First-order separation over countable ordinals*

Thomas Colcombet[1], Sam van Gool[1], and Rémi Morvan[2](✉)

[1] IRIF, Université de Paris & CNRS, Paris, France
{thomas.colcombet,vangool}@irif.fr
[2] École normale supérieure Paris-Saclay, Gif-sur-Yvette, France
fistname.lastname@ens-paris-saclay.fr

Abstract. We show that the existence of a first-order formula separating two monadic second order formulas over countable ordinal words is decidable. This extends the work of Henckell and Almeida on finite words, and of Place and Zeitoun on ω-words. For this, we develop the algebraic concept of monoid (resp. ω-semigroup, resp. ordinal monoid) with aperiodic merge, an extension of monoids (resp. ω-semigroup, resp. ordinal monoid) that explicitly includes a new operation capturing the loss of precision induced by first-order indistinguishability. We also show the computability of FO-pointlike sets, and the decidability of the covering problem for first-order logic on countable ordinal words.

Keywords: Regular languages · Separation, Pointlike sets · Countable Ordinals · First-order logic · Monadic second-order logic

A full version of this paper can be found on arXiv. This document contains internal hyperlinks, and is best read on an electronic device.

1 Introduction

In this paper, we establish the decidability of FO-separability over countable ordinal words:

Theorem 1. *There is an algorithm which, given two regular languages of countable ordinal words K, L, either:*

- *answers 'yes', and outputs an FO-*separator *which is an FO-formula φ which* separates *K from L, i.e. such that $u \models \varphi$ for all $u \in K$, and $v \models \neg\varphi$ for all $v \in L$, or*
- *answers 'no', and outputs a* witness function, *i.e., a computable function taking as input an FO-sentence φ and returning a pair of words $(u, v) \in K \times L$ such that $u \models \varphi$ if and only if $v \models \varphi$.*

* This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC DuaLL, grant agreement No. 670624), and by the DeLTA ANR project (ANR-16-CE40-0007).

P. Bouyer and L. Schröder (Eds.): FoSSaCS 2022, LNCS 13242, pp. 264–284, 2022.
https://doi.org/10.1007/978-3-030-99253-8_14

The decidability of FO-separability was previously only known for finite words [19,2,25,17] and for words of length ω [25]. Countable ordinal words are sequences of letters that are indexed by a countable total well-ordering, *i.e.*, up to isomorphism, by a countable ordinal. There is a natural notion of regular languages over these objects which can be equivalently described in terms of logic (either monadic second-order logic or weak monadic second-order logic), automata (Büchi introduced a notion of automata for countable ordinal words [13], which was studied in more detail by Wojciechowski [39] and which generalises Choueka's automata [15] for words of length at most ω^n—the fact that Choueka's automata can be seen as a restriction of Büchi's automata for countable ordinals was proven by Bedon [5]), rational expressions (introduced by Wojciechowski [40]), or algebra (recognisable by finite ordinal monoids—introduced by Bedon and Carton [8]). A detailed survey of the equivalence between all these notions can be found in Bedon's thesis [6].

Our algorithm follows the approach initiated by Henckell, and constructs the FO-pointlike sets in an ordinal monoid that recognises the two input languages simultaneously. FO-pointlike sets are subsets of a monoid whose elements are inherently indistinguishable by first-order logic. Our completeness proof for the algorithm follows a scheme similar to the one followed by Place and Zeitoun in the context of finite and ω-words [25], which was inspired by Wilke's characterisation of FO-definable languages [38]. We had to make several substantial changes to this approach for the proofs to generalize from finite and ω-words to the setting of countable ordinal words. A seemingly slight modification of the notion of saturation (Definition 8) allows for a careful redesign of several of the core lemmas in the proof of completeness, and in particular the construction of an FO-approximant in Section 5 below.

Related work This work lies in a line of research that aims to obtain a decidable understanding of the expressive power of subclasses of the class of regular languages. The seminal work in this area is the Schützenberger-McNaughton-Papert theorem [34,22] which effectively characterizes the languages of finite words definable in first-order logic as the ones which have an aperiodic syntactic monoid. This theorem was at the origin of a large body of work that studies classes of languages through the corresponding classes of monoids, including for instance Simon's result characterising piecewise-testable languages via $\mathcal{J}$-trivial monoids [36]. FO-pointlike sets are also known in the literature as aperiodic pointlike sets, and were first studied and shown to be computable by Henckell [19], in the context of the Krohn-Rhodes semigroup complexity problem. The computability of pointlike sets was shown to be equivalent to the decidability of the covering problem by Almeida [2]. Alternative proofs of separation and covering problems for FO were given recently in [25,17], and, ever since Henckell's work, the computability of FO-pointlike sets was also extended to pointlike sets for other varieties—for example [4] for the variety of finite groups, [3] for the variety of $\mathcal{J}$-trivial finite semigroups and [18] for varieties of finite semigroups determined by a variety of finite groups; also see [18] for further references. Place and Zeitoun recently used pointlike sets, in the form of covering problems [27], to resolve long-standing open membership problems for the lower levels of the dot-depth and of the Straubing-Thérien hierarchies [26,28,29].

Another, orthogonal, line of research consists in the extension of the notions of regularity (logic/automata/rational expressions/algebra) to models beyond finite words. This is the case for finite or infinite trees [30]. In this paper, we are concerned with words that go beyond finite, such as words of length ω [12,37,24], of countable ordinal length [6,5], of countable scattered[3] length [31,32], or of general countable length [30,35,14].

These two branches have also been studied jointly, and first-order logic was characterised on words of length ω [23], of countable ordinal length [7], of countable scattered length [10] (and in [9] for first-order augmented with quantifiers over Dedekind cuts), and for words of countable length [16] (as well as other logics [16,21,1]). Prior to the current work, the questions of computing the FO-pointlike sets and deciding FO-separation for languages of infinite words had only been investigated for words of length ω [25].

Structure of the document In Section 2, we introduce important definitions for manipulating infinite words in algebraic terms (ordinal monoids and their powerset), and in logical terms (first-order logic and first-order definable maps). In Section 3, we describe the algorithm, and in particular its core, a saturation construction. The correctness of the algorithm is then proved in Section 4, and the completeness in Section 5. In Section 6, we show two stronger results that arise from the same technique: the decidability of the covering problem and the computability of pointlikes. Section 7 concludes.

2 Preliminaries

2.1 Ordinals

A *linear ordering* is a set equipped with a total order. It is *countable* (resp. *finite*) if the underlying set is countable (resp. finite). Let α and β be two linear orderings. A *morphism* from α to β is a monotonic function, and an *isomorphism* between α and β is a bijective morphism. The (ordered) *sum* of two linear orders α and β is denoted by $\alpha + \beta$ and is defined, as usual, on the disjoint union of the linear orders α and β, by further postulating that every element of α is below every element of β. The *product* of two linear orders is denoted by $\alpha \cdot \beta$ and is defined to be the right-to-left lexicographic ordering on the Cartesian product of the two orders, *i.e.*, $(x, y) \leqslant (x', y')$ iff $y < y'$ or $y = y'$ and $x \leqslant x'$. The n-fold product of α with itself is denoted by α^n. A linear ordering is *well-founded* when it does not contain an infinite strictly decreasing sequence. An *ordinal* is a well-founded linear ordering, considered only up to isomorphism of linear orderings. The empty linear ordering, the linear ordering with a single element and the linear ordering of natural numbers are all ordinals, and are denoted 0, 1 and ω, respectively. The class of all ordinals is itself totally ordered by the *embedding*

[3] A linear ordering is *scattered* if it does not contain a dense subordering.

relation: $\alpha \preccurlyeq \beta$ means that there exists an injective monotonic function from α to β. The relation $\prec$ denotes the strict ordering associated with $\preccurlyeq$. An ordinal is a *successor ordinal* if it has a maximum, and a *limit ordinal* otherwise.

2.2 Ordinal words

Given a set X, a *word* w over X is a map from some linear ordering to X. The linear ordering is called the *domain* of w, and denoted $\mathrm{dom}(w)$. A word is *countable* (resp. *finite*, resp. *scattered*, resp. *ω-word*), if its domain is countable (resp. finite, resp. scattered, resp. ω). In this paper, a *countable ordinal word* is a word that has a countable and ordinal domain (hence, the countability assumption in silently assumed throughout the paper). The set of all finite words over X is denoted by X^*, and the collection of all countable ordinal words over X is denoted by X^{ord}. Similarly, the set of finite non-empty words is denoted by X^+ and the collection of non-empty countable ordinal words is denoted by $X^{\mathrm{ord}+}$. The concatenation of two countable ordinal words u and v over X is the word $u \cdot v : \mathrm{dom}(u) + \mathrm{dom}(v) \to X$ over X defined by $(u \cdot v)_\iota := u_\iota$ if $\iota \in \mathrm{dom}(u)$ and $(u \cdot v)_\iota := v_\iota$ if $\iota \in \mathrm{dom}(v)$. If w is a countable ordinal word, we define its *omega iteration*, denoted by w^ω, as the word with domain $\mathrm{dom}(w) \cdot \omega$ defined by $(w^\omega)_{(\iota,n)} := w_\iota$ for every $\iota \in \mathrm{dom}(w)$ and $n \in \omega$. For example, if $a, b \in X$, then the omega iteration $(ab)^\omega$ of the two-letter word ab is the word $ababab\cdots$ with domain $2 \cdot \omega = \omega$.

2.3 Ordinal monoids

A *semigroup* is a set S equipped with an associative binary product, denoted by $\cdot$. A *monoid* is a semigroup with a distinguished neutral element for the product, denoted as 1. An element $x \in S$ is called *idempotent* if $x^2 = x$. In a finite finite semigroup S, every element $x \in S$ has a unique *idempotent power*, denoted by[4] x^{idem}, which we recall is the limit of the ultimately constant series $n \mapsto x^{n!}$. We also denote $x^{\mathrm{idem}+k}$, for k integer, the limit of the ultimately constant series $n \mapsto x^{n!+k}$. Note that x^{idem} is the identity element of the unique maximal group inside the subsemigroup generated by x. A finite semigroup is *aperiodic* (we equivalently write *group-trivial*) if $a^{\mathrm{idem}} = a^{\mathrm{idem}+1}$ for all of its elements a.

We now extend the notion of monoid to obtain an algebraic structure in which one can evaluate a product indexed by any countable ordinal. Let Σ be any set, and α a countable ordinal. For any word $(w_\iota)_{\iota<\alpha}$ over the set Σ^{ord} of countable ordinal words—i.e. $(w_\iota)_{\iota<\alpha}$ is a word whose letters are words over Σ— we define $\mathrm{flat}(w_\iota \mid \iota < \alpha)$ to be the word over Σ with domain $\sum_{\iota<\alpha} \mathrm{dom}(w_\iota)$, which has the letter $(w_\iota)_\kappa \in \Sigma$ at position (ι, κ), for every $\iota \in \alpha$ and $\kappa \in \mathrm{dom}(w_\iota)$.

[4] The standard notation is x^ω, but this notation conflicts with the linear ordering ω. It is sometimes denoted x^π or $x^!$ when in the context of infinite words. We find the notation x^{idem} more self-explanatory.

Definition 2. *An* ordinal monoid[5] *is a pair* $\mathcal{M} = (M, \pi)$ *where* M *is a set and* $\pi : M^{\mathrm{ord}} \to M$ *is a function, called* generalised product*, such that:*

- $\pi(x) = x$ *for every* $x \in M$*, and*
- $\pi((\pi(u_\iota))_{\iota<\alpha}) = \pi(\mathrm{flat}((u_\iota)_{\iota<\alpha}))$ *for every word* $(u_\iota)_{\iota<\alpha} \in (M^{\mathrm{ord}})^{\mathrm{ord}}$.

The second axiom is called *generalised associativity*. An *ordinal monoid morphism* is a map between ordinal monoids preserving the generalised product. An ordinal monoid is *ordered* if it is equipped with an order $\leqslant$ that makes π monotonic, *i.e.* such that $u \leqslant v$ implies $\pi(u) \leqslant \pi(v)$, in which $\leqslant$ is extended letter-by-letter to words in M^{ord}.

Given a set Σ (the *alphabet*), an ordinal monoid $\mathcal{M} = (M, \pi)$, a letter-to-letter map $\sigma\colon \Sigma \to M$ extended to $\sigma^{\mathrm{ord}}\colon \Sigma^{\mathrm{ord}} \to M^{\mathrm{ord}}$, and $F \subseteq M$, the language $L \subseteq \Sigma^{\mathrm{ord}}$ *recognised* by $(\mathcal{M}, \sigma, F)$ is

$$L = \{u \in \Sigma^{\mathrm{ord}} \ : \ \pi(\sigma^{\mathrm{ord}}(u)) \in F\},$$

and a language $L \subseteq \Sigma^{\mathrm{ord}}$ is called *recognisable* if it is recognised by some such tuple $(\mathcal{M}, \sigma, F)$. We recall that recognisable languages of ordinal words coincide with the ones definable in monadic second-order logic, or definable by suitable automata. These languages are called *regular*. Example 9 below will illustrate this concept.

We now recall a finite presentation of finite ordinal monoids (originally for ordinal semigroups), first given by Bedon [6] by extending a similar result established by Perrin and Pin [24, prop II.5.2] for ω-semigroups[6]. Let (S, π) be an ordinal monoid. We define the constant $\underline{1}$ and two functions $\underline{\cdot} : S \times S \to S$ and $-^{\underline{\omega}} : S \to S$ by

$$\underline{1} := \pi(\varepsilon) \qquad x \,\underline{\cdot}\, y := \pi(xy) \qquad \text{and} \qquad x^{\underline{\omega}} := \pi(x^{\omega}) = \pi(\overbrace{xxx\cdots}^{\omega \text{ times}}) \ .$$

The following proposition lets us interchangeably regard an ordinal monoid $\mathcal{M}$ as either a pair (M, π) or as a quadruple $(M, \underline{1}, \underline{\cdot}, -^{\underline{\omega}})$, that we refer to as its *presentation*.

Proposition 3 ([6, Thm. 3.5.6], originally for ordinal semigroups). *In a finite ordinal monoid the generalised product is uniquely determined by the operations* $\underline{1}, \underline{\cdot}$ *and* $-^{\underline{\omega}}$.

An important construction on which our proof relies is the *power ordinal monoid*: given an ordinal monoid (M, π), we equip the powerset $\mathcal{P}(M)$ of M with a generalised product $\pi : \mathcal{P}(M)^{\mathrm{ord}} \to \mathcal{P}(M)$ defined by

$$\pi((X_\iota)_{\iota<\kappa}) := \{\pi((x_\iota)_{\iota<\kappa}) \mid x_\iota \in X_\iota \text{ for all } \iota < \kappa\}$$
$$\text{for all words } (X_\iota)_{\iota<\kappa} \in (\mathcal{P}(M))^{\mathrm{ord}}.$$

[5] The object should probably be called a ‘countable ordinal monoid’ since its intent is to model countable ordinal words. However the naming becomes clumsy for ‘finite countable ordinal monoids’...

[6] The finitary reprensation of ω-semigroups is usually called a Wilke algebra, which is the algebraic structure introduced by Wilke in [37] to recognise regular ω-languages.

Observe that if M is a finite ordinal monoid, then so is $\mathcal{P}(M)$. We can compute a finite representation of the power ordinal monoid $\mathcal{P}(M)$ of M from a finite representation of M. Indeed,

$$\underline{1} = \{\underline{1}\}, \quad X \mathbin{\underline{\cdot}} Y = \{x \cdot y \mid x \in X,\, y \in Y\}, \quad \text{and} \qquad X^{\underline{\omega}} = \{u \cdot v^{\underline{\omega}} \mid u, v \in X^{+}\}$$

for all $X, Y \in \mathcal{P}(M)$. The two first properties are trivial while the third one can be proven using the infinite Ramsey's theorem—this is a classical argument used to give finite representation of infinite structures, see e.g. [24, Theorem II.2.1]. Note that this power ordinal monoid is indeed an ordinal monoid. It is even an ordered ordinal monoid when equipped with the inclusion ordering.

2.4 First-order logic

Over a fixed (finite) alphabet Σ, we define the set of *first-order logic formulæ* or FO-*formulæ* for short, by the grammar:

$$\varphi ::= \exists x.\,\varphi \quad|\quad \forall x.\,\varphi \quad|\quad \varphi \wedge \varphi \quad|\quad \varphi \vee \varphi \quad|\quad \neg\varphi \quad|\quad x < y \quad|\quad a(x)$$

where x, y range over some fixed infinite set of variables, and a over Σ. *Free variables* are defined as usual, and an FO-*sentence* is a formula with no free variables. In our setting, a model is a countable ordinal word, and a *valuation* over this model is a total map from variables to the domain of the word. We define, for any word w and any valuation ν, the semantic relation $w, \nu \models \varphi$ of first-order logic on countable ordinal words by structural induction on the FO-formula φ, by interpreting variables as positions in the word and propositions of the form $a(x)$ as "the letter at position x is an a". If φ is an FO-sentence, then the semantics of φ over a word w does not depend on the valuation, and thus we write $w \models \varphi$ or $w \not\models \varphi$. When $w \models \varphi$ we say that w *satisfies* φ, or also that φ *accepts* w.

A language $L \subseteq \Sigma^{\text{ord}}$ is said to be FO-*definable* if $L = \{w \in \Sigma^{\text{ord}} \mid w \models \varphi\}$ for some FO-sentence φ. For example, the language of words over the alphabet $\{a, b, c\}$ such that every 'a' is at a finite distance from a 'b' is defined by the FO-sentence $\forall x.a(x) \rightarrow \exists y.b(y) \wedge \text{finite}(x, y)$, where:

$$\text{isSuccessor}(z) ::= \exists y.y < z \wedge (\forall x.\, x < z \rightarrow x \leqslant y)$$
$$\text{finite}(x, y) ::= \forall z.(x < z \leqslant y \vee y < z \leqslant x) \rightarrow \text{isSuccessor}(z)\ .$$

Bedon [7] extended the Schützenberger-McNaughton-Papert theorem [34,22] to countable ordinal words.

Proposition 4 (*Bedon's theorem* [7, Theorem 3.4]). *A language of countable ordinal words is* FO*-definable if and only if it is recognised by a finite aperiodic ordinal monoid.*

Let $L \subseteq \Sigma^{\text{ord}}$. A function $f : L \rightarrow X$ whose codomain X is a finite set is said to be FO-*definable* when every preimage $f^{-1}[x]$, with $x \in X$, is an FO-definable language. Note that if f is FO-definable, then its domain L is necessarily an FO-definable language.

For example, the function $\Sigma^* \to \mathbb{Z}/2\mathbb{Z}$, sending a word $w \in \Sigma^*$ to its length modulo 2, is not FO-definable. On the other hand, for a fixed letter $a \in \Sigma$, the total function sending a word $w \in \Sigma^{\text{ord}+}$ to $\top$ if w contains the letter 'a' and to $\bot$ otherwise is FO-definable.

A useful tool to manipulate words is the notion of condensation — see, e.g., [33, §4] for an introduction to the subject. A *condensation* of a countable ordinal α is an equivalence relation $\sim$ over α whose equivalence classes are convex. Note that the quotient of an countable ordinal by a condensation is still a countable ordinal.

A *condensation formula* $\varphi(x, y)$ is a formula which is interpreted as a condensation of the domain over all countable ordinal words, *i.e.* for every word $w \in \Sigma^{\text{ord}}$, the relation defined on $\text{dom}(w)$ by $\iota \sim_\varphi \kappa$ if and only if $w, [x \mapsto \iota, y \mapsto \kappa] \models \varphi(x, y)$ is a condensation. A condensation formula $\varphi(x, y)$ induces a map:

$$\hat{\varphi} \colon \Sigma^{\text{ord}} \to (\Sigma^{\text{ord}+})^{\text{ord}}$$

where for every $u \in \Sigma^{\text{ord}}$, $\hat{\varphi}(u)$ is a word whose domain is $\text{dom}(w)/{\sim_\varphi}$, and such that for every class $I \in \text{dom}(w)/{\sim_\varphi}$, the I-th letter of $\hat{\varphi}(u)$ is the word $(u_\iota)_{\iota \in I}$—hence $\text{flat}(\hat{\varphi}(u)) = u$.

For example, the formula $\text{finite}(x, y)$ is a condensation formula, called *finite condensation*. The function $\hat{\varphi}_{\text{finite}} \colon \Sigma^{\text{ord}} \to (\Sigma^{\text{ord}})^{\text{ord}}$ that it induces sends the word $ababab\cdots cdcdcd\cdots abc \in \Sigma^{\text{ord}}$ of length $\omega \cdot 2 + 3$ to the 3-letter word $(ababab\cdots)(cdcdcd\cdots)(abc)$. Observe that for every word $w \in \Sigma^{\text{ord}}$, every letter of $\hat{\varphi}_{\text{finite}}(w)$ is a word of length ω, except possibly for the last letter (if the word has one), which can be finite.

Given two FO-definable functions—one that describes "local transformations" and another that described how to glue these local transformations together—the following lemma allows us to build a new FO-definable function. It is one of the key ingredients in our proof of Theorem 1.

Lemma 5. *Let A, B, C be finite sets. Let $\varphi(x, y)$ be a condensation FO-formula over A, let $f \colon A^{\text{ord}+} \to B$ and $g \colon B^{\text{ord}} \to C$ be FO-definable functions. Then, the map*

$$\begin{aligned} g \circ_\varphi f \colon A^{\text{ord}} &\to \quad C \\ u &\mapsto g\left(\prod_{i \in dom(\hat{\varphi}(u))} f(\hat{\varphi}(u)_i)\right) \end{aligned}$$

is FO-definable.

3 The algorithm

In this section we describe the algorithm behind Theorem 1. We first introduce the key notion of saturation in Section 3.1, and formalise the algorithm in Section 3.2.

3.1 The saturation construction

Until the end of Section 3.1, we fix a finite ordinal monoid $\mathcal{M} = (M, \cdot, 1, -^{\omega})$.

The saturation construction is at the heart of the algorithm, both in this paper, and in previous work. We introduce the necessary definitions. Note however that in our case, we do not close the definition under subsets as is usually done. This change, which may look minor, is in fact key for our proof to go through in the case of countable ordinals, and we find it also simplifies some points in the setting of finite words. We first recall an essential operation on $\mathcal{P}(M)$ that we denote $-^{\mathrm{grp}}$. Applied to a set $X \subseteq M$, it computes the union of all the elements that belong to the maximal group in the subsemigroup of $\mathcal{P}(M)$ generated by X.

Definition 6. *Let $X \subseteq M$. Define*

$$X^{\mathrm{grp}} = \bigcup_{k \in \mathbb{N}} X^{\mathrm{idem}+k} =^{\star} \bigcap_{n \in \mathbb{N}} \bigcup_{m \geq n} X^{m}.$$

Note that the $\star$ equality holds: Left to right inclusion comes from the fact that $X^{\mathrm{idem}+k} = X^m$ holds for infinitely many values of m, while the other inclusion stems from the fact that X^m can be written as $X^{\mathrm{idem}+k}$ for some k whenever m is sufficiently large.

Some important properties of this operation are the following.

Lemma 7. *The operation $-^{\mathrm{grp}}$ is monotonic, and for all $A, B \subseteq M$, and all integers k,*

$$A^{\mathrm{idem}+k} \subseteq A^{\mathrm{grp}}, \qquad (A \cdot B)^{\mathrm{grp}} = A \cdot (B \cdot A)^{\mathrm{grp}} \cdot B\ ,$$

and $$A^{\mathrm{grp}} \cdot A^{\mathrm{grp}} = (A^{\mathrm{grp}})^{\mathrm{grp}} = A^{\mathrm{grp}}.$$

The core of the algorithm computes the closure under $-^{\mathrm{grp}}$ and all the operations of the algebra of the images of the letters.

Definition 8. *Let $A \subseteq \mathcal{P}(M)$. The set $\langle A \rangle^{\mathrm{grp,ord}} \subseteq \mathcal{P}(M)$ is defined to be the least set containing A, $\{1\}$, and closed under $\cdot$, $^{\mathrm{grp}}$ and $^{\omega}$.*[7]

This definition is close in spirit to what is called saturation in previous works, with the difference that we do not take the downward closure, and that we close under the operation $-^{\omega}$. Despite this difference, we sometimes call $\langle A \rangle^{\mathrm{grp,ord}}$ the *saturation.*

Observe that the ordinal monoid $\mathcal{M}$ is aperiodic if and only if

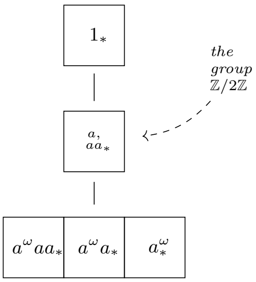
$$\langle \{\{x\} \mid x \in \mathcal{M}\} \rangle^{\mathrm{grp,ord}} = \{\{x\} \mid x \in \mathcal{M}\}\ .$$

[7] Recall that we showed that in a power ordinal monoid, the operation $-^{\omega}$ is computable.

3.2 The algorithm

We are now ready to describe the core of the algorithm that is claimed to exist in Theorem 1. Let K and L be two regular languages of countable ordinal words over the alphabet Σ. The *algorithm* is:

1. Let $\mathcal{M}$, σ, F_K, F_L be such that K is recognised by $(\mathcal{M}, \sigma, F_K)$ and L by $(\mathcal{M}, \sigma, F_L)$.
2. Compute $\mathrm{Sat} := \langle \{\{\sigma(a)\} \mid a \in \Sigma\} \rangle^{\mathrm{grp,ord}}$ (inside $\mathcal{P}(\mathcal{M})$).
3. If $F_K \cap X \neq \varnothing$ and $F_L \cap X \neq \varnothing$ for some $X \in \mathrm{Sat}$, answer 'no'. Otherwise answer 'yes'.



$\cdot$	1	a	aa	a^ω	$a^\omega a$	$a^\omega aa$
1	1	a	aa	a^ω	$a^\omega a$	$a^\omega aa$
a	a	aa	a	a^ω	$a^\omega a$	$a^\omega aa$
aa	aa	a	aa	a^ω	$a^\omega a$	$a^\omega aa$
a^ω	a^ω	$a^\omega a$	$a^\omega aa$	a^ω	$a^\omega a$	$a^\omega aa$
$a^\omega a$	$a^\omega a$	$a^\omega aa$	$a^\omega a$	a^ω	$a^\omega a$	$a^\omega aa$
$a^\omega aa$	$a^\omega aa$	$a^\omega a$	$a^\omega aa$	a^ω	$a^\omega a$	$a^\omega aa$
$\_^\omega$	1	a^ω	a^ω	a^ω	a^ω	a^ω

$$\langle \{\{a\}\} \rangle^{\mathrm{grp,ord}} = \{\{1\}, \{a\}, \{aa\}, \{a, aa\}, \{a^\omega\}, \{a^\omega a\}, \{a^\omega aa\}, \{a^\omega a, a^\omega aa\}\}$$

Fig. 1. Egg-box diagram of a finite ordinal monoid $\mathcal{M}$ recognising J, K and L (left), multiplication table and ω-iteration of $\mathcal{M}$ (right) and saturation (bottom).

Example 9. We illustrate the saturation construction and the algorithm on the following three languages over the singleton alphabet $\{a\}$:

$$\begin{aligned} J &= \{\text{infinite words whose longest finite suffix has even length}\}, \\ K &= \{\text{infinite words whose longest finite suffix has odd length}\}, \\ \text{and} \quad L &= \{\text{words that do not have a last letter}\}. \end{aligned}$$

It is classical that J and K are not FO-definable, while L is defined by the formula $\forall x. \exists y. y > x$. We can build a finite ordinal monoid $\mathcal{M}$ recognising all three languages: it has six elements, 1, a, aa, a^ω, $a^\omega a$ and $a^\omega aa$. Its presentation its described Figure 1. Naturally, the letter a is mapped to $\sigma(a) = a$. Then J, K and L are recognised by $F_J := \{a^\omega, a^\omega aa\}$, by $F_K := \{a^\omega a\}$ and by $F_L := \{1, a^\omega\}$, respectively.

The languages K and L are FO-separable: in fact L is an FO-separator of K and L. On the other hand, J and K are not FO-separable, as witnessed by the saturation algorithm. Indeed, the saturation $\langle\{\{\sigma(a)\} \mid a \in \Sigma\}\rangle^{\text{grp,ord}}$ contains all singletons, and furthermore $\{a, aa\} = \{a\}^{\text{grp}}$. As a consequence, it also contains $\{a^\omega a, a^\omega aa\} = \{a\}^{\underline{\omega}} \cdot \{a, aa\}$. This last set intersects both F_J and F_K.

The rest of the paper is dedicated to establishing the validity of this approach. In Section 4, we prove Proposition 12 stating that if the algorithm answers 'no', then the languages cannot be separated, as described in Theorem 1. In Section 5, we prove Corollary 16 stating that if the algorithm answers 'yes', then it is possible to construct an FO-separator sentence as described in Theorem 1. In Section 6, we shall package the results of Sections 4 and 5 differently, concluding that we have in fact computed the pointlike sets, and that we can also decide the more general covering problem.

4 When the algorithm says 'no'

In this section, we establish the correctness of the algorithm, i.e., when the algorithm answers 'no', we have to prove that the two input languages cannot be separated by an FO-definable language, and that we can produce a witness function. This is established in Proposition 12. The proof follows standard arguments.

The *quantifier depth*, a.k.a. quantifier rank, of an FO-formula is the maximal number of nested quantifiers in the formula. Two words $u, v \in \Sigma^{\text{ord}}$ are said to be FO_k*-equivalent*, denoted by $u \equiv_{\text{FO}_k} v$, if every FO-sentence of quantifier depth at most k accepts u if and only if it accepts v.

Proposition 10. *Let $k \in \mathbb{N}$.*

- *For $u, u', v, v' \in \Sigma^{\text{ord}}$, if $u \equiv_{\text{FO}_k} u'$ and $v \equiv_{\text{FO}_k} v'$ then $uv \equiv_{\text{FO}_k} u'v'$,*
- *for all Σ^{ord}-valued sequences $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$, if $u_n \equiv_{\text{FO}_k} v_n$ for all $n \in \mathbb{N}$, then $\text{flat}(u_n \mid n \in \mathbb{N}) \equiv_{\text{FO}_k} \text{flat}(v_n \mid n \in \mathbb{N})$, and*
- *for all $n \geqslant 2^k - 1$, for all $u \in \Sigma^{\text{ord}}$, $u^n \equiv_{\text{FO}_k} u^{n+1}$.*

This can be proved, for example, by using Ehrenfeucht-Fraïssé games—see e.g. [33, Lemma 6.5 & Corollary 6.9] for a proof of the first and third items ; the proof of the second item is similar[8]. Note that the first two items are also immediate corollaries of the Feferman-Vaught theorem [20, Theorem 1.3]. Note that the third property can be used to prove that every FO-definable language is recognised by an aperiodic finite ordinal monoid—this is the easy direction of Bedon's theorem [7].

Throughout the rest of this section, we fix K and L, two regular languages of countable ordinal words over an alphabet Σ. Recall that the algorithm computes the subset $\text{Sat} := \langle\{\{\sigma(a)\} \mid a \in \Sigma\}\rangle^{\text{grp,ord}}$ of $\mathcal{P}(\mathcal{M})$, where $\mathcal{M}$ is a finite ordinal monoid recognizing both K and L.

[8] Moreover, note that the first item can be deduced from the second item by taking $u_n = v_n = \varepsilon$ for $n \geq 2$.

We begin with a lemma which states that to all sets that belong to Sat can be effectively associated witnesses of indisinguishability (we shall see in Proposition 30 that what we have proved is that the elements in Sat are pointlike sets).

Lemma 11. *There exists a computable function which takes as input a number $k \in \mathbb{N}$ and an element $X \in$ Sat, and produces an X-indexed sequence of ordinal words $(u_x)_{x\in X} \in (\Sigma^{\text{ord}})^X$ such that,*

- $\pi(\sigma^{\text{ord}}(u_x)) = x$ *for all* $x \in X$, *and*
- $u_x \equiv_{\text{FO}_k} u_{x'}$ *for all* $x, x' \in X$.

The proof is by structural induction on the definition of Sat, making use of the two first items of Proposition 10 for composing witnesses, and of furthermore the third item for treating the $-^{\text{grp}}$ operation.

From the above lemma, one can easily deduce that when the algorithm answers 'no', there is indeed an obstruction to the fact that K and L can be FO-separated.

Proposition 12. *Assume that the algorithm answers 'no' when run with input languages K and L. Then there is a witness function which computes, for any FO-sentence φ, a pair of words $(u, u') \in K \times L$ such that $u \models \varphi$ if and only if $u' \models \varphi$. In particular, K and L cannot be FO-separated.*

Proof. Since the algorithm answered 'no', pick a pair $(x, x') \in F_K \times F_L$ such that $x, x' \in X$ for some $X \in$ Sat. Now, for any FO-sentence φ, using the function of Lemma 11 with k the quantifier depth of φ, we can compute a sequence $(u_x)_{x\in X}$ of ordinal words. Now define $u := u_x$ and $u' := u_{x'}$. Then $u \equiv_{\text{FO}_k} u'$, so that $u \models \varphi$ if and only if $u' \models \varphi$. Also, $\pi(\sigma^{\text{ord}}(u)) = x \in F_K$ and $\pi(\sigma^{\text{ord}}(u')) = x' \in F_L$, so $u \in K$ and $u' \in L$. $\square$

Example 13 (Continuing Example 9). Recall that J and K are not FO-separable. Because of the set $\{a^\omega a, a^\omega aa\} \in \langle\{\sigma(a) \mid a \in \Sigma\}\rangle^{\text{grp,ord}}$, the algorithm outputs 'no', and can return, to witness the FO-inseparability of the two languages the computable map $\varphi \mapsto (a^\omega a^{2^k+1}, a^\omega a^{2^k+2}) \in J \times K$, where k denoted the quantifier depth of φ. To prove that $a^\omega a^{2^k+1} \equiv_{\text{FO}_k} a^\omega a^{2^k+2}$, one can simply use the first and third items of Proposition 10.

5 When the algorithm says 'yes'

We now establish the completeness part of the proof of the main theorem, Theorem 1. The goal of this proof is to establish that if the algorithm answers 'yes', it is indeed possible to produce an FO-separator (Corollary 16).

This is the part of the proof that differs most substantially from previous works on separation. In Section 5.1, we abstract the question with the notion of ordinal monoids with merge, and we introduce the notion of FO-approximants which are FO-definable over-approximations of the product. The key result, Lemma 15, states their existence for all finite ordinal monoid with merge. Corollary 16 follows immediately. The proof of Lemma 15 is then established in Section 5.2 for words of finite or ω length. Building on these simpler cases, the general case is the subject Section 5.3.

5.1 Merge operators and FO-approximants

We abstract in this section the ordinal $\mathcal{P}(M)$ equipped with the $-^{\mathrm{grp}}$ operator into a new algebraic structure. A finite *ordinal monoid with merge* $\mathcal{M} = (M, \underline{1}, \leqslant, \underline{\cdot}, \underline{\omega}, ^{\mathrm{grp}})$ consists of:

- a presentation of an ordered ordinal monoid $(M, \underline{1}, \leqslant, \underline{\cdot}, \underline{\omega})$, together with
- a monotonic *merge operator* $-^{\mathrm{grp}} \colon M \to M$ such that for all $a, b \in M$, and all integers k,

$$a^{\mathrm{idem}+k} \leqslant a^{\mathrm{grp}}, \qquad\qquad (a^{\mathrm{idem}})^{\mathrm{grp}} = a^{\mathrm{idem}},$$
$$a^{\mathrm{grp}} \,\underline{\cdot}\, a^{\mathrm{grp}} = (a^{\mathrm{grp}})^{\mathrm{grp}} = a^{\mathrm{grp}}, \qquad \text{and} \qquad (a \,\underline{\cdot}\, b)^{\mathrm{grp}} = a \,\underline{\cdot}\, (b \,\underline{\cdot}\, a)^{\mathrm{grp}} \,\underline{\cdot}\, b\ .$$

The following lemma is an immediate consequence of Lemma 7.

Lemma 14. *Both* $(\mathcal{P}(\mathcal{M}), \{1\}, \subseteq, \underline{\cdot}, \underline{\omega}, ^{\mathrm{grp}})$ *and* $(\mathrm{Sat}, \{1\}, \subseteq, \underline{\cdot}, \underline{\omega}, ^{\mathrm{grp}})$ *are ordinal monoids with merge.*

The idea behind ordinal monoids with merge is that not only there is a product operation as for every ordinal monoid, but also an FO-definable over-approximation for it. This is the concept of FO-approximant that we introduce now. Given a an FO-definable language $L \subseteq M^{\mathrm{ord}}$, an FO*-approximant* of π over L is an FO-definable map $\rho\colon L \to M$ such that:

$$\pi(u) \leqslant \rho(u), \qquad \text{for all } u \in L.$$

The key result concerning ordinal monoids with merge is the existence of a total FO-approximant:

Lemma 15. *There is an* FO*-approximant* ρ *over* M^{ord} *for all ordinal monoids with merge* $\mathcal{M}$.

An example of an FO-approximant can be found in Example 26. Before establishing Lemma 15, let us explain why it is sufficient for concluding the proof of Theorem 1 in the case the algorithm answers 'yes'.

Corollary 16. *If the algorithm answers 'yes', there exists an* FO*-separator.*

Proof. By Lemmas 14 and 15, there exists an FO-approximant $\rho : A^{\mathrm{ord}} \to \langle A\rangle^{\mathrm{grp,ord}}$ over the power ordinal monoid $\mathcal{P}(\mathcal{M})$, where $A = \{\{\sigma(a)\} \mid a \in \Sigma\}$. Now define the language

$$S := \{u \in \Sigma^{\mathrm{ord}} \mid \rho(\tilde{\sigma}^{\mathrm{ord}}(u)) \cap F_K \neq \varnothing\}$$
$$\text{where } \tilde{\sigma}^{\mathrm{ord}}(u) := (\{\sigma(u_i)\})_{i \in \mathrm{dom}(u)} \in A^{\mathrm{ord}} \text{ for all } u \in \Sigma^{\mathrm{ord}}.$$

Note first that since ρ is FO-definable, this language is FO-definable. Let us show that it separates K from L.

For every $u \in K$, $F_K \ni \pi(\sigma^{\text{ord}}(u)) \subseteq \rho(\tilde{\sigma}^{\text{ord}}(u))$, and as a consequence $\rho(\tilde{\sigma}^{\text{ord}}(u)) \cap F_K \neq \varnothing$. We have proved $K \subseteq S$.

Conversely, consider some $u \in L$. We have $F_L \ni \pi(\sigma^{\text{ord}}(u)) \in \rho(\tilde{\sigma}^{\text{ord}}(u)) \in \langle A\rangle^{\text{grp,ord}}$, and thus $\rho(\tilde{\sigma}^{\text{ord}}(u)) \cap F_L \neq \varnothing$. Since the algorithm returns 'yes', this means that there is no set in $\langle A\rangle^{\text{grp,ord}}$ that intersects both F_K and F_L. In our case, this means that $\rho(\tilde{\sigma}^{\text{ord}}(u)) \cap F_K = \varnothing$, proving that $u \notin S$. We have proved $L \cap S = \varnothing$.

Overall, S is an FO-separator for K and L. □

Remark 17. Notice how the "difficult" implication of Bedon's theorem (Proposition 4) can be easily deduced from Lemma 15[9]: recall that this implication consists in showing that a regular language $L \subseteq \Sigma^{\text{ord}}$, recognised by some triplet $(\mathcal{M}, \sigma, F)$ with $\mathcal{M}$ is aperiodic is definable in first-order logic. Indeed, by aperiodicity of $\mathcal{M}$, the operation $^{\text{grp}}$ applied to a singleton $\{a\}$ yields the singleton $\{a^{\text{idem}}\}$. Hence, the set $\langle\{\{\sigma(a)\} \mid a \in \Sigma\}\rangle^{\text{grp,ord}} = \{\{\pi \circ \sigma^{\text{ord}}(u)\} \mid u \in \Sigma^{\text{ord}}\}$ consists only of singletons, and as a consequence, all FO-approximants ρ (and in particular the one constructed in Lemma 15) maps a word u to $\pi(u)$. Hence, π is an FO-definable map, and thus L is an FO-definable language.

The rest of this section is devoted to establishing Lemma 15. The construction is based on subresults showing the existence of FO-approximants over subsets of M^{ord}; first for finite and ω-words in Section 5.2, and finally for words of any countable ordinal length in Section 5.3. But beforehand, we shall introduce some more definitions and elementary results.

In what follows we use the notation $\langle - \rangle^{\text{grp,ord}}$ from Definition 8, interpreted in a generic ordinal monoid with merge, as well as some variants. Let $A \subseteq M$. We define $\langle A\rangle^{+}$ as the closure of A under $\underline{\cdot}$, $\langle A\rangle^{\text{grp}+}$ as the closure of A under $\underline{\cdot}$ and $-^{\text{grp}}$, and $\langle A\rangle^{\text{grp}*}$ as $\langle A\rangle^{\text{grp}+} \cup \{\underline{1}\}$. We define $\langle A\rangle^{\text{grp,ord}+}$ as the closure of A under $\underline{\cdot}$, $^{\text{grp}}$ and $\underline{^{\omega}}$. Note that thanks to the identities of ordinal monoids with merge, we have $\langle A\rangle^{\text{grp,ord}} = \langle A\rangle^{\text{grp,ord}+} \cup \{\underline{1}\}$. Moreover, we have the following identities[10]:

Proposition 18. *Let $\mathcal{M}$ be an ordinal monoid with merge. For every $A \subseteq \mathcal{M}$,*

$$\langle A\rangle^{\text{grp}+} = A\langle A\rangle^{\text{grp}*} = \langle A\rangle^{\text{grp}*}A \qquad \textit{and} \qquad \langle A\rangle^{\text{grp,ord}+} = A\langle A\rangle^{\text{grp,ord}} .$$

Proof. Note, by definition, that $\langle A\rangle^{\text{grp}*} = \langle A\rangle^{\text{grp}+} \cup \{1\}$, so

$$A\langle A\rangle^{\text{grp}*} = A\langle A\rangle^{\text{grp}+} \cup A \subseteq \langle A\rangle^{\text{grp}+}.$$

The converse inclusion $\langle A\rangle^{\text{grp}+} \subseteq A\langle A\rangle^{\text{grp}*}$ is obtained by induction. Let $b \in \langle A\rangle^{\text{grp}+}$. If $b \in A$, then $b \in A\langle A\rangle^{\text{grp}*}$ since $1 \in \langle A\rangle^{\text{grp}*}$. If $c = cd$ with $c, d \in$

[9] Similarly, for finite words, Schützenberger-McNaughton-Papert's theorem is a consequence of Henckell's algorithm for aperiodic pointlikes—see e.g. [25, Corollary 4.8]

[10] Notice the similarity with the (trivial) identities $A^{+} = AA^{*} = A^{*}A$ and $A^{\text{ord}+} = AA^{\text{ord}}$.

$\langle A\rangle^{\mathrm{grp}+}$, then, by induction, $c = ac'$ for some $a \in A$ and $c' \in \langle A\rangle^{\mathrm{grp}*}$, thus $b = a(c'd) \in A\langle A\rangle^{\mathrm{grp}*}$ since $a \in A$ and $c'd \in \langle A\rangle^{\mathrm{grp}*}$. Finally, if $b = c^{\mathrm{grp}}$, then, again by induction, $c = ac'$ for some $a \in A$ and $c' \in \langle A\rangle^{\mathrm{grp}*}$, and thus $b = c^{\mathrm{grp}} = cc^{\mathrm{grp}} = a(c'c^{\mathrm{grp}}) \in A\langle A\rangle^{\mathrm{grp}*}$.

The equality $\langle A\rangle^{\mathrm{grp}+} = \langle A\rangle^{\mathrm{grp}*}A$ is symmetric.

The identity $\langle A\rangle^{\mathrm{grp,ord}+} = A\langle A\rangle^{\mathrm{grp,ord}}$ is similar. The new case in the induction is if some $b \in \langle A\rangle^{\mathrm{grp,ord}+}$ is of the form c^ω, then, by induction hypothesis, $c = ac'$ for some $a \in A$ and $c' \in \langle A\rangle^{\mathrm{grp,ord}}$, and thus $b = c^\omega = cc^\omega = a(c'c^\omega) \in A\langle A\rangle^{\mathrm{grp,ord}}$. ☐

Proposition 19. *If there are FO-approximants over K and L respectively, then there exist effectively FO-approximants over $K \cup L$ and KL.*

5.2 Construction of FO-approximants for words of finite and ω-length

First, we show how to construct FO-approximants for finite words. It serves at the same time as a building block for more complex cases, as a way to show the proof mechanisms in simpler cases, as well as to comment on differences with previous works.

Lemma 20. *Let $A \subseteq M$, then either*

- $a \cdot \langle A\rangle^{\mathrm{grp}+} \subsetneq \langle A\rangle^{\mathrm{grp}+}$, *for some $a \in A$,*
- $\langle A\rangle^{\mathrm{grp}+} \cdot a \subsetneq \langle A\rangle^{\mathrm{grp}+}$, *for some $a \in A$, or*
- $\langle A\rangle^{\mathrm{grp}+}$ *has a maximum.*

Proof. Assume the two first items do not hold. Because of the non-first-one, the map $x \mapsto a \cdot x$ is surjective on $\langle A\rangle^{\mathrm{grp}+}$, for all $a \in A$. Since $\langle A\rangle^{\mathrm{grp}+}$ is finite, this means that it is bijective on $\langle A\rangle^{\mathrm{grp}+}$. Hence it is also bijective on $\langle A\rangle^{+}$. The negation of the second item has a symmetric consequence. Together we get that $\langle A\rangle^{+}$ is a group. Let I be its neutral element. Note first that for all $x \in \langle A\rangle^{+}$, $I = x^k$ for some k, and hence, $I \leqslant x^{\mathrm{grp}}$. Set now $a_1, \ldots, a_n$ to be the elements in A, and define: $M = (a_1^{\mathrm{grp}} \cdot a_2^{\mathrm{grp}} \cdots a_n^{\mathrm{grp}})^{\mathrm{grp}}$.

By the above remark $a_i = I^{i-1} \cdot a_i \cdot I^{n-i} \leqslant a_1^{\mathrm{grp}} \cdot a_2^{\mathrm{grp}} \cdots a_n^{\mathrm{grp}} \leqslant M$ for all i. Since furthermore for all $x, y \leqslant M$, $x \cdot y \leqslant M$ and $x^{\mathrm{grp}} \leqslant M$, it follows that $z \leqslant M$ for all $z \in \langle A\rangle^{\mathrm{grp}+}$. ☐

A similar lemma is used in [25], but concludes with the existence of a pseudo-group as the third item.

Lemma 21. *For all $a \in M$ there exists an FO-approximant from a^+ to $\langle\{a\}\rangle^{\mathrm{grp}+}$.*

Construction. Let k be such that $a^{\mathrm{idem}} = a^k$. Define

$$\rho(\overbrace{a \cdots a}^{\text{length } n}) = \begin{cases} a^n & \text{if } n < k, \\ a^{\mathrm{grp}} & \text{otherwise.} \end{cases}$$ ☐

We can now use this for proving the finite word case.

Lemma 22. *For all $A \subseteq M$ there exists an FO-approximant from A^+ to $\langle A\rangle^{\mathrm{grp}+}$.*

Proof. We use a double induction on $|\langle A\rangle^{\mathrm{grp}+}|$ and $|A|$. The induction is guided by Lemma 20. The base case is $A = \varnothing$, and the nowhere defined FO-approximant proves it.

First case: $a \cdot \langle A\rangle^{\mathrm{grp}+} \subsetneq \langle A\rangle^{\mathrm{grp}+}$ for some $a \in A$. This part of the proof is similar to [25, Lemma 6.7]. Let $B ::= A \setminus \{a\}$.

We first construct an FO-approximant from a^+B^+ to $a \cdot \langle A\rangle^{\mathrm{grp}+}$. Indeed, we know by Lemma 21 that there is an FO-approximant from a^+ to $\langle\{a\}\rangle^{\mathrm{grp}+} \subseteq a \cdot \langle A\rangle^{\mathrm{grp}*}$. We also know by induction[11] that there is an FO-approximant from B^+ to $\langle B\rangle^{\mathrm{grp}+} \subseteq \langle A\rangle^{\mathrm{grp}+}$. Thus by Proposition 19, there exists effectively an FO-approximant τ from a^+B^+ to $a \cdot \langle A\rangle^{\mathrm{grp}*} \cdot \langle A\rangle^{\mathrm{grp}+} \subseteq a \cdot \langle A\rangle^{\mathrm{grp}+}$.

We now provide an FO-approximant for $(a^+B^+)^+$ (which is FO-definable), and for this, define the condensation FO-formula $\varphi(x, y)$ that expresses that "two positions x and y are equivalent if the subword on the interval $[x, y]$ belongs to a^*B^*" (this can be expressed in first-order logic). Over a word $u \in (a^+B^+)^+$, each of the condensation classes belong to a^+B^+ and its image under τ belongs to $a \cdot \langle A\rangle^{\mathrm{grp}+}$. Furthermore, still by induction hypothesis[12], there is an FO-approximant from $(a \cdot \langle A\rangle^{\mathrm{grp}+})^+$ to $\langle A\rangle^{\mathrm{grp}+}$. By Lemma 5, we thus obtain an FO-definable map from $(a^+B^+)^+$ to $\langle A\rangle^{\mathrm{grp}+}$. It is an FO-approximant by construction.

Using the above case and Proposition 19, it can be easily extended to an FO-approximant from $A^+ = AB^*(a^+B^+)^*a^*$ to $\langle A\rangle^{\mathrm{grp}+}$.

Second case: $\langle A\rangle^{\mathrm{grp}+} \cdot a \subsetneq \langle A\rangle^{\mathrm{grp}+}$. This case is symmetric to the first case.

Third case: $\langle A\rangle^{\mathrm{grp}+}$ has a maximum M. Then the constant map that sends every word $u \in A^*$ to M is an FO-approximant over A^*. ☐

Following similar ideas, we can treat the case of ω-words. We define here $\langle A\rangle^{\mathrm{grp},\omega}$ as the elements of the form $\{a \cdot b^{\omega} \mid a, b \in \langle A\rangle^{\mathrm{grp}+}\}$—or, equivalently, $\langle A\rangle^{\mathrm{grp},\omega} = (\langle A\rangle^{\mathrm{grp}+})^{\omega}$.

Lemma 23. *Let M be an ordinal monoid with merge. For all $A \subseteq M$, there exists an FO-approximant from A^ω to $\langle A\rangle^{\mathrm{grp},\omega}$.*

5.3 Construction of FO-approximants for countable ordinal words

As for the finite case, the proof revolves around a carefully designed case distinction. This one is more complex to establish, and makes use of Green's relations and a precise understanding of the properties of ordinal monoids with merge.

Lemma 24 (Trichotomy principle). *Let M be a finite ordinal monoid with merge and $A \subseteq M$, then either*

[11] Indeed, $|B| < |A|$.

[12] This time, we can use the induction hypothesis because $|\langle(a \cdot \langle A\rangle^{\mathrm{grp}+})^+\rangle^{\mathrm{grp}+}| < |\langle A\rangle^{\mathrm{grp}+}|$. Indeed, by Proposition 18, $\langle(a \cdot \langle A\rangle^{\mathrm{grp}+})^+\rangle^{\mathrm{grp}+} \subseteq (a \cdot \langle A\rangle^{\mathrm{grp}+})^+\langle(a \cdot \langle A\rangle^{\mathrm{grp}+})^+\rangle^{\mathrm{grp}*} \subseteq a \cdot \langle A\rangle^{\mathrm{grp}+} \subsetneq \langle A\rangle^{\mathrm{grp}+}$.

– $a \mathbin{\underline{\cdot}} \langle A\rangle^{\mathrm{grp,ord+}} \subsetneq \langle A\rangle^{\mathrm{grp,ord+}}$, *for some* $a \in A$,
– $\langle\langle A\rangle^{\mathrm{grp},\omega}\rangle^{\mathrm{grp,ord+}} \subsetneq \langle A\rangle^{\mathrm{grp,ord+}}$, *or*
– $x \mathbin{\underline{\cdot}} y = y$ *and* $x^{\underline{\omega}} = y^{\underline{\omega}}$, *for all* $x, y \in \langle A\rangle^{\mathrm{grp,ord+}}$.

The above lemma is key in the proof of the existence of an FO-approximant.

Lemma 25. *For all* $a \in \mathcal{M}$, *there exists an* FO-*approximant over* a^{ord}.

The proof follows a similar structure as the one for Lemma 22 for the finite case. This time, Lemma 24 is the key argument that makes the induction progress, playing the same role as Lemma 20 in the finite case. Note, however, that the second items in Lemmas 20 and 24 are very different in structure. And indeed, this entails a different argument for constructing the FO-approximant. It is based on performing in one step the condensation of all the maximal factors of order-type ω.

Example 26 (Continuing Example 13). An FO-approximant ρ of π over a^{ord} in the ordinal monoid defined in Example 9 can be defined for all $u \in \{a\}^{\mathrm{ord}}$ as:

$$\rho(u) := \begin{cases} \{1\} & \text{if } \mathrm{dom}(u) \text{ is empty,} \\ \{a, aa\} & \text{if } \mathrm{dom}(u) \text{ is finite and non-empty,} \\ \{a^\omega\} & \text{if } \mathrm{dom}(u) \text{ is a non-zero limit ordinal,} \\ \{a^\omega a, a^\omega aa\} & \text{if } \mathrm{dom}(u) \text{ is an infinite successor ordinal.} \end{cases}$$

Lemma 27. *For all* $A \subseteq \mathcal{M}$, *there exists an* FO-*approximant from* $A^{\mathrm{ord+}}$ *to* $\langle A\rangle^{\mathrm{grp,ord+}}$.

Proof. We prove the result by induction on $|\langle A\rangle^{\mathrm{grp,ord+}}|$ and $|A^{\mathrm{ord+}}|$. The base case $A = \varnothing$ is trivial. If A is non-empty, following Lemma 24, there are three cases to treat.

First case: There exists $a \in A$ *such that* $a \mathbin{\underline{\cdot}} \langle A\rangle^{\mathrm{grp,ord+}} \subsetneq \langle A\rangle^{\mathrm{grp,ord+}}$. This case is as in the proof for finite words, Lemma 22, using Lemma 25 in place of Lemma 21. The key reason why the proof remains valid is because the hypothesis $a \mathbin{\underline{\cdot}} \langle A\rangle^{\mathrm{grp,ord+}} \subsetneq \langle A\rangle^{\mathrm{grp,ord+}}$ implies $|\langle(a \mathbin{\underline{\cdot}} \langle A\rangle^{\mathrm{grp,ord+}})^{\mathrm{ord+}}\rangle^{\mathrm{grp,ord+}}| < |\langle A\rangle^{\mathrm{grp,ord+}}|$ by Proposition 18[13].

Second case[14]: $\langle\langle A\rangle^{\mathrm{grp},\omega}\rangle^{\mathrm{grp,ord+}} \subsetneq \langle A\rangle^{\mathrm{grp,ord+}}$. By Lemma 23, there is an FO-approximant from A^ω to $\langle A\rangle^{\mathrm{grp},\omega}$. By induction hypothesis[15], we have an FO-approximant from $(\langle A\rangle^{\mathrm{grp},\omega})^{\mathrm{ord+}}$ to $\langle\langle A\rangle^{\mathrm{grp},\omega}\rangle^{\mathrm{grp,ord+}} \subseteq \langle A\rangle^{\mathrm{grp,ord+}}$. Since

[13] More precisely, we are using the property $\langle B\rangle^{\mathrm{grp,ord+}} = B\langle B\rangle^{\mathrm{grp,ord}}$ of Proposition 18. By thinking of elements of $\langle B\rangle^{\mathrm{grp,ord+}}$ as "countable ordinal words with merge", this property is simply saying that every "countable ordinal word with merge" has a first letter. However, countable ordinal words need not have a last letter: this is what makes an hypothesis of the form $\langle A\rangle^{\mathrm{grp,ord+}} \mathbin{\underline{\cdot}} a \subsetneq \langle A\rangle^{\mathrm{grp,ord+}}$ unusable—and this is the motivation behind the trichotomy principle Lemma 24.

[14] Note here that it is different from the second case in the proof of Lemma 22.

[15] Indeed, $\langle\langle A\rangle^{\mathrm{grp},\omega}\rangle^{\mathrm{grp,ord+}} \subsetneq \langle A\rangle^{\mathrm{grp,ord+}}$.

the formula finite(x, y) is a condensation FO-formula, we obtain by Lemma 5 an FO-approximant from $(A^\omega)^{\mathrm{ord}+} \to \langle A\rangle^{\mathrm{grp,ord}+}$. Using Proposition 19 and Lemma 22, we easily extend it to an FO-approximant from $A^{\mathrm{ord}+} = A(A^\omega)^{\mathrm{ord}} A^*$ to $\langle A\rangle^{\mathrm{grp,ord}+}$.

Third case: $x \cdot y = y$ *and* $x^{\underline{\omega}} = y^{\underline{\omega}}$, *for all* $x, y \in \langle A\rangle^{\mathrm{grp,ord}+}$. Then the product over A sends a countable ordinal word $u \in A^{\mathrm{ord}+}$ to its last letter if the word has a last letter, and to the unique omega power of $\langle A\rangle^{\mathrm{grp,ord}+}$ if the word has no last letter. Since the languages of the form $A^{\mathrm{ord}+}a$ where $a \in A$ and $\{u \in A^{\mathrm{ord}+} \mid \mathrm{dom}(u) \text{ is a limit ordinal}\}$ all are FO-definable, it follows that the product over A is FO-definable. □

6 Related problems

In this section, we solve two related problems: the decidability of the covering problem (Proposition 28), and the computability of pointlike sets (Proposition 30). Both are direct applications of the key lemmas presented above.

The FO-*covering problem* asks, given regular languages, in our case of countable ordinal words, $L, K_1, \ldots, K_n$, to determine if there exist FO-definable languages $C_1, \ldots, C_n$ such that $L \subseteq \cup_i C_i$ and $C_i \cap K_i = \varnothing$ for all i—see [27] for more details. In general, separation problems trivially reduce to covering problems, since L and K are separable if and only if there is a solution to the covering problem for the instance (L, K). In the other direction, there is no known example of a variety with decidable separation problem but undecidable covering problem. We show that a further consequence of the above results is that the FO-pointlike sets in a finite ordinal monoid (see Definition 29) are computable, from which we deduce:

Proposition 28. *The* FO-*covering problem for countable ordinal words is decidable.*

Let us now introduce, and explain, the relation with pointlike sets. The FO$_k$-*closure* of a word u is the set $[u]_{\mathrm{FO}_k}$ which contains all words that are FO$_k$-equivalent to u.

Definition 29. *Given a finite ordinal monoid* $\mathcal{M}$ *the* FO-pointlike sets *of a map* $\sigma\colon \Sigma \to M$ *are defined by*

$$\mathrm{Pl}_{\mathrm{FO}}(\sigma) ::= \bigcap_{k\in\mathbb{N}} \downarrow \left\{\pi(\sigma^{\mathrm{ord}}([u]_{\mathrm{FO}_k})) \mid u \in \Sigma^{\mathrm{ord}}\right\},$$

where $\downarrow X$ *denotes the downward closure of* X.

The definition of pointlike sets is in fact more general[16]: given a variety of finite semigroups $\mathbb{V}$ one can define a notion of pointlike sets with respect to this

[16] In the following discussion, we focus on finite words, but the notion of variety—of algebras, or of languages—can be extended to countable ordinal words [8] and many other settings [11, §4].

variety. Almeida observed that the separation problem for the variety $\mathbb{V}$—given two regular languages, can they be separated by a $\mathbb{V}$-recognisable language?—is decidable if and only if the $\mathbb{V}$-pointlikes of size 2 of every morphism are computable [2, Prop. 3.4]. The covering problem also has an algebraic counterpart: it is decidable for the variety $\mathbb{V}$ if and only if, for every morphism, the collection of all $\mathbb{V}$-pointlike sets of this morphism is computable [2, Prop. 3.6][17]. Hence, the fact that FO-covering and FO-separation are decidable for finite words is simply a corollary of Henckell's theorem on aperiodic pointlikes [19, Fact 3.7 & Fact 5.31], stating that they are computable. Place & Zeitoun's simpler proof of the decidability of FO-covering for finite words and for ω-words [25] relies on the same principle.[18] Unsurprisingly, our result can be interpreted in the same way: we are implicitly showing the following property, from which one can immediately deduce the computability of $\mathrm{Pl}_{\mathrm{FO}}(\sigma)$.

Proposition 30. *Given a finite ordinal monoid $\mathcal{M}$ and $\sigma\colon \Sigma \to M$,*

$$\mathrm{Pl}_{\mathrm{FO}}(\sigma) = \downarrow \langle \{\{\sigma(a)\} \mid a \in \Sigma\} \rangle^{\mathrm{grp,ord}}.$$

7 Conclusion

In this paper, we have studied the problem of FO-separation over words of countable ordinal length. Our proof is based on the work of Place and Zeitoun over words of length ω [25]. We build an FO-approximant using essentially the same technique as Place and Zeitoun. However a key difference is that for finite words and ω-words, the proof relies on a case distinction (Lemma 20) which is conceptually similar to the characterisation of groups as semigroups whose translations are bijective. This was no longer sufficient for countable ordinal words because of ω-iterations. In this situation, our new case distinction (Lemma 24) captures the subtle interaction of ω-iteration with groups in finite ordinal monoids. In particular, a difference with previously known algorithms is that we do not close the saturation under subset. This a priori innocuous difference has significant consequences on the proof of completeness, yielding some simplifications in the finite and ω-case, and necessary for the proof to be extendable to all ordinals.

Of course, the next step is to go to longer words, in particular scattered countable words, or even better to all countable words. Here, there are conceptual difficulties, and let us stress also that, starting from scattered countable words, first-order logic and first-order logic with access to Dedekind cuts begin to have a different expressiveness. Thus several notions of separation have to be studied.

References

1. Adsul, B., Sarkar, S., Sreejith, A.V.: First-order logic and its infinitary quantifier extensions over countable words (2021)

[17] Beware: there is a typo in the statement of the first item of the proposition.

[18] There is a difference in terminology: they refer to the $\mathrm{Pl}_{\mathrm{FO}}(\varphi)$ as "optimal imprint with respect to FO on φ".

2. Almeida, J.: Some algorithmic problems for pseudovarieties. Publ. Math. Debrecen **54**(1), 531–552 (1999)
3. Almeida, J., Zeitoun, M.: The pseudovariety J is hyperdecidable. RAIRO-Theoretical Informatics and Applications **31**(5), 457–482 (1997)
4. Ash, C.J.: Inevitable graphs: a proof of the type II conjecture and some related decision procedures. International Journal of Algebra and Computation **1**(01), 127–146 (1991)
5. Bedon, N.: Finite automata and ordinals. Theoretical Computer Science **156**(1), 119–144 (1996). https://doi.org/10.1016/0304-3975(95)00006-2
6. Bedon, N.: Langages reconnaissables de mots indexés par des ordinaux. Theses, Université de Marne la Vallée (Jan 1998), https://tel.archives-ouvertes.fr/tel-00003586
7. Bedon, N.: Logic over words on denumerable ordinals. Journal of Computer and System Sciences **63**(3), 394–431 (2001). https://doi.org/10.1006/jcss.2001.1782
8. Bedon, N., Carton, O.: An Eilenberg theorem for words on countable ordinals. In: Lucchesi, C.L., Moura, A.V. (eds.) LATIN'98: Theoretical Informatics. pp. 53–64. Springer Berlin Heidelberg, Berlin, Heidelberg (1998). https://doi.org/10.1007/BFb0054310
9. Bedon, N., Rispal, C.: Schützenberger and Eilenberg theorems for words on linear orderings. Journal of Computer and System Sciences **78**(2), 517–536 (Mar 2012). https://doi.org/10.1016/j.jcss.2011.06.003
10. Bès, A., Carton, O.: Algebraic Characterization of FO for Scattered Linear Orderings. In: Bezem, M. (ed.) Computer Science Logic (CSL'11) - 25th International Workshop/20th Annual Conference of the EACSL. Leibniz International Proceedings in Informatics (LIPIcs), vol. 12, pp. 67–81. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2011). https://doi.org/10.4230/LIPIcs.CSL.2011.67
11. Bojańczyk, M.: Recognisable languages over monads. In: Potapov, I. (ed.) Developments in Language Theory. pp. 1–13. Springer International Publishing, Cham (2015), https://arxiv.org/abs/1502.04898v1
12. Büchi, J.R.: On a decision method in restricted second order arithmetic. In: Logic, Methodology and Philosophy of Science (Proc. 1960 Internat. Congr .), pp. 1–11. Stanford Univ. Press, Stanford, Calif. (1962)
13. Büchi, J.R.: The monadic second order theory of ω_1, pp. 1–127. Springer Berlin Heidelberg (1973). https://doi.org/10.1007/BFb0082721
14. Carton, O., Colcombet, T., Puppis, G.: An algebraic approach to MSO-definability on countable linear orderings (May 2018). https://doi.org/10.1017/jsl.2018.7
15. Choueka, Y.: Finite automata, definable sets, and regular expressions over ωn-tapes. Journal of Computer and System Sciences **17**(1), 81–97 (1978)
16. Colcombet, T., Sreejith, A.V.: Limited set quantifiers over countable linear orderings. In: Proceedings, Part II, of the 42nd International Colloquium on Automata, Languages, and Programming - Volume 9135. pp. 146–158. ICALP 2015, Springer-Verlag, Berlin, Heidelberg (2015). https://doi.org/10.1007/978-3-662-47666-6_12
17. van Gool, S.J., Steinberg, B.: Merge decompositions, two-sided Krohn–Rhodes, and aperiodic pointlikes. Canadian Mathematical Bulletin **62**(1), 199–208 (2019). https://doi.org/10.4153/CMB-2018-014-8
18. Gool, S., Steinberg, B.: Pointlike sets for varieties determined by groups. Advances in Mathematics **348**, 18–50 (2019). https://doi.org/10.1016/j.aim.2019.03.020
19. Henckell, K.: Pointlike sets: the finest aperiodic cover of a finite semigroup. Journal of Pure and Applied Algebra **55**(1), 85–126 (1988). https://doi.org/10.1016/0022-4049(88)90042-4

20. Makowsky, J.A.: Algorithmic uses of the Feferman–Vaught theorem. Annals of Pure and Applied Logic **126**(1-3), 159–213 (2004). https://doi.org/10.1016/j.apal.2003.11.002
21. Manuel, A., Sreejith, A.V.: Two-variable logic over countable linear orderings. In: Faliszewski, P., Muscholl, A., Niedermeier, R. (eds.) 41st International Symposium on Mathematical Foundations of Computer Science, MFCS 2016, August 22-26, 2016 - Kraków, Poland. LIPIcs, vol. 58, pp. 66:1–66:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2016). https://doi.org/10.4230/LIPIcs.MFCS.2016.66, `https://doi.org/10.4230/LIPIcs.MFCS.2016.66`
22. McNaughton, R., Papert, S.A.: Counter-Free Automata. The MIT Press (1971)
23. Perrin, D.: Recent results on automata and infinite words. In: International Symposium on Mathematical Foundations of Computer Science. pp. 134–148. Springer (1984). https://doi.org/10.1007/BFb0030294
24. Pin, J.E., Perrin, D.: Infinite Words: Automata, Semigroups, Logic and Games. Elsevier (2004), `https://hal.archives-ouvertes.fr/hal-00112831`
25. Place, T., Zeitoun, M.: Separating regular languages with first-order logic. Logical Methods in Computer Science **12** (2016). https://doi.org/10.2168/LMCS-12(1:5)2016
26. Place, T., Zeitoun, M.: The complexity of separation for levels in concatenation hierarchies. In: Ganguly, S., Pandya, P. (eds.) 38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018). Leibniz International Proceedings in Informatics (LIPIcs), vol. 122, pp. 47:1–47:17. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2018). https://doi.org/10.4230/LIPIcs.FSTTCS.2018.47
27. Place, T., Zeitoun, M.: The covering problem. Logical Methods in Computer Science **Volume 14, Issue 3** (Jul 2018). https://doi.org/10.23638/LMCS-14(3:1)2018
28. Place, T., Zeitoun, M.: On all things star-free. In: 46th International Colloquium on Automata, Languages, and Programming (ICALP 2019). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2019), `https://arxiv.org/abs/1904.11863v1`
29. Place, T., Zeitoun, M.: Separation for dot-depth two. Logical Methods in Computer Science **Volume 17, Issue 3** (Sep 2021). https://doi.org/10.46298/lmcs-17(3:24)2021
30. Rabin, M.O.: Decidability of second-order theories and automata on infinite trees. Trans. Amer. Math. Soc. **141**, 1–35 (1969)
31. Rispal, C.: Automates sur les ordres linéaires : Complémentation. Theses, Université de Marne la Vallée (Dec 2004), `https://tel.archives-ouvertes.fr/tel-00720658`
32. Rispal, C., Carton, O.: Complementation of Rational Sets on Countable Scattered Linear Orderings. International Journal of Foundations of Computer Science **16**(4), 767–786 (2005), `https://hal.archives-ouvertes.fr/hal-00160985`
33. Rosenstein, J.G.: Linear orderings. Academic press (1982)
34. Schützenberger, M.: On finite monoids having only trivial subgroups. Information and Control **8**(2), 190–194 (1965). https://doi.org/10.1016/S0019-9958(65)90108-7
35. Shelah, S.: The monadic theory of order. Ann. of Math. (2) **102**(3), 379–419 (1975)
36. Simon, I.: Piecewise testable events. In: Brakhage, H. (ed.) Automata Theory and Formal Languages. pp. 214–222. Springer Berlin Heidelberg, Berlin, Heidelberg (1975)
37. Wilke, T.: An algebraic theory for regular languages of finite and infinite words. International Journal of Algebra and Computation **03**(04), 447–489 (1993). https://doi.org/10.1142/S0218196793000287

38. Wilke, T.: Classifying discrete temporal properties. In: Meinel, C., Tison, S. (eds.) STACS 99. pp. 32–46. Springer Berlin Heidelberg, Berlin, Heidelberg (1999). https://doi.org/10.1007/3-540-49116-3_3
39. Wojciechowski, J.: Classes of transfinite sequences accepted by nondeterministic finite automata. Fundamenta informaticæ **7**(2), 191–223 (1984)
40. Wojciechowski, J.: Finite automata on transfinite sequences and regular expressions. Fundamenta informaticæ **8**(3-4), 379–396 (1985)