finite* FORC.

**Proof**: The only-if part follows from Corollary 21. Suppose $E$ is recognized by a FORC. Since every set $\langle u\rangle$ and $\langle v\rangle_u$ is regular, every finite union of sets of the form $\langle u\rangle(\langle v\rangle_u)^\omega$ is, by definition, $\omega$-regular. ❑

The next step is to define a partial-order among FORCs.

**Definition 8** *Let $\mathcal{R} = (\sim, \{\sim_u\}_{\langle u\rangle \in \Sigma^*/\sim})$ and $\mathcal{R}' = (\sim', \{\sim'_u\}_{\langle u\rangle \in \Sigma^*/\sim'})$ be two* FORCs. *We say that $\mathcal{R}'$ refines $\mathcal{R}$ ($\mathcal{R}' \leq \mathcal{R}$) if*

1. $\forall x, y \in \Sigma^* (x \sim' y \rightarrow x \sim y)$, *and if* $\sim' = \sim$ *then*
2. $\forall u, x, y \in \Sigma^* (x \sim'_u y \rightarrow x \sim_u y)$.

**Definition 9 (Syntactic FORC)** *Let $E$ be a regular $\omega$-language. The syntactic* FORC *associated with $E$ is $\mathcal{R}_E = (\sim_E, \{\approx_u\}_{\langle u\rangle \in \Sigma^*/\sim_E})$ where $\sim_E$ is the syntactic right-congruence of $E$ and for every $u$, $x \approx_u y$ iff*

1. $ux \sim_E uy$ and
2. $\forall v(v \in \Sigma^* \wedge uxv \sim_E u \rightarrow (u(xv)^\omega \in E \iff u(yv)^\omega \in E))$

One can see that $\approx_u$ is coarser than the infinitary congruence $\approx$ in two respects:

1. It does not quantify over all $u$ (just those in $\langle u \rangle$), and
2. it does not quantify over all $v$, only over those for which $xv$ (and hence also $yv$) makes a cycle from $\langle u \rangle$.

**Lemma 23** *Any regular $\omega$-language $E$ is recognized by its syntactic* FORC $\mathcal{R}_E$.

**Proof**: (We prove it similarly to Lemma 2.2 in [Ar85]). Suppose the contrary, i.e., $\langle u \rangle (\langle v \rangle_u)^\omega \cap E \neq \emptyset$ but $\langle u \rangle (\langle v \rangle_u)^\omega \not\subseteq E$ for some $u, v$ satisfying $uv \sim_E u$. Then by regularity there exist $uv^\omega \in E$ and $xy^\omega \in \langle u \rangle (\langle v \rangle_u)^\omega - E$. Due to the finiteness of $y$ there exist some $m, n$ such that $xy^\omega = zx_1 \ldots x_m (y_1 \ldots y_n)^\omega$ with $z \sim_E u$ and $x_i \approx_u y_j \approx_u v$ for every $i \leq m, j \leq n$. This implies that $zx_1 \ldots x_m \sim_E u$ and $y_1 \ldots y_n \approx_u v^n$ and thus by the definition of $\approx_u$, $zx_1 \ldots x_m (y_1 \ldots y_n)^\omega \in E$ if $u(v^n)^\omega \in E$ which means $xy^\omega \in E$, a contradiction. ❑

**Theorem 24** *For every regular $\omega$-language $E$, its syntactic* FORC $\mathcal{R}_E$ *is the largest* FORC *recognizing it.*

**Proof**: Let $\mathcal{R} = (\sim, \{\sim_u\}_{\langle u \rangle \in \Sigma^* / \sim})$ be a FORC recognizing $E$ and let $\mathcal{R} > \mathcal{R}_E$. Then $\sim\ \supset\ \sim_E$, or $\sim\ =\ \sim_E$ and $\sim_u\ \supset\ \approx_u$ for some $u \in \Sigma^*$.
First, suppose that for some $x, y$ we have $x \sim y$ but $x \not\sim_E y$, that is, for some $\alpha \in \Sigma^\omega$, $x\alpha \in E$ but $y\alpha \notin E$. But then $x\alpha$ has a factorization $x\alpha = uv_1v_2, \ldots$ where $x \preceq u$ and $v_i \in \langle v \rangle_u$. Since $x \sim y$, $y\alpha$ has a similar factorization $y\alpha = u'v_1v_2, \ldots$ with $u' \sim u$ and thus we have shown $\langle u \rangle (\langle v \rangle_u)^\omega$ contains both $x\alpha$ and $y\alpha$ contrary to the assumption that $\mathcal{R}$ recognizes $E$.
Suppose now that $\sim = \sim_E$ and for some $u, x, y$ we have $x \sim_u y$ but $x \not\approx_u y$. This means that there is some $z$ such that $uxz \sim uyz \sim u$ and $u(xz)^\omega \in E$ but $u(yz)^\omega \notin E$. Since $\sim_u$ is a right-congruence we also have $xz \sim_u yz$ and thus $\langle u \rangle (\langle xz \rangle_u)^\omega$ contains both members and non-members of $E$, again, contrary to the assumption that $\mathcal{R}$ recognizes $E$. ❑

Applying this result to the FORC associated with an automaton we get:

**Corollary 25** *A Muller automaton $\mathcal{A}$ can accept a regular $\omega$-language $E$ if and only if its associated* FORC $\mathcal{R}_\mathcal{A}$ *refines the syntactic* FORC $\mathcal{R}_E$.

In particular, considering the minimal-state automaton of $E$, $\mathcal{A}_E$, its corresponding FORC can be rephrased as follows:

**Definition 10 (Automatic-Syntactic FORC)** *Let $E$ be a regular $\omega$-language. The automatic-syntactic* FORC *associated with $E$ is $\mathcal{R}_{\mathcal{A}_E} = (\sim_E, \{\sim_u^*\}_{\langle u \rangle \in \Sigma^* / \sim_E})$ where $\sim_E$ is the syntactic right-congruence of $E$ and for every $u$, $x \sim_u^* y$ iff*

1. $ux \sim_E uy$ and
2. $\forall v(v \preceq x \wedge \exists z(uxz \sim_E uv) \to \exists v'(v' \preceq y \wedge uv \sim_E uv'))$ and
3. $\forall v'(v' \preceq y \wedge \exists z(uyz \sim_E uv') \to (\exists v(v \preceq x \wedge uv \sim_E uv')).$

This is just a reformulation of Definition 7 but in an automaton-free manner. As a direct application we can give an exact characterization of those regular $\omega$-languages that can be accepted by their minimal-state automaton.

**Theorem 26** *Let $E$ be a regular $\omega$-language Let $\mathcal{R}_E = (\sim_E, \{\approx_u\}_{\langle u\rangle \in \Sigma^*/\sim_E})$ be its syntactic* FORC (Definition 9) *and let $\mathcal{R}_{\mathcal{A}_E} = (\sim_E, \{\sim_u^*\}_{\langle u\rangle \in \Sigma^*/\sim_E})$ be its automatic-syntactic* FORC (Definition 10). *$E$ can be accepted by the automaton $\mathcal{A}_E$ if and only if for all $u, x, y \in \Sigma^*$, $x \sim_u^* y$ implies $x \approx_u y$.*

**Proof**: It follows from Corollary 25 ❑

As an illustration consider once more $E_3 = \{a,b\}^* a^\omega \cup ca^\omega$. Now we have $a \sim_a^* b$ but $a \not\approx_a b$ and for this reason $\mathcal{A}_{E_3}$ cannot accept $E_3$.
On the other hand consider $E = a^* b\{b \cup aa\}^* ab^\omega$. The methods developed in [Wa79] prove that $E \in F_\sigma - G_\delta$, hence $E \notin F_\sigma \cap G_\delta$. Here the classes of $\sim_E$ are $\langle e\rangle = a^*$, $\langle b\rangle = a^* b\{b \cup ab^* a\}^*$ and $\langle ba\rangle = a^* b\{b \cup ab^* a\}^* ab^*$. The following table depicts the congruence classes of $\{\sim_u^*\}$ and $\{\approx_u\}$ for $\langle u\rangle \in \Sigma^*/\sim_E$ and one can, indeed, see that the condition of Theorem 26 is satisfied, $E = \langle ba\rangle(\langle b\rangle_{ba})^\omega$ and $\mathcal{A}_E$ accepts $E$.

| $\langle u\rangle$ | $\approx_u$ | $\sim_u^*$ |
|---|---|---|
| $\langle e\rangle$ | $a^*$<br>$a^* b\{b \cup ab^* a\}^*$<br>$a^* b\{b \cup ab^* a\}^* ab^*$ | $a^*$<br>$a^* b^+, a^* b(b^* ab^* a)^+ b^*$<br>$a^* b\{b \cup ab^* a\}^* ab^*$ |
| $\langle b\rangle$ | $\{b \cup ab^* a\}^*$<br>$\{b \cup ab^* a\}^* ab^*$ | $b^*, (b^* ab^* a)^+$<br>$\{b \cup ab^* a\}^* ab^*$ |
| $\langle ba\rangle$ | $b^*$<br>$(b^* ab^* a)^+ b^*$<br>$(b^* ab^* a)^* b^* a$ | $b^*$<br>$(b^* ab^* a)^+ b^*$<br>$(b^* ab^* a)^* b^* a$ |

It can be easily verified that for $\omega$-languages having the period exchange property the hypothesis of Theorem 26 is trivially satisfied. Hence in connection with the Lemmas 9 and 11 we obtain an alternative proof of Corollary 16.
Unlike Theorem 10, our Theorem 26 and also Lemma 23 in general do not hold for arbitrary $\omega$-languages: Consider, e.g., the $\omega$-language *Ult* defined above. Since $\sim_{Ult}$ is trivial and *Ult* contains all ultimately periodic $\omega$-words, also the congruences $\sim_u^*$ and $\approx_u$ are trivial. Hence, $x \sim_u^* y$ implies $x \approx_u y$, but $\mathcal{A}_{Ult}$ does not accept *Ult*.
The introduction of the FORC concept may have significance beyond the proof of the above theorem. Up to now the only syntactic characterization of $\omega$-languages was by means of a two-sided congruence and the lack of the other half of a Myhill-Nerode theorem was believed to be an inherent feature of the theory of $\omega$-languages — we have shown that this is not the case. From a practical point of view, although the iteration congruence $\cong_E$ (which is the intersection of $\sim_E$ with $\{\approx_u\}_{\langle u\rangle \in \Sigma^*/\sim_E}$) has a simpler

definition, its size might be exponentially larger, and there are situations[3] where *the right-congruences are the right congruences*.

## Acknowledgement

We would like to thank A. Arnold for not believing a stronger version of Claim 17 and E. Badouel for pointing out that an earlier version of Theorem 26 was weaker than necessary. Finally, our debts are due to N. Klarlund for pointing out a flaw in an earlier version of Definition 8, and to an anonymous referee for some useful suggestions.

# References

[Ar85] A. Arnold, A syntactic congruence for rational $\omega$-languages, *Theoret. Comput. Sci.* 39, 333–335, 1985.

[DL95] Do Long Van, B. LeSaëc and I. Litovsky, Characterizations of rational $\omega$-languages by means of right congruences, *Theoret. Comput. Sci.* 143 (1995) 1, 1–21.
Preliminary version: A syntactic approach to deterministic $\omega$-automata. in D. Krob (Ed.), *Théorie des Automates et Applications*, 133–146, Actes des Deuxièmes Journées Franco-Belges, Rouen, 1991.

[Ei74] S. Eilenberg, *Automata, Languages and Machines, Vol. A*, Academic Press, New York, 1974.

[EH93] J. Engelfriet and H.J. Hoogeboom, X-automata on $\omega$-words. *Theoret. Comput. Sci.* 110, 1–51, 1993.

[HR85] H.J. Hoogeboom and G. Rozenberg, Infinitary Languages – Basic Theory and Applications to Concurrent Systems, in J.W. de Bakker et al. (Eds.), *Current Trends in Concurrency*, 266–342, LNCS 224, Springer-Verlag, Berlin, 1985.

[JT83] H. Jürgensen and G. Thierrin, On $\omega$-languages whose syntactic monoid is trivial, *Intern. J. Comput. Inform. Sci.* 12, 359–365, 1983.

[Ku66] K. Kuratowski, *Topology I*, Academic Press, New York, 1966.

[Le90] B. LeSaëc, Saturating right congruences, *RAIRO Infor. théor. et Appl.* 24, 545–560, 1990.

[LS77] R. Lindner and L. Staiger, *Algebraische Codierungstheorie – Theorie der sequentiellen Codierungen*, Akademie-Verlag, Berlin, 1977.

[MP95] O. Maler and A. Pnueli, On the Learnability of Infinitary Regular Sets, *Information and Computation* 118, 316-326, 1995.
Preliminary version: in L.G. Valiant and M.K. Warmuth (Eds.), Proc. of the 4th annual workshop on *Computational Learning Theory*, Morgan Kaufmann, San Mateo, 1991.

[3] For example, when we want to learn an $\omega$-language from examples as in [MP95].

[Mu63] D.E. Muller, Infinite sequences and finite machines, in Proc. of the 4th Ann. IEEE Symp. *Switching Theory and Logical Design*, Chicago 1963, 3–16.

[PP93] D. Perrin and J.-E. Pin, *Mots Infinis*, Report LITP 93.40, Institut Blaise Pascal, Paris, 1993.

[St83] L. Staiger, Finite-state $\omega$-languages, *J. Comput. System Sci.* 27, 434–448, 1983.

[St87] L. Staiger, Research in the Theory of $\omega$-languages, *J. Inf. Process. Cybern. EIK* 23, 415–439, 1987.

[SW74] L. Staiger and K. Wagner, Automatentheoretische und automatenfreie Charakterisierungen topologischer Klassen regulärer Folgenmengen, *Elektron. Informationsverarb. Kybernet. EIK* 10, 379–392, 1974.

[Th90] W. Thomas, Automata on Infinite Objects, in J. Van Leeuwen (Ed.), *Handbook of Theoretical Computer Science*, Vol. B, 133–191, Elsevier, Amsterdam, 1990.

[Tr62] B.A. Trakhtenbrot, Finite automata and monadic second order logic, *Siberian Math. J.* 3, 103–131, 1962. (Russian; English translation in: AMS Transl. 59, 23–55, 1966.)

[Wa79] K. Wagner, On $\omega$-regular sets, *Inform. Control* 43, 123–177, 1979.

[Wi93] Th. Wilke, An algebraic theory for regular languages of finite and infinite words, *Intern. J. Algebra and Computation* 3, 447–489, 1993.
Preliminary version: An Eilenberg theorem for $\infty$-languages, in J. Albert Leach, B. Monien and M. Rodríguez Artalejo (Eds.), Proc. of the 18th. Intern. Colloquium on *Automata, Languages and Programming*, 588–599, LNCS 510, Springer-Verlag, Berlin 1991.

# A Proof of Lemma 13

As it was announced in Section 5 we give a proof of Lemma 13:
*If $E \subseteq \Sigma^\omega$ is simultaneously an $F_\sigma$- and a $G_\delta$-set then*

$$E = (E \cap E^{\varnothing}) \cup \bigcup_{w \in Pref(E) \wedge E(u) \cap E \neq \varnothing} \{\beta : w \prec \beta \wedge \forall u(u \preceq \beta \rightarrow \exists v(uv \sim_E w))\}.$$

To this end we introduce some notation. We call an $\omega$-language $D \subseteq \Sigma^\omega$ *strongly connected* iff

$$\forall u\big(u \in Pref(D) \rightarrow \exists v(v \in \Sigma^* \wedge uv \sim_D e)\big),$$

that is, for every $u$ which is a finite prefix of some $\omega$-word $\xi \in D$ there is a $v \in \Sigma^*$ such that $D \cap uv\Sigma^\omega = uvD$.
This notion corresponds to the strong connectivity of the partial automaton $\mathcal{A}'_D$ which is obtained from the minimal-state automaton $\mathcal{A}_D$ by deleting the state (dead sink) $\langle \tilde{w} \rangle = \{w : w \notin Pref(D)\}$.
With an arbitrary $\omega$-language $D$ we associate the following $\omega$-language $\tilde{D}$

$$\tilde{D} := \{\xi : \forall u(u \prec \xi \rightarrow \exists v(uv \sim_D e))\}, \tag{5}$$

and its *connected part* $\mathrm{cn}(D)$ (cf. also [St83]).

$$\mathrm{cn}(D) := D \cap \tilde{D}. \tag{6}$$

*Remark*: In (5) we can likewise replace the quantifier $\forall u$ by $\exists^\infty u$ (there are infinitely many $u$)
Since $\tilde{D} = \Sigma^\omega - \{w : \forall v(wv \not\sim_D e)\} \cdot \Sigma^\omega$, the $\omega$-language $\tilde{D}$ is closed.
Moreover, we have the following.

**Fact 27** *Let for $D \subseteq \Sigma^\omega$ the $\omega$-language $\tilde{D}$ be defined as above. Then $w \sim_D w'$ implies $w \sim_{\tilde{D}} w'$ and $w \sim_{\mathrm{cn}(D)} w'$.*

**Proof.** Let $w \sim_D w'$. In order to show $w \sim_{\tilde{D}} w'$, by symmetry it suffices to verify that $w\beta \in \tilde{D}$ implies $w'\beta \in \tilde{D}$. Now let $(u_i)_{i \in \mathbb{N}}$ be an infinite family of finite prefixes of $\beta$ such that $\forall i \exists v_i(wu_iv_i \sim_D e)$. Then in view of $w \sim_D w'$ we have also $\forall i \exists v_i(w'u_iv_i \sim_D e)$. Hence $w'\beta \in \tilde{D}$.
Now $w \sim_{\mathrm{cn}(D)} w'$ follows from (6). ❑

Furthermore the connected part $\mathrm{cn}(D)$ has the following properties (cf. [St83, Lemma 16 and Proposition 17]).

**Lemma 28** 1. *$\mathrm{cn}(D)$ is a strongly connected $\omega$-language.*

2. *If $\mathrm{cn}(D)$ is nonempty and closed then $\mathrm{cn}(D) = \tilde{D}$*

**Proof.** 1. Let $w \in Pref(\mathrm{cn}(D)) = Pref(D \cap \tilde{D})$. Then in view of the definition of $\tilde{D}$ (see (5) above) $wv \sim_D e$ for some $v$. Now Fact 27 shows $wv \sim_{\mathrm{cn}(D)} e$.
2. Assume $\varnothing \neq \mathrm{cn}(D) \subset \tilde{D}$. Since $\tilde{D}$ itself is closed, there is a $w \in \Sigma^*$ such that $\tilde{D} \cap w\Sigma^\omega \neq \varnothing$ and $\mathrm{cn}(D) \subseteq \tilde{D} - w\Sigma^\omega$.

Let $\beta \in \tilde{D}$ such that $w \prec \beta$. Then there is a $v$ satisfying $wv \sim_D e$. According to Fact 27 we have $wv \sim_{\mathrm{cn}(D)} e$. Hence, $\mathrm{cn}(D) \cap wv\Sigma^\omega = wv \cdot \mathrm{cn}(D)$, and $\mathrm{cn}(D) \cap w\Sigma^\omega = \emptyset$ implies $\mathrm{cn}(D) = \emptyset$, a contradiction. ❏

As a next result we need a topological property of strongly connected $\omega$-languages in $F_\sigma \cap G_\delta$ (cf. [St83, Lemma 20]).

**Lemma 29** *Let $D$ be a strongly connected $\omega$-language which is simultaneously in $F_\sigma$ and in $G_\delta$. Then $D$ is already closed.*

**Proof.** From [Ku66] it is known that for every nonempty $D \in F_\sigma \cap G_\delta$ there is a $w \in \Sigma^*$ such that $D \cap w\Sigma^\omega$ is nonempty and closed. Utilizing the strong connectivity of $D$ we obtain a $v \in \Sigma^*$ satisfying $D \cap wv\Sigma^\omega = wvD$. The left hand side of this identity equals $(D \cap w\Sigma^\omega) \cap wv\Sigma^\omega$, thus it is closed. Consequently, $wvD$ and also $D$ are closed. ❏

The assertion of Lemma 13 can be restated now as follows. Observe that $\widetilde{E/w} = \{\xi : \forall u(u \prec \xi \rightarrow \exists v(wuv \sim_E w))\}$ and $E(u) = u \cdot \widetilde{E/u}$.
*If $E \in F_\sigma \cap G_\delta$ and $E \cap w\Sigma^\omega \neq \emptyset$ then*

$$E \supseteq w \cdot \widetilde{E/w} \quad \text{or} \quad E \cap w \cdot \widetilde{E/w} = \emptyset\,.$$

**Proof.** Set $E/w := \{\beta : w\beta \in E\}$, that is $w \cdot E/w = E \cap w\Sigma^\omega$. Hence $E/w$ is also in $F_\sigma \cap G_\delta$. According to Lemma 28.1 and Lemma 29 the set $\mathrm{cn}(E/w)$ is closed.
Assume now $E \cap w \cdot \widetilde{E/w} \neq \emptyset$, that is, $E/w \cap \widetilde{E/w} \neq \emptyset$. Then $\mathrm{cn}(E/w) \neq \emptyset$ and following Lemma 28.2 we have $\mathrm{cn}(E/w) = \widetilde{E/w}$. Now the assertion follows from $\mathrm{cn}(E/w) \subseteq E/w$. ❏

# B Independence examples

Here we show that although the condition of Claim 17 fails to be a sufficient one, it is neither trivially satisfied nor does it necessarily imply one of the conditions '$\simeq_E = \cong_E$' or '$\mathcal{A}_E$ accepts $E$' even in case if $E$ is regular.
First we give an example of an $\omega$-language $E_6$ such that $\sim^*_{E_6} \subseteq \approx_{E_6}$ does not hold true.

**Example 6** *Let $\Sigma := \{a, b\}$ and $E_6 := \{a, b\}^* a^\omega$. Then $\simeq_{E_6}$ is trivial. Hence $\sim^*_{E_6}$ is also trivial, but $a \not\approx_{E_6} b$.*
*Consequently, neither $\simeq_{E_6}$ and $\cong_{E_6}$ coincide nor does $\mathcal{A}_{E_6}$ accept $E_6$.*

In the second example an $\omega$-language $E_7$ is given for which $\sim^*_{E_7} \subseteq \approx_{E_7}$ holds, but neither $\simeq_{E_7}$ and $\cong_{E_7}$ coincide nor does $\mathcal{A}_{E_7}$ accept $E_7$.

**Example 7** *Define $E_7 := (b^* a)^\omega \cup (a^2)^* c a^\omega$. The automaton $\mathcal{A}_{E_7}$ has five states $\langle e \rangle$, $\langle a \rangle$, $\langle b \rangle$, $\langle c \rangle$, and $\langle c \rangle c$, and is given by the following equations:*

$$\begin{aligned}
\langle a \rangle a &= \langle e \rangle \\
\langle a \rangle b &= \langle b \rangle = \langle b \rangle a = \langle b \rangle b \\
\langle a \rangle c &= \langle b \rangle c = \langle c \rangle b = \langle c \rangle c \\
\langle c \rangle a &= \langle c \rangle
\end{aligned}$$

*One can see that $\mathcal{A}_{E_7}$ does not accept $E_7$, and that $(a^2)^*$, $a(a^2)^*$, $a^* b \{a, b\}^*$, $(a^2)^* c a^*$, $a(a^2)^* c a^*$, and $\{a, b\}^* c \{a, b, c\}^* - a^* c a^*$ are the congruence classes of $\simeq_{E_7}$.*
*Now consider the empty word $e$. Since $c(ea)^\omega = ca^\omega \in E_7$ but $c(xa)^\omega \notin E_7$ unless $x \in a^*$, we have that $x \approx_{E_7} e$ implies $x \in a^*$. On the other hand we have $b^\omega \in E_7$ and $(ba^n)^\omega \notin E_7$ for $n > 0$. Hence $x \approx_{E_7} e$ implies that $x \notin a^* - \{e\}$. Thus $aa \simeq_{E_7} e$ but $aa \not\cong_{E_7} e$.*
*Utilizing similar arguments it is easy to verify that $\approx_{E_7}$ has the following congruence classes: $\{e\}$, $a^+$, $a^* b \{a, b\}^*$, and $\{a, b\}^* c \{a, b, c\}^*$.*
*Since $\sim^*_{E_7}$ refines $\simeq_{E_7}$, we obtain $\sim^*_{E_7} \subseteq \approx_{E_7}$ from the observation that $w \sim^*_{E_7} e$ implies $w = e$.*