Conelikes and Ranker Comparisons

Viktor Henriksson[1] and Manfred Kufleitner[2]

[1] Loughborough University, Loughborough, UK
b.v.d.henriksson@lboro.ac.uk

[2] University of Stuttgart, Stuttgart, Germany
kufleitner@fmi.uni-stuttgart.de

For every fixed class of regular languages, there is a natural hierarchy of increasingly more general problems: Firstly, the membership problem asks whether a given language belongs to the fixed class of languages. Secondly, the separation problem asks for two given languages whether they can be separated by a language from the fixed class. And thirdly, the covering problem is a generalization of separation problem to more than two given languages. Most instances of such problems were solved by the connection of regular languages and finite monoids. Both the membership problem and the separation problem were also extended to ordered monoids. The computation of pointlikes can be interpreted as the algebraic counterpart of the separation problem. In this paper, we consider the extension of computation of pointlikes to ordered monoids. This leads to the notion of conelikes for the corresponding algebraic framework.

We apply this framework to the Trotter-Weil hierarchy and both the full and the half levels of the FO^2 quantifier alternation hierarchy. As a consequence, we solve the covering problem for the resulting subvarieties of **DA**. An important combinatorial tool are uniform ranker characterizations for all subvarieties under consideration; these characterizations stem from order comparisons of ranker positions.

1 Introduction

For a given variety of regular languages, there is a hierarchy of decision problems: First, we can ask whether a given regular language is in the variety; this is known as the *membership problem*. Very often, the membership problem is solved by giving an effective characterization. Famous solutions to the membership problem includes Simon's characterization of the piecewise testable languages in terms of $\mathcal{J}$-trivial monoids [22], and Schützenberger's characterization of the star-free languages by aperiodic monoids [20]. Inspired by these results, Eilenberg showed that there exists a one to one correspondence between varieties of regular languages and varieties of finite monoids [7]. This correspondence leads to an important approach for deciding the membership problem: one verifies some equivalent algebraic property of the syntactic monoid. The challenge here, however, is to identify the algebraic property and to prove its equivalence.

1

A more general problem is the *separation problem*. Given two languages, it asks whether there exists a language in the fixed variety which contains the first language and is disjoint with the second language. By applying the separation problem to a language and its complement, we obtain an answer to the membership problem. Thus, the separation problem is more general than the membership problem. Moreover, the separation problem can be used as a tool to solve the membership problem for varieties where this was not previously known; see e.g. [17]. A further generalization is given by the *covering problem* [18]. This problem considers a finite set of languages and a distinguished language, and asks how well the finite set of languages can be separated by a cover of the distinguished language.

As noted by Almeida, the separation problem for regular languages can also be solved via algebra by deciding so-called pointlikes [1]. The problem of deciding pointlikes is well studied, and there are effective characterizations for many varieties, e.g. aperiodics [9], $\mathcal{R}$-trivial monoids [3], $\mathcal{J}$-trivial monoids [4, 23] and finite groups [5].

A well studied fragment of first order logic is two-variable first-order logic FO^2. The languages definable in FO^2 form a variety, with the corresponding monoid variety $\mathbf{DA}$. In the study of FO^2 and $\mathbf{DA}$, two natural hierarchies have emerged: the Trotter-Weil hierarchy defined by a deep connection to the hierarchy of bands, and the quantifier alternation hierarchy. In stark contrast to the full FO quantifier alternation hierarchy, membership of the FO^2 quantifier alternation hierarchy is solved for all levels [14, 11, 8]. In particular, a tight connection between the Trotter-Weil and the quantifier alternation hierarchy has appeared; Weil and the second author showed that the join levels of the quantifier alternation hierarchy (i.e., the FO^2_m levels) correspond to the intersection levels of the Trotter-Weil hierarchy [14], and combining two results from [13] and [8] shows that the join levels of the Trotter-Weil hierarchy correspond to the intersection levels of the quantifier alternation hierarchy.

Rankers have emerged as an important tool in the study of FO^2. These were first introduced by Schwentick, Thérien, and Vollmer by the name of turtle programs [21]. Using comparisons of restricted sets of rankers, Weis and Immerman gave a combinatorial characterization of the full levels of the quantifier alternation hierarchy [24]. This approach was extended to the half-levels of the quantifier alternation hierarchy in the PhD-thesis of Lauser [16]. The corners of the Trotter-Weil hierarchy also admit ranker characterizations using the concept of condensed rankers [15].

This article solves the covering problem (and thus also separation problem) for all levels of the Trotter-Weil hierarchy and quantifier alternation hierarchy inside FO^2. For this, we rely on two main tools, *conelikes* and *ranker comparisons*. Conelikes are introduced in Section 3. They extend pointlikes to ordered monoids, and are algebraic versions of the imprints used by Place and Zeitoun; see e.g. [18]. Thus, they have a strong connection to the covering problem; an algorithm for computing the conelikes with respect to a monoid variety can be used to solve the covering problem for the corresponding language variety and vice versa.

Sections 4 and 5 deals with ranker comparisons. In Section 4, we give a framework for ranker comparisons using general sets of ranker pairs. We show that any set of pairs of rankers which is closed under ranker subwords gives rise to a stable relation and thus defines a monoid.

In Section 5, we use this framework to give uniform characterizations for all levels of the Trotter-Weil and quantifier alternation hierarchy. In particular, we give a characterization of the corners of the Trotter-Weil hierarchy in terms of ranker comparisons. Together, these sets of ranker comparisons form a natural hierarchy, the *ranker comparison hierarchy* which encompasses both the quantifier alternation hierarchy and the Trotter-Weil hierarchy.

The rest of the article is devoted to the solution of the covering problem. In Section 6, we

present sets of subsets of a monoid which can be computed effectively. Our main theorem states that these sets coincide with the conelikes (or the pointlikes for the unordered varieties). Our main theorem also provides optimal separators: relational morphisms such that the conelikes with respect to these morphisms are the same as the conelikes with respect to the corresponding variety. The co-domains in these morphisms are defined using ranker comparisons.

After handling two special cases in Section 7, we use Section 8 to show that being in the introduced sets implies being conelike with respect to the relevant variety. The more difficult direction is showing that every conelike lies in the introduced sets. In Section 9, we show how languages defined by ranker comparisons can be factored in a convenient way, and in Section 10, we use these factorizations to complete the circle. We show that given a monoid defined by the proper set of ranker comparisons, the conelikes with respect to this monoid are all in the corresponding set. In particular, this means that every conelike with respect to the corresponding monoid variety is in the set.

2 Preliminaries

2.1 Words and Monoids

If M is a semigroup and $e \in M$ satisfies $ee = e$, then e is *idempotent*. Given a semigroup M there exists a (smallest) number ω_M such that u^{ω_M} is idempotent for each $u \in M$. If M is clear from context, we write ω for this number. For sets $S, T \subseteq M$, we have

$$ST = \{st \in M \mid s \in S, t \in T\}.$$

Note that 2^M is a monoid under this operation.

An important tool in semigroup theory are the *Green's relations*, out of which we introduce the following three. Given a monoid M and $s, t \in M$, we define

- $s \leq_{\mathcal{R}} t$ if $sM \subseteq tM$,

- $s \leq_{\mathcal{L}} t$ if $Ms \subseteq Mt$,

- $s \leq_{\mathcal{J}} t$ if $MsM \subseteq MtM$.

We define $s \mathrel{\mathcal{R}} t$ if $s \leq_{\mathcal{R}} t$ and $t \leq_{\mathcal{R}} s$ and we define $s \mathrel{\mathcal{L}} t$ and $s \mathrel{\mathcal{J}} t$ correspondingly. We say that $s <_{\mathcal{R}} t$ if $s \leq_{\mathcal{R}} t$ but not $s \mathrel{\mathcal{R}} t$ and equivalently for $\mathcal{L}$ and $\mathcal{J}$. Let $u \in A^*$ and $\mu : A^* \to M$. Then there is a unique factorization $u = u_1 a_1 \ldots u_{n-1} a_{n-1} u_n$ such that $\mu(u_1 a_1 \ldots a_i) \mathrel{\mathcal{R}} \mu(u_1 a_1 \ldots a_i u_{i+1}) <_{\mathcal{R}} \mu(u_1 a_1 \ldots a_i u_{i+1} a_{i+1})$. This is the $\mathcal{R}$*-factorization of u with respect to μ*. The $\mathcal{L}$*-factorization of u with respect to μ* is defined symmetrically.

Given a monoid M with a binary relation $\preceq$, we say that $\preceq$ is *stable* if for all $s, t, x, y \in M$, $s \preceq t$ implies $xsy \preceq xty$. We say that a monoid is *ordered* if it is equipped with a stable order. A *congruence* is a stable equivalence relation. In particular, any stable preorder $\preceq$ induces the congruence given by $s \sim t$ if and only if $s \preceq t$ and $t \preceq s$. If M is ordered, and $s \in M$, then $\uparrow s = \{t \in M \mid s \leq t\}$. If M is a monoid, and $\preceq$ is a stable preorder, then $M/\preceq$ is the monoid whose elements are the equivalence classes of the induced congruence, the multiplication is that induced by the multiplication in M and where, for $s, t \in M$ with $[s], [t]$ the corresponding equivalence classes, we have $[s] \leq [t]$ if and only if $s \preceq t$. Given a language L, the *syntactic preorder* is the relation $u \leq_L v$ if and only if $xuy \in L \Rightarrow xvy \in L$ for all $x, y \in A^*$. Let

3

$\mu : A^* \to A^*/\leq_L$ be the canonical projection, then π is the *syntactic morphism* and $A^*/\leq_L$ the *syntactic monoid* of L. A language is *regular* if and only if the syntactic monoid is finite.

Let M and N be (possibly ordered) monoids. A *relational morphism* is a relation $\tau : M \to N$ (or mapping $M \to 2^N$) which satisfies

(i) $1_N \in \tau(1_M)$,

(ii) for all $s \in M$, $\tau(s) \neq \emptyset$,

(iii) for all $s, t \in M$, $\tau(s)\tau(t) \subseteq \tau(st)$.

If there is a relational morphism $\tau : M \to N$ such that $\tau(s) \cap \tau(s') \neq \emptyset$ implies $s = s'$ we say that M *divides* N. A *division of ordered monoids* is a division where we also assume $t \leq t'$ for some $t \in \tau(s)$, $t' \in \tau(s')$ implies $s \leq s'$. A *variety of monoids* is a collection of monoids closed under division and finite direct products. A collection of ordered monoids is a *positive variety* if it is closed under finite direct products and division of ordered monoids.

For a relational morphism $\tau : M \to N$, a set $S \subseteq M$ such that $\bigcap_{s \in S} \tau(s) \neq \emptyset$ is *pointlike* with respect to τ. If $t \in \bigcap_{s \in S} \tau(s)$, then t is a *witness* of S being pointlike. If $\mathbf{V}$ is a variety and S is pointlike for every relational morphism $\tau : M \to N \in \mathbf{V}$, then S is pointlike with respect to $\mathbf{V}$. The set of all pointlikes in M with respect to τ is $\mathsf{PL}_\tau(M)$, and the set of all pointlikes with respect to $\mathbf{V}$ is $\mathsf{PL}_{\mathbf{V}}(M)$.

A useful way to define varieties is through the use of *ω-identities* and *ω-relations*. An *ω-term* is either x where x is taken from some (usually infinite) set of variables X, or tt' or t^ω where t and t' are ω-terms. An ω-identity is given by $t = t'$ or $t \leq t'$ where t and t' are ω-terms. Given a monoid M, an *interpretation* of ω-terms is any extension of a map $\chi : X \to M$ for which $\chi(tt') = \chi(t)\chi(t')$ and $\chi(t^\omega) = \chi(t)^{\omega_M}$. We say that a monoid M *satisfy* an ω-identity $t = t'$ if $\chi(t) = \chi(t')$ for all interpretations χ. It similarly satisfies $t \leq t'$ if $\chi(t) \leq \chi(t')$ for all interpretations. If $R_1, \ldots, R_n$ are ω-identities or -relations, then $[\![R_1, \ldots, R_n]\!]$ denotes the collection of all monoids which satisfy all R_i. Some varieties that are of importance in this text are

- $\mathbf{DA} = [\![(xzy)^\omega = (xzy)^\omega z(xzy)^\omega]\!]$.

- $\mathbf{J} = [\![(st)^\omega s(xy)^\omega = (st)^\omega y(xy)^\omega]\!]$, or equivalently all monoids whose $\mathcal{J}$-classes are trivial.

- $\mathbf{J}_1 = [\![x^2 = x, xy = yx]\!]$,

- $\mathbf{J}^+ = [\![1 \leq z]\!]$.

Let A be a collection of symbols, called an *alphabet*. We denote by A^* the set of concatenations of symbols in A. In other words, A^* is the *free monoid* of A. An element $u \in A^*$ is a *word* and a subset $L \subseteq A^*$ a *language*. We will denote the empty word by ε. A *(scattered) subword* of u is a word $v = a_1 \ldots a_n$ such that $u = u_1 a_1 \ldots u_n a_n u_{n+1}$ for some (possibly empty) words u_i. Let $u = u_1 u_2 u_3$ for some (possibly empty) words u_1, u_2, u_3. Then u_1 is a *prefix* and u_2 is a *factor* of u.

If $u = u_1 \ldots u_n$ where each u_i is a word, then $u_1 \ldots u_n$ is a *factorization* of u. This concept extends to subsets of A^*; if $U \subseteq A^*$, a factorization of U is $U_1 \ldots U_n$ where each $u \in U$ can be factored as $u = u_1 \ldots u_n$ in such a way that $u_i \in U_i$. This definition does not coincide with the monoid operation on subsets given above.[1] To resolve this ambiguity, we always consider concatenation to mean factorization when dealing with A^*, unless otherwise specified.

---

[1] Indeed, $\{aa, bb\}$ can be factored as $\{a, b\}\{a, b\}$, but $ab, ba \in \{a, b\}\{a, b\}$ if seen as a multiplication.

For an alphabet A, let J_A denote the monoid whose elements are subsets of A and whose operation is the union operation. This monoid has a natural ordering defined by $U \leq V$ if $U \subseteq V$ for $U, V \subseteq A$. Let $\mathsf{alph}_A : A^* \to J_A$ be the extension of $\mathsf{alph}_A(a) = \{a\}$ for each $a \in A$. We will drop the subscript when A is clear from context. Given a surjective homomorphism $\mu : A^* \to M$, a morphism $\alpha : M \to J_A$ is called a *content morphism* if $\mathsf{alph}_A = \alpha \circ \mu$.

Let τ be a relation, and let $G = \{(s,t) \in M \times N \mid t \in \tau(s)\}$ be its graph. Then τ is a relational morphism if and only if G is a submonoid of $M \times N$ and the projection on M is surjective. Any relational morphism can be thought of as pulling the elements of M back to some free monoid A^* where A generates both M and N, and then pushing the words to their corresponding elements in N. More formally, we have the following lemma, which follows easily by choosing A to be a generating set of G.

Lemma 1. *For any relational morphism, there exists A and $\mu : A^* \to M$, $\nu : A^* \to N$ such that μ is surjective and $\tau = \nu \circ \mu^{-1}$, as in the following diagram:*

$$
\begin{array}{ccc}
 & A^* & \\
 \swarrow{\scriptstyle \mu} & & \searrow{\scriptstyle \nu} \\
 M & \xrightarrow{\quad \tau \quad} & N.
\end{array}
$$

Since any such diagram also gives rise to a relational morphism, this means that the relational morphisms between M and N are exactly given by diagrams of this form.

2.2 Logic

We will consider FO[<], first order logic using the following syntax:

$$\varphi ::= \top \mid \bot \mid a(x) \mid x = y \mid x < y \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists x \varphi.$$

Here $a \in A$ for some fixed alphabet A, and $\varphi \in$ FO[<]. We interpret formulae in FO[<] over words as follows. If $i, j \in \mathbb{N}$, then $u, i, j \vDash x < y$ if and only if $i < j$, and $u, i \vDash a(x)$ if and only if $u[i] = a$. The logical connectives and existential quantifier are interpreted as usual. We use the macro $x \leq y$ to mean $x < y \vee x = y$ and the macro $\forall x \varphi$ to mean $\neg \exists x \neg \varphi$. If φ is a formula without free variables over the alphabet A, we define $L(\varphi) = \{u \in A^* \mid u \vDash \varphi\}$. If $\mathscr{F}$ is a collection of formulae, we say that $L \subseteq A^*$ is *definable in $\mathscr{F}$* if there exists $\varphi \in \mathscr{F}$ such that $L = L(\varphi)$.

In particular, we are interested in $\text{FO}^2[<]$, i.e. the fragment of FO[<] where we only allow the use (and reuse) of two variable names. Thus

$$\exists x : a(x) \wedge (\exists y : y > x \wedge b(y) \wedge (\exists x : x > y \wedge c(x)))$$

is allowed in $\text{FO}^2[<]$ whereas

$$\exists x : a(x) \wedge (\exists y : y > x \wedge b(y) \wedge (\exists z : z > x \wedge y > z \wedge c(z)))$$

is not. It is well known that $\text{FO}^2[<]$ is a proper fragment of FO[<]. We will primarily be interested in some fragments of $\text{FO}^2[<]$. Consider the syntax

$$\varphi_0 ::= \top \mid \bot \mid a(x) \mid x = y \mid x < y \mid \neg\varphi_0 \mid \varphi_0 \vee \varphi_0 \mid \varphi_0 \wedge \varphi_0$$

$$\varphi_m ::= \varphi_{m-1} \mid \neg\varphi_{m-1} \mid \varphi_m \vee \varphi_m \mid \varphi_m \wedge \varphi_m \mid \exists x \varphi_m$$

The collection of formulae $\varphi_m[<]$ is denoted by $\Sigma^2_m[<]$, the collection of negations of formulae in $\Sigma^2_m[<]$ is $\Pi^2_m[<]$ and the Boolean closure of $\Sigma^2_m[<]$ is $\text{FO}^2_m[<]$. In what follows, we drop the reference to the predicate symbol <, and assume it to be understood from context.

2.3 Ramsey Numbers

A *graph* is a pair $\mathcal{G} = (V, E)$ where V is a set of *vertices* and $E \subseteq \left\{ S \subseteq 2^V \mid |S| = 2 \right\}$ is a set of *edges*. An *edge-coloring* is a map $c : E \to C$ where C is some set of colors. A graph is *complete* if $E = \left\{ S \subseteq 2^V \mid |S| = 2 \right\}$, i.e. if there is an edge between any two elements. A set $F \subseteq E$ of edges is *monochrome* if $c(e) = c(e')$ for all $e \in F$. A *triangle* is a set of three distinct edges $e_1, e_2, e_3 \in E$ where $e_i \cap e_j \neq \emptyset$ for $1 \leq i, j \leq 3$. The following theorem is a special case of Ramsey's Theorem [19].

Theorem 1. *Let C be a finite set of colours. Then there exists a number R, called the* Ramsey number *of C such that any complete graph $\mathcal{G} = (V, E)$ with $R \leq |V|$ contains a monochrome triangle.*

2.4 Hierarchies Inside DA

A variety of special importance for this article is **DA**. This monoid variety has a natural correspondence to FO^2 since a language is definable in the latter if and only if its syntactic monoid is in **DA**.

We are interested in hierarchies of subvarieties of **DA**. One important such hierarchy is the *Trotter-Weil hierarchy*. Its original motivation comes from an intimate relation with the hierarchy of bands, but here we give a more explicit definition.

Definition 1. Let M be a monoid, and let $s, t \in M$. Then

- $s \sim_{\mathbf{K}} t$ if for all idempotents $e \in M$, either $ev, eu <_{\mathcal{J}} e$ or $ev = eu$,

- $s \sim_{\mathbf{D}} t$ if for all idempotents $f \in M$, either $vf, uf <_{\mathcal{J}} f$ or $vf = uf$.

The join of these relations is $\sim_{\mathbf{KD}}$.

It is straight-forward to check that these relations are congruences (see e.g. [12]). Let $\mathbf{R}_1 = \mathbf{L}_1 = \mathbf{J}_1$, and let $M \in \mathbf{R}_m$ if $M/\sim_{\mathbf{K}} \in \mathbf{L}_{m-1}$ and $M \in \mathbf{L}_m$ if $M/\sim_{\mathbf{D}} \in \mathbf{R}_{m-1}$. When defining $\mathbf{R}_m$ and $\mathbf{L}_m$ for $m \geq 2$, starting with $\mathbf{J}_1$ yields the same result as starting with $\mathbf{J}$. For our purposes, starting with $\mathbf{J}_1$ is more natural.

The varieties $\mathbf{R}_m$ and $\mathbf{L}_m$ are all contained in **DA**. Together with their joins and intersections they make up the Trotter-Weil hierarchy shown in Figure 1 on page 13.

There is an intimate connection between the quantifier alternation hierarchy, also shown in Figure 1, and the Trotter-Weil hierarchy. Indeed, it was shown by Weil and the second author that the languages definable in FO^2_m are exactly those whose syntactic monoid is in $\mathbf{R}_{m+1} \cap \mathbf{L}_{m+1}$ [14]. Furthermore, combining the results in [13] and [8] gives the following proposition.

Proposition 1. *A language is definable in both Σ^2_m and Π^2_m if and only if its syntactic monoid is in $\mathbf{R}_m \vee \mathbf{L}_m$.*

The *corners* of the quantifier alternation hierarchy, Σ^2_m and Π^2_m, also have algebraic characterizations, given by Fleischer, Kufleitner and Lauser [8]. We will define these recognizing varieties using the stable relation $\preceq_{\mathbf{KD}}$ introduced by the authors [10].

Definition 2. Let M be a monoid, and let $s, t \in M$. We say that $s \preceq_{\mathbf{KD}} t$ if for all $x, y \in M$, the following holds:

(i) If $x \mathcal{R} xty$, then $x \mathcal{R} xsy$,

(ii) If $xty \mathcal{L} y$, then $xsy \mathcal{L} y$,

(iii) If $x \mathcal{R} xt$ and $ty \mathcal{L} y$, then $xsy \leq xty$.

If $u \leq_{\mathbf{KD}} v$ and $v \leq_{\mathbf{KD}} u$, we say that $u \equiv_{\mathbf{KD}} v$.[2]

Let $\mathbf{Si}_1 = \mathbf{J}^+ = [\![1 \leq z]\!]$ and let $M \in \mathbf{Si}_m$ if $M/{\preceq_{\mathbf{KD}}} \in \mathbf{Si}_{m-1}$. For every m, the collection $\mathbf{Si}_m$ is a positive variety. A language is definable in Σ^2_m if and only if its syntactic monoid is in $\mathbf{Si}_m$. We say that an ordered monoid M is in $\mathbf{Pi}_m$ if and only if M with the order reversed, is in $\mathbf{Si}_m$. It is clear that a language is definable in Π^2_m if and only if its syntactic monoid is in $\mathbf{Pi}_m$.

The following property of $\mathbf{DA}$ is standard (see e.g. [6]), and is used throughout the article.

Lemma 2. *Let $M \in \mathbf{DA}$ and suppose that M has a content morphism α. Let $s, t \in M$ and suppose that $\alpha(s) \leq \alpha(t)$, then $s^\omega t s^\omega = s^\omega$.*

3 Conelikes and the Covering Problem

In this section, we introduce the main problems of the article, the *separation problem* and the *covering problem*. Given a variety $\mathcal{V}$, the (asymmetric) separation problem is defined as follows:

> Given $L, L' \subseteq A^*$, determine if there is a language $K \in \mathcal{V}$ such that $L \subseteq K$ and $L' \cap K = \emptyset$.

If there exists such a K, we say that L is $\mathcal{V}$-separable from L'. The symmetric separation problem is to determine whether both L is $\mathcal{V}$-separable from L' and L' is $\mathcal{V}$-separable from L. If $\mathcal{V}$ is a full variety, i.e. closed under complements, these two problems are equivalent (just choose $A^* \setminus K$ to separate L' from L).

There is a strong connection between the (symmetric) separation and the problem of deciding pointlikes [1]. In this section, we introduce the more general covering problem, together with a generalization of pointlikes which works well with the asymmetric setting. This generalization, which we call *conelikes*, is folklore. However, to the knowledge of the authors they have not been made precise in the algebraic setting.[3]

Let $\mathbf{K}$ be a set of languages, and $\mathbf{L}$ a finite set of languages. Then $\mathbf{K}$ is *separating* for $\mathbf{L}$ if for all $K \in \mathbf{K}$, there exists $L' \in \mathbf{L}$ such that $K \cap L' = \emptyset$. We only consider situations when $\mathbf{K}$ is a *cover* of some language L, i.e. such that $L \subseteq \bigcup \mathbf{K}$.

Definition 3. Let $\mathcal{V}$ be a (positive) variety. The *covering problem* for $\mathcal{V}$ is defined as follows:

> Given $L \subseteq A^*$ and $\mathbf{L} \subseteq 2^{A^*}$ where $\mathbf{L}$ is finite, determine if there is $\mathbf{K} \subseteq \mathcal{V}$ which covers L and is separating for $\mathbf{L}$.

---

[2] The name $\preceq_{\mathbf{KD}}$ was originally inspired by the relation $\sim_{\mathbf{KD}}$ since they share some properties. However, it should be noted that the relation $\equiv_{\mathbf{KD}}$ is not the same as $\sim_{\mathbf{KD}}$. As a counter example, note the syntactic monoid of $a^+b^+cA^*da^+b^+$ where the equivalence class of ab is $\equiv_{\mathbf{KD}}$-related to all elements in the minimal $\mathcal{J}$-class, whereas it is not $\sim_{\mathbf{K}}$- or $\sim_{\mathbf{D}}$-related to anything.

[3] The imprints used by Place and Zeitoun in [18] yield is a corresponding object in the language setting

If such a **K** exists, we say that $(L, \mathbf{L})$ is $\mathcal{V}$-coverable. For $\mathbf{L} = \{L'\}$, this reduces to the separation problem. Whenever $\mathcal{V}$ is a full variety, it is equivalent to answer the covering problem for $(L, \mathbf{L})$ and $(A^*, \{L\} \cup \mathbf{L})$ [18], and for regular languages this is in turn equivalent to computing the **V**-pointlikes of a finite monoid recognizing all languages of **L** [1].

Example 1. Consider the variety $\mathcal{J}^+$ of languages whose syntactic monoid is in $\mathbf{J}^+$, and let $L = (ab)^+$, $L_1 = b(ab)^*$, $L_2 = (ab)^*a$. We note that in order to cover L, we need the language $A^*aA^*bA^*$. However, this language also contains words from L_1 and L_2, showing that $(L, \{L_1, L_2\})$ is not $\mathcal{J}^+$-coverable.

Example 2. Suppose that $L \in \mathcal{V}$ can be written as a disjoint union of a finite number of sets $\mathbf{K} \subseteq \mathcal{V}$. Suppose that **K** has at least two elements. Then **K** is separating for itself (this is the case for all sets with magnitude greater than one), and thus $(L, \mathbf{K})$ is $\mathcal{V}$-coverable. However, it is clear that L is not $\mathcal{V}$-separable from any $L_i \in \mathbf{K}$.

We want to use algebraic methods to solve the covering problem. However, the covering problem is relevant for positive varieties, whereas pointlikes do not take orders into account. This motivates the following generalization of pointlikes.

Definition 4. Let M be a monoid, and let $\tau : M \to N$ be a relational morphism. For $s \in M$, $S \subseteq M$, we say that (s, S) is *conelike* with respect to τ if there exists an element $x \in \tau(s)$ such that $S \subseteq \tau^{-1}(\uparrow x)$. We call x a *witness* of (s, S) being conelike. As with pointlikes we say that a pair (s, S) is conelike with respect to a variety **V** if it is conelike for any $\tau : M \to N \in \mathbf{V}$. We denote by $\mathsf{Cone}_\tau(M)$ the conelikes of M with respect to τ, and by $\mathsf{Cone}_\mathbf{V}(M)$ the conelikes of M with respect to **V**.

Note that if N is unordered, we can define an order $u \leq v$ if and only if $u = v$. In this case, a pair (s, S) is conelike if and only if S is pointlike and $s \in S$. In particular, this means that for non-positive varieties, calculating the pointlikes and the conelikes is the same problem.

The concept of pointlikes is in general not expressive enough to solve the covering problem. However, it is still possible to define pointlikes for a variety of ordered monoids, and such pointlikes are used throughout the article. Note that if $S \subseteq M$ is pointlike with respect to some variety of ordered monoids, then (s, S) is conelike for any $s \in S$.

For regular languages, which is what is considered in this contribution, the problem of finding pointlikes is equivalent to the covering problem. We require the following two lemmas. The first is standard when dealing with pointlikes (see e.g. [1]) and shows that when finding the optimal conelikes, we need only consider relational morphisms through a fixed alphabet A.

Lemma 3. *Let M be a monoid, and $\mu : A^* \to M$ a surjective morphism. Let **V** be a variety. Then $(s, S) \in M \times 2^M$ is conelike with respect to **V** if and only if (s, S) is conelike with respect to all $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N \in \mathbf{V}$ where ν is a homomorphism.*

Proof. It is clear that in order to be conelike with respect to **V**, it must in particular be conelike with respect to the morphisms factoring through A^*. For the other direction, suppose (s, S) is conelike with respect to all relational morphisms on the desired form, and let $\rho : M \to N \in \mathbf{V}$ be arbitrary. By Lemma 1 we can factor this as $\rho : M \xleftarrow{f} B^* \xrightarrow{g} N$ where f and g are homomorphisms, and μ' is surjective.

We define the homomorphism $h : A^* \to B^*$ as follows. For each $a \in A$, choose $u \in B^*$ such that $f(u) = \mu(a)$ and extend this to a homomorphism. The composition $\varphi^{-1} : g \circ h \circ h^{-1} \circ f^{-1}$ is

a relational morphism and factors through A^* in the desired way. Thus, there exists $x \in N$ such that $x \in \varphi(s)$ and $S \subseteq \varphi^{-1}(\uparrow x)$. This in particular implies that $x \in \rho(s)$ and $S \subseteq \rho^{-1}(\uparrow x)$, showing that (s, S) is conelike with respect to ρ. Since ρ was arbitrary with co-domain in $\mathbf{V}$, the result follows. $\square$

The next lemma shows that for all varieties $\mathbf{V}$ and monoids M, there is a relational morphism $\tau : M \to N \in \mathbf{V}$ which is *optimal* for determining conelikes. Furthermore, this relational morphism can also be assumed to factor through A^*.

Lemma 4. *Let $\mathbf{V}$ be a variety, and let M be a finite monoid. Then there exists $\tau : MN \in \mathbf{V}\mu\nu$ such that $\mathsf{Cone}_{\mathbf{V}}(M) = \mathsf{Cone}_{\tau}(M)$. Furthermore, if $\mu : A^* \to M$ is a surjective homomorphism, we can assume that τ can be written as $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N$ where μ and ν are homomorphisms and μ is surjective.*

Proof. Let $\tau_1 : M \to N_1$ be arbitrary. We will define a chain $\tau_1, \ldots, \tau_n$ where $\mathsf{Cone}_{\tau_{i+1}}(M) \subsetneq \mathsf{Cone}_{\tau_i}(M)$, using the following procedure. If $\mathsf{Cone}_{\tau_i}(M) = \mathsf{Cone}_{\mathbf{V}}(M)$, we set $i = n$ and are done. If not, there exists $\tau' : M \to N' \in \mathbf{V}$ such that $\mathsf{Cone}_{\tau_i}(M) \not\subseteq \mathsf{Cone}_{\tau'}(M)$. We define $\tau_{i+1} : M \to N_i \times N'$ by $\tau_{i+1}(s) = \{(t, t')\}_{t \in \tau_i(s), t' \in \tau'(s)}$. It is clear that this is a relational morphism, and that $(s, S) \in \mathsf{Cone}_{\tau_{i+1}}(M)$ if and only if $(s, S) \in \mathsf{Cone}_{\tau_i}(M) \cap \mathsf{Cone}_{\tau'}(M)$. By choice of τ', this shows that τ_{i+1} has the desired properties. Since M is finite, this process must eventually stop giving the desired result.

For the final part of the lemma, suppose that τ_n has the desired properties. Using the same argument as in Lemma 3, we get a relational morphism $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N$ such that every conelike with respect to τ is also conelike with respect to τ_n and thus also with respect to $\mathbf{V}$. $\square$

Proposition 2. *Let $\mathbf{L} = \{L_i\}$ be a finite collection of regular languages, and for each L_i let $\mu_i : A^* \to M_i$ be its syntactic monoid. Let $\mu : A^* \to M_1 \times \cdots \times M_n$ be defined by $\mu(a) = (\mu_1(a), \ldots, \mu_n(a))$ and let $M = \mu(A^*)$. Let $\mathbf{V}$ be a variety of monoids recognizing a variety $\mathcal{V}$ of languages. Then the following are equivalent*

*(i) $(L, \mathbf{L})$ is **not** $\mathcal{V}$-coverable,*

(ii) there exists a conelike (s, S) with respect to $\mathbf{V}$ such that $L \cap \mu^{-1}(s) \neq \emptyset$ and for all $L' \in S$ there exists $s' \in S$ such that $L' \cap \mu^{-1}(s') \neq \emptyset$.

Proof. Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N$ be a relational morphism with $N \in \mathbf{V}$ chosen to have the property in Lemma 4. Let

$$\mathbf{P} = \left\{ (L, \nu^{-1}(\uparrow x)) \mid x \in \nu(L) \right\}.$$

It is clear that the set $\bigcup_{x \in \nu(L)} \nu^{-1}(\uparrow x)$ is a cover of L. If $(L, \mathbf{L})$ is not $\mathcal{V}$-coverable, then there exists $x \in \nu(L)$ such that for all $L' \in \mathbf{L}$, we have $L' \cap \nu^{-1}(x) \neq \emptyset$. Choose $s \in \tau^{-1}(x) \cap \mu(L)$ arbitrary; such s exist since we only consider x in the image of L. Then $(s, \tau^{-1}(\uparrow x))$ is a conelike with property *(ii)*.

For the other direction, let $\mathbf{K} \subseteq \mathcal{V}$ cover L and be separating for $\mathbf{L}$ and let $\nu : A^* \to N$ recognize all $K_i \in \mathbf{K}$. Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N$ be the natural relational morphism. For contradiction, assume that there is a conelike with property *(ii)*. Then we find $u \in L$ and $v_i \in L_i$ for each $L_i \in U$, such that $(\mu(u), \{\mu(v_i)\}_i)$ is conelike with respect to τ. In particular, we can assume the u and v_i to be the representatives for which $\nu(v_i) \in \uparrow \nu(u)$. Now, $\uparrow \nu(u)$ intersects every language in $\mathbf{L}$, and thus $\mathbf{K}$ is not separating for $\mathbf{L}$. $\square$

Example 3. Let us return to our example with the languages $(ab)^+$, $b(ab)^*$ and $(ab)^*a$. These are all recognized by the syntactic monoid of $(ab)^+$. The above proposition tells us that $(ab, \{b, a\})$ is conelike, where we identify the languages with representatives of their equivalence classes. In particular, an optimal relational morphism is the natural morphism into the syntactic monoid of $A^*aA^*bA^*$.

Before leaving the topic of conelikes, we introduce a common tool for determining pointlikes and solving the covering problem (see e.g. [18]). The idea is to construct sets of subsets of M which have closure properties analogous to those of pointlikes and conelikes.

Definition 5. Given a monoid M, a subset of 2^M is *closed* if it contains the singletons and is closed under multiplication and subsets. Similarly, a set $\mathcal{C} \subseteq M \times 2^M$ is closed if it has the following closure properties:

- $(s, \{s\}) \in \mathcal{C}$ for all elements $s \in M$,

- $(s, S), (t, T) \in \mathcal{C}$ implies $(st, ST) \in \mathcal{C}$,

- $(s, S) \in \mathcal{C}$ implies $(s, S') \in \mathcal{C}$ for all subsets $S' \subseteq S$.

In Sections 6 to 10, we introduce computable closed sets for all varieties of interest and show that these sets coincide with the pointlikes and conelikes.

4 A Framework for Ranker Comparisons

One of the main techniques in this paper is ranker comparisons. This concept has close connections to fragments of FO^2, a connection we explore in Section 5 (see. [24, 16]). In this section, we introduce a general framework and give sufficient conditions for instances of this framework to define a monoid.

Definition 6. Let A be some alphabet. A *ranker* over A is a nonempty word over $\{\mathsf{X}_a, \mathsf{Y}_a\}_{a \in A}$, which can be interpreted as a partial function from A^* to $\mathbb{N}$. The interpretation is defined inductively as follows:

- $\mathsf{X}_a(u) = \inf \{n \in \mathbb{N} \mid u[n] = a\}$ if the infimum is finite, and undefined otherwise,

- $\mathsf{Y}_a(u) = \sup \{n \in \mathbb{N} \mid u[n] = a\}$ if the supremum is finite, and undefined otherwise,

- $r\mathsf{X}_a(u) = \inf \{n \in \mathbb{N} \mid n > r(u), u[n] = a\}$ if $r(u)$ is defined and the infimum is finite, and undefined otherwise,

- $r\mathsf{Y}_a(u) = \sup \{n \in \mathbb{N} \mid n < r(u), u[n] = a\}$ if $r(u)$ is defined and the supremum is finite, and undefined otherwise.

Note that we read rankers from left to right (as opposed to function composition). Thus $\mathsf{X}_a\mathsf{Y}_b(bab) = 1$, whereas $\mathsf{Y}_b\mathsf{X}_a(bab)$ is undefined. If $p = a_1 \cdots a_n$, we define $\mathsf{X}_p = \mathsf{X}_{a_1} \cdots \mathsf{X}_{a_n}$ and $\mathsf{Y}_p = \mathsf{Y}_{a_1} \cdots \mathsf{Y}_{a_n}$ for compactness.

We define the following collections of rankers:

$$R_1^X = \{X_a\}_{a \in A}^+, \qquad R_1^Y = \{Y_a\}_{a \in A}^+, \qquad R_{m+1}^X = \{X_a\}_{a \in A}^* R_m^Y, \qquad R_{m+1}^Y = \{Y_a\}_{a \in A}^* R_m^X,$$

10

Here the juxtaposition on the left denotes concatenation. We furthermore define $R_m = R_m^X \cup R_m^Y$, and $R = \bigcup_m R_m$. Note that these sets depend on the alphabet A although this dependence is not written out explicitly. We will always assume that the alphabet is clear from context. Given a ranker r, the *alternation depth* of r is the smallest m such that $r \in R_m$. The *depth* of r is the length of r as a word.

Since rankers are $\mathbb{N}$-valued functions, there is a natural way of comparing them given a speficied word u. In other words, given $u \in A^*$ and rankers r, s, we are interested in whether $r(u) \leq s(u)$ and $r(u) < s(u)$ hold. The following definition introduces a general framework, inspired by the comparisons of Weis and Immerman [24] and Lauser [16].

Definition 7. Let A be some alphabet, and let $\mathscr{C} \subseteq R \times R$ be some set of pairs of rankers over A. We define $[\mathscr{C}] = \bigcup_{(r,s) \in \mathscr{C}} \{r, s\}$, i.e. the set of rankers that occurs on some position in some pair of $\mathscr{C}$. We say that $u \leq^{\mathscr{C}} v$ if:

(i) The same set of rankers in $[\mathscr{C}]$ are defined on u and v,

(ii) For each $(r, s) \in \mathscr{R}$ such that r and s are defined on u and v, we have $r(u) \leq s(u) \Rightarrow r(v) \leq s(v)$, and $r(u) < s(u) \Rightarrow r(v) < s(v)$.

If $u \leq^{\mathscr{C}} v$ and $v \leq^{\mathscr{C}} u$, we say that $u \equiv^{\mathscr{C}} v$.

For rankers r, s, and words $u, v \in A^*$ we have $r(u) \leq s(u) \Rightarrow r(v) \leq s(v)$ if and only if $s(v) < r(v) \Rightarrow s(u) < r(u)$. In particular, this means that if $\mathscr{C}$ is symmetric, i.e. $(r, s) \in \mathscr{C} \Leftrightarrow (s, r) \in \mathscr{C}$, then $\leq^{\mathscr{C}}$ and $\equiv^{\mathscr{C}}$ are equivalent relations. Note that for a certain choice of symmetric $\mathscr{C}$ we get the relation introduced in [24]. We say that a language L is *definable* by $\mathscr{C}$ if L is an ideal under the relation $\leq^{\mathscr{C}}$. Furthermore, we say that a language is an $\mathscr{C}$-*set* if it is a subset of such an ideal.

For a language A and some sets of rankers $\mathscr{C} \subseteq R \times R$, we want to consider the monoid $A^*/\leq^{\mathscr{C}}$, which is a well defined monoid only when $\leq^{\mathscr{C}}$ is stable. For general $\mathscr{C}$, this is not the case. However, Proposition 3 provides a large class of sets for which it does hold.[4]

Proposition 3. *Let R be some collection of rankers and let $\mathscr{C} \subseteq R \times R$ be closed under subwords, i.e. be such that $(r, s) \in \mathscr{C}$ implies $(r', s') \in \mathscr{C}$ for any subwords r' of r and s' of s. Then the preorder $\leq^{\mathscr{C}}$ is stable.*

Proof. By symmetry, it is enough to show that $u \leq^{\mathscr{C}} v$ implies $xu \leq^{\mathscr{C}} xv$. Suppose $r \in [\mathscr{C}]$ is defined on xu. Without loss of generality, we assume that r starts with an X-modality.

We will factor $r = s_1 t_1 \ldots s_n t_n$ in the following way; let s_1 be the longest (possibly empty) prefix of r which is defined on x, and factor $r = s_1 r'$. Next, let t_1 be the longest prefix of r' which is defined on u, and continue this process. Since t_i is defined on u, it is also defined on v, and since no longer factor is defined on u, no longer factor is defined on v. Thus, the factorization of r would be the same if taken with respect to xv, and in particular, r is defined on xv.

Suppose next that $r(xu) < r'(xu)$. We factor $r = s_1 t_1 \ldots s_k t_k$ and $r' = s_1' t_1' \ldots s_\ell' t_\ell'$ as before where we allow s_1, s_1', t_k and t_ℓ' to be empty. By the above argument, the t_i and t_j' are also the longest factors being defined also on v. Suppose that t_k is empty, but t_ℓ' is not. Then

---

[4]As an example on when it does not hold, consider the singleton $\{(X_{aa}, Y_{aa})\}$. We note that neither X_{aa} nor Y_{aa} are defined on ε nor on a. Thus $\varepsilon \leq^{\mathscr{C}} a$. However, $a \not\leq^{\mathscr{C}} aa$. The following proposition gives a condition on $\mathscr{C}$ which implies that $\leq^{\mathscr{C}}$ is stable.

$r(xv) \leq |x| < |x| + 1 \leq r'(xv)$. Note that we can not have t_k nonempty while t'_ℓ is empty since this would contradict $r(xu) < r'(xu)$. Finally, if t_k and t'_ℓ are both nonempty. Then $t_k(u) < t'_\ell(u)$, and we get $r(xv) = |x| + t_k(v) < |x| + t'_\ell(v) = r'(xv)$. The case when both t_k and t'_ℓ are empty is handled similarly. Thus, we get $r(xv) < r'(xv)$. The case dealing with the non-strict order is analogous. $\square$

If $\mathscr{C}$ furthermore is finite, then $A^*/\leq^\mathscr{C}$ is a finite monoid. This monoid can be constructed explicitly; given representatives u and v for some elements, one can check which ranker comparisons uv satisfy.

We also mention the two following lemmas which are trivial and have been written down without a proof.

Lemma 5. *Let $r = r'\mathsf{X}_p$ and s be rankers. If $r(u)$ and $s(u)$ are defined, then*

$$r(u) \leq s(u) \Leftrightarrow r'(u) \leq s\mathsf{Y}_{\overline{p}}(u) \qquad\qquad r(u) < s(u) \Leftrightarrow r'(u) < s\mathsf{Y}_{\overline{p}}(u)$$
$$s(u) \leq r(u) \Leftrightarrow s\mathsf{Y}_{\overline{p}}(u) \leq r'(u) \qquad\qquad s(u) < r(u) \Leftrightarrow s\mathsf{Y}_{\overline{p}}(u) < r'(u).$$

where in the second line the implications only holds if $s\mathsf{Y}_{\overline{p}}(u)$ is defined. In the first line, $s\mathsf{Y}_{\overline{p}}(u)$ being defined is also implied. Symmetrically, if $r(u) = r'\mathsf{Y}_u(u)$, and $r(u)$ and $s(u)$ are defined, then

$$r(u) \leq s(u) \Leftrightarrow r'(u) \leq s\mathsf{X}_{\overline{p}}(u) \qquad\qquad r(u) < s(u) \Leftrightarrow r'(u) < s\mathsf{X}_{\overline{p}}(u)$$
$$s(u) \leq r(u) \Leftrightarrow s\mathsf{X}_{\overline{p}}(u) \leq r'(u) \qquad\qquad s(u) < r(u) \Leftrightarrow s\mathsf{X}_{\overline{p}}(u) < r'(u).$$

where this time definedness is only implied in the second line.

Lemma 6. *Let $u = xavby$ where $u, v, x, y \in A^*$, $a, b \in A$. If there are rankers r, s such that $r(u) = |xa|$, $s(u) = |xavb|$, then a ranker t starting with an X-modality is defined on v if and only if rt is defined on u and $r(u) < rt'(u) < s(u)$ for all prefixes t' of t. Symmetrically, if t starts with a Y-modality, then t is defined on v if and only if st is defined on u and $r(u) < st'(u) < s(u)$ for all prefixes t' of t.*

5 The Ranker Comparison Hierarchy

Rankers and ranker comparisons have a long tradition in the study of fragments of FO^2. Indeed, rankers were first introduced by Schwentick, Thérien and Vollmer as a characterization of FO^2 itself [21]. Ranker comparisons were used by Weis and Immerman to give a characterization of the languages definable in FO^2_m [24]; this was later expanded to the full alternation hierarchy in the PhD thesis of Lauser [16]. A ranker characterization of the corners of the Trotter-Weil hierarchy is also known, using so called condensed rankers [15].

In this section, we place these results into our general framework. In particular, we rephrase the characterization of the Trotter-Weil corners in terms of ranker comparisons. This leads to a natural hierarchy containing both the Trotter-Weil and quantifier alternation hierarchies: the *ranker comparison hierarchy*, shown in Figure 1.

The levels of the ranker comparison hierarchy are built using the following collections of ranker comparisons. We note that they are finite and closed under subwords, and thus define finite monoids by Proposition 3. For $m \geq 1$:

$$\mathrm{XX}_{m,n} = \left\{ (r, s) \in R^X_m \times R^X_m \mid |r|, |s| \leq n \right\}, \quad \mathrm{YY}_{m,n} = \left\{ (r, s) \in R^Y_m \times R^Y_m \mid |r|, |s| \leq n \right\}$$

Figure 1: The Ranker Comparison Hierarchy surrounded by the Trotter-Weil hierarchy (left) and the Quantifier Alternation Hierarchy (right)

and for $m \geq 2$:

$$\text{XY}_{m,n} = \left\{ (r,s) \in R_m^X \times R_m^Y \mid |r|, |s| \leq n \right\}, \quad \text{YX}_{m,n} = \left\{ (r,s) \in R_m^Y \times R_m^X \mid |r|, |s| \leq n \right\}.$$

We also consider unions of these sets. We will use the notation $\leq_{m,n}^{\text{XY}}$ instead of $\leq^{\text{XY}_{m,n}}$ and similarly for the other sets. Note that we have $u \leq_{m,n}^{\text{XY}} v$ if and only if $v \leq_{m,n}^{\text{YX}} u$ and $v \leq_{m,n}^{\text{XY} \cup \text{YX}} u$ if and only if both $v \leq_{m,n}^{\text{XY}} u$ and $v \leq_{m,n}^{\text{YX}} u$.[5] We also give the following names for the induced monoids.

Definition 8. We define

$$N_{m,n}^{\text{XX}} = A^* / \leq_{m,n}^{\text{XX}} \qquad N_{m,n}^{\text{YY}} = A^* / \leq_{m,n}^{\text{YY}} \qquad N_{m,n}^{\text{XX} \cup \text{YY}} = A^* / \leq_{m,n}^{\text{XX} \cup \text{YY}}$$
$$N_{m,n}^{\text{XY}} = A^* / \leq_{m,n}^{\text{XY}} \qquad N_{m,n}^{\text{YX}} = A^* / \leq_{m,n}^{\text{YX}} \qquad N_{m,n}^{\text{XY} \cup \text{YX}} = A^* / \leq_{m,n}^{\text{XY} \cup \text{YX}}.$$

The fragment $\text{FO}_{1,n}^2$ is characterized by existence of subwords. Since our framework require the same subwords up to a certain length to be present in both words of interest (condition *(i)* in Definition 7), this means that our framework can not properly handle this case. Therefore, we make the following special definition.

Definition 9. Let A be some alphabet, we say that $u \leq_{1,n}^{\text{XY}} v$ if any ranker in $R_{1,2n}^X \cup R_{1,2n}^Y$ which is defined on u is also defined on v. Equivalently, $u \leq_{1,n}^{\text{XY}} v$ if every subword of length $2n$ which exists in u also exists in v. We say that $u \leq_{1,n}^{\text{YX}} v$ if $v \leq_{1,n}^{\text{XY}} u$ and $u \leq_{1,n}^{\text{XY} \cup \text{YX}} v$ if $u \equiv_{1,n}^{\text{XY}} v$.

A language is definable by $\text{XY}_{1,n}$ (resp. $\text{YX}_{1,n}$ or $\text{XY}_{1,n} \cup \text{YX}_{1,n}$) if it is an ideal under the relation $\leq_{1,n}^{\text{XY}}$ (resp. $\leq_{1,n}^{\text{YX}}$ or $\leq_{1,n}^{\text{YX} \cup \text{YX}}$). This does, however, *not* mean that we can interpret

---

[5]Note that although the relation $\leq_{m,n}^{\text{XY} \cup \text{YX}}$ is similar to the relation introduced by Weis and Immerman [24], there is a slight different regarding the variable n. The length of rankers allowed by Weis and Immerman is made to ensure correspondence with the depth of formulae in FO^2. The results of this article does not consider depths of FO^2 formulae, and thus this difference is not important here.

13

$XY_{1,n}$ as a set of rankers and definability in the sense of Definition 7. We also define $N_{1,n}^{XY}$ and $N_{1,n}^{YX}$ to be the monoids induced by the respective (stable) preorders. This special definition ensures that we can define all levels of the quantifier alternation hierarchy using consistent terminology. Furthermore, it makes the following lemma true for all m.

Lemma 7. *Let $m \geq 2$ and let $u, u' \in A^*$ such that*

$$u = xavby \leq_{m,n}^{YX} x'av'by' = u'$$

where $x, v, y, x', v', y' \in A^*$ and $a, b \in A$. If there exists rankers X_p, Y_q with $|p|, |q| \leq k < n$ such that $X_p(u) = |xavb|$, $X_p(u') = |x'av'b|$, $Y_q(u) = |xa|$ and $Y_q(u') = |x'a|$, then $v \leq_{m-1,n-k}^{XY} v'$. Similarly, if $u \leq_{m,n}^{XY} u'$, then $v \leq_{m-1,n-k}^{YX} v'$.

In other words, the implication of true ranker comparisons carries over from u and u' to v and v' at the cost of one quantifier alternation and the depth needed to reach the positions a and b.

Proof. We start by considering the case $m \geq 3$, since $m = 2$ needs to be treated specifically. Let $(r, s) \in XY_{m-1,n-k}$.

We first assume that one of the rankers, say r, is defined on v and show that this implies that it is also defined on v'. Suppose r starts with an X-modality, and let r' be an arbitrary prefix of r. We have $Y_q r'(u) < X_p(u)$ which implies $Y_q r'(u') < X_p(u')$ since $(Y_q r', X_p) \in YX_{m,n}$.

We also need to show $Y_q(u') < Y_q r'(u')$, however, since the alternation depth of r' is possibly $m-1$, the comparison $(Y_q, Y_q r')$ might not be in $YX_{m,n}$. Suppose r' ends with a Y-modality, say $r' = r'' Y_t$, where again r'' has a strictly smaller alternation depth. Then

$$
\begin{aligned}
Y_p(u) \leq Y_q r'' Y_t(u) &\Rightarrow Y_p X_{\bar{t}}(u) \leq Y_q r''(u) && \text{By Lemma 5} \\
&\Rightarrow Y_p X_{\bar{t}}(u') \leq Y_q r''(u') && \text{since } u \leq_{m,n}^{YX} u' \\
&\Rightarrow Y_p(u') \leq Y_q r'' Y_t(u') && \text{By Lemma 5}
\end{aligned}
$$

where in the second implication, we rely on the fact that r'' has alternation depth at most $m-2$, and thus $Y_q r''$ has alternation depth at most $m-1$. If r' ends with an X-modality, we get the same chain of equivalences. Thus we have $Y_q(u') < Y_q r'(u') < X_p(u)$ for all prefixes of r, which by Lemma 6 implies that r is defined on v'. The case when r starts with a Y-modality is symmetric.

Assume next that r is defined on v', and let r' be an arbitrary prefix. We will use induction on the alternation depth of r' to show that every such prefix is defined on r. We consider the case when $r' = r'' X_t$ where r'' is either empty or has strictly smaller alternation depth. We also assume r starts with an X-modality. The other cases are similar.

By induction r'' is either empty or defined on v, and it thus suffices to show $Y_q r'' X_t(u) < X_p(u)$. Suppose this is not the case, then

$$
\begin{aligned}
X_p(u) \leq Y_q r'' X_t(u) &\Rightarrow X_p Y_{\bar{t}}(u) \leq Y_q r''(u) && \text{By Lemma 5} \\
&\Rightarrow X_p Y_{\bar{t}}(u') \leq Y_q r''(u') && \text{since } u \leq_{m,n}^{YX} u' \\
&\Rightarrow X_p(u') \leq Y_q r'' X_t(u') && \text{By Lemma 5}
\end{aligned}
$$

where in the second implication, we both rely on the fact that $m \geq 3$ and the alternation depth of r'' is at most $m-2$. This contradicts r' being defind on v'. Thus, we must have $Y_q r'(u) < X_p(u)$ which means r' is defined on v.

14

Finally, suppose that $r(u) \leq s(u)$. We suppose r starts with an X-modality and s with a Y-modality. Then $Y_q r(u) \leq X_p s(u)$ holds, giving $Y_q r(u') \leq X_p s(u')$. Since r and s are defined on v', this implies $r(v') \leq s(v')$. The other cases are similar, using the fact that if r starts with a Y-modality, or s with an X-modality, then their alternation depths are at most $m - 2$.

We finally turn to the case $m = 2$, and let $X_s X_t \in R^X_{1,2(n-k)}$ with $|s|, |t| \leq n - k$. Clearly, $X_s X_t$ being defined on a word w is equivalent to X_s and $Y_{\bar{t}}$ being defined and $X_s(w) < Y_{\bar{t}}(w)$ holding. Suppose that $X_s(v) < Y_{\bar{t}}(v)$, then since $u \leq^{YX}_{2,n} u'$, we have

$$Y_q X_s(u) < X_p Y_{\bar{t}}(u) \Rightarrow Y_q X_s(u') < X_p Y_{\bar{t}}(u'). \tag{1}$$

In particular, $Y_q X_{s'}(u) < X_p$ for all prefixes s' of s. Hence X_s is defined on u' and analogously, so is $Y_{\bar{t}}$. Thus (1) in particular implies $X_s(v') < Y_{\bar{t}}(v')$ which implies that $X_s X_t$ is defined on u. By a symmetrical argument, the same holds for rankers in $R^Y_{1,2n}$. $\square$

This lemma also has a counterpart for the XX and YY-levels. Note that in this case, we lose depth but no alternation depth when considering factors.

Lemma 8. *Let $m \geq 2$ and let $u, u' \in A^*$ such that*

$$u = xavby \equiv^{XX}_{m,n} x'av'by' = u'$$

If there exists rankers X_p, X_q with $|p|, |q| \leq k < n$ such that $X_p(u) = |xa|$, $X_p(u') = |x'a|$, $X_q(u) = |xavb|$ and $X_q(u') = |x'av'b|$, then $v \equiv^{XX}_{m,n-k} v'$. We also allow xa and $x'a$ (resp. by and by') to be empty, in which case we interpret X_p (resp. X_q) to be the empty word, $|p| = 0$ (resp. $|q| = 0$) and $|X_p(u)| = |X_p(v)| = 0$ (resp. $|X_p(u)| = |u| + 1$, $|X_q(v)| = |v| + 1$).

Proof. By symmetry, it is enough to show that a ranker defined on v is also defined on v' and if $r(v) \leq s(v)$, then $r(v') \leq s(v')$.

Suppose $r \in [XX_{m,n-k}]$ starts with an X-modality. If r is defined on v, then if follows from Lemma 6 that $X_p r$ is defined on u and $X_p(u) < X_p r'(u) < X_q(u)$ for all nonempty prefixes of r. Since $u \equiv^{XX}_{m,n} u'$, we get that $X_p r$ is defined on u' and $X_p(u') < X_p r'(u') < X_q(u')$. Again by Lemma 6, it follows that r is defined on v'. If r starts with a Y-modality, we make the same argument using $X_q r'$ instead of $X_p r'$.

Let $(r, s) \in XX_{m,n-k}$ where both r and s are defined on v (and hence on v') such that $r(v) \leq s(v)$. Suppose both rankers start with X-modalities. Then

$$\begin{aligned} r(v) < s(v) &\Rightarrow X_p r(u) < X_p s(u) \\ &\Rightarrow X_p r(u') < X_p s(u') \qquad \text{Since } u \equiv^{XX}_{m,n} u' \\ &\Rightarrow r(v') < s(v') \end{aligned}$$

The other cases are similar. $\square$

We now restate the known correspondences between rankers and the quantifier alternation hierarchy in our framework. The following characterization of the FO^2_m levels is due to Weis and Immerman (*(i)* and *(ii)* [24]) and Kufleitner and Weil (*(i)* and *(iii)* [14]).

Proposition 4. *Given a language L, the following are equivalent:*

(i) L is definable in FO^2_m,

15

(ii) L is definable by $\mathrm{XY}_{m,n} \cup \mathrm{YX}_{m,n}$ *for some n,*

(iii) the syntactic morphism of L is in $\mathbf{R}_{m+1} \cap \mathbf{L}_{m+1}$.

Similarly, we have the following characterization of the Σ_m^2 levels, due to Fleischer et al. (*(i)* and *(iii)* [8]) and Lauser (*(i)* and *(ii)* [16, Thm. 11.3]). One easily gets the symmetric characterization of the Π_m^2-levels.

Proposition 5. *Given a language L, the following are equivalent:*

(i) L *is definable in* Σ_m^2,

(ii) L *is definable by* $\mathrm{XY}_{m,n}$ *for odd m and* $\mathrm{YX}_{m,n}$ *for even m for some n,[6]*

(iii) the syntactic monoid of L is in $\mathbf{Si}_m$.

Proof. We only need to show that the ranker comparisons defined by Lauser is equivalent to ours up to the depth of the rankers. Recall Definition 11.2 in [16], defining $\leq_{m,n}^R$ as follows. If $m = 0$ or $n = 0$, then $u \leq_{m,n}^R v$ for all $u, v \in A^*$. Otherwise, $u \leq_{m,n}^R v$ if $v \leq_{m-1,n}^R u$ and:[7]

(i) for $r \in R_m$ such that $|r| \leq n$, if $r(v)$ is defined then $r(u)$ is defined,

(ii) if $r \in R_m^X$, $s \in R_{m-1}^X$ such that $|r| \leq n$, $|s| \leq n-1$, then if m is odd, then $r(v) \leq s(v)$ implies $r(u) \leq s(u)$ and $r(v) < s(v)$ implies $r(u) < s(u)$ and if m is even, then $r(v) \geq s(v)$ then $r(u) \geq s(u)$ and if $r(v) > s(v)$, then $r(u) > s(u)$,

(iii) if $r \in R_m^Y$, $s \in R_{m-1}^Y$ such that $|r| \leq n$, $|s| \leq n-1$, then if m is odd, then $r(v) \geq s(v)$ implies $r(u) \geq s(u)$ and $r(v) > s(v)$ implies $r(u) > s(u)$ and if m is even, then $r(v) \leq s(v)$ then $r(u) \leq s(u)$ and if $r(v) < s(v)$, then $r(u) < s(u)$,

(iv) if $r \in R_m^Y$, $s \in R_m^Y$ such that $|r| + |s| < 2n$, then if m is odd, then $r(v) \leq s(v)$ implies $r(u) \leq s(u)$ and $r(v) < s(v)$ implies $r(u) < s(u)$ and if m is even, then $r(v) \geq s(v)$ then $r(u) \geq s(u)$ and if $r(v) > s(v)$, then $r(u) > s(u)$.

We claim that for $m \geq 2$, we can without loss of generality strengthen *(i)* to read

(i') for $r \in R_m$ such that $|r| \leq n$, we have $r(v)$ defined if and only if $r(u)$ is defined,

in which case the result is obvious.

Without loss of generality, assume that $r = r'\mathsf{X}_p\mathsf{Y}_q$ where r' has alternation depth $m - 2$ or is empty. Since $u \leq_{m,n}^R v$ implies $v \leq_{m-1,n}^R u$ by definition, we have that $r'\mathsf{X}_p$ and $\mathsf{X}_{\overline{q}}$ are both defined on v. If $r(v)$ is not defined, we must have $r'\mathsf{X}_p(v) < \mathsf{X}_{\overline{q}}(v)$ which implies $r'(v) < \mathsf{X}_{\overline{q}}\mathsf{Y}_{\overline{p}}(v)$ (or if r' is empty, it implies that $\mathsf{X}_{\overline{q}}\mathsf{Y}_{\overline{p}}(v)$ is defined). It follows that $r'(u) < \mathsf{X}_{\overline{q}}\mathsf{Y}_{\overline{p}}(u)$ which contradicts $r(u)$ being defined. □

The corners of the Trotter-Weil hierarchy has a ranker characterization in terms of condensed rankers [15]. We reformulate this result in terms of ranker comparisons. The ranker comparison characterization of the join levels then follow directly. We also use Proposition 1 to relate the join levels to the intersection levels of the quantifier alternation hierarchy.

---

[6]The relation in [16] has only one-sided inclusion of definedness of rankers with the maximum number of alternations for all m as an explicit assumption. However, for $m \geq 2$, two sided inclusion follow implicitly.

[7]Note that the direction of $\leq_{m,n}^R$ is reversed compared to Definition 7. In the former, if $u \leq^{\mathscr{C}} v$, then $r(u) \leq s(u)$ implies $r(v) \leq s(v)$ for $r, s \in [\mathscr{C}]$, while if $u \leq_{m,n}^R v$, then $r(v) \leq s(v)$ implies $r(u) \leq s(u)$ for the relevant r and s.

Proposition 6. *Let $m \geq 1$. Then L is definable by $\mathrm{XX}_{m,n}$ (resp. $\mathrm{YY}_{m,n}$) for some n if and only if its syntactic monoid M is in $\mathbf{R}_{m+1}$ (resp. $\mathbf{L}_{m+1}$). Furthermore, the following are equivalent:*

(i) *L is definable by $\mathrm{XX}_{m,n} \cup \mathrm{YY}_{m,n}$ for some n,*

(ii) *the syntactic monoid of L is in $\mathbf{R}_{m+1} \vee \mathbf{L}_{m+1}$,*

(iii) *the syntactic monoid of L is in $\mathbf{Pi}_m \cap \mathbf{Si}_m$,*

(iv) *L is definable in Σ_m^2 and in Π_m^2.*

Proof. Let $\mu : A^* \to M$ be the syntactic morphism of L. We use induction on m. We consider the case $\mathbf{R}_{m+1}$ with $\mathbf{L}_{m+1}$ being symmetric. Suppose first that L is definable by $\mathrm{XX}_{m,n}$, and M is its syntactic monoid. Since M is a quotient of $N_{m,n}^{\mathrm{XX}}$, it is enough to show $N_{m,n}^{\mathrm{XX}} \in \mathbf{R}_{m+1}$. Let $\nu : A^* \to N_{m,n}^{\mathrm{XX}}$ and let $\pi : N_{m,n}^{\mathrm{XX}} \to N_{m,n}^{\mathrm{XX}}/{\sim_{\mathbf{K}}}$ be the natural projection. Let $u, v \in A^*$ be such that $u \equiv_{m-1,n}^{\mathrm{YY}} v$ if $m \geq 2$, or $\mathsf{alph}(u) = \mathsf{alph}(v)$ for $m = 1$. We show that $\mu(u) \sim_{\mathbf{K}} \mu(v)$.

Let e be idempotent such that $e \mathcal{J} e\mu(u)$. Then there exists $\hat{e} \in \mu^{-1}(e)$ with $\mathsf{alph}(u) = \mathsf{alph}(v) \subseteq \mathsf{alph}(\hat{e})$. Since e is idempotent, we have $\hat{e}^{2n} \in \mu^{-1}(e)$. We want to show $\hat{e}^{2n}u \equiv_{m,n}^{\mathrm{XX}} \hat{e}^{2n}v$.

We consider rankers r and s. Suppose first that both start with an X-modality. In particular, for $m = 1$ this is the only possibility. We note that by the length of $\hat{e}^n$ and the alphabetic conditions, r and s are defined on $\hat{e}^{2n}u$ if and only if they are defined on $\hat{e}^n$. The same holds for $\hat{e}^{2n}v$. In particular, they are defined on $\hat{e}^{2n}u$ if and only if $\hat{e}^{2n}v$. Furthermore,

$$r(\hat{e}^{2n}u) \leq s(\hat{e}^{2n}v) \Leftrightarrow r(\hat{e}^n) \leq s(\hat{e}^n) \Leftrightarrow r(\hat{e}^{2n}v) \leq s(\hat{e}^{2n}v)$$

and analogously for the strict relation.

Next, suppose r and s are rankers which start with a Y-modality. Then $(r, s) \in \mathrm{YY}_{m-1,n}$. Since $u \equiv_{m-1,n}^{\mathrm{YY}} v$ it follows by stability that $\hat{e}^{2n}u \equiv_{m-1,n}^{\mathrm{YY}} \hat{e}^{2n}v$. In particular, the same such rankers are defined on u and v.

Finally, suppose r starts with an X-modality, and s starts with a Y-modality. It again follows from the length of $\hat{e}^n$ and the alphabetic conditions that $r(\hat{e}^{2n}u) < s(\hat{e}^{2n}u)$ if and only if both are defined on $\hat{e}^{2n}u$ and similarly for $\hat{e}^{2n}v$. For the same reason, none of the words can satisfy $r \leq s$ if r starts with a Y-modality and s starts with an X-modality.

Since this covers all possible cases, we get $\hat{e}^{2n}u \equiv_{m,n}^{\mathrm{XX}} \hat{e}^{2n}v$. This implies $e\nu(u) = e\nu(v)$ and since e was arbitrary such that $e \mathcal{J} e\nu(u)$, it follows that $\nu(u) \sim_{\mathbf{K}} \nu(v)$. For $m = 1$, this implies that $\pi \circ \nu$ factors through J_A and for for $m \geq 2$ that $\pi \circ \nu$ factors through $N_{m-1,n}^{\mathrm{YY}}$. In either case, $N_{m,n}^{\mathrm{XX}}/{\sim_{\mathbf{K}}} \in \mathbf{L}_m$ (for $m = 1$ since $J_A \in \mathbf{L}_1$ and for $m \geq 2$ since $N_{m,n}^{\mathrm{YY}} \in \mathbf{L}_{m+1}$ by induction). Thus $N_{m,n}^{\mathrm{XX}} \in \mathbf{R}_{m+1}$.

For the other direction, suppose $\pi : M \to M/{\sim_{\mathbf{K}}} \in \mathbf{L}_m$ is the canonical projection. We will use induction to show that $u \equiv_{m,m|M|+1}^{\mathrm{XX}} v$ implies $\mu(u) = \mu(v)$. Suppose that $u \equiv_{m,m|M|+1}^{\mathrm{XX}} v$ and let $u = u_0 a_1 u_1 \cdots a_n u_n$ be the $\mathcal{R}$-factorization of u. We factor $v = v_0 a_1 v_1 \cdots a_n v_n$ where $a_i \notin v_{i-1}$. Such a factorization exists since $u \equiv_{m,m|M|+1}^{\mathrm{XX}} v$. Furthermore, we have

$$b \in \mathsf{alph}(u_i) \Leftrightarrow \mathsf{X}_{a_1 \ldots a_i}(u) < \mathsf{X}_{a_1 \ldots a_i b}(u) < \mathsf{X}_{a_1 \ldots a_i a_{i+1}}(u)$$
$$\Leftrightarrow \mathsf{X}_{a_1 \ldots a_i}(v) < \mathsf{X}_{a_1 \ldots a_i b}(v) < \mathsf{X}_{a_1 \ldots a_i a_{i+1}}(v)$$
$$\Leftrightarrow b \in \mathsf{alph}(v_i).$$

Thus, $\mathsf{alph}(u_i) = \mathsf{alph}(v_i)$ for all i (the argument for $i = 0$ and $i = n$ is similar, but uses definedness and one comparison instead of two comparisons). Furthermore, for $m \geq 2$, Lemma 8 shows that $u_i \equiv^{\mathrm{XX}}_{m,(m-1)|M|+1} v_i$ which in particular implies $u_i \equiv^{\mathrm{YY}}_{m-1,(m-1)|M|+1} v_i$. By induction, this implies $\mu(u_i) \sim_{\mathbf{K}} \mu(v_i)$ for all m. Let x_i be such that $\mu(u_0 a_1 \dots a_i u_i x_i) = \mu(u_0 a_1 \dots a_i)$. We get

$$
\begin{aligned}
\mu(u_0 a_1 u_1 \dots a_n u_n) &= \mu(v_0 a_1 u_1 \dots a_n (u_n x_n)^{\omega} u_n) \\
&= \mu(u_0 a_1 u_1 \dots a_n (u_n x_n)^{\omega} v_n) \\
&= \mu(u_0 a_1 u_1 \dots a_n v_n) \\
&\ \ \vdots \\
&= \mu(v_0 a_1 v_1 \dots a_n v_n).
\end{aligned}
$$

Thus μ factors through $N^{\mathrm{XX}}_{m,n+|M|}$ which gives the desired result.

We now turn to the join levels. The language L has its syntactic monoid in $\mathbf{R}_{m+1} \vee \mathbf{L}_{m+1}$ if it is a divisor of some $M_1 \times \dots \times M_n$ where $M_i \in \mathbf{R}_{m+1} \cup \mathbf{L}_{m+1}$. In particular, the language is recognized by some $\mu : A^* \to M_1 \times \dots \times M_n$. Let $\mu_i = \pi_i \circ \mu$ where π_i is the projection on the ith monoid. Let $u = (u_1, \dots, u_n) \in M_1 \times \dots \times M_n$. By the previous part, $\mu_i^{-1}(u_i)$ is defineable by $\mathrm{XX}_{m,n} \cup \mathrm{YY}_{m,n}$ for some n. Since $\mu^{-1}(u) = \bigcap_i \mu_i^{-1}(u_i)$, so is $\mu^{-1}(u)$. Since it is the union of such sets, it follows that L is definable by $\mathrm{XX}_{m,n} \cup \mathrm{YY}_{m,n}$.

For the other direction, suppose L is definable in $\mathrm{XX}_{m,n} \cup \mathrm{YY}_{m,n}$. Then L is a Boolean combination of languages L_i definable in XX_m or YY_m. Let $\mu_i : A^* \to M_i$ be the syntactic morphism of L_i, and define $\mu : A^* \to M_1 \times \dots \times M_n$ by $\mu(u) = (\mu_1(u), \dots, \mu_n(u))$. It is clear that L is the preimage of a union of elements in $M_1 \times \dots \times M_n$, and it follows that the syntactic monoid of L is in $\mathbf{R}_{m+1} \vee \mathbf{L}_{m+1}$.

Finally, the equivalence between *(iv)* and *(ii)* is Proposition 1 and the equivalence between *(iii)* and *(iv)* is an obvious consequence of Proposition 5. □

Taken together, these three propositions gives us a new way of considering the Trotter-Weil hierarchy and the quantifier alternation hierarchy together, as a ranker comparison hierarchy; see Figure 1.

6 Saturations for Fragments of FO^2

In this section, we present computable closed sets for all levels of the Trotter-Weil and quantifier alternation hierarchies, in other words for all levels of the ranker comparison hierarchy. We also state our main results: that these sets agree with the corresponding sets of pointlikes. The proof thereof is the subject of the subsequent sections.

The sets presented below relies on the monoids having content morphisms (intuitively on the monoid elements having a fixed alphabet). This is not true for all monoids; consider for example $M = \{1, a\}$ with $aa = 1$. However, it is always possible to *alphabetize* a monoid by explicitly distinguishing elements with different alphabets. If M is a monoid with a generating set A, then the submonoid of $M \times J_A$ generated by $(a, \{a\})_{a \in A}$ has a content morphism. It also has a surjective morphism onto M. The following Lemma shows that this is enough to deduce the pointlikes of M.[8]

---

[8] A variant of the Lemma can be found in [2].

Lemma 9. *Let M, M' be finite monoids and suppose that there is a surjective morphism $\pi : M' \to M$. Then $P \subseteq M$ (resp. $(s, S) \in M \times 2^M$) is pointlike (resp. conelike) with respect to a variety $\mathbf{V}$ if and only if there exists a pointlike $P' \subseteq M'$ (resp. conelike $(s', S') \in M' \times 2^{M'}$) with respect to $\mathbf{V}$ such that $P \subseteq \pi(P')$ (resp. $S \subseteq \pi(S')$, $s = \pi(s')$).*

Proof. We give the proof for pointlikes; the proof for conelikes is analogous. Let $\tau : M \to N \in \mathbf{V}$ be a relational morphism. Then $\tau' = \tau \circ \pi^{-1}$ is also a relational morphism. If P' is pointlike with respect to $\mathbf{V}$, then in particular it is pointlike with respect to τ'. It follows that $\pi(P')$ and thus P must be pointlike with respect to τ. Since τ was arbitrary, this is true for $\mathbf{V}$.

For the other direction, let $P \subseteq M$ and let $P_1, \ldots, P_n$ be the collection of all subsets of M' satisfying $P \subseteq \pi(P_i)$. If P_i is not pointlike with respect to $\mathbf{V}$, then there exists $\tau_i : M' \to N_i \in \mathbf{V}$ such that P_i is not pointlike with respect to τ_i. Suppose that no P_i is pointlike. We define $\tau : M' \to N_1 \times \cdots \times N_n$ by

$$\tau(s) = \{(x_1, \ldots, x_n) \mid x_i \in \tau_i(s)\} .$$

It is straightforward to check that τ is a relational morphism. We let $\tau' : \tau \circ \pi^{-1}$. For contradiction, suppose that P is pointlike with respect to $\mathbf{V}$, then in particular P is pointlike with respect to τ'. Thus, there exists $(x_1, \ldots, x_n)$ such that $P \subseteq \tau'^{-1}(x_1, \ldots, x_n)$. Let $U = \tau^{-1}(x_1, \ldots, x_n)$. It is clear that $P \subseteq \pi(U)$, and thus $U = P_i$ for some i. Since $(x_1, \ldots, x_i, \ldots, x_n) \in \tau(s)$ for all $s \in P_i$, we must have $x_i \in \tau_i(s)$ for all $s \in P_i$, a contradiction. $\square$

We now introduce the relevant closed sets. Note that for our purposes, $\mathbf{R}_1 = \mathbf{L}_1 = \mathbf{J}_1$. We first give the sets for the corners of the Trotter-Weil hierarchy. These are important building blocks for the other sets.

Definition 10. Let M be a monoid with a content morphism α. We define:

- $\mathsf{Sat}_{\mathbf{J}_1}(M) = \mathsf{Sat}_{\mathbf{R}_1}(M) = \mathsf{Sat}_{\mathbf{L}_1}(M) = \{S \subseteq M \mid \alpha(s) = \alpha(t) \text{ for all } s, t \in S\}$,

- for $m \geq 2$, $\mathsf{Sat}_{\mathbf{R}_m}(M)$ is the smallest closed set of M such that if $Z \in \mathsf{Sat}_{\mathbf{L}_{m-1}}(M)$, $U \in \mathsf{Sat}_{\mathbf{R}_m}(M)$, $\alpha(Z) \leq \alpha(U)$ and U is idempotent in 2^M, then $UZ \in \mathsf{Sat}_{\mathbf{R}_m}(M)$

- for $m \geq 2$, $\mathsf{Sat}_{\mathbf{L}_m}(M)$ is the smallest closed set of M such that if $Z \in \mathsf{Sat}_{\mathbf{R}_{m-1}}(M)$, $V \in \mathsf{Sat}_{\mathbf{L}_m}(M)$, $\alpha(Z) \leq \alpha(V)$ and V is idempotent in 2^M, then $ZV \in \mathsf{Sat}_{\mathbf{L}_m}(M)$

The definition inductively ensures that for any W in any of the introduced sets, we have $\alpha(w) = \alpha(w')$ for all $w, w' \in W$. Thus, $\alpha(W)$ is a well defined element of J_A, making the comparisons $\alpha(Z) \leq \alpha(U)$ and $\alpha(Z) \leq \alpha(V)$ meaningful.

The other closed sets build on so-called $\mathbf{RL}_m$-factors. If one think of the elements of a monoid as the languages they represent, one can think of $\mathbf{RL}_m$-factors as collections of languages which can not be distinguished from any side using rankers of alternation depth at most m, while containing words of arbitrary length.

Definition 11. Let M be a monoid with a content morphism α. Let $S, E \in \mathsf{Sat}_{\mathbf{R}_m}(M)$, $T, F \in \mathsf{Sat}_{\mathbf{L}_m}(M)$ with $\alpha(S), \alpha(T) \leq \alpha(E) = \alpha(F)$ and E, F idempotent in 2^M. Let W be such that $\alpha(w) \leq \alpha(E)$ for all $w \in W$. Then $SEWFT$ is an $\mathbf{RL}_m$-*factor*.

Since $\mathsf{Sat}_{\mathbf{R}_m}(M)$ and $\mathsf{Sat}_{\mathbf{L}_m}(M)$ can be constructed for each m, the $\mathbf{RL}_m$-factors can also be effectively constructed. Note that the alphabet of an $\mathbf{RL}_m$-factor is well defined. Using these factors, we construct the following sets.

Definition 12. Let M be a monoid. Then

- $\mathsf{ConeSat}_{\mathbf{Si}_1}(M)$ is the smallest closed set for which $(1, S) \in M$ for all $S \subseteq M$.

- $\mathsf{ConeSat}_{\mathbf{Pi}_1}(M)$ is the smallest closed set for which $(s, \{1, s\}) \in M$ for all $s \in M$.

Suppose further that M has a content morphism α, then for $m \geq 2$:

- $\mathsf{Sat}_{\mathbf{J}}(M)$, is the smallest closed set such that

$$XEY \in \mathsf{Sat}_{\mathbf{J}}(M)$$

  for all $X, Y, E \in 2^M$ where E is idempotent, $\alpha(s) = \alpha(t)$ for all $s, t \in E$, and $\alpha(w) \subseteq \alpha(E)$ for all $w \in X \cup Y$.

- $\mathsf{Sat}_{\mathbf{R}_{m+1} \cap \mathbf{L}_{m+1}}(M)$ is the smallest closed set which for all n contain the product

$$U_1 V_1 U_2 \dots V_{n-1} U_n$$

  where every U_i is an $\mathbf{RL}_m$-factor while $V_i \in \mathsf{Sat}_{\mathbf{R}_m \cap \mathbf{L}_m}(M)$ (or $V_i \in \mathsf{Sat}_{\mathbf{J}}(M)$ for $m = 2$), and $\alpha(v_i') \leq \alpha(U_i), \alpha(U_{i+1})$ for all $v_i' \in V_i$,

- $\mathsf{Sat}_{\mathbf{R}_m \vee \mathbf{L}_m}(M)$ is the smallest closed set which for all n contain the product

$$U_1 V_1 U_2 \dots V_{n-1} U_n$$

  where every U_i is an $\mathbf{RL}_m$-factor while $V_i \in \mathsf{Sat}_{\mathbf{R}_{m-1} \vee \mathbf{L}_{m-1}}(M)$ (or $V_i \in 2^M$ for $m = 2$), and $\alpha(v_i') \leq \alpha(U_i), \alpha(U_{i+1})$ for all $v_i' \in V_i$,

- $\mathsf{ConeSat}_{\mathbf{Si}_m}(M)$ is the smallest closed set which for all n contain the product

$$(u_1, U_1)(v_1, V_1)(u_2, U_2) \dots (v_{n-1}, V_{n-1})(u_n, U_n)$$

  where for all i, we have $u_i \in U_i$ and U_i is an $\mathbf{RL}_m$-factor while $(v_i, V_i) \in \mathsf{ConeSat}_{\mathbf{Si}_{m-1}}(M)$, and $\alpha(v_i') \leq \alpha(U_i), \alpha(U_{i+1})$ for all $v_i' \in V_i$,

- $\mathsf{ConeSat}_{\mathbf{Pi}_m}(M)$ is the smallest closed set which for all n contain the product

$$(u_1, U_1)(v_1, V_1)(u_2, U_2) \dots (v_{n-1}, V_{n-1})(u_n, U_n)$$

  where for all i, we have $u_i \in U_i$ and U_i is an $\mathbf{RL}_m$-factor while $(v_i, V_i) \in \mathsf{ConeSat}_{\mathbf{Pi}_{m-1}}(M)$, and $\alpha(v_i') \leq \alpha(U_i), \alpha(U_{i+1})$ for all $v_i' \in V_i$,

We now state our main theorem. Apart from giving the pointlikes of the different levels, it also provides *separators*. These are monoids with relational morphisms which are optimal in $\mathbf{V}$ for separating the elements of M. In other words, the relational morphisms τ satisfy $\mathsf{PL}_\tau(M) = \mathsf{PL}_{\mathbf{V}}(M)$. The theorem states only the monoids explicitly; the relational morphisms are the natural relational morphisms, obtained by mapping every element in M to their preimage in A^* and projecting onto the relevant monoids.

Theorem 2. *Let M be a finite monoid, and let $n = \lceil R/2 \rceil - 1$ where R is the Ramsey number of M. Then*

(i) $\mathsf{Cone}_{\mathbf{Si}_1}(M) = \mathsf{ConeSat}_{\mathbf{Si}_1}(M)$ *with separator* $N_{1,n}^{\mathrm{XY}}$,

(ii) $\mathsf{Cone}_{\mathbf{Pi}_1}(M) = \mathsf{ConeSat}_{\mathbf{Pi}_1}(M)$ *with separator* $N_{1,n}^{\mathrm{YX}}$,

Furthermore, suppose M has a content morphism $\alpha: M \to J_A$, and let $n = (m + |A|)(R-1)$ and $n' = (m - 1 + 3|A|)(R-1) + |A|$ where R is the Ramsey number of 2^M.

(iii) $\mathsf{PL}_{\mathbf{J}_1}(M) = \mathsf{Sat}_{\mathbf{J}_1}(M)$ *with the separator* J_A,

(iv) $\mathsf{PL}_{\mathbf{R}_m}(M) = \mathsf{Sat}_{\mathbf{R}_m}(M)$ *with the separator* $N_{m,n}^{\mathrm{XX}}$,

(v) $\mathsf{PL}_{\mathbf{L}_m}(M) = \mathsf{Sat}_{\mathbf{L}_m}(M)$ *with the separator* $N_{m,n}^{\mathrm{YY}}$,

(vi) $\mathsf{PL}_{\mathbf{J}}(M) = \mathsf{Sat}_{\mathbf{J}}(M)$ *with separator* $N_{1,|A|R+R-1}^{\mathrm{XY}\cup\mathrm{YX}}$,[9]

(vii) $\mathsf{PL}_{\mathbf{R}_{m+1}\cap\mathbf{L}_{m+1}}(M) = \mathsf{Sat}_{\mathbf{R}_{m+1}\cap\mathbf{L}_{m+1}}(M)$ *with separator* $N_{m,n'}^{\mathrm{XX}\cup\mathrm{YY}}$,

(viii) $\mathsf{PL}_{\mathbf{R}_m\vee\mathbf{L}_m}(M) = \mathsf{Sat}_{\mathbf{R}_m\vee\mathbf{L}_m}(M)$ *with separator* $N_{m,n'}^{\mathrm{XX}\cup\mathrm{YY}}$,

(ix) $\mathsf{Cone}_{\mathbf{Si}_m}(M) = \mathsf{ConeSat}_{\mathbf{Si}_m}(M)$ *with separator $N_{m,n'}^{\mathrm{XY}}$ for odd m and $N_{m,n'}^{\mathrm{YX}}$ for even m,*

(x) $\mathsf{Cone}_{\mathbf{Pi}_m}(M) = \mathsf{ConeSat}_{\mathbf{Pi}_m}(M)$ *with separator $N_{m,n'}^{\mathrm{YX}}$ for odd m and $N_{m,n'}^{\mathrm{XY}}$ for even m,*

The following is an immediate corollary, given Proposition 2.

Corollary 1. *The covering problem has a solution for all language varieties associated with the levels of the quantifier alternation hierarchy. In particular, this implies solutions to the separation problems for all of these varieties.*

Example 4. Let M be the syntactic monoid of $(ab)^+$; see Figure 2a. Note that this monoid does not have a content morphism since (e.g.) a and aba maps to the same element in M. Thus, we instead consider the alphabetized monoid $M' = M \times J_{\{a,b\}}$ shown in Figure 2b.

In Example 3, we note that $(ab, \{aba, bab\})$ is conelike with respect to $\mathbf{J}^+ = \mathbf{Si}_1$. To see that this follows from Theorem 2, we show that $(ab, \{aba, bab\}) \in \mathsf{ConeSat}_{\mathbf{Si}_1}(M')$. Indeed, we have $(1, \{1, a\}), (1, \{1, b\}) \in \mathsf{ConeSat}_{\mathbf{Si}_1}(M')$ and thus by closure under multiplication we get $(1, \{1, b\})(ab, \{ab\})(1, \{1, a\}) = (ab, \{ab, bab, aba, ba\}) \in \mathsf{ConeSat}_1(M')$. It follows that $(ab, \{bab, aba\}) \in \mathsf{ConeSat}_1(M')$ by closure under subsets in the second entry.

Next, we consider the pointlikes with respect to $\mathbf{R}_2$ and $\mathbf{L}_2$, which in particular helps us to calculate the $\mathbf{RL}_m$-factors. We have $\{aba, ab, ba, bab, aab\} \in \mathsf{Sat}_{\mathbf{J}_1}(M')$ and, for example, $\{ab\} \in \mathsf{Sat}_{\mathbf{R}_1}(M')$. Thus $\{ab\}\,\{aba, ab, ba, bab, aab\} = \{aba, ab, aab\} \in \mathsf{Sat}_{\mathbf{R}_1}(M')$. Similarly, we get $\{ba, bab, aab\} \in \mathsf{Sat}_{\mathbf{R}_1}(M')$. The intuition here is that $\mathbf{R}_2$ can distinguish the possible beginnings of words in the respective languages, but no other details. Analogously, $\mathbf{L}_2$ can distinguish the possible endings, and we get $\{ab, bab, aab\}, \{aba, ba, aab\} \in \mathsf{Sat}_{\mathbf{L}_1}(M')$.

Following this intuition, one can consider the $\mathbf{RL}_1$-factors to be the sets where we can distinguish both the beginning and the end. Indeed, calculating the $\mathbf{RL}_1$-factors yields (apart from the singletons) the sets $\{aba, aab\}, \{ab, aab\}, \{ba, aab\}$ and $\{bab, aab\}$. A straightforward calculation shows that the $\mathbf{RL}_m$ factors are exactly the $\mathbf{RL}_1$-factors independent of m.

---

[9]See. [2]. We reprove it in order to get a separator defined using rankers.

a The monoid M

b The monoid $M' = M \times J_{\{a,b\}}$

Figure 2

Since the products in Definition 12 are also valid for $n = 1$, these $\mathbf{RL}_m$-factors are themselves pointlike for all levels above $\mathbf{R}_2 \vee \mathbf{L}_2$. Furthermore, we note that

$$\{aba, aab\}\, M'\, \{ab, aab\} = \{ab, aab\}$$

and similarly for the other combinations. Thus the $\mathbf{RL}_m$-factors are the only pointlikes for all levels above $\mathbf{R}_m \vee \mathbf{L}_m$.

Using Lemma 9, we see that there is a surjective morphism $\pi : M' \to M$ such that (e.g.) $\{aba, aab\} \subseteq M'$ maps to $\{a, aab\} \subseteq M$. Thus for each level above $\mathbf{R}_2 \vee \mathbf{L}_2$, the pointlikes of M are the singletons and $\{a, aa\}$, $\{ab, aa\}$, $\{b, aa\}$ and $\{ba, aa\}$.

7 The Cases $\mathbf{Si}_1$ and $\mathbf{Pi}_1$

The closed sets for $\mathbf{Si}_1$ and $\mathbf{Pi}_1$ do not rely on alphabetic properties, and the proof that they coincide with the pointlikes uses techniques different from the other levels. Thus we devote this section to these instances. The inclusions of the sets in the conelikes are trivial.

Lemma 10. *Let M be a monoid. We have* $\mathsf{ConeSat}_{\mathbf{Si}_1}(M) \subseteq \mathsf{Cone}_{\mathbf{Si}_1}(M)$.

Proof. It is clear that the closure properties for closed sets also hold for conelikes. Thus we only need to show that $(1, S)$ is conelike for any $S \subseteq M$. This follows since $\uparrow 1 = N$ for any $N \in \mathbf{J}^+$. $\qquad\square$

Lemma 11. *Let M be a monoid. We have* $\mathsf{ConeSat}_{\mathbf{Pi}_1}(M) \subseteq \mathsf{Cone}_{\mathbf{Pi}_1}(M)$.

Proof. Analogously to the previous proof, we need only show that $(s, \{s, 1\})$ is conelike for all s. For every $N \in \mathbf{Pi}_1$, and every $x \in N$ we have $1 \in\, \uparrow x$. Thus $\{1, s\} \subseteq \tau^{-1}(\uparrow x)$ for each $x \in \tau(s)$ giving the desired result. $\qquad\square$

For the other direction, the idea is to find minimal representatives for each element in the monoid, and use knowledge about these to separate the languages. We call these representatives minors.

Definition 13. Let $\mu : A^* \to M$ be a morphism. Given $u \in A^*$, we say that $v \in A^*$ is a μ-minor of u if v is a subword of u, $\mu(u) = \mu(v)$ and v does not have any strict subwords w such that $\mu(w) = \mu(v)$. In other words, v is a subword of u minimal with respect to the subword relation such that $\mu(v) = \mu(u)$.

Lemma 12. *Let $\mu : A^* \to M$ be a morphism, and suppose that the Ramsey number of M is R. If v is a μ-minor (of some u), then $|v| \leq R - 2$.*

Proof. We argue by contradiction. Suppose $v = v_1 \cdots v_n$ is a μ-minor where $n \geq R - 1$ and each v_i is nonempty. Let $\mathcal{G} = (V, E)$ be the complete graph with $V = \{1, \cdots, n+1\}$. The word v induces an M-coloring of $\mathcal{G}$ by setting $c(\{i, i'\}) = \mu(v_i \cdots v_{i'-1})$. By Theorem 1, there is a monocrome triangle. Say for instance that $\mu(v_j \cdots v_{k-1}) = \mu(v_k \cdots v_{\ell-1}) = \mu(v_j \cdots v_{\ell-1})$. We have

$$
\begin{aligned}
\mu(v) &= \mu(v_1) \ldots \mu(v_n) \\
&= \mu(v_1) \ldots \mu(v_{j-1}) \mu(v_j \ldots v_{\ell-1}) \mu(v_\ell) \ldots \mu(v_n) \\
&= \mu(v_1) \ldots \mu(v_{j-1}) \mu(v_j \ldots v_{k-1}) \mu(v_\ell) \ldots \mu(v_n) \\
&= \mu(v')
\end{aligned}
$$

where $v' = v_1 \cdots v_{k-1} v_\ell \cdots v_n$. Since v' is a strict subword of v, we get the desired contradiction. $\square$

Lemma 13. *Let M be a monoid with generating set A, and let $n = \lceil R/2 \rceil - 1$ where R is the Ramsey number of M. If $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N_{1,n}^{XY}$ is the natural relational morphism, then $\mathsf{Cone}_\tau(M) \subseteq \mathsf{ConeSat}_{\mathbf{Si}_1}(M)$.*

Proof. Let $x \in N_{1,n}^{XY}$ and suppose that $u \in \nu^{-1}(x)$. Let $\tilde{u} = a_1 \ldots a_k$ be a μ-minor of u. Since $|\tilde{u}| \leq R - 2$ and $\tilde{u}$ is a subword of u, we must have that $\tilde{u}$ is a subword of v for any $v \in \nu^{-1}(\uparrow x)$. Thus we can factor $\nu^{-1}(\uparrow x) = U_0 a_1 U_1 \cdots a_k U_k$. We have

$$
\begin{aligned}
(\mu(u), \mu(\nu^{-1}(\uparrow x))) &= (\mu(\tilde{u}), \mu(U_0 a_1 U_1 \ldots a_k U_k)) \\
&= (1, \mu(U_0))(\mu(a_1), \{\mu(a_1)\})(1, \mu(U_1)) \cdots (\mu(a_k) \{\mu(a_k)\}) (1, \mu(U_k)) \\
&\in \mathsf{ConeSat}_{\mathbf{Si}_1}(M)
\end{aligned}
$$

giving the desired result. $\square$

Lemma 14. *Let M be a monoid with generating set A, and let $n = \lceil R/2 \rceil - 1$ where R is the Ramsey number of M. If $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N_{1,n}^{YX}$ is the natural relational morphism, then $\mathsf{Cone}_\tau(M) \subseteq \mathsf{ConeSat}_{\mathbf{Pi}_1}(M)$.*

Proof. Note that $N_{1,n}^{YX}$ is $N_{1,n}^{XY}$ with the order reversed. In other words, if $u, v \in A^*$, then $\nu(u) \leq \nu(v)$ if and only if every subword of length $2\lceil R/2 \rceil - 2$ in v is also in u.

Let $s = \mu(u)$ where $u = a_1 \ldots a_n$ for some $a_i \in A^*$. Then

$$
(\mu(a_1), \{1, \mu(a_1)\}) \ldots (\mu(a_n), \{1, \mu(a_n)\}) = (s, S) \in \mathsf{ConeSat}_{\mathbf{Pi}_1}(M)
$$

where

$$
S = \{t \in M \mid t = \mu(v), \ v \text{ is a subword of } u\}.
$$

We claim that every conelike with respect to τ is contained in such a pair. Indeed, let $x = \nu(u)$, and let $w \in \nu^{-1}(\uparrow x)$. Let $\tilde{w}$ be a μ-minor of w. Since $\tilde{w}$ is a subword of v it is also a subword of u and thus $\mu(w) = \mu(\tilde{w}) \in S$. Since u was arbitrary, the result follows. $\square$

8 From Saturations to Conelikes

In this section, we prove that all sets in the sets of subsets in Theorem 2 are pointlikes and conelikes respectively. The characterization of the pointlikes for $\mathbf{J}_1$ follows immediately using Lemma 3. Because it is trivial, we show both directions directly.

Proof of Theorem 2 (iii). The set of pointlikes contains the singletons and is closed under multiplication and subsets. Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N$ where A is the alphabet corresponding to the content morphism of M. By Lemma 3, we need only consider relational morphisms of this form. Furthermore, we can assume that N has a content morphism $\beta : N \to J_A$. Indeed, we can if necessary consider $\tau' : M \xleftarrow{\mu} A^* \xrightarrow{\nu'} N \times J_A$ where $\nu'(a) = (\nu(u), \mathsf{alph}(u))$ for all $u \in A^*$. It is clear that if $S \subseteq M$ is pointlike with respect to τ', it is also pointlike with respect to τ.

Let $S \in \mathsf{Sat}_{\mathbf{J}_1}(M)$. Since $\alpha(s) = \alpha(t)$ for all $s, t \in S$, we have $\beta(x) = \beta(y)$ for all $x \in \tau(s)$, $y \in \tau(t)$. In other words, x and y are generated by the same elements in A^*. In $\mathbf{J}_1$, this implies that they are indeed the same, and thus S is pointlike with respect to τ. For the other direction, we note the trivial inclusions $\mathsf{PL}_{\mathbf{J}_1}(M) \subseteq \mathsf{PL}_\alpha(M) = \mathsf{Sat}_{\mathbf{J}_1}(M)$. $\square$

Next, we consider our final special case, that of $\mathbf{J}$. The idea here is that inside $\mathbf{J}$, any two idempotents with the same alphabet are the same element. Furthermore, any element which has an idempotent as factor is itself idempotent.

Lemma 15. *Let M be a monoid with a content morphism $\alpha : M \to J_A$. Then $\mathsf{Sat}_{\mathbf{J}}(M) \subseteq \mathsf{PL}_{\mathbf{J}}(M)$.*

Proof. We use induction over the construction of sets in $\mathsf{Sat}_{\mathbf{J}}(M)$. Similar to the proof of Lemma 16, all we need to show is that $XEY \in \mathsf{PL}_{\mathbf{J}}(M)$ where X, Y and E has the properties in the definition of $\mathsf{Sat}_{\mathbf{J}}(M)$.

Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N \in \mathbf{J}$ be a relational morphism. As in the proof of Theorem 2 *(iii)*, we can assume that there is a content morphism $\beta : N \to J_A$. Let R be the Ramsey number of N and consider $u \in E$. Since E is idempotent, there exists a factorisation $u = u_1 \dots u_{R-1}$ where $u_i \in E$. In particular, $\alpha(u_i) = \alpha(u)$ for each i. Choose $v_i \in \tau(u_i)$. By an argument similar to that in Lemma 12, there exists a factor $e = v_i \dots v_j$ which is idempotent. We have

$$(v_1 \dots v_{i-1} e v_{j+1} \dots v_R)(v_1 \dots v_{i-1} e v_{j+1} \dots v_R) = v_1 \dots v_{i-1} e v_{j+1} \dots v_R$$

by Lemma 2. Thus $e' = v_1 \dots v_{i-1} e v_{j+1} \dots v_R \in \tau(u)$ is also idempotent.

Let $u' \in E$, and let $f \in \tau(u')$ be an idempotent obtained as above. Let $x \in X$, $y \in Y$ with $\hat{x} \in \mu^{-1}(x)$, $\hat{y} \in \mu^{-1}(y)$. Again by Lemma 2, we get $\nu(\hat{x})fe'f\nu(\hat{y}) = \nu(\hat{x})f\nu(\hat{y})$ and $e'\nu(\hat{x})f\nu(\hat{y})e' = e'$ showing that $\nu(\hat{x})f\nu(\hat{y}) \mathcal{J} e'$ which implies $\nu(\hat{x})f\nu(\hat{y}) = e'$ by $\mathcal{J}$-triviality. Since x, u' and y were arbitrary in their respective sets, we get $XEY \subseteq \tau^{-1}(e')$ showing that XEY is pointlike. $\square$

We continue with the inductive cases. The varieties $\mathbf{R}_m$ and $\mathbf{L}_m$ are handled first since the subsequent cases depend on these results.

Lemma 16. *Let M be a monoid with a content morphism $\alpha : M \to J_A$, and let $m \geq 2$. Then $\mathsf{Sat}_{\mathbf{R}_m}(M) \subseteq \mathsf{PL}_{\mathbf{R}_m}(M)$ and $\mathsf{Sat}_{\mathbf{L}_m}(M) \subseteq \mathsf{PL}_{\mathbf{L}_m}(M)$.*

Proof. We will proceed by induction over m and by structural induction over the construction of sets in $\mathsf{Sat}_{\mathbf{R}_m}(M)$ and $\mathsf{Sat}_{\mathbf{L}_m}(M)$. We only show the result for $\mathsf{Sat}_{\mathbf{R}_m}(M)$; the result for $\mathsf{Sat}_{\mathbf{L}_m}(M)$ follows by symmetry. Since pointlikes are closed under subsets and multiplication, all we need to show is that if $U \in \mathsf{Sat}_{\mathbf{R}_m}(M)$ is idempotent and pointlike with respect to $\mathbf{R}_m$, $Z \in \mathsf{Sat}_{\mathbf{L}_{m-1}}(M)$ is pointlike with respect to $\mathbf{L}_{m-1}$ and $\alpha(Z) \leq \alpha(U)$ then UZ is pointlike with respect to $\mathbf{R}_m$

Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N \in \mathbf{R}_m$ be a relational morphism, and suppose that $\beta : N \to J_A$ is a content morphism. As by the previous proofs, this is all relational morphisms which we need to consider.

By structural induction, we have $x \in N$ witnessing that U is pointlike. Since U is idempotent, have $\tau^{-1}(x^{\omega_N}) \supseteq (\tau^{-1}(x))^{\omega_N} \supseteq U^{\omega_N} = U$ and thus we can assume that the witness x is idempotent.

By induction over m, we have an equivalence class Y over $\sim_{\mathbf{K}}$ such that $Z \subseteq \tau^{-1}(Y)$. Let $s \in U$, $t \in \tau^{-1}(y)$ where $y \in Y$. We have $\beta(y) = \alpha(t) \leq \alpha(s) = \beta(x)$. By Lemma 2, we have $xyx = x$ and since $y_i \sim_{\mathbf{K}} y$ for all $y_i \in Y$ it follows that $xy_i = xy$ for all such y_i. Hence UZ is pointlike with xy as a witness. $\square$

For the other levels, we recall that the introduced closed sets build on $\mathbf{RL}_m$ factors. Our first step is to show that these are pointlike with respect to $\mathbf{R}_{m+1} \cap \mathbf{L}_{m+1}$ which in particular implies that they are pointlike with respect to $\mathbf{Si}_m$, $\mathbf{Pi}_m$ and $\mathbf{R}_m \vee \mathbf{L}_m$.

Lemma 17. *Let M be a monoid with a content morphism $\alpha : M \to J_A$, and let $U = XEWFY$ be an $\mathbf{RL}_m$-factor of M. Then $U \in \mathsf{PL}_{\mathbf{R}_{m+1} \cap \mathbf{L}_{m+1}}(M)$ with an idempotent witness.*

Proof. Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N \in \mathbf{R}_{m+1} \cap \mathbf{L}_{m+1}$ be a relational morphism and suppose $\beta : N \to J_A$ is a content morphism. Since $N \in \mathbf{L}_{m+1}$ and $E \in \mathsf{Sat}_{\mathbf{R}_m}(M)$, there exists $D \subseteq N$ such that D is a conjugacy class over $\sim_{\mathbf{D}}$ and $E \subseteq \tau^{-1}(D)$. We let

$$X = \left\{ x = x_1 \cdots x_n \mid n \geq 1, x_i \in D, x^2 = x \right\}.$$

Note that X is also a conjugacy class over $\sim_{\mathbf{D}}$. Indeed, if $x, x_i \in X$, then $x^{n\omega_N} \sim_{\mathbf{D}} (x_1 \cdots x_n)^{\omega_N}$ by stability. Let $s \in \tau^{-1}(D)$. We have $s = s_1 \cdots s_n$ where n is arbitrary and $s_i \in E$. Using a Ramsey argument similar to that of Lemma 12, we get an idempotent factor $s' = s_j \cdots s_k$ of s. Since $s_i \in E$ and E is idempotent, it follows that $s' \in E$. Thus there exist $x \in \tau(s') \cap D$. We get $x^{\omega_N} \in \tau(s')^{\omega_N} \subseteq \tau(s')$. Furthermore, let $x_i \in \tau(s_i)$. Since $\beta(x_i) = \beta(x^{\omega_N})$, it follows from Lemma 2 that

$$x_1 \cdots x_{j-1} x^{\omega_N} x_{k+1} \cdots x_n x_1 \cdots x_{j-1} x^{\omega_N} x_{k+1} \cdots x_n = x_1 \cdots x_{j-1} x^{\omega_N} x_{k+1} \cdots x_n$$

Thus $x_1 \cdots x_{j-1} x^{\omega_N} x_{k+1} \cdots x_n \in \tau(s)$ is idempotent. Since s was arbitrary, we have $E \subseteq \tau^{-1}(X)$. We also get a conjugacy class P over $\sim_{\mathbf{D}}$ such that $S \subseteq \tau^{-1}(P)$.

Similarly, we get a conjugacy class Y over $\sim_{\mathbf{K}}$ such that Y contains only idempotents and such that $F \subseteq \tau^{-1}(Y)$, and we get a conjugacy class Q over $\sim_{\mathbf{K}}$ such that $T \subseteq \tau^{-1}(Q)$. Finally, let $Z = \tau(W)$. Let $x, x' \in X$, $y, y' \in Y$, $p, p' \in P$, $s, s' \in Q$ and $z, z' \in Z$. Since $\beta(p), \beta(q), \beta(z) \leq \beta(x) = \beta(y)$ we get

$$pxzyq = pxyxzyq = pxyq = pxy'q' = p'x'y'q' = p'x'z'y'q'$$

where we used Lemma 2 and the fact that $x \sim_{\mathbf{D}} x'$, $y \sim_{\mathbf{K}} y'$, $p \sim_{\mathbf{D}} p'$ and $q \sim_{\mathbf{K}} q'$ Since these choices were arbitrary, we get that $PXZYQ$ contains a single element, showing that $SEWFT$

is pointlike with the unique element of $PXZYQ$ as witness. Furthermore, again by Lemma 2, $pxzyqpxzyq = pxzyq$, so the witness is idempotent. $\square$

We are now ready to prove inclusion of the introduced closed sets in the pointlikes/conelikes. We give the Lemma only for $\mathbf{Si}_m$; analogous lemmas for the other levels follows analogously.

Lemma 18. *Let M be a monoid with a content morphism $\alpha : M \to J_A$ and let $m \geq 2$. Then* $\mathsf{ConeSat}_{\mathbf{Si}_m}(M) \subseteq \mathsf{Cone}_{\mathbf{Si}_m}(M)$.

Proof. Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N \in \mathbf{Si}_m$ be a relational morphism and let $\beta : N \to J_A$ be a content morphism. We need to show that

$$(u_1, U_1)(v_1, V_1)(u_2, U_2) \ldots (v_{n-1}, V_{n-1})(u_n, U_n) \in \mathsf{Cone}_\tau(M)$$

where U_i is an $\mathbf{RL}_m$-factor, $u_i \in U_i$, $(v_i, V_i) \in \mathsf{ConeSat}_{\mathbf{Si}_{m-1}}(M)$ and $\alpha(v_i) \leq \alpha(U_i), \alpha(U_{i+1})$ for all $v_i \in V_i$. We proceed by induction on m.

Since $N \in \mathbf{Si}_m$, it follows in particular that N with the order removed is in $\mathbf{R}_{m+1} \cap \mathbf{L}_{m+1}$. Thus Lemma 17 shows that for all $1 \leq i \leq n$, there exists e_i such that $U_i \subseteq \tau^{-1}(e_i)$. By induction, we have that for all $1 \leq i \leq n-1$ there exists an element $z_i \in \tau(v_i)$ and a set $Z_i \subset N$ such that $V_i \subseteq \tau^{-1}(Z_i)$ and $z_i \preceq_{\mathbf{KD}} z_i'$ for all $z_i' \in Z_i$.

Let $x' \in U_1 V_1 U_2 \ldots V_{n-1} U_n$, and let $z_i' \in Z_i$ such that $x' \in \tau^{-1}(e_1 z_1' e_2 \ldots e_{n-1} z_{n-1}' e_n)$. Since $\beta(z_i') \leq \beta(e_i), \beta(e_{i+1})$ it follows from Lemma 2 that $e_i z_i' e_i = e_i$ and $e_{i+1} z_i' e_{i+1} = e_{i+1}$. It follows from the definition of $\preceq_{\mathbf{KD}}$ that

$$
\begin{aligned}
e_1 z_1 e_2 z_2 \ldots z_{n-1} e_n &\leq e_1 z_1' e_2 z_2 \ldots z_{n-1} e_n \\
&\leq e_1 z_1' e_2 z_2' \ldots z_{n-1} e_n \\
&\;\;\vdots \\
&\leq e_1 z_1' e_2 z_2' \ldots z_{n-1}' e_n.
\end{aligned}
$$

Since x' was arbitrary, this shows that $U_1 V_1 \ldots U_{n-1} V_{n-1} U_n$ is conelike with the witness $e_1 z_1 e_2 z_2 \ldots z_{n-1} e_n$. $\square$

9 On the Structure of Comparison Definable Sets

In this section, we explore some structural properties of particular sets defined by ranker comparisons. We will consider $\mathrm{RR}_{1,n} = \mathrm{XY}_{2,n} \cap \mathrm{YX}_{2,n} = \{(r, s) \in R_1 \times R_1 \mid |r|, |s| \leq n\}$, i.e. all comparisons where the rankers has alternation depth 1 and length n. The properties of these sets are used in subsequent sections to show that some $\mathrm{XY}_{m,n}$ or $\mathrm{YX}_{m,n}$-sets are in the relevant closed set. We also consider a special case which does not contain $\mathrm{RR}_{1,n}$, namely $\mathrm{XY}_{1,n} \cup \mathrm{YX}_{1,n}$.

We will show how these languages consist of certain *long* factors, intuitively factors which are long enough so that the relevant rankers can not see the whole factor. We use the following notation: If $u = a_1 \ldots a_n$ and $1 \leq i \leq j \leq n$, then $u[i] = a_i$, and $u[i,j] = a_i \ldots a_{j-1}$. Note in particular that $u[i,i]$ is the empty factor.

Definition 14. Let $u \in A^*$. We say that a factor $u[i,j]$ is *n-long* if every word $v \in \mathsf{alph}(u[i,j])^*$ such that $|v| \leq n$ is a subword of $u[i,j]$. Note in particular that this definition vacuously implies that empty factors $u[i,i]$ are *n*-long. Given a word u, we call a factor $u[i,j]$ *maximal n-long* if it is *n*-long and for every *n*-long $u[i',j']$ such that $i' \leq i \leq j \leq j'$ we have $i' = i$ and $j' = j$.

In other words, an n-long factor is maximal if it is not properly contained in any other n-long factor.

Definition 15. Let $u \in A^*$. We define

$$\mathsf{R}_n(u) = \{0\} \cup \{i \in \mathbb{N} \mid u[j, i) \text{ is a maximal } n\text{-long factor for some } j\}$$

Note that since the factor $u[j, i)$ does not contain the position i, the set $\mathsf{R}_u(u)$ consists of the positions just after some maximal n-long factor. Symmetrically, we define

$$\mathsf{L}_n(u) = \{|u| + 1\} \cup \{i \in \mathbb{N} \mid u[i + 1, j) \text{ is a maximal } n\text{-long factor for some } j\}.$$

Note that the maximality of $u[j, i)$ implies that $u[i] \notin \mathsf{alph}(u[j, i))$ in the definition of $\mathsf{R}_n(u)$ and symmetrically $u[i] \notin \mathsf{alph}(u[i + 1, j))$ in the definition of $\mathsf{L}_n(u)$. For $n = 1$, any factor is long, and thus $\mathsf{R}_n(u) = \mathsf{L}_n(u) = \{0, |u| + 1\}$. In this case, the following results are trivial, and in order to avoid dealing explicitly with it, we always assume $n \geq 2$.

We want to show that in between each pair of positions in $\mathsf{R}_n(u)$, there is a position in $\mathsf{L}_n(u)$ (possibly coinciding with one of the positions in $\mathsf{R}_n(u)$) and vice versa. First, we prove the following Lemma.

Lemma 19. *Suppose* $\{i, i + 1, \ldots, j - 1, j\} \cap \mathsf{R}_n(u) > 1$, *then* $u[i + 1, j + 1)$ *is not n-long.*

Proof. For every n-long factor, we can consider the maximal n-long factor containing it. Thus, we lose no generality in proving the statement only for maximal n-long factors. Assume $k, \ell \in \mathsf{R}_n(u)$ with $i \leq k < \ell \leq j$ and assume $u[i + 1, j + 1)$ is maximal n-long. We have a maximal n-long $u[\ell', \ell)$. Since $\ell < j + 1$, we must have $\ell' < i + 1$; otherwise $u[\ell', \ell)$ would be properly contained in $u[i + 1, j + 1)$ and thus could not be maximal n-long. In particular, this leads to a contradiction if $i = 0$, since $\ell' < 1$ is impossible.

If $i \geq 1$, then there exists a maximal n-long $u[k', k)$. In particular, $k' < \ell'$. Since $u[k] \in \mathsf{alph}(u[\ell', \ell)) \setminus \mathsf{alph}(u[k', k))$ we must have $\mathsf{alph}(u[k + 1, \ell)) = \mathsf{alph}(u[\ell', \ell))$, since the subword $u[k]q$ with $q \in \mathsf{alph}(u[\ell', \ell))^{n-1}$ must exist in $u[\ell', \ell)$. Thus $\mathsf{alph}(u[\ell', \ell)) \subseteq \mathsf{alph}(u[i + 1, j + 1))$. Since these factors intersect and are maximally n-long, they must coincide which contradicts $\ell \leq j$. $\square$

The above Lemma says that even if maximal n-long factors are not necessarily disjoint (note e.g. $(ab)^n a(ac)^n$), no n-long factor can cover more than one R_n-marker.[10] This gives us the following Lemma.

Lemma 20. *Let* $i, j \in \mathsf{R}_n(u)$ *with* $i < j$. *Then there exists* $k \in \mathsf{L}_n(u)$ *such that* $i < k \leq j$. *Symmetrically, if* $i, j \in \mathsf{L}_n(u)$ *with* $i < j$, *then there exists* $k \in \mathsf{R}_n(u)$ *with* $i \leq k < j$.

Proof. Let $a[k + 1, \ell)$ be a maximal n-long factor containing $a[j + 1, j + 1)$. By definition, $k \leq j$ and by Lemma 19, we must have $i < k$, giving the desired result. The other direction is symmetrical. $\square$

We will now turn to the factorizations of $\mathrm{RR}_{1,n}$-sets. We start with factorizing a single word into the desired form.

---

[10]In fact, it says something slightly stronger. Even if we include the position just before an n-long factor, this set of positions can not cover more than one R_n-marker.

Lemma 21. *Given $u \in A^*$ and $n \geq 1$, there is a factorization*

$$u = u_1 b_1 v_1 a_1 \ldots u_{k-1} b_{k-1} v_{k-1} a_{k-1} u_k, \tag{2}$$

where v_i can be empty, a_i and b_i can coincide, such that for all i

(i) u_i is n-long,

(ii) $\mathsf{alph}(v_i) \subsetneq \mathsf{alph}(u_i b_i) \cap \mathsf{alph}(a_i u_{i+1})$,

(iii) a_i is reachable by a ranker X_{p_i} and b_i by a ranker Y_{q_i} where $|p_i|, |q_i| \leq (n+1)|A \setminus \mathsf{alph}(v_i)|$,

(iv) either the markers a_i and b_i coincide or $a_i \in \mathsf{alph}(u_{i+1})$, $b_i \in \mathsf{alph}(u_i)$.

Proof. By Lemma 20, we can choose a_i to mark the positions from $\mathsf{R}_{n+2}(u)$ and b_i to mark the positions from $\mathsf{L}_{n+2}(u)$, and get a factorization with the markers a_i and b_i interlaced, or possibly coinciding, as in (2). We want to show that conditions (i - iv) hold for this factorization.

Let i be fixed. Since u_i contains no position in $\mathsf{R}_{n+2}(u)$ or $\mathsf{L}_{n+2}(u)$, it is contained in some maximal $(n+2)$-long factor, say w_i. By the definition of $\mathsf{R}_{n+2}(u)$ and $\mathsf{L}_{n+2}(u)$, the marker b_{i-1} must be directly to the left of w_i, and the marker a_i need to be directly to the right of w_i. In other words, we have one of the following cases:

(a) $w_i = u_i$, if b_{i-1} coincides with a_{i-1} or $i = 1$ and b_i coincides with a_i or $i = k$,

(b) $w_i = u_i b_i v_i$, if b_{i-1} and a_{i-1} coincide or $i = 1$ while a_i and b_i do not coincide,

(c) $w_i = v_{i-1} a_{i-1} u_i$, if b_i and a_i coincide or $i = k$ while a_{i+1} and b_{i+1} does not coincide,

(d) $w_i = v_{i-1} a_{i-1} u_i b_i v_i$ otherwise.

We will show properties *(i),(ii)* and *(iv)* for case *(d)*. The others are handled similarly. Since $1 < i$, we have that w_{i-1} exist. Since it is maximal $(n+2)$-long, we must have $a_{i-1} \notin \mathsf{alph}(w_{i-1}) \supseteq \mathsf{alph}(v_{i-1})$. In particular, it follows from the fact that w_i is $(n+2)$-long that $u_i b_i v_i$ is $(n+1)$-long with $\mathsf{alph}(w_i) = \mathsf{alph}(u_i b_i v_i)$. Applying the same argument from the right, shows that u_i is n-long with $\mathsf{alph}(u_i) = \mathsf{alph}(w_i)$. This shows *(i)* and *(iv)*. Furthermore, the fact that $b_i \notin \mathsf{alph}(w_{i+1}) \supseteq \mathsf{alph}(v_i)$ shows condition *(ii)*.

Thus, all that is left is to show condition *(iii)*. By symmetry, we only show that a_i is reachable by a ranker X_{p_i} with $|p_i| \leq (n+1)|A \setminus \mathsf{alph}(v_i)|$. We will use induction on the size of the alphabet.

If $A = \{a\}$ then either $u = u_1$ or $u = a^m$ where $m \leq n+1$. In the former case, there are no markers a_i, and thus nothing to show. In the latter case, we note that every position is both an a_i and a b_i marker. Thus, every v_i is empty. If a_i is the ℓth position of u, then X_{a^ℓ} is a ranker of the desired length.

Let $|A| \geq 2$ and let c mark the rightmost position which is not to the right of a_i and which is reachable by a ranker X_p where $|p| \leq n+1$. Let a_j be the leftmost marker occurring after or at c, let s be the factor of u up to the marker c and let t be the factor between c and a_j. We define $u' = t a_j \ldots u_i b_i v_i$.

Note that $a_i \notin \mathsf{alph}(w_i)$, $b_i \notin \mathsf{alph}(w_{i+1})$. If these markers do not coincide, then $a_i \in \mathsf{alph}(w_{i+1})$ and $b_i \in \mathsf{alph}(w_i)$, showing that $\mathsf{alph}(v_i) \leq |A| - 2$. Furthermore, if a_i and b_i do coincide, then v_i is empty. Since $|A| \geq 2$, we have $|A \setminus \mathsf{alph}(v_i)| \geq 2$ in both cases. In particular, if $a_i \notin \mathsf{alph}(u')$, then the ranker $\mathsf{X}_{p a_i}$ has the desired properties.

Next, let us assume that $a_i \in \mathsf{alph}(u')$. In order to use induction, we need to show that $\mathsf{alph}(u') \subsetneq \mathsf{alph}(u)$. Since $a_i \in \mathsf{alph}(u')$ and $a_i \notin \mathsf{alph}(w_i)$, the factor u' must begin before the factor w_i. In particular, this means that scu' is not $(n+2)$-long (since w_i is maximally $n+2$-long). Thus, there exists $r \in A^*$, $d \in A$ with $|r| \le n+1$ such that $\mathsf{X}_r(u) \le |scu'|$ and either $|scu'| < \mathsf{X}_{rd}(u)$ or X_{rd} is undefined on u. By the definition of p, we have, $\mathsf{X}_r(u) \le \mathsf{X}_p(u) = |sc|$. Thus, $d \notin \mathsf{alph}(u')$.

Since w_i is contained in u, it must be maximal $(n+2)$-long in u'. Since $a_i \notin \mathsf{alph}(w_i)$, the word w_i is maximally $(n+2)$-long also in $u'a_i$. It follows that a_i marks a position in $\mathsf{R}_{n+2}(u'a_i)$. By induction, there exists $\mathsf{X}_{p'}$ where $|p'| \le (n+1)|\mathsf{alph}(u') \setminus \mathsf{alph}(v_i)|$ such that $\mathsf{X}_{p'}(u'a_i) = |u'a_i|$. Thus the ranker $\mathsf{X}_{pp'}$ has the desired properties. $\quad\square$

Thus, we can obtain a factorization with some desired properties for each word u. The following proposition shows that these factorizations can be combined to give a factorization of a $\mathrm{RR}_{m,(n+1)|A|+n}$-set.

Lemma 22. *Let $n \in \mathbb{N}$ and let $U \in A^*$ be a $\mathrm{RR}_{m,(n+1)|A|+n}$-set. Then it is possible to find a factorization*

$$U = U_1 b_1 V_1 a_1 \dots U_{k-1} b_{k-1} V_{k-1} a_{k-1} U_k, \tag{3}$$

where

(i) Each U_i is n-long

(ii) $\mathsf{alph}(V_i) \subsetneq \mathsf{alph}(U_i b_i) \cap \mathsf{alph}(a_i U_{i+1})$,

(iii) For each $u \in U$, the position marked by a_i is reachable by X_{p_i} and the position marked by b_i is reachable by Y_{q_i} where $|p_i|, |q_i| \le (n+1)|A \setminus \mathsf{alph}(V_i)|$,

(iv) For each i, either the markers a_i and b_i coincide, or $a_i \in \mathsf{alph}(U_{i+1})$, $b_i \in \mathsf{alph}(U_i)$.

Proof. Suppose $u \in U$ and factor it as Lemma 21. We note that $\mathsf{Y}_{q_i}(u) \le \mathsf{X}_{p_i}(u) < \mathsf{Y}_{q_{i+1}}(u)$ which implies that the same must be true on u'. In particular, this gives the factorization

$$u' = u_1' b_1 v_1' a_1 \dots u_{k-1}' b_{k-1} v_{k-1}' a_{k-1} u_k'$$

We combine these factorizations for all $u' \in U$ to get a factorization of the form in (3). This factorization satisfies *(iii)* by definition.

Let $1 \le i \le k-1$ and let $r \in \mathsf{alph}(u_i)^*$ with $|r| \le n$. Since every u_i in the factorization of u is n-long, we have that $\mathsf{X}_{p_{i-1}r}(u) < \mathsf{Y}_{q_i}(u)$ (if $i=1$ we set $p_{i-1} = \varepsilon$). Thus the same is true for u'. Furthermore, if $a \notin \mathsf{alph}(u_i)$, then either $\mathsf{X}_{p_{i-1}a}$ is not defined on u or $\mathsf{Y}_{q_i}(u) \le \mathsf{X}_{p_{i-1}a}(u)$. Again, the same is true on u'. It follows that $\mathsf{alph}(u_i) = \mathsf{alph}(u_i')$ and u_i' is n-long. For $i = k$, we consider the rankers X_{p_k} and $\mathsf{Y}_{\overline{r}}$ and get the same result. Hence the factorization satisfies *(i)*.

Next, let $1 \le i < k-1$. We have that a_i and b_i coincide on u if and only if $\mathsf{X}_{p_i}(u) \le \mathsf{Y}_{q_i}(u)$ and $\mathsf{Y}_{q_i}(u) \le \mathsf{X}_{p_i}(u)$. Thus, this is true on u if and only if it is true on u'. If a_i and b_i does not coincide, we get $a_i \in \mathsf{alph}(u_{i+1}) = \mathsf{alph}(U_{i+1})$ and $b_i \in \mathsf{alph}(u_i) = \mathsf{alph}(U_i)$ by Lemma 21. Thus the factorization satisfy *(iv)*.

If a_i and b_i does not coincide and $c \in \mathsf{alph}(v_i') \subseteq \mathsf{alph}(u_i' b_i v_i')$, then $\mathsf{X}_{p_i c}(u') < \mathsf{X}_{p_{i+1}}(u')$ which is also satisfied on u. Thus $c \in \mathsf{alph}(u_i b_i v_i)$, and by property *(ii)* of Lemma 21, we get $c \in \mathsf{alph}(u_i) = \mathsf{alph}(U_i)$. Symmetrically, we get that if $c \in \mathsf{alph}(v_i') \subseteq \mathsf{alph}(v_i' a_i u_{i+1}')$, then $c \in \mathsf{alph}(U_{i+1})$. This shows *(ii)*. $\quad\square$

We also provide the following factorization which deals with the $\mathrm{XY}_{1,n} \cup \mathrm{YX}_{1,n}$-case.

Lemma 23. *Given an alphabet A and integer n, and an $\mathrm{XY}_{1,(n+1)|A|+n} \cup \mathrm{YX}_{1,(n+1)|A|+n}$-set U, it is possible to find a factorization*

$$U = U_1 \lambda_1 U_2 \ldots \lambda_{k-1} U_k \lambda_k U_{k+1}$$

where $\lambda_i \in A \cup \{\varepsilon\}$ i.e. the λ_i are either letters or empty, and where each U_i is n-long.

Proof. Choose $u \in U$ and factor it according to Lemma 21. Let X_{p_i} and Y_{q_i} be the rankers from the Lemma, and let $u' \in U$. We factor

$$u' = u'_1 b_1 v'_1 a_1 \ldots u'_{k-1} b_{k-1} v'_{k-1} a_{k-1} u'_k$$

where a_i marks the position of X_{p_i} and b_i marks the position of Y_{q_i}. For each i we have words p_i and $\overline{q}_{i+1}$ such that a_i marks the last letter of the first occurrence of p_i and b_{i+1} marks the first letter of the last occurrence of $\overline{q}_i$. We set $p_0 = \varepsilon$ and $\overline{q}_k = \varepsilon$.

For $1 \le i < k - 1$, suppose $p_i = p'_i a_i$ and $q_i = b_i q'_i$. We have that a_i and b_i coincide if and only if they are the same letter, $p'_i a_i q'_i$ is a subword of u and $p'_i b_i a_i q'_i$ is not a subword of u. In particular, a_i and b_i coincide on u' if and only if they coincide on u.

Let $1 \le i \le k$ and let $r \in \mathsf{alph}(u_i)^*$ with $|r| \le n$. Since every u_i in the factorisation of u is n-long, we have that $p_{i-1} r \overline{q}_i$ is a subword of u. Thus the same is true for u'. Furthermore, if $a \notin \mathsf{alph}(u_i)$, then $p_{i-1} a \overline{q}_i$ is not a subword of u and hence neither of u'. It follows that $\mathsf{alph}(u_i) = \mathsf{alph}(u'_i)$ and u'_i is n-long.

Suppose that a_i and b_i does not coincide. We factor $a_i v'_i b_i = s_i t_i$ where s_i is the longest factor such that $\mathsf{alph}(s_i) \subseteq \mathsf{alph}(u_i)$. Let c be the first letter of t_i, and let $d \in \mathsf{alph}(u_i) \setminus \mathsf{alph}(u_{i+1})$. By the alphabetic requirements, we have that $p_{i-1} c d \overline{q}_i$ can not be a subword of u, and thus neither of u'. Hence $\mathsf{alph}(t_i) \subseteq \mathsf{alph}(u_{i+1})$.

If a_i and b_i do coincide, we put $\lambda_i = a_i$, and $s_i = t_i = \varepsilon$. This way, we get a factor $t_{i-1} u'_i s_i$ from each $u' \in U$ and each i. Setting U_i to be the union of these factors gives us a factorisation of the desired form. $\square$

10 From Conelikes to Saturations

We now have everything we need to show that we can find monoids in the relevant varieties whose pointlikes/conelikes coincide with the introduced closed sets. Since the pointlikes/conelikes with respect to the variety is contained in the pointlikes/conelikes with respect to any particular monoid in the variety, this gives our desired result.

We start with the variety $\mathbf{J}$, where the result is obtained by two lemmas. First, we show that anything $(R-1)$-long is in $\mathsf{Sat}_{\mathbf{J}}(M)$. The result then follows trivially using the factorization in Lemma 23.

Lemma 24. *Let $\mu : A^* \to M$ be a homomorphism, and let R be the Ramsey number of M. If U is $(R-1)$-long, then $\mu(U) \in \mathsf{Sat}_{\mathbf{J}}(M)$.*

Proof. Let $u \in U$. Since u is $(R-1)$-long, we can factor $u = u_1 \ldots u_{R-1}$ where $\mathsf{alph}(u_i) = \mathsf{alph}(U)$. By an argument similar to that of Lemma 12, we find a factor $u_i \ldots u_j$ which is idempotent. Setting $e = u_i \ldots u_j$, we get a factorisation $u = xey$. Combining these factorisations for all $u \in U$, we get $U = XEY$ satisfying the alphabetic properties.

Let ω be the idempotent power of 2^M. Since every element in E is idempotent, we have $XEY \subseteq XE^\omega Y$, where E^ω is clearly idempotent. $\square$

Lemma 25. *Let M be a monoid and let R be the Ramsey number of M. Let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N_{1,n}^{\mathrm{XY} \cup \mathrm{YX}} \in \mathbf{J}$ with $n = |A|R + R - 1$. Then $\mathsf{PL}_\tau(M) \subseteq \mathsf{Sat_J}(M)$.*

Proof. Let $s \in N_{1,n}^{\mathrm{XY} \cup \mathrm{YX}}$. We want to show that $\mu(\nu^{-1}(s)) \in \mathsf{Sat_J}(M)$. To this end, we factor

$$\nu^{-1}(s) = U_1 \lambda_1 \ldots U_{k-1} \lambda_{k-1} U_k$$

according to Lemma 23 for $R - 1$. We recall that each a_i is either a single letter or empty. Thus, by closure under multiplication, it is enough to show that each $U_i \in \mathsf{Sat_J}(M)$. Since each U_i is $(R-1)$-long, this follows from Lemma 24. □

For the varieties $\mathbf{R}_m$, $\mathbf{L}_m$ and for the $\mathbf{RL}_m$-sets, the idea is to choose the depth of the rankers long enough so that we can pump in the monoid 2^M. It turns out that instead of asking for idempotents in 2^M, a weaker condition is sufficient: being *subidempotent*.

Definition 16. Let M be a monoid. A set $U \subseteq M$ is *subidempotent* if $U \subseteq U^2$.

We note that every subidempotent is, by definition, a subset of U^θ where $\theta = \omega_{2^M}$ is the idempotent power of 2^M. Since pointlikes and conelikes are closed under subsets, this means that any mention of idempotent in Definition 10 and 11 can be substituted with subidempotent without changing the result. Furthermore, subidempotents are easily obtained by long enough factorizations.

Lemma 26. *Let M be a monoid, and let R be the Ramsey number of the set 2^M. Let $U \subseteq A^*$ have the factorization $U = U_1 \ldots U_n$, with $n \geq R - 1$. Then there exists $1 \leq i, j \leq n$ such that $\mu(U_i U_{i+1} \ldots U_{j-1} U_j)$ is subidempotent.*

Proof. The empty set is subidempotent, and thus the statement is true for the degenerate case $U = \emptyset$. We assume $U \neq \emptyset$ which implies $U_i \neq \emptyset$ for all i.

Let $\mathcal{G} = (V, E)$ be the complete graph with vertices $V = \{1, \ldots, n+1\}$. The map μ induces an edge-coloring c of size $2^{|M|}$ on $\mathcal{G}$ by $c(\{i, j\}) = \mu(U_i \ldots U_{j-1})$. By Theorem 1, there exists a monochrome triangle. In other words, there exists i, j, k with the property that $\mu(U_i \ldots U_{j-1}) = \mu(U_j \ldots U_{j-1}) = \mu(U_i \ldots U_{k-1})$. Let us call this set S. Since $S = \mu(U_i \ldots U_{k-1}) \subseteq \mu(U_i \ldots U_{j-1})\mu(U_j \ldots U_{k-1}) = S^2$ it follows that S is subidempotent. □

We now prove the desired result for $\mathbf{R}_m$ and $\mathbf{L}_m$. By obtaining a long enough factorization, we find an idempotent using the above Lemma. Everything before the idempotent is handled using Lemma 8 and induction on the alphabet, whereas everything after is handled using the same lemma and induction on m.

Lemma 27. *Let M be a monoid with a content morphism $\alpha : M \to J_A$. Let $m \geq 1$, and let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N_{m,n}^{\mathrm{XX}}$ be the natural relational morphism where $n = (m + |A|)(R - 1)$ depends on the Ramsey number R of 2^M. Then $\mathsf{PL}_\tau(M) \subseteq \mathsf{Sat}_{\mathbf{R}_{m+1}}(M)$. Dually, if $\tau' : M \leftarrow A^* \to N_{m,n}^{\mathrm{YY}}$, then $\mathsf{PL}_{\tau'}(M) \subseteq \mathsf{Sat}_{\mathbf{L}_{m+1}}(M)$.*

Proof. By symmetry, we need only show the result for $N_{m,n}^{\mathrm{XX}}$. We show that if U is an $\mathrm{XX}_{m,n}$-set, then $\mu(U) \in \mathsf{Sat}_{\mathbf{R}_{m-1}}(M)$. In particular, this shows that $\mu(\nu^{-1}(s)) \in \mathsf{Sat}_{\mathbf{R}_{m+1}}(M)$ for all $s \in N_{m,n}^{\mathrm{XX}}$, which is the desired result. We use induction over $(m, |A|)$ ordered alphabetically. First, we show that we can factor U into one of the following forms

$$U = U_1 a_1 U_2 a_2 \cdots U_k a_k U_{k+1}$$
$$U = U_1 a_1 U_2 a_2 \cdots U_{(R-1)} a_{(R-1)} V$$

subject to the conditions

(i) $k < (R-1)$

(ii) $\mathsf{alph}(U_i) \subsetneq \mathsf{alph}(U_i a_i) = A$ for all i (if U does not contain all letters in A, then $k = 0$).

Indeed, suppose we already factored $U = U_1 a_1 \cdots U_i a_i W$ such that $i < (R-1)$ and condition *(ii)* holds. In particular, a_i is reachable by the ranker $\mathsf{X}_{a_1 \ldots a_i}$ on all $v \in U$. If $i = R-1$, we choose $V = W$ and if $\mathsf{alph}(W) \subsetneq A$, we choose $U_{k+1} = W$. In both these cases, we are done. Thus, we suppose that $\mathsf{alph}(W) = A$ and that $i < (R-1)$.

For some $u \in U$, choose the letter a_{i+1} such that $\mathsf{X}_{a_1 \ldots a_i b}(u) \le \mathsf{X}_{a_1 \ldots a_i a_{i+1}}(u)$ for all $b \in A$. It is clear that this choice is unique and since U is an $\mathrm{XX}_{m,n}$-set, it is independent of the choice of u. We factor $U = U a_1 \cdots U_i a_i U_{i+1} a_{i+1} W'$ where a_{i+1} marks the first occurrence of a_{i+1} after the position marked by a_i. By choice of a_{i+1}, we have $\mathsf{alph}(U_{i+1}) = A \setminus \{a_{i+1}\}$ and so this factorization satisfies *(ii)* and is thus a factorization of the desired form.

Since a_i is reachable by the ranker $\mathsf{X}_{a_1 \ldots a_i}$, which in particular has depth at most $R - 1$, it follows from Lemma 8 that U_i is an $\mathrm{XX}_{m,n'}$-set where $n' = (m + A - 1)(R-1) = (m + |\mathsf{alph}(U_i)|)(R-1)$. Let M_i be the submonoid of M generated by $\mathsf{alph}(U_i)$. Since the Ramsey number of 2^{M_i} is at most R, it follows from induction on the size of the alphabet that $\mu(U_i) \in \mathsf{Sat}_{\mathbf{R}_{m+1}}(M_i) \subseteq \mathsf{Sat}_{\mathbf{R}_{m+1}}(M)$. Since closed sets are closed under multiplication, we get $\mu(U_i a_i \cdots U_j a_j) \in \mathsf{Sat}_{\mathbf{R}_{m+1}}(M)$ for every factor $U_i a_i \ldots U_j a_j$ of U. In particular, this gives $\mu(U) \in \mathsf{Sat}_{\mathbf{R}_{m+1}}(M)$ when the factorization is of the first form.

For the second form of the factorization, we know by Lemma 26 that there exists $U_i a_i \cdots U_j a_j$ such that $\mu(U_i a_i \cdots U_j a_j)$ is subidempotent. By the argument above, it follows that we have $\mu(U_1 a_1 \ldots U_{i-1} a_{i-1}), \mu(U_i a_i \ldots U_j a_j) \in \mathsf{Sat}_{\mathbf{R}_{m+1}}(M)$. Let $E = \mu(U_i a_i \cdots U_j a_j)^\theta$ where $\theta = \omega_{2^M}$ is the idempotent power of 2^M. By subidempotency, we have $U_i a_i \cdots U_j a_j \subseteq E$.

Suppose $m \ge 2$. Since a_j is reachable by a ranker X_{p_i} of length at most $R - 1$, it follows by Lemma 8 that $U_{j+1} a_{j+1} \ldots U_{(R-1)} a_{(R-1)} V$ is a $\mathrm{XX}_{m,n''}$-set where $n'' = (m - 1 + |A|)(R-1)$. This in particular implies that it is a $\mathrm{YY}_{m-1,n''}$-set. If $m = 1$, we can use definedness of rankers $\mathsf{X}_{a_1 \ldots a_i b}$ for $b \in A$ to conclude that $U_{j+1} a_{j+1} \cdots U_{(R-1)} a_{(R-1)} V$ has a well defined alphabet. Thus in both cases $\mu(U_{j+1} a_{j+1} \ldots U_R a_R V) \in \mathsf{Sat}_{\mathbf{L}_m}(M)$ (in the former case by induction, and in the latter by definition). We get

$$\mu(U) \subseteq \mu(U_1 a_1 \ldots U_i a_i) \mu(U_i a_i \ldots U_j a_j) \mu(U_{j+1} a_{j+1} \ldots U_{(R-1)} a_{(R-1)} V)$$
$$\subseteq \mu(U_1 a_1 \ldots U_i a_i) E \mu(U_{j+1} a_{j+1} \ldots U_{(R-1)} a_{(R-1)} V) \in \mathsf{Sat}_{\mathbf{R}_m}(M)$$

which is the desired result. $\square$

Proof of Theorem 2 (iv) and (v). We prove the result for *(iv)*. The result for *(v)* is symmetric. From Lemma 16 it follows that $\mathsf{Sat}_{\mathbf{R}_m}(M) \subseteq \mathsf{PL}_{\mathbf{R}_m}(M)$ and from Lemma 27 it follows that $\mathsf{PL}_\tau(M) \subseteq \mathsf{Sat}_{\mathbf{R}_m}(M)$. By Proposition 6, we have $N_{m,n}^{\mathrm{XX}} \in \mathbf{R}_{m+1}$ and thus $\mathsf{PL}_{\mathbf{R}_m}(M) \subseteq \mathsf{PL}_\tau(M)$. In particular, $\mathsf{PL}_{\mathbf{R}_m}(M) = \mathsf{PL}_\tau(M) = \mathsf{Sat}_{\mathbf{R}_m}(M)$ which gives the desired result. $\square$

To prove what is left of Theorem 2, we start by showing that every long enough $\mathrm{XX}_{m-1,n} \cup \mathrm{YY}_{m-1,n}$-set maps to an $\mathbf{RL}_m$-factor. The idea is analogous to the $\mathbf{R}_m$ and $\mathbf{L}_m$ cases.

Lemma 28. *Let $\mu : A^* \to M$ be a surjective morphism, and suppose that M has a content morphism. Let R be the Ramsey number of 2^M, let $m \ge 2$ and let $n = (m - 1 + |A|)(R-1)$. If $U \subseteq A^*$ is $2(R-1)$-long and a $\mathrm{XX}_{m-1,n} \cup \mathrm{YY}_{m-1,n}$-set, then $\mu(U) \subseteq V$ where V is an $\mathbf{RL}_m$-factor.*

Proof. Since U is $2(R-1)$-long, we can factor

$$U = S_1 a_1 \ldots S_{R-1} a_{R-1} Z b_1 T_1 \ldots b_{R-1} T_{R-1}$$

where $\mathsf{alph}(S_i), \mathsf{alph}(T_i) \subsetneq \mathsf{alph}(S_i a_i) = \mathsf{alph}(b_i T_i) = \mathsf{alph}(U)$. Since each a_i is reachable by a ranker $\mathsf{X}_{a_1 \ldots a_i}$ of depth at most $R-1$, it follows from Lemma 8 that every S_i is an $\mathrm{XX}_{m-1,n'}$-set where $n' = (m + |A| - 2)(R-1) = (m - 1 + |\mathsf{alph}(S_i)|)(R-1)$. It follows from Lemma 27 that $\mu(S_i) \in \mathsf{Sat}_{\mathbf{R}_m}(M)$. In particular, this implies that any product of S_i and a_j is in $\mathsf{Sat}_{\mathbf{R}_m}(M)$. We get the symmetric result for products of T_i and b_j; they are in $\mathsf{Sat}_{\mathbf{L}_m}(M)$.

By Lemma 26, there exists $S_j a_j \ldots S_k a_k$ such that $\mu(S_j a_j \ldots S_k a_k)$ is subidempotent. We set $S = \mu(S_1 a_1 \ldots S_{j-1} a_{j-1})$ and note that $S \in \mathsf{Sat}_{\mathbf{R}_m}(M)$ by the argument above. Let $E = \mu(S_i a_i \ldots S_{i'} a_{i'})^\theta$ where $\theta = \omega_{2^M}$ is the idempotent power of 2^M. Since $\mu(S_i a_i \ldots S_{i'} a_{i'}) \in \mathsf{Sat}_{\mathbf{R}_m}(M)$ and closed sets are closed under multiplication, we get $E \in \mathsf{Sat}_{\mathbf{R}_m}(M)$. By subidempotency, it follows that $\mu(S_i a_i \ldots S_{i'} a_{i'}) \subseteq E$. Symmetrically, we find j', k' such that $\mu(b_{j'} T_{j'} \ldots b_{k'-1} T_{k'-1}) \subseteq F$ where $F \in \mathsf{Sat}_{\mathbf{L}_m}(M)$ is idempotent and such that $T = \mu(b_{k'} T_{k'} \ldots b_{R-1} T_{R-1}) \in \mathsf{Sat}_{\mathbf{L}_m}(M)$. Finally, we set $W = S_{k+1} \ldots S_R a_R Z b_1 T_1 \ldots T_{j'-1}$. It follows that $U \subseteq SEWFT$ where $SEWFT = V$ is an $\mathbf{RL}_m$-factor. $\qquad\square$

The desired result now follows using the factorizations obtained in Lemma 22. We give the result for the monoids $N_{m,n}^{\mathrm{XY}}$, but the proof generalizes directly to the other cases.

Lemma 29. *Let M be a monoid with a content morphism $\alpha : M \to J_A$. Let $m \geq 2$ and let $\tau : M \xleftarrow{\mu} A^* \xrightarrow{\nu} N_{m,n}^{\mathrm{XY}}$ where $n = (m - 1 + 3|A|)(R-1) + |A|$ depends on the Ramsey number R of 2^M. If m is odd, then $\mathsf{Cone}_\tau(M) \subseteq \mathsf{ConeSat}_{\mathbf{Si}_m}(M)$ and if m is even, then $\mathsf{Cone}_\tau(M) \subseteq \mathsf{ConeSat}_{\mathbf{Pi}_m}(M)$.*

Proof. Let $U \subseteq A^*$ and $u \in U$ be such that $u \leq_{m,n}^{\mathrm{XY}} u'$ for all $u' \in U$ (in other words, U is a $\mathrm{XY}_{m,n}$-set with minimal element u). We will prove $(\mu(u), \mu(U)) \in \mathsf{ConeSat}_{\mathbf{Si}_m}(M)$ for odd m, and $(\mu(u), \mu(U)) \in \mathsf{ConeSat}_{\mathbf{Pi}_m}(M)$ for even m. Given $x \in N$, we can choose $u \in \nu^{-1}(x)$ such that $\nu^{-1}(\uparrow x)$ is a $\mathrm{XY}_{m,n}$-set with minimal element u, and thus this implies the desired result. We will use induction over $(m, |A|)$ ordered alphabetically. We consider the cases when m is odd, with the even cases handled symmetrically.

Let $m \geq 2$. We apply Lemma 22 with $2(R-1)$ to obtain a factorization

$$U = U_1 b_1 V_1 a_1 \ldots U_{k-1} b_{k-1} V_{k-1} a_{k-1} U_k$$

with the properties specified in the theorem. In particular, this gives a factorization

$$u = u_1 b_1 v_1 a_1 \ldots u_{k-1} b_{k-1} v_{k-1} a_{k-1} u_k.$$

Every position marked by a_i or b_i in every u' can be reached by a ranker of depth at most $(2(R-1) + 1)|A \setminus \mathsf{alph}(V_i)|$. Since

$$(m - 1 + |A|)(R-1) = (m - 1 + 3|A|)(R-1) + |A| - (2(R-1) + 1)|A|$$

it follows that every U_i is an $\mathrm{XX}_{m-1,(m-1+|A|)(R-1)} \cup \mathrm{YY}_{m-1,(m-1+|A|)(R-1)}$-set. By Lemma 28 we get an $\mathbf{RL}_m$-set S_i such that $\mu(U_i) \subseteq S_i$.

Let $A_i = \mathsf{alph}(V_i) \subsetneq A_i$, let $M_i = \mu(A_i^*)$ and let $\mu_i : A_i \to M_i$ be the restriction of μ. For $m \geq 3$, we define

$$\begin{aligned} n_i &= (m - 2 + 3|A_i|)(R_i - 1) + |A_i| \\ &\leq (m - 1 + 3|A|)(R-1) + |A| - 2(R-1)|A \setminus A_i|, \end{aligned}$$

where R_i is the Ramsey number of 2^{M_i} and for $m = 2$, we define

$$n_i = \lceil R_i'/2 \rceil - 1$$
$$\leq (m - 1 + 3|A|)(R - 1) + |A| - 2(R - 1)|A \setminus A_i|,$$

where R_i' is the Ramsey number of M_i. We note that every V_i is a $\mathrm{XY}_{m-1,n_i} \cup \mathrm{YX}_{m-1,n_i}$-set by Lemma 7. If $|A| = 0$, then $(\mu(v_i), \mu(V_i)) = (\varepsilon, \{\varepsilon\})$ which is in any closed set. Otherwise, induction on $|A|$ gives $(t_i, T_i) \in \mathsf{ConeSat}_{\mathbf{Si}_{m-1}}(M_i) \subseteq \mathsf{ConeSat}_{\mathbf{Si}_{m-1}}(M)$ such that $\mu(v_i) = t_i$, $\mu(V_i) \subseteq T_i$. We define

$$W = S_1\mu(b_1)T_1\mu(a_1)\ldots S_{k-1}\mu(b_{k-1})T_{k-1}\mu(a_{k-1})S_k,$$
$$w = s_1\mu(b_1)t_1\mu(a_1)\ldots s_{k-1}\mu(b_{k-1})t_{k-1}\mu(a_{k-1})s_k.$$

Suppose a_j and b_j conincide, and a_ℓ and b_ℓ coincide but that a_i and b_i does not coincide for any $j < i < \ell$. We claim that

$$(\mu(u_jb_j), S_j\mu(b_j))(t_j, T_j) \cdots (t_{\ell-1}, T_{\ell-1})(\mu(a_{\ell-1}u_\ell), \mu(a_{\ell-1})U_\ell)$$

is of the form required for $\mathsf{ConeSat}_{\mathbf{Si}_m}(M)$ (we can use the same argument for $j = 1$ and $\ell = k$). Indeed, since $a_{i-1}, b_i \in \alpha(S_i) = \mathsf{alph}(U_i)$ and S_i is an $\mathbf{RL}_m$ factor, it follows that $S_j\mu(b_j)$, $\mu(a_{i-1})S_i\mu(b_i)$ and $\mu(a_{\ell-1})S_\ell$ are $\mathbf{RL}_m$-factors. Since $\mathsf{alph}(v_i') \subseteq \mathsf{alph}(U_i) \cap \mathsf{alph}(U_{i+1})$ for all $v_i' \in V_i$, this gives a factorization of the desired form.

Since (w, W) is a product of such factors and pairs $(\mu(a_i), \{\mu(a_i)\})$, it follows that $(w, W) \in \mathsf{ConeSat}_{\mathbf{Si}_m}(M)$. Since $\mu(u) = w$ $\mu(U) \subseteq W$ and closed sets are closed under subsets, this implies the desired result. $\qquad \square$

Proof of Theorem 2 (vii), (viii), (ix) and (x). We prove the result for *(ix)*; the other results are analogous.

From Lemma 18 it follows that $\mathsf{ConeSat}_{\mathbf{Si}_m}(M) \subseteq \mathsf{Cone}_{\mathbf{Si}_m}(M)$ and from Lemma 29 it follows that $\mathsf{PL}_\tau(M) \subseteq \mathsf{Sat}_{\mathbf{R}_m}(M)$ for m odd. We get a similar result for m even. In particular, $\mathsf{PL}_{\mathbf{R}_m}(M) = \mathsf{PL}_\tau(M) = \mathsf{Sat}_{\mathbf{R}_m}(M)$ which gives the desired result. $\qquad \square$

11 Conclusion and Outlook

We considered conelikes, an algebraic counterpart to the covering problem. In particular, solving the conelike problem yields solutions to the separation problem for positive varieties. For full varieties, this problem coincides with the pointlike problem. We provided solutions to the conelike (resp. pointlike) problem for all levels of the Trotter-Weil and the quantifier alternation hierarchy (Theorem 2). This was done by providing computable subsets of $M \times 2^M$ (resp. 2^M) and showing that these coincided with the conelikes (resp. pointlikes).

Furthermore, we considered comparisons of rankers. We showed that any set of ranker comparisons closed under subwords gives rise to a stable preorder, and thus a monoid. The quantifier alternation hierarchy has previously been given a characterization in terms of ranker comparisons. We extended this to a characterization using ranker comparisons for the corners of the Trotter-Weil hierarchy. Apart from giving a unifying picture of the two hierarchies as a ranker comparison hierarchy (Figure 1), this also served as a tool for the result on pointlikes. Having a unified formalism made moving up in the hierarchy much more uniform. We also used ranker comparisons to find separators for the conelike problem, i.e. relational morphisms

such that a set is conelike with respect to that relational morphism if and only if it is conelike with respect to the variety.

Given a separator, one can try all possible quotients in order to find an *optimal separator*, i.e., the smallest monoid acting as a separator. However, the monoids provided here are doubly exponential in the size of M (exponential in the size n of the rankers which in turn is exponential in M), making such an approach computationally hardly feasable. Thus, in future work, it would be interesting to get a better understanding of optimal separators.

References

[1] J. Almeida. Some algorithmic problems for pseudovarieties. *Publ. Math. Debrecen*, 54(1):531–552, 1999.

[2] J. Almeida, J. C. Costa, and M. Zeitoun. Pointlike sets with respect to **R** and **J**. *J. Pure Appl. Algebra*, 212(3):486–499, 2008.

[3] J. Almeida and P. V. Silva. SC-hyperdecidability of **R**. *Theoret. Comput. Sci.*, 255(1-2):569–591, 2001.

[4] J. Almeida and M. Zeitoun. The pseudovariety **J** is hyperdecidable. *RAIRO Inform. Théor. Appl.*, 31(5):457–482, 1997.

[5] C. J. Ash. Inevitable graphs: a proof of the type II conjecture and some related decision procedures. *Internat. J. Algebra Comput.*, 1(1):127–146, 1991.

[6] V. Diekert, P. Gastin, and M. Kufleitner. A survey on small fragments of first-order logic over finite words. *Internat. J. Found. Comput. Sci.*, 19(3):513–548, 2008.

[7] S. Eilenberg. *Automata, languages, and machines. Vol. B.* Academic Press, 1976.

[8] L. Fleischer, M. Kufleitner, and A. Lauser. Block products and nesting negations in FO^2. In *Computer science—theory and applications*, volume 8476 of *Lecture Notes in Comput. Sci.*, pages 176–189. Springer, 2014.

[9] K. Henckell. Pointlike sets: the finest aperiodic cover of a finite semigroup. *J. Pure Appl. Algebra*, 55(1-2):85–126, 1988.

[10] V. Henriksson and M. Kufleitner. Nesting negations in FO^2 over infinite words. *CoRR*, abs/2012.01309, 2020.

[11] A. Krebs and H. Straubing. An effective characterization of the alternation hierarchy in two-variable logic. *ACM Trans. Comput. Log.*, 18(4):30:1–30:22, 2017.

[12] K. Krohn, J. L. Rhodes, and B. Tilson. Homomorphisms and semilocal theory. In *Algebraic Theory of Machines, Languages, and Semigroups*, chapter 8, pages 191–231. Academic Press, 1968.

[13] M. Kufleitner and A. Lauser. The join levels of the Trotter-Weil hierarchy are decidable. In *Mathematical foundations of computer science 2012*, volume 7464 of *Lecture Notes in Comput. Sci.*, pages 603–614. Springer, 2012.

[14] M. Kufleitner and P. Weil. The FO^2 alternation hierarchy is decidable. In *Proc. CSL'12*, volume 16 of *LIPIcs*, pages 426–439. Dagstuhl Publishing, 2012.

[15] M. Kufleitner and P. Weil. On logical hierarchies within FO^2-definable languages. *Log. Methods Comput. Sci.*, 8(3):3:11, 30, 2012.

[16] A. Lauser. *Formal Language Theory of Logic Fragments*. PhD thesis, University of Stuttgart, 2014.

[17] T. Place and M. Zeitoun. Going higher in the first-order quantifier alternation hierarchy on words. In *Automata, languages, and programming. Part II*, volume 8573 of *Lecture Notes in Comput. Sci.*, pages 342–353. Springer, 2014.

[18] T. Place and M. Zeitoun. The covering problem. *Log. Methods Comput. Sci.*, 14(3), 2018.

[19] F. P. Ramsey. On a Problem of Formal Logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929.

[20] M.-P. Schützenberger. Sur le produit de concaténation non ambigu. *Semigroup Forum*, 13(1):47–75, 1976.

[21] T. Schwentick, D. Thérien, and H. Vollmer. Partially-ordered two-way automata: A new characterization of DA. In *International Conference on Developments in Language Theory*, pages 239–250. Springer, 2001.

[22] I. Simon. Piecewise testable events. In *Automata theory and formal languages (Second GI Conf., Kaiserslautern, 1975)*, volume 33 of *Lecture Notes in Comput. Sci.*, pages 214–222. Springer, 1975.

[23] B. Steinberg. On pointlike sets and joins of pseudovarieties. *Internat. J. Algebra Comput.*, 8(2):203–234, 1998. With an addendum by the author.

[24] P. Weis and N. Immerman. Structure theorem and strict alternation hierarchy for FO^2 on words. *Log. Methods Comput. Sci.*, 5(3):3:4, 23, 2009.