Ehrenfeucht-Fraïssé Games on Omega-Terms

Martin Huschenbett[1] **and Manfred Kufleitner**[2]

1 **Institut für Theoretische Informatik**
Technische Universität Ilmenau, Germany
martin.huschenbett@tu-ilmenau.de

2 **Institut für Formale Methoden in der Informatik***
Universität Stuttgart, Germany
kufleitner@fmi.uni-stuttgart.de

Abstract

Fragments of first-order logic over words can often be characterized in terms of finite monoids or finite semigroups. Usually these algebraic descriptions yield decidability of the question whether a given regular language is definable in a particular fragment. An effective algebraic characterization can be obtained from identities of so-called omega-terms. In order to show that a given fragment satisfies some identity of omega-terms, one can use Ehrenfeucht-Fraïssé games on word instances of the omega-terms. The resulting proofs often require a significant amount of book-keeping with respect to the constants involved. In this paper we introduce Ehrenfeucht-Fraïssé games on omega-terms. To this end we assign a labeled linear order to every omega-term. Our main theorem shows that a given fragment satisfies some identity of omega-terms if and only if Duplicator has a winning strategy for the game on the resulting linear orders. This allows to avoid the book-keeping.

As an application of our main result, we show that one can decide in exponential time whether all aperiodic monoids satisfy some given identity of omega-terms, thereby improving a result of McCammond (Int. J. Algebra Comput., 2001).

1998 ACM Subject Classification F.4.1 Mathematical Logic, F.4.3 Formal Languages

Keywords and phrases regular language, first-order logic, finite monoid, Ehrenfeucht-Fraïssé games, pseudoidentity

Digital Object Identifier 10.4230/LIPIcs.STACS.2014.374

1 Introduction

By combining a result of McNaughton and Papert [12] with Schützenberger's characterization of star-free languages [16], a given language over finite words is definable in first-order logic if and only if its syntactic monoid is finite and aperiodic. The implication from left to right can be shown using Ehrenfeucht-Fraïssé games, see e.g. [17]. A similar result for two-variable first-order logic FO^2 was obtained by Thérien and Wilke [19]: A language is definable in FO^2 if and only if its syntactic monoid belongs to the variety **DA**. Both the variety **DA** and the class of finite aperiodic monoids can be defined using identities of omega-terms. Roughly speaking, omega-terms are words equipped with an additional operation, the ω-power. If M is a finite monoid, then there exists a positive integer ω_M such that $u^{\omega_M} = (u^{\omega_M})^2$ for all $u \in M$. We call u^{ω_M} the *idempotent* generated by u. Every mapping $h : \Lambda \to M$ uniquely extends to omega-terms over the alphabet Λ by setting $h(uv) = h(u)h(v)$ and $h(u^\omega) = h(u)^{\omega_M}$. Now,

* The second author was supported by the German Research Foundation (DFG) under grant DI 435/5-1 and by the Technische Universität München, Germany.

31st Symposium on Theoretical Aspects of Computer Science (STACS'14).
Editors: Ernst W. Mayr and Natacha Portier; pp. 374–385

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

the monoid M satisfies an identity $u = v$ of omega-terms u and v over Λ if for every mapping $h : \Lambda \to M$ we have $h(u) = h(v)$. A finite monoid is *aperiodic* if and only if it satisfies $a^\omega = a^\omega a$, and it is in **DA** if and only if it satisfies $(abc)^\omega b(abc)^\omega = (abc)^\omega$, see e.g. [15]. Showing that some first-order fragment $\mathcal{F}$ satisfies an identity $u = v$ of omega-terms u, v usually works as follows. Suppose $\mathcal{F}$ does not satisfy $u = v$. Then there exists a formula $\varphi \in \mathcal{F}$ such that the syntactic monoid of $L(\varphi)$ does not satisfy $u = v$. The depth n of the formula φ defines an n-round Ehrenfeucht-Fraïssé game on instances of u and v (i.e., on finite words which are obtained by replacing the ω-powers by fixed positive integers depending on n). Giving a winning strategy for Duplicator yields a contradiction, thereby showing that $\mathcal{F}$ satisfies $u = v$. Usually, playing the game on u and v involves some non-trivial book-keeping since one has to formalize intuitive notions such as positions being near to one another or being close to some border. For first-order logic and for FO^2 the book-keeping is still feasible [17, 5] whereas for other fragments such as the quantifier alternation inside FO^2 this task becomes much more involved (and therefore other techniques are applied [10, 18]).

Instead of defining new instances of a given omega-term depending on the fragment and the number of rounds in the Ehrenfeucht-Fraïssé game, we give a single instance which works for all fragments of first-order logic and any number of rounds. In addition, we allow an infinite number of rounds. The fragments we consider in this paper rely on an abstract notion of logical fragments as introduced in [9]. We show that a fragment $\mathcal{F}$ satisfies an identity of omega-terms if and only if Duplicator has a winning strategy for the Ehrenfeucht-Fraïssé game for $\mathcal{F}$ on the instances of the omega-terms. These instances are labeled linear orders which, in general, are not finite words.

An obvious application of our main result is the simplification of proofs showing that some fragment $\mathcal{F}$ satisfies a given identity of omega-terms. The main reason is that with this new approach one can avoid the book-keeping. It is slightly less straightforward that one can use this approach for solving word problems for omega-terms over varieties of finite monoids. Let $\mathbf{V}$ be a variety of finite monoids. Then the word problem for omega-terms over $\mathbf{V}$ is the following: Given two omega-terms u and v, does every monoid in $\mathbf{V}$ satisfy the identity $u = v$? This problem was solved for various varieties, see e.g. [2, 11, 13]. Using our main result, one approach to solving such word problems is as follows. First, find a logical fragment for $\mathbf{V}$. Second, find a winning strategy for Duplicator on omega-terms satisfied by this fragment. Third, use this winning strategy for finding the desired decision algorithm. In the case of aperiodic monoids, we use this scheme for improving the decidability result of McCammond [11] by showing that the word problem for omega-terms over aperiodic monoids is solvable in exponential time.

Historically, the greek letter ω is used for two different things which are frequently used thoughout this paper: First, the idempotent power of an element and second, the smallest infinite ordinal. In order to avoid confusion in our presentation, we chose to follow the approach of Perrin and Pin [14] by using π instead of ω to denote idempotent powers. In particular, we will use the exponent π in omega-terms which is why we will call them π-terms in the remainder of this paper.

2 Preliminaries

As mentioned above, one of the central notions in this paper are so called π-terms. In order to make their interpretation by several semantics possible in a uniform way, we follow an algebraic approach. A *π-algebra* is a structure $(U, \cdot, {}^\pi)$ comprised of an associative binary operation $\cdot$ and a unary operation ${}^\pi$ on a carrier set U. The application of $\cdot$ is usually written

as juxtaposition, i.e., $uv = u \cdot v$, and the application of $^\pi$ as u^π. A *π-term* is an arbitrary element of the free π-algebra T_Λ generated by Λ, where Λ *is a countably infinite set which is fixed for the rest of this paper*. We also use this set as a universe for letters (of alphabets).

Monoids as π-Algebras. Let M be a monoid. For any $k \geq 1$ we extend M to a π-algebra, called *k-power algebra on M*, by defining $u^\pi = u^k$ for $u \in M$. Suppose that M is finite. An element $u \in M$ is *idempotent* if $u^2 = u$. We extend M to another π-algebra, called *idempotency algebra on M*, by defining u^π for $u \in M$ to be the unique idempotent element in the set $\{\, u^k \mid k \geq 1 \,\}$. In fact, there are infinitely many $k \geq 1$, called *idempotency exponents of M*, such that for each $u \in M$ the element u^k is idempotent, i.e., the k-power algebra and the idempotency algebra on M coincide. An identity $s = t$ of π-terms $s, t \in T_\Lambda$ *holds in M* if every π-algebra morphism h from T_Λ into the idempotency algebra on M satisfies $h(s) = h(t)$.

The set of all *finite words* over an alphabet $A \subseteq \Lambda$ is A^*. Let $L \subseteq A^*$ be a language over a finite alphabet $A \subseteq \Lambda$. The *syntactic congruence* of L is the equivalence relation $\equiv_L$ on A^* defined by $u \equiv_L v$ if $xuy \in L$ is equivalent to $xvy \in L$ for all $x, y \in A^*$. In fact, $\equiv_L$ is a monoid congruence on A^*. The quotient monoid $M_L = A^*/\equiv_L$ is called *syntactic monoid* of L. It is finite precisely if L is regular. Suppose that L is regular and let $k \geq 1$ be an idempotency exponent of M_L. Then the map sending each $w \in A^*$ to its $\equiv_L$-class is a π-algebra morphism from the k-power algebra on A^* onto the idempotency algebra on M_L. Thus, any identity $s = t$ of π-terms $s, t \in T_\Lambda$ holds in M_L if and only if every π-algebra morphism h from T_Λ into the k-power algebra on A^* satisfies $h(s) \equiv_L h(t)$.

Generalized Words. The third semantic domain we consider is the class of generalized words. A *generalized word* (over Λ) is a triple $u = (P_u, \leq_u, \ell_u)$ comprised of a (possibly empty) linear ordering $(P_u, \leq_u)$ being labeled by a map $\ell_u \colon P_u \to \Lambda$. The set $\mathrm{dom}(u) = P_u$ is the *domain* of u, its elements are called *positions* of u. We write $u(p)$ instead of $\ell_u(p)$ for $p \in P_u$. The *order type* of u is the isomorphism type of $(P_u, \leq_u)$. We regard any finite word $w = a_1 \dots a_n \in \Lambda^*$ as a generalized word by defining $\mathrm{dom}(w) = [1, n]$, $\leq_w$ as the natural order on $[1, n]$ and $w(k) = a_k$ for $k \in [1, n]$. On that view, generalized words indeed generalize finite words. As of now, we mean "generalized word" when writing just "word". Two words u and v are *isomorphic* if there exists an isomorphism f of linear orderings from $(\mathrm{dom}(u), \leq_u)$ to $(\mathrm{dom}(v), \leq_v)$ such that $u(p) = v(f(p))$ for all $p \in \mathrm{dom}(u)$. *We identify isomorphic words.* We denote the set of all (isomorphism classes of) *countable* words by Λ^{LO}. The exponent LO is for *linear order*. We regard Λ^* as a subset of Λ^{LO}.

Let $u, v \in \Lambda^{\mathsf{LO}}$ be two words. Their *concatenation* is the word $uv \in \Lambda^{\mathsf{LO}}$ defined by $\mathrm{dom}(uv) = \mathrm{dom}(u) \uplus \mathrm{dom}(v)$, $\leq_{uv}$ makes all positions of u smaller than those of v and retains the respective orders inside u and inside v, and $(uv)(p)$ is $u(p)$ if $p \in \mathrm{dom}(u)$ and $v(p)$ if $p \in \mathrm{dom}(v)$. The set Λ^{LO} with concatenation forms a monoid. On finite words this concatenation coincides with the usual definition and hence Λ^* is a submonoid of Λ^{LO}.

It is customary to regard $n \in \mathbb{N}$ also as the order type of the natural linear ordering on $[1, n]$. We extend the notion of the n-power algebra on Λ^{LO} to arbitrary countable order types τ as follows. Let $(T, \leq_T)$ be a linear ordering of isomorphism type τ. The *τ-power* of any word $u \in \Lambda^{\mathsf{LO}}$ is the word $u^\tau \in \Lambda^{\mathsf{LO}}$ defined by $\mathrm{dom}(u^\tau) = \mathrm{dom}(u) \times T$, $(p, t) \leq_{u^\tau} (p', t')$ if $t <_T t'$ or if $t = t'$ and $p \leq_u p'$, and $(u^\tau)(p, t) = u(p)$. We extend the monoid Λ^{LO} to a π-algebra, called *τ-power algebra on Λ^{LO}*, by defining $u^\pi = u^\tau$ for $u \in \Lambda^{\mathsf{LO}}$. We denote by $[\![\cdot]\!]_\tau$ the unique π-algebra morphism from T_Λ into this π-algebra mapping each $a \in \Lambda$ to the word consisting of a single position which is labeled by a. Finally, notice that there are two definitions of the n-power algebra on Λ^{LO} around, but actually they coincide.

Logic over Words. *For the rest of this paper, we fix a countably infinite set $\mathbb{V}$ of (first-order) variables $x, y, z, \ldots$.* The syntax of first-order logic over words is given by

$$\varphi ::= \top \mid \bot \mid \mathsf{empty} \mid x = y \mid \lambda(x) = a \mid x < y \mid x \leq y \mid \mathsf{suc}(x, y) \mid \mathsf{min}(x) \mid \mathsf{max}(x) \mid \neg\varphi \mid (\varphi \vee \varphi) \mid (\varphi \wedge \varphi) \mid \exists x\, \varphi \mid \forall x\, \varphi\,,$$

where $x, y \in \mathbb{V}$ and $a \in \Lambda$. The set of all formulae is denoted by FO. Brackets can be omitted when originating no ambiguity. The *free variables* $\mathrm{FV}(\varphi)$ of a formula $\varphi \in \mathsf{FO}$ are defined as usual. A *sentence* is a formula φ with $\mathrm{FV}(\varphi) = \emptyset$.

We only give a brief sketch of the semantics of formulae. Let $\mathbb{X} \subseteq \mathbb{V}$ be a *finite* set of variables. An $\mathbb{X}$-*valuation on* u is a pair $\langle u, \alpha \rangle$ consisting of a word $u \in \Lambda^{\mathsf{LO}}$ and a map $\alpha\colon \mathbb{X} \to \mathrm{dom}(u)$. It is a model of a formula $\varphi \in \mathsf{FO}$ with $\mathrm{FV}(\varphi) \subseteq \mathbb{X}$, in symbols $\langle u, \alpha \rangle \models \varphi$, if u satisfies the formula φ under the following assumptions:

- variables range over positions of u and free variables are interpreted according to α,
- $\top$ is always satisfied, $\bot$ never, and empty only in case $\mathrm{dom}(u) = \emptyset$,
- the function symbol λ is interpreted by the labeling map $\ell_u\colon \mathrm{dom}(u) \to \Lambda$ and
- the predicates $<$, $\leq$, suc, min and max are evaluated in the linear ordering $(\mathrm{dom}(u), \leq_u)$, where $\mathsf{suc}(x, y)$ means that y is the immediate successor of x.

We identify any word $u \in \Lambda^{\mathsf{LO}}$ with the only $\emptyset$-valuation on u, namely $\langle u, \emptyset \rangle$ with $\emptyset$ also denoting the empty map. Thus, for sentences φ the meaning of $u \models \varphi$ is well-defined. Let $A \subseteq \Lambda$ be a finite alphabet and $\varphi \in \mathsf{FO}$ a sentence. Due to the result of Büchi, Elgot, and Trakhtenbrot [4, 7, 20], the *language over A defined by* φ, namely $L_A(\varphi) = \{\, w \in A^* \mid w \models \varphi \,\}$, is regular. A language $L \subseteq A^*$ is *definable in* a class $\mathcal{F} \subseteq \mathsf{FO}$ of formulae if there exists a sentence $\varphi \in \mathcal{F}$ such that $L = L_A(\varphi)$.

Fragments. We reintroduce (a slight variation of) the notion of a fragment as a class of formulae obeying natural syntactic closure properties [9]. A *context* is a formula μ with a unique occurrence of an additional constant predicate $\circ$ which is intended to be a placeholder for another formula $\varphi \in \mathsf{FO}$. The result of replacing $\circ$ in μ by φ is denoted by $\mu(\varphi)$. Unfortunately, the notion of a fragment as defined in [9, Definition 1] is slightly too weak for our purposes. We require one more *natural* syntactic closure property, namely condition **4.** in Definition 2.1 below. Condition **6.** is missing in the exposition in [9]. Nevertheless, since we only add requirements, every fragment in our sense is still a fragment in the sense of [9].

▶ **Definition 2.1.** A *fragment* is a non-empty set of formulae $\mathcal{F} \subseteq \mathsf{FO}$ such that for all contexts μ, formulae $\varphi, \psi \in \mathsf{FO}$, $a \in \Lambda$ and $x, y \in \mathbb{V}$ the following conditions are satisfied:

1. If $\mu(\varphi) \in \mathcal{F}$, then $\mu(\top) \in \mathcal{F}$, $\mu(\bot) \in \mathcal{F}$, and $\mu(\lambda(x) = a) \in \mathcal{F}$.
2. $\mu(\varphi \vee \psi) \in \mathcal{F}$ if and only if $\mu(\varphi) \in \mathcal{F}$ and $\mu(\psi) \in \mathcal{F}$.
3. $\mu(\varphi \wedge \psi) \in \mathcal{F}$ if and only if $\mu(\varphi) \in \mathcal{F}$ and $\mu(\psi) \in \mathcal{F}$.
4. If $\mu(\neg\neg\varphi) \in \mathcal{F}$, then $\mu(\varphi) \in \mathcal{F}$.
5. If $\mu(\exists x\, \varphi) \in \mathcal{F}$ and $x \notin \mathrm{FV}(\varphi)$, then $\mu(\varphi) \in \mathcal{F}$.
6. If $\mu(\forall x\, \varphi) \in \mathcal{F}$ and $x \notin \mathrm{FV}(\varphi)$, then $\mu(\varphi) \in \mathcal{F}$.

It is *closed under negation* if the following condition is satisfied:

7. If $\varphi \in \mathcal{F}$, then $\neg\varphi \in \mathcal{F}$.

It is *order-stable* if the following condition is satisfied:

8. $\mu(x < y) \in \mathcal{F}$ if and only if $\mu(x \leq y) \in \mathcal{F}$.

It is suc-*stable* if the following two conditions are satisfied:

9. If $\mu(\mathsf{suc}(x, y)) \in \mathcal{F}$, then $\mu(x = y) \in \mathcal{F}$, $\mu(\mathsf{max}(x)) \in \mathcal{F}$ and $\mu(\mathsf{min}(y)) \in \mathcal{F}$.
10. If $\mu(\mathsf{min}(x)) \in \mathcal{F}$ or $\mu(\mathsf{max}(x)) \in \mathcal{F}$, then $\mu(\mathsf{empty}) \in \mathcal{F}$.

Table 1 A single round of the $\mathcal{F}$-game in configuration $S = (\mathcal{G}, \langle u, \alpha\rangle, \langle v, \beta\rangle)$.

1. Spoiler chooses $\mathsf{Q}x$	**2.** Spoiler chooses q in	**3.** Duplicator chooses r in	**4.** resulting configuration $S[\mathsf{Q}x, q, r]$
$\mathsf{Q}x = \exists x$	$\mathrm{dom}(u)$	$\mathrm{dom}(v)$	$(\mathcal{G}/\exists x, \langle u, \alpha[x/q]\rangle, \langle v, \beta[x/r]\rangle)$
$\mathsf{Q}x = \forall x$	$\mathrm{dom}(v)$	$\mathrm{dom}(u)$	$(\mathcal{G}/\forall x, \langle u, \alpha[x/r]\rangle, \langle v, \beta[x/q]\rangle)$
$\mathsf{Q}x = \neg\exists x$	$\mathrm{dom}(v)$	$\mathrm{dom}(u)$	$(\mathcal{G}/\neg\exists x, \langle v, \beta[x/q]\rangle, \langle u, \alpha[x/r]\rangle)$
$\mathsf{Q}x = \neg\forall x$	$\mathrm{dom}(u)$	$\mathrm{dom}(v)$	$(\mathcal{G}/\neg\forall x, \langle v, \beta[x/r]\rangle, \langle u, \alpha[x/q]\rangle)$

Examples for fragments in this sense include all classes of formulae which are obtained from full first-order logic FO by limiting the quantifier depth (e.g., FO_n), the number of quantifier alternations (e.g., Σ_n and Π_n), the number of quantified variables (e.g., FO^m), the available predicates (e.g., first-order logic $\mathsf{FO}[<]$ without min, max, suc) or combinations of those.

The *quantifier depth* $\mathrm{qd}(\varphi)$ of a formula $\varphi \in \mathsf{FO}$ is defined as usual. A fragment $\mathcal{F}$ has *bounded quantifier depth* if there is an $n \in \mathbb{N}$ such that $\mathrm{qd}(\varphi) \leq n$ for all $\varphi \in \mathcal{F}$. For any $n \in \mathbb{N}$ and every fragment $\mathcal{F}$ the set $\mathcal{F}_n = \{\, \varphi \in \mathcal{F} \mid \mathrm{qd}(\varphi) \leq n \,\}$ is a fragment of bounded quantifier depth. Moreover, the fragment $\mathcal{F}_n$ is order-stable (respectively suc-stable) in case $\mathcal{F}$ has the according property.

3 Ehrenfeucht-Fraïssé Games for Arbitrary Fragments

In this section, we introduce an Ehrenfeucht-Fraïssé game for arbitrary fragments of first-order logic on generalized words and develop its basic theory. Before we can describe this game, we need to define some notation. In the following, we call the "negated quantifiers" $\neg\exists$ and $\neg\forall$ also *quantifiers*. The set of all quantifiers (in this sense) is $\mathcal{Q} = \{\exists, \forall, \neg\exists, \neg\forall\}$. For a quantifier $\mathsf{Q} \in \mathcal{Q}$ and a variable $x \in \mathbb{V}$, the *reduct* of $\mathcal{F}$ by $\mathsf{Q}x$ is the set

$$\mathcal{F}/\mathsf{Q}x = \{\, \varphi \in \mathsf{FO} \mid \mathsf{Q}x\,\varphi \in \mathcal{F} \,\} \ .$$

Whenever this set is not empty, it is a fragment as well.

Now, let $\mathcal{F}$ be a fragment and u, v two words over Λ. We are about to describe the *$\mathcal{F}$-game on (u, v)*. A *configuration* of this game is a triple $S = (\mathcal{G}, \langle u, \alpha\rangle, \langle v, \beta\rangle)$ comprised of a non-empty, iterated reduct $\mathcal{G}$ of $\mathcal{F}$ and $\mathbb{X}$-valuations $\langle u, \alpha\rangle$ and $\langle v, \beta\rangle$ on u and v for the same arbitrary *finite* subset $\mathbb{X} \subseteq \mathbb{V}$. To emphasize the set $\mathbb{X}$, we also speak of an *$\mathbb{X}$-configuration*. The game starts in the $\emptyset$-configuration $(\mathcal{F}, u, v)$ and goes on for an arbitrary—possibly infinite—number of rounds. Assuming that the game is currently in configuration $S = (\mathcal{G}, \langle u, \alpha\rangle, \langle v, \beta\rangle)$, a single round proceeds as follows (see Table 1 for a summary of this procedure):

1. Spoiler chooses a quantifier $\mathsf{Q} \in \mathcal{Q}$ and a variable $x \in \mathbb{V}$ such that $\mathcal{G}/\mathsf{Q}x \neq \emptyset$.
2. Spoiler chooses a position q (like "quest") in the domain of u if $\mathsf{Q} \in \{\exists, \neg\forall\}$ or in the domain of v if $\mathsf{Q} \in \{\forall, \neg\exists\}$.
3. Duplicator chooses a position r (like "reply") in the domain of the other word.
4. The resulting configuration $S[\mathsf{Q}x, q, r]$ consists of the reduct $\mathcal{G}/\mathsf{Q}x$ and the extension of the valuations $\langle u, \alpha\rangle$ and $\langle v, \beta\rangle$ by variable x at positions q and r, accordingly. Whenever Q is a negated quantifier, the role of the two extended valuations is additionally interchanged (see the last column of Table 1 for a formal definition of $S[\mathsf{Q}x, q, r]$).

Whenever a player cannot perform a choice because $\mathcal{G}$ contains no more quantified formulae or the domain of the according word is empty, the game immediately stops and

the other player wins. Besides the inability of Duplicator to move, the winning condition for Spoiler is to reach an $\mathbb{X}$-configuration $(\mathcal{G}, \langle u, \alpha\rangle, \langle v, \beta\rangle)$ such that there exists a literal $\varphi \in \mathcal{G}$ with $\mathrm{FV}(\varphi) \subseteq \mathbb{X}$ and $\langle u, \alpha\rangle \models \varphi$ but $\langle v, \beta\rangle \not\models \varphi$; in this case the game immediately stops. Duplicator's goal is simply to prevent Spoiler from winning. In particular, Duplicator wins all games that go on forever. Due to this circumstance, the $\mathcal{F}$-game is determined, i.e., either Spoiler or Duplicator has a winning strategy on (u, v).

▶ Remark. The $\mathcal{F}$-game is quite asymmetric since Spoiler is not allowed to choose before his first move whether he wants to play on (u, v) or on (v, u). This may lead to the situation that he has a winning strategy on (u, v) but not on (v, u) or vice versa. This asymmetry is owed to the circumstance that $\mathcal{F}$ might not be closed under negation. As soon as $\mathcal{F}$ is assumed to be closed under negation this phenomenon disappears and Spoiler has a winning strategy on (u, v) if and only if he has a winning strategy on (v, u). We also note that, in general, the winning condition for Spoiler can be asymmetric since it does not rely on any notion of isomorphism. ◀

If the quantifier depth of a fragment $\mathcal{F}$ is bounded by $n \in \mathbb{N}$, the $\mathcal{F}$-game lasts at most n rounds. In particular, for any fragment $\mathcal{F}$ the $\mathcal{F}_n$-game can be regarded as an n-round version of the $\mathcal{F}$-game. For instance, the FO_n-game resembles the classical n-round Ehrenfeucht-Fraïssé game. The following result is an adaption of the Ehrenfeucht-Fraïssé Theorem to the $\mathcal{F}$-game for fragments of bounded quantifier depth.

▶ **Theorem 3.1.** *Let $\mathcal{F}$ be a fragment of bounded quantifier depth. For all words $u, v \in \Lambda^{\mathsf{LO}}$ the following are equivalent:*

1. *$u \models \varphi$ implies $v \models \varphi$ for all sentences $\varphi \in \mathcal{F}$ and*
2. *Duplicator has a winning strategy in the $\mathcal{F}$-game on (u, v).*

A proof of this theorem can easily be achieved along the lines of a proof of the classical version, cf. [8]. In fact, such a proof reveals that the implication "**2.** $\Rightarrow$ **1.**" even holds if the quantifier depth of $\mathcal{F}$ is not bounded. In contrast, the implication "**1.** $\Rightarrow$ **2.**" substantially relies the boundedness of the quantifier depth of $\mathcal{F}$. If ζ denotes the order type of the integers $\mathbb{Z}$, then Duplicator has a winning strategy in the FO_n-game on $(a^\zeta, a^{\zeta+\zeta})$ for each $n \in \mathbb{N}$ and hence $a^\zeta \models \varphi$ implies $a^{\zeta+\zeta} \models \varphi$ for all sentences $\varphi \in \mathsf{FO}$, but Spoiler has a winning strategy in the infinite FO-game on $(a^\zeta, a^{\zeta+\zeta})$.

The objective of the remainder of this section is to identify additional requirements on $\mathcal{F}$ and/or u, v such that the boundedness of the quantifier depth can be omitted. It turns out that the property introduced in Definition 3.2 below in combination with suc-stability of the fragment is sufficient for this purpose and still allows for the applications in Section 4. The order types of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{Z}_{<0}$ ordered naturally are denoted by ω, ζ, η and ω^*, respectively. Then $\omega + \zeta \cdot \eta + \omega^*$ is the order type of the word $a^\omega\big(a^\zeta\big)^\eta a^{\omega^*}$, where $a \in \Lambda$.

▶ **Definition 3.2.** Let $\varrho = \omega + \zeta \cdot \eta + \omega^*$. A word $u \in \Lambda^{\mathsf{LO}}$ is *ϱ-rational* if it can be constructed from the finite words in Λ^{LO} using the operations of concatenation and ϱ-power only or, equivalently, if $u = [\![t]\!]_\varrho$ for some π-term $t \in T_\Lambda$.

▶ **Theorem 3.3.** *Let $\mathcal{F}$ be a suc-stable fragment. For all ϱ-rational words $u, v \in \Lambda^{\mathsf{LO}}$ the following are equivalent:*

1. *$u \models \varphi$ implies $v \models \varphi$ for all sentences $\varphi \in \mathcal{F}$ and*
2. *Duplicator has a winning strategy in the $\mathcal{F}$-game on (u, v).*

As already mentioned, the implication "**2.** $\Rightarrow$ **1.**" can be shown using the very same proof as for the according implication of Theorem 3.1. The key idea behind proving the implication

"**1. ⇒ 2.**" is as follows: Theorem 3.1 provides us for each $n \in \mathbb{N}$ with a winning strategy for Duplicator in the $\mathcal{F}_n$-game on (u, v). A winning strategy in the $\mathcal{F}$-game is obtained by defining a limit of all those strategies. This limit process relies on the ϱ-rationality of the underlying words and is formalized by Lemma 3.6 below. A major ingredient of its proof is Proposition 3.4.

In order to keep notation concise, we abbreviate the circumstance that Duplicator has a winning strategy in a configuration $S = (\mathcal{F}, \langle u, \alpha\rangle, \langle v, \beta\rangle)$ by $\langle u, \alpha\rangle \lesssim_{\mathcal{F}} \langle v, \beta\rangle$. Since the $\mathcal{F}$-game is determined, $\langle u, \alpha\rangle \not\lesssim_{\mathcal{F}} \langle v, \beta\rangle$ hence means that Spoiler has a winning strategy in S. The relation $\lesssim_{\mathcal{F}}$ is reflexive and transitive, i.e., a preorder on the set of all configurations. It induces an equivalence $\approx_{\mathcal{F}}$ defined by $\langle u, \alpha\rangle \approx_{\mathcal{F}} \langle v, \beta\rangle$ if $\langle u, \alpha\rangle \lesssim_{\mathcal{F}} \langle v, \beta\rangle$ and $\langle v, \beta\rangle \lesssim_{\mathcal{F}} \langle u, \alpha\rangle$.

▶ **Proposition 3.4.** *Let $\mathcal{F}$ be a* suc*-stable fragment, $k \in \mathbb{N}$ and $\langle u_i, \alpha_i\rangle, \langle v_i, \beta_i\rangle$ $\mathbb{X}_i$-valuations with mutually disjoint $\mathbb{X}_i$ for $i \in [1, k]$. If $\langle u_i, \alpha_i\rangle \lesssim_{\mathcal{F}} \langle v_i, \beta_i\rangle$ for each $i \in [1, k]$, then $\langle u_1 \cdots u_k, \alpha_1 \cup \cdots \cup \alpha_k\rangle \lesssim_{\mathcal{F}} \langle v_1 \cdots v_k, \beta_1 \cup \cdots \cup \beta_k\rangle$.* ◀

▶ **Lemma 3.5.** *Let $\mathcal{F}$ be a* suc*-stable fragment with quantifier depth bounded by $n \in \mathbb{N}$ and $u, v \in \Lambda^{\mathsf{LO}}$. If $u \lesssim_{\mathcal{F}} v$, then $u^m \lesssim_{\mathcal{F}} v^\varrho$ and $u^\varrho \lesssim_{\mathcal{F}} v^m$ for all $m \geq 2^{n+1} - 1$.* ◀

The following lemma formalizes the limit process mentioned above.

▶ **Lemma 3.6.** *Let $\mathcal{F}$ be a* suc*-stable fragment, $x \in \mathbb{V}$ and $\langle u, \alpha\rangle$ an $\mathbb{X}$-valuation on a ϱ-rational word $u \in \Lambda^{\mathsf{LO}}$. For every infinite sequence $(q_i)_{i\in\mathbb{N}} \in \mathrm{dom}(u)^{\mathbb{N}}$ there exists a position $q \in \mathrm{dom}(u)$ such that for all $n \in \mathbb{N}$ there are arbitrarily large $i \in \mathbb{N}$ with $\langle u, \alpha[x/q_i]\rangle \lesssim_{\mathcal{F}_n} \langle u, \alpha[x/q]\rangle$.*

Proof. To simplify notation, we call a position q with the property above a $\langle u, \alpha\rangle$*-limit point* of the sequence $(q_i)_{i\in\mathbb{N}}$ (w.r.t. to $\mathcal{F}$ and x). Using this terminology, we have to show that every sequence $(q_i)_{i\in\mathbb{N}} \in \mathrm{dom}(u)^{\mathbb{N}}$ possesses a $\langle u, \alpha\rangle$-limit point. Since neither $\alpha[x/q_i]$ nor $\alpha[x/q]$ would depend on $\alpha(x)$, we may simply assume that $x \notin \mathbb{X}$. We proceed by induction on the ϱ-rational construction of u.

Base case: u is finite. Since $\mathrm{dom}(u)$ is finite, there exists a $q \in \mathrm{dom}(u)$ such that $q = q_i$ for infinitely many $i \in \mathbb{N}$. Thus, q is a $\langle u, \alpha\rangle$-limit point of $(q_i)_{i\in\mathbb{N}}$.

Inductive step 1: $u = v_1 v_2$ with ϱ-rational words v_1, v_2. The valuation $\langle u, \alpha\rangle$ splits into valuations $\langle v_1, \beta_1\rangle$ and $\langle v_2, \beta_2\rangle$ such that $\alpha = \beta_1 \cup \beta_2$. For either $\ell = 1$ or $\ell = 2$ we have $q_i \in \mathrm{dom}(v_\ell)$ for infinitely many $i \in \mathbb{N}$. Let I be the set of these i. By the induction hypothesis, there is a $\langle v_\ell, \beta_\ell\rangle$-limit point $q \in \mathrm{dom}(v_\ell)$ of the subsequence $(q_i)_{i\in I}$. Proposition 3.4 implies that q is also a $\langle u, \alpha\rangle$-limit point of $(q_i)_{i\in\mathbb{N}}$.

We split the inductive step for ϱ-powers in two parts, one for $\mathbb{X} = \emptyset$ and another for $\mathbb{X} \neq \emptyset$.

Inductive step 2: $u = v^\varrho$ with a ϱ-rational v and $\mathbb{X} = \emptyset$. Let $(P, \leq_P)$ be a linear ordering of isomorphism type ϱ such that $\mathrm{dom}(u) = \mathrm{dom}(v) \times P$. For each $i \in \mathbb{N}$ we write $q_i = (s_i, p_i)$. For every $p \in P$ let $\overleftarrow{\tau}_p$ and $\vec{\tau}_p$ be the order types of the suborders of $(P, \leq_P)$ induced by the open intervals $(-\infty, p)$ and $(p, +\infty)$, respectively. Then $\varrho = \overleftarrow{\tau}_p + 1 + \vec{\tau}_p$. Due to the nature of ϱ, each of $\overleftarrow{\tau}_p$ and $\vec{\tau}_p$ is either finite or equals ϱ. However, the case that $\overleftarrow{\tau}_p$ and $\vec{\tau}_p$ both are finite at the same time cannot occur. Accordingly, we distinguish three cases:

Case 1: $\overleftarrow{\tau}_{p_i} = \vec{\tau}_{p_i} = \varrho$ for infinitely many $i \in \mathbb{N}$. Let I be the set of these i. By the induction hypothesis, there exists a $\langle v, \emptyset\rangle$-limit point $s \in \mathrm{dom}(v)$ of the subsequence $(s_i)_{i\in I}$. We pick some $j \in I$. Proposition 3.4 reveals that $q = (s, p_j)$ is a $\langle u, \alpha\rangle$-limit point of $(q_i)_{i\in\mathbb{N}}$.

Case 2: $\overleftarrow{\tau}_{p_i}$ is finite and $\vec{\tau}_{p_i} = \varrho$ for infinitely many $i \in \mathbb{N}$. Let I be the set of these i. If there is an order type which occurs infinitely often among the $\overleftarrow{\tau}_{p_i}$ with $i \in I$, the same

argumentation as in Case 1 applies. Henceforth, we assume that such an order type does not exist. By the induction hypothesis, the subsequence $(s_i)_{i\in I}$ possesses a $\langle v, \emptyset\rangle$-limit point $s \in \mathrm{dom}(v)$. Let $p \in P$ be arbitrary with $\overleftarrow{\tau}_p = \vec{\tau}_p = \varrho$. We show that $q = (s, p)$ is a $\langle u, \alpha\rangle$-limit point of $(q_i)_{i\in\mathbb{N}}$.

Let $n \in \mathbb{N}$. Due to the choice of I and s, there are arbitrarily large $i \in I$ such that $\overleftarrow{\tau}_{p_i}$ is of size at least $2^{n+1} - 1$ and $\langle v, \emptyset[x/s_i]\rangle \lesssim_{\mathcal{F}_n} \langle v, \emptyset[x/s]\rangle$. Lemma 3.5 then implies $v^{\overleftarrow{\tau}_{p_i}} \lesssim_{\mathcal{F}_n} v^\varrho$. Since also $v^{\vec{\tau}_{p_i}} \lesssim_{\mathcal{F}_n} v^\varrho$, Proposition 3.4 yields $\langle u, \emptyset[x/q_i]\rangle \lesssim_{\mathcal{F}_n} \langle u, \emptyset[x/q]\rangle$.

Case 3: $\overleftarrow{\tau}_{p_i} = \varrho$ and $\vec{\tau}_{p_i}$ is finite for infinitely many $i \in \mathbb{N}$. Symmetric to Case 2.

Inductive step 3: $u = v^\varrho$ with a ϱ-rational v and $\mathbb{X} \neq \emptyset$. Let $(P, \leq_P)$ be as above. Recall that $\mathbb{X}$ is supposed to be finite. Let $\tilde{p}_1 <_P \cdots <_P \tilde{p}_k$ be an enumeration of all positions $p \in P$ for which there exists a variable $y \in \mathbb{X}$ with $\alpha(y) \in \mathrm{dom}(v) \times \{p\}$. We consider the open intervals $P_0 = (-\infty, \tilde{p}_1)$, $P_\ell = (\tilde{p}_\ell, \tilde{p}_{\ell+1})$ for $\ell \in [1, k-1]$, and $P_k = (\tilde{p}_k, +\infty)$ in $(P, \leq_P)$. For $\ell \in [0, k]$ we let τ_ℓ be the order type of the suborder induced by P_ℓ. Then $\varrho = \tau_0 + 1 + \tau_1 + 1 + \cdots + 1 + \tau_k$ and hence $u = v^{\tau_0} v v^{\tau_1} v \cdots v v^{\tau_k}$. Due to the nature of ϱ, each τ_ℓ is either finite or equals ϱ. Since for every finite τ_ℓ the word v^{τ_ℓ} is the concatenation of τ_ℓ copies of v, the factorization of u above is an alternative ϱ-rational construction of u. This construction has the additional property that α does not map into the ϱ-powers v^ϱ but only in the individual intermediate copies of v. Thus, the induction hypothesis and the inductive steps 1 and 2 above yield the claim. ◀

Now, we are prepared to prove the remaining implication of Theorem 3.3.

Proof of Theorem 3.3, "1. ⇒ 2.". We show that Duplicator can maintain the invariant of staying in configurations which are *good* for her. A configuration $(\mathcal{G}, \langle u, \alpha\rangle, \langle v, \beta\rangle)$ of the $\mathcal{F}$-game on (u, v) is considered to be *good* for Duplicator if $\langle u, \alpha\rangle \lesssim_{\mathcal{G}_n} \langle v, \beta\rangle$ for every $n \in \mathbb{N}$. Statement **1.** and Theorem 3.1 imply that the initial configuration $(\mathcal{F}, u, v)$ is good. Moreover, good configurations do not meet Spoiler's winning condition as they particularly satisfy $\langle u, \alpha\rangle \lesssim_{\mathcal{G}_0} \langle v, \beta\rangle$. Consequently, it suffices to provide a strategy for Duplicator which never leaves the set of good configurations since such a strategy is a winning strategy.

Suppose Spoiler chooses the quantifier $\mathsf{Q}x$ and the quest q in a good configuration $(\mathcal{G}, \langle u, \alpha\rangle, \langle v, \beta\rangle)$. We only demonstrate the case $\mathsf{Q} = \exists$, where $q \in \mathrm{dom}(u)$. For every $i \in \mathbb{N}$ we have $\langle u, \alpha\rangle \lesssim_{\mathcal{G}_{i+1}} \langle v, \beta\rangle$ and hence there exists $r_i \in \mathrm{dom}(v)$ such that $\langle u, \alpha[x/q]\rangle \lesssim_{\mathcal{G}_{i+1}/\exists x} \langle v, \beta[x/r_i]\rangle$. Since $\mathcal{G}_{i+1}/\exists x = (\mathcal{G}/\exists x)_i$, this is the same as $\langle u, \alpha[x/q]\rangle \lesssim_{(\mathcal{G}/\exists x)_i} \langle v, \beta[x/r_i]\rangle$. Due to Lemma 3.6 applied to the sequence $(r_i)_{i\in\mathbb{N}}$, there exists $r \in \mathrm{dom}(v)$ such that, for every $n \in \mathbb{N}$, there are arbitrarily large $i \in \mathbb{N}$ with $\langle v, \beta[x/r_i]\rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r]\rangle$. We show that the configuration $S[\exists x, q, r] = (\mathcal{G}/\exists x, \langle u, \alpha[x/q]\rangle, \langle v, \beta[x/r]\rangle)$ is good again.

Let $n \in \mathbb{N}$. Due to the choice of r, there is an $i \geq n$ with $\langle v, \beta[x/r_i]\rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r]\rangle$. Above, the position r_i was chosen such that $\langle u, \alpha[x/q]\rangle \lesssim_{(\mathcal{G}/\exists x)_i} \langle v, \beta[x/r_i]\rangle$. Since $n \leq i$, this implies $\langle u, \alpha[x/q]\rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r_i]\rangle$ and in turn $\langle u, \alpha[x/q]\rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r]\rangle$. ◀

4 Ehrenfeucht-Fraïssé Games on Identities

Identities play an important role in the study of the expressive power of first-order fragments. A recurring problem is to show that a certain identity of π-terms holds in the syntactic monoid/semigroup of every language definable in the fragment under consideration. Theorems 4.1 and 4.2 below can remarkably simplify this task, as demonstrated at the end of this section. In fact, the two theorems are just slight variations of one another and the sole

reason for having two theorems is that the suc-predicate does not play well with syntactic monoids but only with syntactic semigroups.

▶ **Theorem 4.1.** *Let $\mathcal{F}$ be an order-stable fragment not containing the predicates* suc, min, max *and* empty. *For all π-terms $s, t \in T_\Lambda$ the following are equivalent:*
1. *The identity $s = t$ holds in the syntactic monoid of every language definable in $\mathcal{F}$.*
2. *Duplicator has winning strategies in the $\mathcal{F}$-games on $([\![s]\!]_\varrho, [\![t]\!]_\varrho)$ and $([\![t]\!]_\varrho, [\![s]\!]_\varrho)$.*

▶ **Theorem 4.2.** *Let $\mathcal{F}$ be a* suc*-stable and order-stable fragment. For all π-terms $s, t \in T_\Lambda$ the following are equivalent:*
1. *The identity $s = t$ holds in the syntactic semigroup of every language definable in $\mathcal{F}$ over non-empty words.*
2. *Duplicator has winning strategies in the $\mathcal{F}$-games on $([\![s]\!]_\varrho, [\![t]\!]_\varrho)$ and $([\![t]\!]_\varrho, [\![s]\!]_\varrho)$.* ◀

The main ingredients of the proofs of both theorems are Theorem 3.3 and [9, Proposition 2] which is restated as Proposition 4.3 below.

▶ **Proposition 4.3.** *Let $\mathcal{F}$ be a fragment, $A, B \subseteq \Lambda$ finite alphabets and h a monoid morphism from A^* into B^*. Suppose the following:*
1. *If $\mathcal{F}$ contains the predicate $\leq$ or $<$, then $\mathcal{F}$ is order-stable or $h(A) \subseteq B \cup \{\varepsilon\}$.*
2. *If $\mathcal{F}$ contains the predicate* suc, min, max *or* empty, *then $\varepsilon \notin h(A)$.*

Then $h^{-1}(L)$ is $\mathcal{F}$-definable whenever $L \subseteq B^*$ is $\mathcal{F}$-definable.

Applying this proposition to $\mathcal{F}$-games yields that monoid morphisms satisfying the two conditions above preserve the existence of winning strategies for Duplicator.

▶ **Corollary 4.4.** *Let $\mathcal{F}$, A, B and h be as in Proposition 4.3 satisfying conditions **1.** and **2.**. Moreover, let $\mathcal{F}$ be a* suc*-stable. Then $u \lesssim_\mathcal{F} v$ implies $h(u) \lesssim_\mathcal{F} h(v)$ for all $u, v \in A^*$.*

Proof. Let $u, v \in A^*$ with $u \lesssim_\mathcal{F} v$. Since finite words are ϱ-rational and due to Theorem 3.3, it suffices to show that $h(u) \models \varphi$ implies $h(v) \models \varphi$ for all sentences $\varphi \in \mathcal{F}$. Consider a sentence $\varphi \in \mathcal{F}$. By Proposition 4.3, there is a sentence $\psi \in \mathcal{F}$ such that $L_A(\psi) = h^{-1}\big(L_B(\varphi)\big)$. Altogether, $h(u) \models \varphi$ implies $u \models \psi$ and since $u \lesssim_\mathcal{F} v$ this implies $v \models \psi$ which in turn implies $h(v) \models \varphi$. ◀

The following corollary is an immediate consequence of Proposition 3.4 and Lemma 3.5.

▶ **Corollary 4.5.** *Let $\mathcal{F}$ be a* suc*-stable fragment whose quantifier depth is bounded by $n \in \mathbb{N}$ and let $t \in T_\Lambda$ be a π-term. Then $[\![t]\!]_\varrho \approx_\mathcal{F} [\![t]\!]_m$ for all $m \geq 2^{n+1} - 1$.* ◀

The previous results allow us to show Theorems 4.1 and 4.2. However, since their proofs are as similar as their statements, we only demonstrate the first one.

Proof of Theorem 4.1. Let $A \subseteq \Lambda$ be the finite set containing all $a \in \Lambda$ appearing in s or t. We show both implications separately.

"1. $\Rightarrow$ 2.". By Theorem 3.3, it suffices to show for every sentence $\varphi \in \mathcal{F}$ that $[\![s]\!]_\varrho \models \varphi$ if and only if $[\![t]\!]_\varrho \models \varphi$. Consider a sentence $\varphi \in \mathcal{F}$ and put $n = \mathrm{qd}(\varphi)$. We put $L = L_A(\varphi)$ and let $k \geq 2^{n+1} - 1$ be an idempotency exponent of M_L. We consider an arbitrary π-algebra morphism h from T_A into the k-power algebra on A^* with $h(a) = a$ for each $a \in A$. Because $s = t$ holds in M_L, we have $h(s) \equiv_L h(t)$. Since $h(s) = [\![s]\!]_k$ as well as $h(t) = [\![t]\!]_k$ and by Corollary 4.5, we obtain $h(s) \approx_{\mathcal{F}_n} [\![s]\!]_\varrho$ and $h(t) \approx_{\mathcal{F}_n} [\![t]\!]_\varrho$. Altogether, we conclude that $[\![s]\!]_\varrho \models \varphi$ if and only if $h(s) \models \varphi$ if and only if $h(t) \models \varphi$ if and only if $[\![t]\!]_\varrho \models \varphi$.

"2. $\Rightarrow$ 1.". Let $B \subseteq \Lambda$ be a finite alphabet and $L \subseteq B^*$ a language defined by a sentence $\varphi \in \mathcal{F}$. Let $n = \mathrm{qd}(\varphi)$ and $k \geq 2^{n+1} - 1$ be an idempotency exponent of M_L. We have to show that every π-algebra morphism g from T_A into the k-power algebra on B^* satisfies $g(s) \equiv_L g(t)$. Consider such a morphism g and let h be the unique monoid morphism from A^* into B^* defined by $h(a) = g(a)$ for each $a \in A$. Then $g(s) = h([\![s]\!]_k)$ and $g(t) = h([\![t]\!]_k)$. Corollary 4.5 and the assumption $[\![s]\!]_\varrho \approx_{\mathcal{F}} [\![t]\!]_\varrho$ yield $[\![s]\!]_k \approx_{\mathcal{F}_n} [\![s]\!]_\varrho \approx_{\mathcal{F}_n} [\![t]\!]_\varrho \approx_{\mathcal{F}_n} [\![t]\!]_k$. We conclude $g(s) \approx_{\mathcal{F}_n} g(t)$ by Corollary 4.4. By Proposition 3.4, we obtain $ug(s)v \approx_{\mathcal{F}_n} ug(t)v$ for all $u, v \in B^*$. Since $\varphi \in \mathcal{F}_n$, this finally implies $g(s) \equiv_L g(t)$. ◀

In the remainder of this section, we demonstrate two applications of Theorem 4.1 by providing quite short proofs of two well-known results. The following corollary can be obtained by combining a result of McNaughton and Papert [12] with Schützenberger's characterization of star-free languages [16]. A more direct proof can, for instance, be found in [17]. A finite monoid M is called *aperiodic* if the identity $a^\pi a = a^\pi$ holds in M.

▶ **Corollary 4.6.** *The syntactic monoid of every first-order definable language is aperiodic.*

Proof. The predicates suc, min, max and empty can be expressed in $\mathsf{FO}[<]$. By Theorem 4.1, it suffices to show $[\![a^\pi a]\!]_\varrho \approx_{\mathsf{FO}[<]} [\![a^\pi]\!]_\varrho$. The property $\varrho + 1 = \varrho$ of the order type ϱ implies $[\![a^\pi a]\!]_\varrho = [\![a^\pi]\!]_\varrho$ and the claim follows. ◀

The second application relates definability in $\mathsf{FO}^2[<]$ to the class $\mathbf{DA}$. The fragment $\mathsf{FO}^2[<]$ consists of all formulae not containing the predicates suc, min, max and empty which quantify over two fixed variables $x_1, x_2 \in \mathbb{V}$ only. The class $\mathbf{DA}$ consists of all finite monoids in which the identity $(abc)^\pi b (abc)^\pi = (abc)^\pi$ holds. A significant amount of book-keeping is involved when showing that the syntactic monoid of every $\mathsf{FO}^2[<]$-definable language is in $\mathbf{DA}$ by applying the classical Ehrenfeucht-Fraïssé game approach, see e.g. [5][1]. On the other hand, the abstract idea of this proof is very simple: Duplicator copies every move near the left and near the right border, and he does not need to care in the center. We now show that this idea can easily be formalized when using Theorem 4.1.

▶ **Corollary 4.7.** *The syntactic monoid of any language definable in* $\mathsf{FO}^2[<]$ *is in* $\mathbf{DA}$.

Proof. Let $s = (abc)^\pi b (abc)^\pi$ and $t = (abc)^\pi$. Again by Theorem 4.1, it suffices to show $[\![s]\!]_\varrho \approx_{\mathsf{FO}^2[<]} [\![t]\!]_\varrho$. With $u = (abc)^\omega (abc)^{\zeta \cdot \eta}$ and $v = (abc)^{\zeta \cdot \eta} (abc)^{\omega^*}$ we obtain

$$[\![s]\!]_\varrho = u\,(abc)^{\omega^*} b (abc)^\omega\, v \qquad \text{and} \qquad [\![t]\!]_\varrho = u\,(abc)^{\omega^*} (abc)^\omega\, v\,.$$

Since $\mathsf{FO}^2[<]$ is closed under negation and due to Proposition 3.4, it further suffices to show that Duplicator has a winning strategy in the $\mathsf{FO}^2[<]$-game on

$$\left((abc)^{\omega^*} b (abc)^\omega,\ (abc)^{\omega^*} (abc)^\omega\right).$$

The strategy is to choose a reply that is labeled by the same letter as the request and such that the positions corresponding to x_1 and x_2 are in the same order in both words. This is always possible, since in both words there are always infinitely many positions to the left (respectively to the right) of any position which are labeled by a given letter from a, b, c. ◀

[1] Actually, the proof given in [5] does not use the language of Ehrenfeucht-Fraïssé games, but it can easily be restated this way.

5 The Word Problem for π-Terms over Aperiodic Monoids

The word problem for π-terms over aperiodic monoids was solved by McCammond [11] by computing normal forms. In the process of computing these normal forms the intermediate terms can grow and, to the best of our knowledge, neither the worst-case running time nor the maximal size of the intermediate terms has been estimated (and it seems to be difficult to obtain such results). In this section we give an exponential algorithm for solving the word problem for π-terms over aperiodic monoids. Our algorithm does not compute normal forms as π-terms; instead we show that the evaluation under $[\![\cdot]\!]_\varrho$ can be used as a normal form for π-terms.

▶ **Theorem 5.1.** *Given two π-terms $s, t \in T_\Lambda$, one can decide whether the identity $s = t$ holds in every aperiodic monoid in time exponential in the size of s and t.*

The proof is a reduction to the isomorphism problem for regular words, cf. [3]. These generalized words particularly include all ϱ-rational words and can be described by expressions similar to π-terms but using ω-power, ω^*-power and dense shuffle instead of the π-power. Due to [3, Theorem 79], one can decide in polynomial time whether two such expressions describe isomorphic words. The characterization underlying the reduction is as follows:

▶ **Proposition 5.2.** *For all π-terms $s, t \in T_\Lambda$ the following conditions are equivalent:*
1. *The identity $s = t$ holds in every aperiodic finite monoid.*
2. $[\![s]\!]_\varrho = [\![t]\!]_\varrho$.

Proof. "1. $\Rightarrow$ 2.". The results in [11] imply that the identity $s = t$ can be deduced from the following list of axioms, where $n \geq 1$:

$$(uv)w = u(vw) \qquad (u^\pi)^\pi = u^\pi \qquad (u^n)^\pi = u^\pi$$
$$u^\pi u^\pi = u^\pi \qquad u^\pi u = uu^\pi = u^\pi \qquad (uv)^\pi u = u(vu)^\pi .$$

As a matter of fact, the ϱ-power algebra on Λ^{LO} satisfies these axioms as well. Consequently, $[\![s]\!]_\varrho = [\![t]\!]_\varrho$ can be proved along a deduction of the identity $s = t$ from the axioms.

"2. $\Rightarrow$ 1.". Due to Eilenberg's Variety Theorem [6], the pseudovariety of aperiodic monoids is generated by the class of syntactic aperiodic monoids. The latter are precisely the syntactic monoids of first-order definable languages [12, 16]. By Theorem 4.1 the identity $s = t$ holds in the syntactic monoid of every such language. ◀

Proof of Theorem 5.1. In order to apply the decision procedure from [3, Theorem 79], we have to translate s and t into expressions generating the same words and which do not use ϱ-power but ω-power, ω^*-power and dense shuffle instead. Such a translation can be based on the identity $u^\varrho = u^\omega \left(u^{\omega^*} u^\omega\right)^\eta u^{\omega^*}$ which holds for all words $u \in \Lambda^{\mathsf{LO}}$. Therein, the η-power is a special case of the dense shuffle. Since this translation leads to a blow-up which is exponential in the number of nested applications of π-powers within s and t, we can decide $[\![s]\!]_\varrho = [\![t]\!]_\varrho$ in time at most exponential in the size of s and t. ◀

6 Summary

For every π-term t we define a labeled linear order $[\![t]\!]_\varrho$, and every first-order fragment $\mathcal{F}$ over finite words naturally yields a (possibly infinite) Ehrenfeucht-Fraïssé game on labeled linear orders. The important property of these constructions is that $\mathcal{F}$ satisfies an identity

$s = t$ of π-terms s and t if and only if Duplicator has a winning strategy in the $\mathcal{F}$-game on $[\![s]\!]_\varrho$ and $[\![t]\!]_\varrho$. We note that $[\![t]\!]_\varrho$ does not depend on $\mathcal{F}$. Usually showing that a fragment $\mathcal{F}$ satisfies an identity $s = t$ requires a significant amount of book-keeping which in most cases is not part of the actual proof idea. Our main results Theorem 4.1 and Theorem 4.2 allow to formalize such proof ideas without further book-keeping, see e.g. Corollary 4.7. A probably less obvious application of our main result are word problems for π-terms over varieties of finite monoids. We show that the word problem for π-terms over aperiodic finite monoids is solvable in exponential time (Theorem 5.1), thereby improving a result of McCammond [11].

Several possible extensions of our result come to mind: Other implicit operations (see [1] for further details on implicit operations), logical fragments beyond classical first-order logic, and other structures such as infinite words, trees or Mazurkiewicz traces.

References

1 J. Almeida. *Finite Semigroups and Universal Algebra*. World Scientific, 1994.
2 J. Almeida and M. Zeitoun. An automata-theoretic approach to the word problem for ω-terms over $\mathbf{R}$. *Theoret. Comput. Sci.*, 370(1–3):131–169, 2007.
3 S. L. Bloom and Z. Ésik. The equational theory of regular words. *Information and Computation*, 197(1–2):55–89, 2005.
4 J. R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.
5 V. Diekert, P. Gastin, and M. Kufleitner. A survey on small fragments of first-order logic over finite words. *Int. J. Found. Comput. Sci.*, 19(3):513–548, 2008.
6 S. Eilenberg. *Automata, Languages, and Machines*, vol. B. Academic Press, 1976.
7 C. C. Elgot. Decision problems of finite automata design and related arithmetics. *Trans. Amer. Math. Soc.*, 98:21–51, 1961.
8 N. Immerman and D. Kozen. Definability with bounded number of bound variables. *Information and Computation*, 83(2):121–139, Nov. 1989.
9 M. Kufleitner and A. Lauser. Lattices of logical fragments over words. In *ICALP 2012, Proceedings Part II*, volume 7392 of *LNCS*, pp. 275–286. Springer, 2012.
10 M. Kufleitner and A. Lauser. Quantifier alternation in two-variable first-order logic with successor is decidable. In *STACS 2013, Proceedings*, vol. 20 of *LIPIcs*, pages 305–316. Dagstuhl Publishing, 2013.
11 J. P. McCammond. Normal forms for free aperiodic semigroups. *Int. J. Algebra Comput.*, 11(5):581–625, 2001.
12 R. McNaughton and S. Papert. *Counter-Free Automata*. The MIT Press, 1971.
13 A. Moura. The word problem for ω-terms over $\mathbf{DA}$. *Theoret. Comput. Sci.*, 412(46):6556–6569, 2011.
14 D. Perrin and J.-É. Pin. *Infinite words*, volume 141 of *Pure and Applied Mathematics*. Elsevier, 2004.
15 J.-É. Pin. *Varieties of Formal Languages*. North Oxford Academic, 1986.
16 M. P. Schützenberger. On finite monoids having only trivial subgroups. *Inf. Control*, 8:190–194, 1965.
17 H. Straubing. *Finite Automata, Formal Logic, and Circuit Complexity*. Birkhäuser, 1994.
18 H. Straubing. Algebraic characterization of the alternation hierarchy in $\mathrm{FO}^2[<]$ on finite words. In *CSL 2011, Proc.*, vol. 12 of *LIPIcs*, pp. 525–537. Dagstuhl Publishing, 2011.
19 D. Thérien and Th. Wilke. Over words, two variables are as powerful as one quantifier alternation. In *STOC 1998, Proceedings*, pp. 234–240. ACM Press, 1998.
20 B. A. Trakhtenbrot. Finite automata and logic of monadic predicates (in Russian). *Dokl. Akad. Nauk SSSR*, 140:326–329, 1961.