Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Computer Science 348 (2005) 277 – 293
Theoretical Computer Science
www.elsevier.com/locate/tcs

Regular solutions of language inequalities and well quasi-orders ☆

Michal Kunc

Department of Mathematics, Masaryk University, Janáčkovo nám. 2a, 662 95 Brno, Czech Republic

Abstract

By means of constructing suitable well quasi-orders of free monoids we prove that all maximal solutions of certain systems of language inequalities are regular. This way we deal with a wide class of systems of inequalities where all constants are languages recognized by finite simple semigroups. In a similar manner we also demonstrate that the largest solution of the inequality $XK \subseteq LX$ is regular provided the language L is regular.

MSC: 68Q45; 20M35; 68Q70; 68R05

Keywords: Language equation; Regular language; Well quasi-order; Syntactic semigroup; Finite simple semigroup

1. Introduction

Systems of language equations and inequalities were intensively studied especially in connection with context-free languages, since these languages can be elegantly described as components of least solutions of systems of explicit polynomial equations. Much less attention was devoted to implicit language equations and to equations employing other operations than union and concatenation. Only little research has been done also on maximal solutions of language equations. Such issues were first addressed by Conway [4], who observed that inequalities of the form $E \subseteq L$, where E is a regular function of variables and L is a regular language, possess only finitely many maximal solutions, all of them are regular and computable. More precisely, every component of a maximal solution of such an equation is a union of certain classes of the syntactic congruence of L. In particular, this leads to an algorithm for calculating best approximations of a given regular language by other given languages.

In his book Conway also formulated several conjectures concerning for instance maximal solutions of commutation equations $XL = LX$, where L is a regular language, and so-called semi-linear inequalities. Problems of commutation of languages were revisited in the past few years in a series of articles (e.g. [3,9]), where it was proved that in certain special cases the largest language commuting with a given regular language is again regular (see [10] for a survey and simplified proofs). On the other hand, recently the author demonstrated that the largest language commuting with a given finite language even need not be recursively enumerable [14].

☆ Supported by the project MSM 143100009 of the Ministry of Education of the Czech Republic.
E-mail address: kunc@math.muni.cz
URL: http://www.math.muni.cz/∼kunc/.

0304-3975/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2005.09.018

Regular solutions of systems of inequalities generalizing regular grammars were studied for example by Leiss [15]. Baader and Küsters [1] used largest solutions of systems of linear equations, i.e. equations of the form

$$K_0 + K_1X_1 + \cdots + K_nX_n = L_0 + L_1X_1 + \cdots + L_nX_n,$$

where $K_0, \dots , K_n$ and $L_0, \dots , L_n$ are regular languages, for dealing with unification of concept descriptions; they proved that the computation of the largest solution, which is always regular, is an ExpTime-complete problem. An attempt to initiate development of a unified theory of general language equations has been made by Okhotin [16]; in particular, he describes classes of languages definable as components of unique, smallest and largest solutions of systems of language inequalities using all Boolean operations.

In this paper we introduce a new method of demonstrating regularity of maximal solutions of language inequalities based on the concept of well quasi-orders of free monoids. Well quasi-orders already proved to be a very useful tool in many areas of mathematics and computer science [13]. In the theory of formal languages well quasi-orders are frequently applied to obtain regularity conditions. The most important result of this kind is a generalization of Myhill theorem due to Ehrenfeucht et al. [7] stating that a language is regular if and only if it is upward closed with respect to a monotone well quasi-order. A number of results on regularity of languages based on well quasi-orders can be found for instance in [6].

This article deals with two different classes of language inequalities. First, we consider systems of inequalities of a very general form (involving even infinitary union and intersection operations) and prove that regularity of maximal solutions of such systems is guaranteed when only constant languages recognized by finite simple semigroups are allowed (this in particular covers the case of group languages). In the second part of the paper we prove that the largest solution of the inequality $XK \subseteq LX$ is regular provided the language L is regular. This contrasts with the fact that the largest solution of the equation $XL = LX$, where L is a regular language, is not always recursively enumerable. In both situations studied in this paper the result is achieved by constructing a suitable well quasi-order of the free monoid and demonstrating that every solution of our system is in fact contained in some solution upward closed with respect to this quasi-order.

Basic notions employed in our considerations are recalled in the following section. For a more comprehensive introduction to formal languages, to semigroup theory and to well quasi-orders the reader is referred to [18,8,6], respectively.

2. Preliminaries

2.1. Languages

We denote the sets of positive and non-negative integers by $\mathbb{N}$ and $\mathbb{N}_0$, respectively. For any set S, by 2^S we mean the set of all subsets of S. Throughout the paper we consider a finite alphabet A and an infinite set of variables $\mathcal{V}$. As usual, we write A^+ for the set of all non-empty finite words over A, and A^* for the set obtained from A^+ by adding the empty word ε. We use the same symbols A^+ and A^* to denote the free semigroup and the free monoid, respectively, which arise from these sets when we equip them with the operation of concatenation.

If we have $w = xyz$ for some words $w, x, y, z \in A^*$, then the words x, y and z are called a *prefix*, a *factor* and a *suffix* of w, respectively. For two words $u, v \in A^*$, the notation $u^{-1}v$ stands for the suffix w of v satisfying $v = uw$. The number of occurrences of letters from a set $B \subseteq A$ in a word $w \in A^*$ is written as $|w|_B$.

Languages over an alphabet A are arbitrary subsets of A^* and a language $L \subseteq A^*$ is called *ε-free* if $\varepsilon \notin L$. Let us recall that states of the *minimal automaton* of a language L over A can be identified with classes of the Nerode equivalence relation $\sim_L$ on A^* defined for $u, v \in A^*$ by the formula

$$u \sim_L v \iff (\forall x \in A^*)(ux \in L \iff vx \in L).$$

This relation is a right congruence of the monoid A^* (i.e. $u \sim_L v \Longrightarrow uw \sim_L vw$ for all $u, v, w \in A^*$) and the transition function of the automaton is defined by the rule $\delta(u \sim_L, w) = (uw) \sim_L$ for $u, w \in A^*$. *Regular languages* are languages whose minimal automaton is finite.

Let $\mathfrak{S} = (S, *)$ be a semigroup and let $\sigma : A^+ \to \mathfrak{S}$ be a semigroup homomorphism. We say that an ε-free language $L \subseteq A^+$ is *recognized* by the homomorphism σ if $\sigma^{-1}\sigma(L) = L$, i.e. if there exists a subset $T \subseteq S$ such that $L = \sigma^{-1}(T)$. The *syntactic congruence* $\equiv_L$ of a language $L \subseteq A^+$ is the congruence of the free semigroup A^+ defined by the condition

$$u \equiv_L v \iff (\forall x, y \in A^*)(xuy \in L \iff xvy \in L).$$

In other words, the relation $\equiv_L$ is the largest congruence of A^+ such that the corresponding projection homomorphism recognizes L. The factor semigroup $A^+/\equiv_L$ is called the *syntactic semigroup* of L and denoted $\mathcal{S}(L)$; the projection homomorphism $\sigma_L : A^+ \to \mathcal{S}(L)$ is referred to as the *syntactic homomorphism* of L. It is well-known that an ε-free language is regular if and only if its syntactic semigroup is finite.

The *syntactic monoid* $\mathcal{M}(L)$ of an arbitrary language $L \subseteq A^*$ is defined analogously as the syntactic semigroup of an ε-free language. A language whose syntactic monoid is a finite group is called a *group language*.

Regular operations on languages over a given alphabet are union, concatenation $K \cdot L = \{uv \mid u \in K,\ v \in L\}$ and Kleene star $L^* = L^+ \cup \{\varepsilon\}$, where $L^+ = \bigcup_{m \in \mathbb{N}} L^m$. Kleene's theorem states that regular languages are exactly those languages which can be obtained from finite languages using regular operations.

Some results of this paper will be formulated also for inequalities employing other than regular operations. Let E be an arbitrary expression built from languages over A (called constants) and variables from $\mathcal{V}$ using some symbols for language operations and let $\alpha : \mathcal{V} \to 2^{A^*}$ be a mapping assigning to each variable a language over A. Then $\alpha(E)$ denotes the language obtained by replacing each occurrence of every variable $X \in \mathcal{V}$ in E with the language $\alpha(X)$ and evaluating the resulting expression. A language inequality is a formal inequality $E \subseteq F$ of two expressions over constant languages and variables. A solution of the inequality $E \subseteq F$ is any mapping $\alpha : \mathcal{V} \to 2^{A^*}$ satisfying $\alpha(E) \subseteq \alpha(F)$. We call a solution α regular if all the languages $\alpha(X)$, for $X \in \mathcal{V}$, are regular. Solutions of a given system of language inequalities are partially ordered by componentwise inclusion

$$\alpha \leqslant \beta \iff \forall X \in \mathcal{V} : \alpha(X) \subseteq \beta(X)$$

and we are mainly interested in solutions maximal with respect to this ordering (notice that if a variable $X \in \mathcal{V}$ does not occur in the system, then $\alpha(X) = A^*$ for every maximal solution α).

2.2. Semigroups

Let us now proceed to fix basic notation from semigroup theory and to state necessary facts about simple semigroups and chains of simple semigroups.

Let $\mathfrak{S} = (S, *)$ be an arbitrary semigroup. The monoid obtained from $\mathfrak{S}$ by adding a new neutral element 1 will be written as $\mathfrak{S}^1 = (S^1, *)$. We denote by $\mathrm{eval}_{\mathfrak{S}} : S^+ \to \mathfrak{S}$ the evaluation homomorphism from the free semigroup over S defined by the rule $\mathrm{eval}_{\mathfrak{S}}(s) = s$ for all $s \in S$.

An element $s \in S$ of a semigroup $\mathfrak{S} = (S, *)$ which satisfies $s * s = s$ is referred to as an *idempotent*. Note that if $\mathfrak{S}$ is finite, then there exists $n \in \mathbb{N}$ such that s^n is an idempotent for every $s \in S$; for instance $n = |S|!$ can be used.

A *null semigroup* is a semigroup $(S, *)$ containing a zero element 0 such that $s * t = 0$ for every $s, t \in S$. A *semilattice* is a commutative semigroup whose every element is an idempotent. Any semilattice $(S, *)$ is completely determined by the partial order on S defined by the rule $s \leqslant t \iff s * t = s$; the operation $*$ is just the meet (greatest lower bound) operation in the resulting partially ordered set $(S, \leqslant)$. A semilattice is called a *chain* if the corresponding ordering $\leqslant$ of S is total, in other words, if for all elements $s, t \in S$ either $s * t = s$ or $s * t = t$.

An *ideal* of a semigroup $\mathfrak{S} = (S, *)$ is a non-empty subset $I \subseteq S$ such that for all $s \in I$ and $t \in S$ we have $s * t \in I$ and $t * s \in I$. The ideal of $\mathfrak{S}$ generated by a given element $s \in S$ is equal to $S^1 * s * S^1 = \{p * s * q \mid p, q \in S^1\}$. For the semigroup $\mathfrak{S}$, the quasi-order $\leqslant_{\mathcal{J}}$ is defined for any $s, t \in S$ by the rule $s \leqslant_{\mathcal{J}} t \iff s \in S^1 * t * S^1$. The Green relation $\mathcal{J}$ of the semigroup $\mathfrak{S}$ is the equivalence relation on S associated with the quasi-order $\leqslant_{\mathcal{J}}$, i.e. two elements of $\mathfrak{S}$ are $\mathcal{J}$-equivalent if they generate the same ideal.

A semigroup is called *simple* if it has no proper ideal, i.e. if all elements of the semigroup are $\mathcal{J}$-equivalent. Now we recall the well-known construction of matrix semigroups over groups, which gives a complete description of finite simple semigroups. Let I and J be arbitrary finite sets, $\mathfrak{G} = (G, \cdot)$ a finite group and $P : J \times I \to G$ any mapping (this mapping can be understood as a $J \times I$-matrix with entries in G). The Rees matrix semigroup $\mathfrak{M}(I, J, \mathfrak{G}, P)$ is defined on the set $I \times G \times J$ by the multiplication formula

$$(i, g, j) * (i', g', j') = (i, g \cdot P(j, i') \cdot g', j').$$

Proposition 1 (*Suschkewitsch [19]*). *A finite semigroup is simple if and only if it is isomorphic to a Rees matrix semigroup.*

Every finite simple semigroup $\mathfrak{S} = (S, *)$ is in fact a disjoint union of its maximal subgroups and we denote for every element $s \in S$ by s^0 the identity element of the subgroup of $\mathfrak{S}$ containing s and by s^{-1} the inverse of s in this subgroup. Equivalently, the element s^0 can be defined as the unique idempotent of $\mathfrak{S}$ which is a power of s. In the semigroup $\mathfrak{M}(I, J, \mathfrak{G}, P)$ maximal subgroups are precisely subsets of the form $\{(i, g, j) \mid g \in G\}$ for any $i \in I$, $j \in J$ and we have $(i, g, j)^0 = (i, P(j, i)^{-1}, j)$ and $(i, g, j)^{-1} = (i, (P(j, i) \cdot g \cdot P(j, i))^{-1}, j)$. Further, in every finite simple semigroup the equality $(s * t * s)^0 = s^0$ holds for every $s, t \in S$.

If a semigroup $\mathfrak{S} = (S, *)$ possesses a congruence relation $\equiv$ such that the factor semigroup $\mathfrak{S}/\equiv$ is a chain, then every congruence class $(s \equiv)$, for $s \in S$, is a subsemigroup of $\mathfrak{S}$ and the semigroup $\mathfrak{S}$ is called a *chain of semigroups* $(s \equiv)$. The following lemma provides us with a simple criterion for verifying that a semigroup is a chain of simple semigroups.

Lemma 2. *For any semigroup $\mathfrak{S} = (S, *)$ the following conditions are equivalent*:
(i) *The semigroup $\mathfrak{S}$ is a chain of simple semigroups.*
(ii) *For every $s, t \in S$ either $s * t \mathrel{\mathcal{J}} s$ or $s * t \mathrel{\mathcal{J}} t$.*
(iii) *The Green relation $\mathcal{J}$ of $\mathfrak{S}$ is a congruence, the factor semigroup $\mathfrak{S}/\mathcal{J}$ is a chain and every $\mathcal{J}$-class is a simple semigroup.*

Proof. (i) $\Longrightarrow$ (ii). Let $\equiv$ be a congruence relation of $\mathfrak{S}$ such that $\mathfrak{S}/\equiv$ is a chain and each $\equiv$-class is a simple semigroup. Then for every $s, t \in S$ their product $s * t$ is $\equiv$-equivalent to one of the elements s and t, say to the former one. Because $(s \equiv)$ is a simple semigroup, we have $s \in (s \equiv)^1 * s * t * (s \equiv)^1$, which shows that s and $s * t$ are $\mathcal{J}$-equivalent in $\mathfrak{S}$.

(ii) $\Longrightarrow$ (iii). In order to verify that the relation $\mathcal{J}$ is a congruence, let s, t, p and q be elements of $\mathfrak{S}$ satisfying $s \mathrel{\mathcal{J}} p$ and $t \mathrel{\mathcal{J}} q$. Without loss of generality assume that $s * t \mathrel{\mathcal{J}} s$. This in particular means that $s \leqslant_{\mathcal{J}} s * t \leqslant_{\mathcal{J}} t$ and consequently also $p \leqslant_{\mathcal{J}} q$. Because we trivially have $p * q \leqslant_{\mathcal{J}} p$ and $p * q$ is $\mathcal{J}$-equivalent to one of p and q, we immediately obtain $p * q \mathrel{\mathcal{J}} p \mathrel{\mathcal{J}} s \mathrel{\mathcal{J}} s * t$, hence $\mathcal{J}$ is a congruence.

Clearly, the semigroup $\mathfrak{S}/\mathcal{J}$ is a chain. It remains to prove that each $\mathcal{J}$-class of $\mathfrak{S}$ is a simple semigroup. Let $s, t \in S$ be any elements satisfying $s \mathrel{\mathcal{J}} t$. Since $(s\mathcal{J})$ is a subsemigroup of $\mathfrak{S}$, we have $t \mathrel{\mathcal{J}} t^3$ and so there exist $p, q \in S^1$ such that $s = p * t^3 * q$. Then $p * t \mathrel{\mathcal{J}} t * q \mathrel{\mathcal{J}} s$ because one trivially gets $s \leqslant_{\mathcal{J}} p * t \leqslant_{\mathcal{J}} t \leqslant_{\mathcal{J}} s$ and $s \leqslant_{\mathcal{J}} t * q \leqslant_{\mathcal{J}} t \leqslant_{\mathcal{J}} s$. Therefore s belongs to $(s\mathcal{J}) * t * (s\mathcal{J})$, which shows that $(s\mathcal{J})$ is a simple semigroup.

(iii) $\Longrightarrow$ (i) is trivial. $\square$

Remark 3. It is clear that any homomorphic image of a simple semigroup is again simple and using condition (ii) of Lemma 2 one can also easily verify that homomorphic images of chains of simple semigroups are again chains of simple semigroups. Further notice that if $\mathfrak{S} = (S, *)$ is a finite simple semigroup and $T \subseteq S$ its subsemigroup, then for arbitrary elements $s, t \in T$ we have $s = s * t * (s * (s * t * s)^{-1} * s)$ and therefore $s \leqslant_{\mathcal{J}} t$ holds in this subsemigroup. This means that any subsemigroup of a finite simple semigroup is simple and any subsemigroup of a chain of finite simple semigroups is a chain of simple semigroups.

Similarly as for groups, inverses in maximal subgroups can be used in the case of chains of finite simple semigroups to cancel elements from products.

Lemma 4. *Let $\mathfrak{S} = (S, *)$ be a chain of finite simple semigroups and let $r, s \in S$ be elements satisfying $s \leqslant_{\mathcal{J}} r$. Then there exists $t \in S$ such that $s * r * t = s$. In particular, if $p * s * r = q * s * r$ holds for some $p, q \in S$, then $p * s = q * s$.*

Proof. By Lemma 2 we know that $r * s$ and s belong to the same $\mathcal{J}$-class of $\mathfrak{S}$, which is a finite simple semigroup. Therefore $(s * r * s * s)^0 = s^0$ and setting

$$t = s * s * (s * r * s * s)^{-1} * s,$$

we immediately obtain $s * r * t = s$. $\square$

2.3. Well quasi-orders

The rest of this section is devoted to recalling the concept of well quasi-orders and an important result about quasi-orders defined on the set of all words by means of context-free productions.

A *quasi-order* $\leqslant$ on a set S is a reflexive and transitive binary relation. We say that a subset T of S is *upward closed* with respect to $\leqslant$ if for every $t \in T$ and $s \in S$, the inequality $t \leqslant s$ implies $s \in T$.

Definition 5. A quasi-order $\leqslant$ on a set S is called a *well quasi-order* if the following equivalent conditions are satisfied:

(i) There exists neither an infinite strictly descending sequence in S nor an infinite sequence of mutually incomparable elements of S.

(ii) If $(s_n)_{n \in \mathbb{N}}$ is an infinite sequence of elements of S, then there exist $m, n \in \mathbb{N}$ such that $m < n$ and $s_m \leqslant s_n$.

(iii) For every subset $T \subseteq S$ there exists a finite subset U of T such that for each $t \in T$ there exists some element $s \in U$ satisfying $s \leqslant t$.

(iv) There does not exist an infinite sequence of upward closed subsets of S strictly ascending with respect to inclusion.

The following basic fact can be easily verified using condition (ii) of Definition 5.

Lemma 6. *Let* $(S, \leqslant)$ *and* $(T, \sqsubseteq)$ *be quasi-ordered sets and let* $\varphi : S \to T$ *be an arbitrary mapping satisfying the condition*

$$(\forall s_1, s_2 \in S)(\varphi(s_1) \sqsubseteq \varphi(s_2) \implies s_1 \leqslant s_2).$$

If the relation $\sqsubseteq$ *is a well quasi-order on* T, *then* $\leqslant$ *is a well quasi-order on* S.

Derivation relations of context-free production systems which are well quasi-orders on A^* can be characterized using the notion of unavoidable languages. Recall that a context-free rule over an alphabet A is a pair (a, u), where $a \in A$ and $u \in A^*$. And the *derivation relation* associated with a set of context-free rules $R \subseteq A \times A^*$ is the reflexive and transitive closure of the relation $\{(vaw, vuw) \mid v, w \in A^*,\ (a, u) \in R\}$ on A^*. We say that a language L is *unavoidable* over an alphabet A if all but finitely many words over A contain some word from L as a factor.

Proposition 7 (*Bucher et al. [2]*). *Let* $\leqslant$ *be a derivation relation on* A^* *associated with some finite set of context-free rules over* A. *Then* $\leqslant$ *is a well quasi-order if and only if the language* $\{awa \mid a \in A,\ w \in A^*,\ a \leqslant awa\}$ *is unavoidable over* A.

3. Decomposition quasi-orders

Let us start this section by describing systems of inequalities which will be considered here. Let $\mathfrak{L}$ be a finite set of ε-free languages over A. We say that an inequality $E \subseteq F$ is an $\mathfrak{L}$-*inequality* if the expression E is a product of variables and arbitrary constants and the expression F is built from variables and languages belonging to the set $\mathfrak{L} \cup \{\{\varepsilon\}\}$ using symbols for the operations of concatenation, arbitrary (possibly infinite) union and arbitrary (possibly infinite) intersection.

Let $\sigma : A^+ \to \mathfrak{S}$ be a homomorphism onto a finite semigroup $\mathfrak{S}$. We define a quasi-order $\leqslant_\sigma$ on A^* by setting $v \leqslant_\sigma u$ if and only if $v = a_1 \cdots a_n$, where $a_1, \ldots, a_n \in A$, and $u = u_1 \cdots u_n$, where $u_j \in A^+$ and $\sigma(u_j) = \sigma(a_j)$ for $j = 1, \ldots, n$. This quasi-order is monotone, i.e. from $v_1 \leqslant_\sigma u_1$ and $v_2 \leqslant_\sigma u_2$ it follows that $v_1 v_2 \leqslant_\sigma u_1 u_2$. Notice that if $v \leqslant_\sigma u$ then either $u = v = \varepsilon$ or $u, v \in A^+$ and $\sigma(v) = \sigma(u)$.

The following theorem states that all maximal solutions of arbitrary systems of $\mathfrak{L}$-inequalities are regular provided there exists a homomorphism σ recognizing all languages from $\mathfrak{L}$ for which the relation $\leqslant_\sigma$ is a well quasi-order.

Theorem 8. *Let* $\mathfrak{L}$ *be a finite set of* ε*-free languages over* A *and let* $\sigma : A^+ \to \mathfrak{S}$ *be a homomorphism onto a finite semigroup* $\mathfrak{S}$ *recognizing all languages in* $\mathfrak{L}$, *and such that* A^* *is well quasi-ordered by* $\leqslant_\sigma$. *Let* I *be an arbitrary* (*possibly infinite*) *set and let* $\Sigma = \{E_i \subseteq F_i \mid i \in I\}$ *be a system of* $\mathfrak{L}$*-inequalities. Then every solution of* Σ *is contained in a regular solution of* Σ; *in particular*, *every maximal solution of the system* Σ *is regular. If only finitely many* *variables occur in Σ, then every solution of Σ is contained in a maximal one. The same conclusions hold true if only ε-free solutions of Σ are considered.*

Proof. Let α be a solution of Σ. For every $X \in \mathcal{V}$ define the language

$$\beta(X) = \{u \in A^* \mid \exists v \in \alpha(X) : v \leqslant_\sigma u\}.$$

It is clear that $\alpha(X) \subseteq \beta(X)$ and that $\varepsilon \in \beta(X)$ if and only if $\varepsilon \in \alpha(X)$ since the empty word is incomparable with the other elements of A^*. We are going to show that β is a regular solution of Σ.

First observe that because the quasi-order $\leqslant_\sigma$ is monotone, if a word u belongs to the language $\beta(E_i)$, there exists $v \in \alpha(E_i)$ such that $v \leqslant_\sigma u$. We prove by induction with respect to the structure of the expression F_i that if $v \in \alpha(F_i)$ and $v \leqslant_\sigma u$ for some words u and v, then $u \in \beta(F_i)$, which is enough to conclude that β is a solution of Σ. So assume a word v belongs to $\alpha(e)$ for some subexpression e of F_i and $v \leqslant_\sigma u$. If e is a variable, we have $u \in \beta(e)$ by the definition of β. In the case e is a language from $\mathfrak{L}$, one obtains $u \in \beta(e)$ from the fact $\sigma(u) = \sigma(v)$. For $e = \{\varepsilon\}$, the only possibility is $u = v = \varepsilon \in \beta(e)$. If the expression e is of the form $\bigcup_{k\in K} e_k$ or $\bigcap_{k\in K} e_k$ for some set K, then $u \in \beta(e)$ is clear from the induction hypothesis. Finally, consider the case $e = e_1 \cdot e_2$. Then $v = v_1 \cdot v_2$, where $v_1 \in \alpha(e_1)$ and $v_2 \in \alpha(e_2)$. From $v \leqslant_\sigma u$ we deduce $v_1 = a_1 \cdots a_m$ and $v_2 = a_{m+1} \cdots a_n$, where $0 \leqslant m \leqslant n$ and $a_j \in A$, and $u = u_1 \cdots u_n$ for some words $u_1, \ldots, u_n \in A^+$ satisfying $\sigma(u_j) = \sigma(a_j)$ for $j = 1, \ldots, n$. Therefore, $v_1 \leqslant_\sigma u_1 \cdots u_m$ and $v_2 \leqslant_\sigma u_{m+1} \cdots u_n$ and we can apply the induction hypothesis to these words. Hence $u \in \beta(e)$.

In order to prove that $\beta(X)$ is a regular language, observe that $\beta(X)$ is upward closed with respect to the well quasi-order $\leqslant_\sigma$, therefore it can be generated by finitely many elements of A^* due to condition (iii) of Definition 5, i.e. $\beta(X)$ is a union of finitely many languages of the form $\langle v \rangle = \{u \in A^* \mid v \leqslant_\sigma u\}$ for a word $v \in A^*$. And it is easy to see that for arbitrary letters $a_1, \ldots, a_n \in A$ we have

$$\langle a_1 \cdots a_n \rangle = (\sigma^{-1}\sigma(a_1)) \cdots (\sigma^{-1}\sigma(a_n)),$$

which shows that each language $\langle v \rangle$ is regular.

We have already proved that every solution of Σ is contained in a regular solution whose every component is a language upward closed with respect to the well quasi-order $\leqslant_\sigma$. Because there is no infinite strictly ascending sequence of such upward closed sets by condition (iv) of Definition 5, this immediately implies that if there are only finitely many variables, every solution is in fact contained in a maximal solution. □

Remark 9. Observe that existence of a maximal solution above every solution follows immediately from Zorn's Lemma (even if there are infinitely many variables) since all operations in our inequalities are monotone and left-hand sides employ only finitary operations. In contrast, our proof of this fact in the case of finitely many variables avoids the Axiom of Choice, although even for regular solutions of simple inequalities it does not provide us with an algorithm for computing such a maximal solution.

Further notice that the relation $\leqslant_\sigma$ in the proof is a monotone well quasi-order on A^* and therefore the languages $\beta(X)$ are regular due to the result of Ehrenfeucht et al. [7]; we give a direct proof of their regularity because it also provides us with some information on how maximal solutions are related to constant languages occurring in the system.

Now we are going to precisely characterize those homomorphisms $\sigma : A^+ \to \mathfrak{S}$ for which the relation $\leqslant_\sigma$ is a well quasi-order on A^* and therefore Theorem 8 can be applied. The proof of the fact that $\leqslant_\sigma$ is a well quasi-order will be performed for the partial order $\leqslant_{\mathrm{eval}_\mathfrak{S}}$ on S^*, which is sufficient because every quasi-order $\leqslant_\sigma$ can be embedded into this one; note that the relation $\leqslant_{\mathrm{eval}_\mathfrak{S}}$ is really a partial order since from the validity of both $u \leqslant_{\mathrm{eval}_\mathfrak{S}} v$ and $v \leqslant_{\mathrm{eval}_\mathfrak{S}} u$ it follows that the words u and v are of the same length and if $\mathrm{eval}_\mathfrak{S}(s) = \mathrm{eval}_\mathfrak{S}(t)$ holds for some letters $s, t \in S$, then in fact $s = t$.

Theorem 10. *For an arbitrary homomorphism $\sigma : A^+ \twoheadrightarrow \mathfrak{S}$ onto a finite semigroup $\mathfrak{S} = (S, *)$ the following conditions are equivalent*:

(i) *The relation $\leqslant_\sigma$ is a well quasi-order on A^*.*
(ii) *The semigroup $\mathfrak{S}$ is a chain of simple semigroups.*
(iii) *The relation $\leqslant_{\mathrm{eval}_\mathfrak{S}}$ is a well partial order on S^*.*

Proof. (i) $\Longrightarrow$ (ii). Let $\leqslant_\sigma$ be a well quasi-order. Consider arbitrary elements $s, t \in S$ and let us prove that either $s * t \;\mathcal{J}\; s$ or $s * t \;\mathcal{J}\; t$. Take arbitrary words $u, v \in A^+$ such that $\sigma(u) = s$ and $\sigma(v) = t$. Because $\leqslant_\sigma$ is a well quasi-order, by condition (ii) of Definition 5 the sequence $w_n = (uv)^{2^n}$ for $n \in \mathbb{N}$ contains some elements satisfying $m < n$ and $w_m \leqslant_\sigma w_n$. Since the number of copies of uv in w_n is at least the same as the number of copies of u and copies of v in w_m together, it is easy to see that some occurrence of either u or v in w_m corresponds to a factor of w_n containing a whole copy of uv, i.e. either $u \leqslant_\sigma xuvy$ or $v \leqslant_\sigma xuvy$ for certain words $x, y \in A^*$. Therefore, the product $\sigma(x) * s * t * \sigma(y)$ is equal to either s or t, and by Lemma 2 this means that $\mathfrak{S}$ is a chain of simple semigroups.

(ii) $\Longrightarrow$ (iii). Let $\mathfrak{S}$ be a chain of simple semigroups. In particular, due to Lemma 2 we know that $\mathcal{J}$-classes of $\mathfrak{S}$ form a chain.

First observe that the partial order $\leqslant_{\mathrm{eval}_\mathfrak{S}}$ on S^* is equal to the derivation relation associated with the set of context-free rules $\{(s * t, st) \mid s, t \in S\}$ over S, where each element of S can be rewritten to any word in S^* which evaluates to this element. By Proposition 7, in order to prove that $\leqslant_{\mathrm{eval}_\mathfrak{S}}$ is a well partial order, it is enough to verify that the language

$$U = \{sws \mid s \in S,\ w \in S^*,\ \mathrm{eval}_\mathfrak{S}(sws) = s\}$$

is unavoidable over S. Let us denote the cardinality of S by m and let $J_1 < \cdots < J_n$ be all the $\mathcal{J}$-classes of $\mathfrak{S}$. Consider an arbitrary word $u \in S^*$ of length at least $(m^2 + 1)^n$.

Let us show by contradiction that there exists a factor v of the word u such that some element of S, which has at least $m + 1$ occurrences in v, is $\mathcal{J}$-minimal among all elements occurring in v. So suppose that for every factor v of u, each element $s \in S$, which is minimal with respect to $\leqslant_\mathcal{J}$ among all elements occurring in v, has at most m occurrences in v. Then for every $k \in \{1, \ldots, n\}$ there are at most $|u|_{J_1 \cup \cdots \cup J_{k-1}} + 1$ maximal factors of u without elements of $J_1 \cup \cdots \cup J_{k-1}$, and due to our assumption none of them contains more than m occurrences of any element of J_k. This means that

$$|u|_{J_1 \cup \cdots \cup J_k} \leqslant m^2 \cdot (|u|_{J_1 \cup \cdots \cup J_{k-1}} + 1) + |u|_{J_1 \cup \cdots \cup J_{k-1}},$$

giving us a recursive formula for estimating the number of occurrences of elements from any given ideal of $\mathfrak{S}$ in u. Altogether we find that the length of u is at most $(m^2 + 1)^n - 1$, which contradicts the choice of u.

Therefore, we can find a factor v of u where a certain element $s \in S$, minimal with respect to $\leqslant_\mathcal{J}$ among all elements which occur in v, has at least $m + 1$ occurrences. By the pigeon-hole principle, there are two suffixes sx and $swsx$ of v, where $w, x \in S^*$, which satisfy $\mathrm{eval}_\mathfrak{S}(sx) = \mathrm{eval}_\mathfrak{S}(swsx)$. Because the element s is $\mathcal{J}$-minimal in v and $\mathcal{J}$-classes of $\mathfrak{S}$ form a chain, we have $s \leqslant_\mathcal{J} \mathrm{eval}_\mathfrak{S}(x)$. Hence, we can apply Lemma 4 to obtain $s = \mathrm{eval}_\mathfrak{S}(sws)$. This shows that u contains a factor belonging to the language U and therefore U is unavoidable.

(iii) $\Longrightarrow$ (i). Let us assume that $\leqslant_{\mathrm{eval}_\mathfrak{S}}$ is a well partial order on S^* and consider the literal monoid homomorphism $\varphi : A^* \to S^*$ defined by the rule $\varphi(a) = \sigma(a)$ for every letter $a \in A$. Since $\leqslant_{\mathrm{eval}_\mathfrak{S}}$ is a well partial order on S^*, we can use Lemma 6 to prove that $\leqslant_\sigma$ is a well quasi-order on A^*. Take arbitrary words $u, v \in A^*$ satisfying $\varphi(v) \leqslant_{\mathrm{eval}_\mathfrak{S}} \varphi(u)$ and let us verify that $v \leqslant_\sigma u$ holds. Because the homomorphism φ is literal, decompositions of $\varphi(u)$ and $\varphi(v)$ verifying $\varphi(v) \leqslant_{\mathrm{eval}_\mathfrak{S}} \varphi(u)$ are of the form $\varphi(v) = \varphi(a_1) \cdots \varphi(a_n)$ and $\varphi(u) = \varphi(u_1) \cdots \varphi(u_n)$, where $a_i \in A$ and $u_i \in A^+$, for $i = 1, \ldots, n$, are such that $v = a_1 \cdots a_n$ and $u = u_1 \cdots u_n$, and $\mathrm{eval}_\mathfrak{S}(\varphi(u_i)) = \varphi(a_i)$. Then the decomposition $u = u_1 \cdots u_n$ verifies $v \leqslant_\sigma u$ since for every $i \in \{1, \ldots, n\}$ we have $\sigma(u_i) = \mathrm{eval}_\mathfrak{S}(\varphi(u_i)) = \varphi(a_i) = \sigma(a_i)$. $\square$

Remark 11. Let us mention that semigroups which are chains of simple semigroups play an important role also with respect to the finite power property of regular languages, namely, they are precisely those semigroups which recognize only languages possessing the finite power property. It follows directly from the characterization given by Kirsten [11] that these semigroups really have this feature. On the other hand, if a semigroup $\mathfrak{S} = (S, *)$ is not a chain of simple semigroups, then by Lemma 2 there exist elements $s, t \in S$ such that $s * t <_\mathcal{J} s$ and $s * t <_\mathcal{J} t$. Therefore, the homomorphism $\sigma: \{a, b\}^+ \to \mathfrak{S}$ defined by the rules $\sigma(a) = s$ and $\sigma(b) = t$ recognizes the language $\{a, b\}$ which does not have the finite power property.

When we are concerned with regular languages, the most interesting inequalities are those built using regular operations. Systems of such inequalities are in fact a special case of systems considered in this section: the star operation is constructed from the operations of concatenation and infinite union and as our systems are allowed to be infinite, we can actually use the operation of infinite union (and consequently the star operation) also on left-hand sides of inequalities. Because all languages from a given set $\mathfrak{L}$ can be recognized by the product of semigroups recognizing individual languages and any product of simple semigroups is again simple, the following result is an immediate consequence of Theorems 8 and 10.

Corollary 12. *Let Σ be a finite system of inequalities of the form $E_i \subseteq F_i$, where E_i are arbitrary regular expressions and F_i are regular expressions over variables, the language $\{\varepsilon\}$ and regular languages recognizable by homomorphisms to finite simple semigroups. Then every solution of Σ is contained in a maximal solution and every maximal solution of Σ is regular.*

Remark 13. In the system Σ of Corollary 12 one can prescribe whether a given variable X contains the empty word or not since the inequalities $\{\varepsilon\} \subseteq X$ and $X \subseteq A^+$ are of the required form.

If only one constant language occurs on the right-hand sides of inequalities, then in order to apply Theorems 8 and 10 it is sufficient to know that the language is recognized by a chain of finite simple semigroups. Note that unlike for languages recognized by groups or simple semigroups, recognizability of a regular language by a chain of simple semigroups is independent of the underlying alphabet since additional letters not employed by the language form a zero element in the syntactic semigroup, which becomes the least element of the chain.

Corollary 14. *Let $L \subseteq A^+$ be a regular language recognized by a chain of finite simple semigroups. Let Σ be an arbitrary system of inequalities of the form $E_i \subseteq F_i$ over finitely many variables, where E_i are arbitrary regular expressions and F_i are regular expressions over variables and the languages L and $\{\varepsilon\}$. Then every solution of Σ is contained in a maximal solution and every maximal solution of Σ is regular.*

Before we proceed to demonstrate results of this section on examples, let us describe some equivalent characterizations of regular languages recognized by simple semigroups and chains of simple semigroups. First, observe that by Remark 3 the classes of finite simple semigroups and finite chains of finite simple semigroups are closed under taking homomorphic images and subsemigroups and therefore a language is recognized by a semigroup from one of these classes if and only if its syntactic semigroup belongs to that class. Recall that group languages are precisely those regular languages whose minimal automaton is codeterministic, i.e. contains no distinct states p and q such that $\delta(p,a) = \delta(q,a)$ for some $a \in A$. This condition can be transformed into a condition corresponding to the case of simple semigroups by considering codeterminism for two-letter words instead of single letters.

Lemma 15. *The syntactic semigroup of a regular ε-free language L over A is simple if and only if its minimal automaton contains no states p and q which for some letters $a, b \in A$ satisfy $\delta(p,a) \neq \delta(q,a)$ and $\delta(p,ab) = \delta(q,ab)$.*

Proof. Reformulating the statement of the lemma using the relation $\sim_L$ instead of the minimal automaton, we are going to prove that $\mathcal{S}(L)$ is simple if and only if

$$\forall v, w \in A^* \; \forall a, b \in A : \; vab \sim_L wab \Longrightarrow va \sim_L wa. \tag{1}$$

First assume that $\mathcal{S}(L)$ is a simple semigroup. Let us consider arbitrary words $v, w \in A^*$ and letters $a, b \in A$ satisfying $vab \sim_L wab$. By Lemma 4 there exists a word $x \in A^+$ such that $abx \equiv_L a$ and therefore

$$va \sim_L vabx \sim_L wabx \sim_L wa.$$

In order to prove the converse implication, assume that (1) holds and consider arbitrary non-empty words $u, v \in A^+$. Let $n \in \mathbb{N}$ be such that s^n is an idempotent for every element s of $\mathcal{S}(L)$. Because $x(uv)^n \equiv_L x(uv)^{2n}$ is true for any $x \in A^*$, condition (1) can be applied several times to deduce $xu \sim_L x(uv)^n u$. As this holds for every word $x \in A^*$, we have $u \equiv_L (uv)^n u$, which implies $\sigma_L(u) \leqslant_{\mathcal{J}} \sigma_L(v)$ and proves that $\mathcal{S}(L)$ is simple. $\square$

Remark 16. A characterization similar to the one given in Lemma 15 was proved by Zhang [20]. Actually, these conditions on the automaton can be equivalently reformulated as follows: for any letters $a, b \in A$, the image of the transformation of states determined by a forms a set of representatives for the kernel of the transformation determined by b.

The following lemma shows that an analogous condition on minimal automata can be formulated also in the case of chains of simple semigroups.

Lemma 17. *The syntactic semigroup of a regular ε-free language L over A is a chain of simple semigroups if and only if there exists a total ordering $\leqslant$ of A such that the minimal automaton of L contains no states p and q for which there exist a word $u \in A^*$ and letters $a, b \in A$ satisfying $a \leqslant b$, $\delta(p, au) \neq \delta(q, au)$ and $\delta(p, aub) = \delta(q, aub)$.*

Proof. We have to prove that $\mathcal{S}(L)$ is a chain of simple semigroups if and only if there exists a total ordering $\leqslant$ of A satisfying

$$\forall u, v, w \in A^* \; \forall a, b \in A\colon (a \leqslant b \;\&\; vaub \sim_L waub) \Longrightarrow vau \sim_L wau. \tag{2}$$

Let $\mathcal{S}(L)$ be a chain of simple semigroups and define a total ordering of A in such a way that the implication $a \leqslant b \Longrightarrow \sigma_L(a) \leqslant_{\mathcal{J}} \sigma_L(b)$ holds, i.e. we use the total ordering of $\mathcal{J}$-classes and in addition we order letters whose σ_L-images belong to the same $\mathcal{J}$-class arbitrarily. Consider any words $u, v, w \in A^*$ and letters $a, b \in A$ satisfying $a \leqslant b$ and $vaub \sim_L waub$. Then $\sigma_L(au) \leqslant_{\mathcal{J}} \sigma_L(b)$, therefore Lemma 4 provides us with $x \in A^+$ such that $aubx \equiv_L au$ and we get $vau \sim_L vaubx \sim_L waubx \sim_L wau$.

Let us now deal with the converse implication. Assuming that $\leqslant$ is a total ordering of A such that (2) holds, we are going to verify condition (ii) of Lemma 2, which is sufficient to show that $\mathcal{S}(L)$ is a chain of simple semigroups. Consider arbitrary words $u, v \in A^+$ and let us prove that $\sigma_L(uv)$ is $\mathcal{J}$-equivalent to either $\sigma_L(u)$ or $\sigma_L(v)$. Take any $n \in \mathbb{N}$ such that s^n is an idempotent for every element s of $\mathcal{S}(L)$. Let $a \in A$ be the smallest letter occurring in at least one of the words u and v. First assume that a occurs in u. Since we have $x(uv)^n \equiv_L x(uv)^{2n}$ for any $x \in A^*$, thanks to our choice of a we can apply (2) several times to obtain $xu \sim_L x(uv)^n u$. Because this holds for every word x, it gives us $u \equiv_L (uv)^n u$, which means that $\sigma_L(uv) \;\mathcal{J}\; \sigma_L(u)$. Similarly, if a occurs in the word v, we use the fact that $x(vu)^n \equiv_L x(vu)^{2n}$ holds for every $x \in A^*$, and apply (2) several times to get $xv \sim_L xv(uv)^n$. Therefore $v \equiv_L v(uv)^n$, verifying $\sigma_L(uv) \;\mathcal{J}\; \sigma_L(v)$. $\square$

As already indicated by Lemma 15, recognizing using simple semigroups is basically just recognizing by groups where instead of reading letters of a word one by one we consider two neighbouring letters at the same time. In order to state this fact formally, let us consider the mapping $\rho: \; A^+ \to (A \times A)^*$ defined by the rule

$$\rho(a_1 \cdots a_n) = (a_1, a_2) \cdot (a_2, a_3) \cdots (a_{n-1}, a_n),$$

for every $a_1, \ldots, a_n \in A$. We are going to show that a language $L \subseteq A^+$ is recognizable by a finite simple semigroup if and only if for every $a, b \in A$ the ρ-image of the restriction of L to words starting with a and ending in b can be obtained as the restriction of some group language over the alphabet $A \times A$ to $\rho(aA^* \cap A^*b)$. Observe that the language $\rho(aA^* \cap A^*b)$ consists precisely of words from $\rho(A^+)$ whose first letter is of the form (a, c) and the last letter of the form (d, b) for some $c, d \in A$; in the case $a = b$, the language $\rho(aA^* \cap A^*a)$ contains in addition the empty word. Further notice that the restriction of the mapping ρ to each of the languages $aA^* \cap A^*b$ is injective.

Lemma 18. *A regular ε-free language L over A can be recognized by a finite simple semigroup if and only if for every $a, b \in A$ there exists a group language $L_{a,b}$ over the alphabet $A \times A$ such that*

$$\rho(L \cap aA^* \cap A^*b) = L_{a,b} \cap \rho(aA^* \cap A^*b). \tag{3}$$

Proof. First assume that L is recognizable by a finite simple semigroup. Due to Proposition 1 we have a finite Rees matrix semigroup $\mathfrak{M}(I, J, \mathfrak{G}, P)$ over a group $\mathfrak{G} = (G, \cdot)$ and a homomorphism $\sigma: \; A^+ \to \mathfrak{M}(I, J, \mathfrak{G}, P)$ recognizing L. Let us denote for every letter $a \in A$ its σ-image by (i_a, g_a, j_a). Define a homomorphism $\tau: \; (A \times A)^* \to \mathfrak{G}$ by setting $\tau((a, b)) = g_a \cdot P(j_a, i_b)$ for every $a, b \in A$. Then we can choose languages $L_{a,b}$ by the rule $L_{a,b} = \tau^{-1}\tau\rho(L \cap aA^* \cap A^*b)$. It remains to verify (3). One can directly calculate that the homomorphisms σ and τ are related via the formula

$$\forall n \in \mathbb{N} \; \forall a_1, \ldots, a_n \in A: \; \sigma(a_1 \cdots a_n) = (i_{a_1}, \tau\rho(a_1 \cdots a_n) \cdot g_{a_n}, j_{a_n}).$$

Because σ recognizes L, this formula shows that a word $a_1 \cdots a_n$ belongs to L if and only if

$$\tau\rho(a_1 \cdots a_n) \cdot g_{a_n} \in \tau\rho(L \cap a_1A^* \cap A^*a_n) \cdot g_{a_n},$$

which is in turn equivalent to

$$\rho(a_1 \cdots a_n) \in \tau^{-1}\tau\rho(L \cap a_1A^* \cap A^*a_n),$$

in other words $\rho(a_1 \cdots a_n) \in L_{a_1,a_n}$. Therefore (3) is valid.

Conversely, assume we have for every $a, b \in A$ a language $L_{a,b} \subseteq (A \times A)^*$ which satisfies (3) and whose syntactic monoid $\mathcal{M}(L_{a,b})$ is a finite group. Let $\mathfrak{S}$ be the Rees matrix semigroup

$$\mathfrak{S} = \mathfrak{M}\left(A, A, \prod_{a,b \in A} \mathcal{M}(L_{a,b}), P\right),$$

where the mapping P is defined for every $c, d \in A$ by the rule

$$P(c, d) = \left(\sigma_{L_{a,b}}((c, d))\right)_{a,b \in A} \in \prod_{a,b \in A} \mathcal{M}(L_{a,b}).$$

We claim that the homomorphism $\sigma : A^+ \to \mathfrak{S}$ defined by setting $\sigma(a) = (a, 1, a)$, for each $a \in A$, recognizes L. It is easy to verify that σ maps all words from A^+ according to the formula

$$\sigma(a_1 \cdots a_n) = \left(a_1, (\sigma_{L_{a,b}}(\rho(a_1 \cdots a_n)))_{a,b \in A}, a_n\right), \tag{4}$$

where $n \in \mathbb{N}$ and $a_1, \ldots, a_n \in A$. Because (3) is valid and ρ is injective on $a_1A^* \cap A^*a_n$, a word $a_1 \cdots a_n$ belongs to L if and only if $\rho(a_1 \cdots a_n) \in L_{a_1,a_n}$, and by (4) this holds if and only if

$$\sigma(a_1 \cdots a_n) \in \{a_1\} \times \prod_{a,b \in A} T_{a,b}^{a_1,a_n} \times \{a_n\},$$

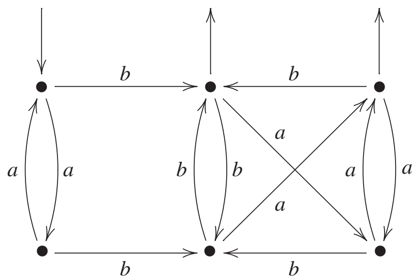
where

$$T_{a,b}^{a_1,a_n} = \begin{cases} \sigma_{L_{a,b}}(L_{a,b}) & \text{if } a = a_1 \text{ and } b = a_n, \\ \mathcal{M}(L_{a,b}) & \text{otherwise.} \end{cases}$$

This shows that L is recognized by σ. □

The following example illustrates Theorem 8 on a concrete non-trivial inequality with one constant language recognized by a chain of simple semigroups.

Example 19. Let L be the ε-free language over the alphabet $A = \{a, b\}$ whose minimal automaton is the following:



The language L consists exactly of those words $u \in A^+$ which contain some occurrence of b and where the difference between the length of u and the number of blocks of occurrences of b in u is even. The syntactic semigroup of L is defined by the relations $a^3 = a$, $b^3 = b$, $ab^2 = a^2b$, $ba^2 = b^2a$ and $bab = b^2$; it is a chain of two simple semigroups whose elements are represented by the words a, a^2 and b, b^2, ab, ab^2, ba, b^2a, aba, ab^2a, respectively. If we want to apply Lemma 17 to the minimal automaton of the language L, we have to set $b < a$.

Let us consider the inequality $aXaXa \subseteq LXL$ with one variable X. It is easy to verify that this inequality possesses a largest solution, namely the regular language $(a^2)^*ab^2a(a^2)^* \cup A^*bA^+bA^*$. In the proof of Theorem 8 we have seen that this solution is upward closed with respect to the well partial order $\leqslant_{\sigma_L}$. In fact, there are precisely 87 minimal elements in this solution with respect to $\leqslant_{\sigma_L}$:

$$\{ab^2a\} \cup \big(\{\varepsilon, a, a^2\} \cdot (\{b, b^2\}\{a, a^2\}\{b, b^2\} \cup \{b^3, b^4, babab\}) \cdot \{\varepsilon, a, a^2\}\big)$$
$$\setminus (\{\varepsilon, a, a^2\}b^2a^2b^2\{\varepsilon, a, a^2\} \cup \{ab^4a, ab^2ab^2a, aba^2b^2a, ab^2a^2ba\}).$$

Let us now give a few simple examples showing that if languages in the set $\mathfrak{L}$ cannot be recognized by a chain of simple semigroups, then the conclusion of Theorem 8 often does not hold. In our examples we deal with the simplest semigroups which are not of this form, namely with null semigroups and semilattices. First, we look at what happens in the presence of infinite unions.

Example 20. Let $a \in A$ and let $\mathfrak{L}$ contain only the one-element language $\{a\}$, whose syntactic semigroup is a two-element null semigroup. Then if we take any non-regular set $N \subseteq \mathbb{N}$, the largest solution of the inequality $X \subseteq \bigcup_{n\in N} a^n$ is not regular.

A similar situation arises for $\mathfrak{L} = \{a^+, b^+\}$, where $a, b \in A$. Both languages a^+ and b^+ are recognized by a homomorphism to a three-element semilattice with a zero element and two incomparable elements corresponding to letters a and b. In this case, the largest solution of the inequality $X \subseteq \bigcup_{n\in N}(a^+b^+)^n$ is not regular provided $N \subseteq \mathbb{N}$ is a non-regular set of positive integers.

The following examples demonstrate that even if no infinitary operations are allowed to occur in our inequalities, the restriction to chains of simple semigroups is essential.

Example 21. Let $\mathfrak{L} = \{\{a\}, \{b\}\}$, where $a, b \in A$. To recognize these two languages we need a three-element null semigroup, and the largest solution of the inequality $X \subseteq aXa \cup \{b\}$ is the non-regular language $\{a^nba^n \mid n \in \mathbb{N}_0\}$.

Analogously, for the set of languages $\mathfrak{L} = \{a^+, b^+, c^+\}$, where $a, b, c \in A$, it is clear that the largest solution of the inequality $X \subseteq a^+b^+Xa^+b^+ \cup c^+$ is not regular, namely equal to the language $\bigcup_{n\in\mathbb{N}_0}(a^+b^+)^n c^+(a^+b^+)^n$. And in order to recognize the languages of $\mathfrak{L}$, one can use a four-element semilattice with a zero element and three mutually incomparable elements.

4. Semi-commutation

Let K and L be languages over the alphabet A and consider the inequality $XK \subseteq LX$. It is easy to see that the union of arbitrarily many solutions of this inequality is again its solution. In particular, this means that this inequality possesses the largest solution, namely the union of all solutions. In this section we show that the largest solution of the inequality $XK \subseteq LX$ is always regular provided L is a regular language. With this aim we introduce another well quasi-order on A^*. But this time we have to consider more involved structures than just plain sequences as we did in Section 3.

The basic idea of the proof is to think of the inequality $XK \subseteq LX$ as a game of two players, the attacker and the defender. The language K determines possible actions of the attacker and the language L determines possible actions of the defender. A position of the game is an arbitrary word w from A^*. At each step of the game, both players successively modify the word according to the following rules. When the game is in a position w, the attacker chooses any element v of K and appends it to w. If no word from L is a prefix of wv, the attacker wins. Otherwise the defender removes any word belonging to L from the beginning of wv. The resulting word is a new position of the game. The defender wins the game if and only if he manages to continue playing forever.

Observe that if the defender has a winning strategy for a given initial position $w \in A^*$, then the set of all positions of the game reachable from w in some scenario corresponding with a chosen winning strategy forms a solution of the inequality $XK \subseteq LX$ containing w. Conversely, given any solution $M \subseteq A^*$ of $XK \subseteq LX$, one can easily construct winning strategies of the defender for all elements of M. Therefore, the largest solution of the inequality $XK \subseteq LX$ is exactly the set of all positions of the game where the defender has a winning strategy. The main result of this section can then be reformulated as follows: if the set of possible actions of the defender is regular, then the set of all winning positions of the defender is regular no matter what actions are available to the attacker.

Given an initial position $w \in A^*$, we consider all possible sequences of actions of the defender which can be performed without removing any letters previously added by the attacker. In other words, we deal with all sequences

$(w_1, \dots, w_n)$ of elements of L whose concatenation $w_1 \cdots w_n$ is a prefix of the word w. We arrange these sequences into the form of a tree expressing the order of actions, i.e. the node $(w_1, \dots, w_n)$ of the tree will be an immediate successor of the node $(w_1, \dots, w_{n-1})$. Then every move of the defender can be seen as a choice of one of the immediate successors of the node representing the current position of the game. In addition, we have to consider for each node $(w_1, \dots, w_n)$ the suffix u of w satisfying $w = w_1 \cdots w_n u$. This word u can be removed by the defender in the following turn together with several letters previously added by the attacker. The only information the defender needs to know when removing u is which words can be appended to u in order to get a word from L. This is uniquely determined by the $\equiv_L$-class of u, and therefore it is sufficient to label the node $(w_1, \dots, w_n)$ with the element $\sigma_L(u) \in \mathcal{S}(L)$. Actually, it is not essential for the defender to know exactly which node of the tree has the desired label, because for any given successor of the current node the defender can reach a position of the game corresponding to this successor by removing several words belonging to L. That is why in the following construction we only indicate for each node which elements of $\mathcal{S}(L)$ occur as labels of its successors, i.e. we assign to each node a set of elements of $\mathcal{S}(L)$.

In this way, we construct a labelled tree for every $w \in A^*$. Then we introduce a well quasi-order on the set of such trees expressing possibility of using winning strategies for one initial position of the game also for another one and prove that the largest solution of the inequality $XK \subseteq LX$ is upward closed with respect to the quasi-order induced on A^*.

Let us now describe the construction in detail. Let $L \subseteq A^+$ be a regular language and let $\sigma_L : A^+ \to \mathcal{S}(L)$ be its syntactic homomorphism. We extend this homomorphism to a monoid homomorphism $\sigma_L : A^* \to \mathcal{S}(L)^1$ by defining $\sigma_L(\varepsilon) = 1$. By an *L-tree* we mean a quadruple $\tau = (N_\tau, r_\tau, \pi_\tau, \ell_\tau)$, where

- N_τ is a finite set of nodes of τ,
- $r_\tau \in N_\tau$ is a distinguished node called the root of τ,
- the mapping $\pi_\tau : N_\tau \setminus \{r_\tau\} \to N_\tau$ maps each node to its predecessor,
- for every $v \in N_\tau$ there exists a non-negative integer $d_\tau(v)$ called the distance of v such that $\pi_\tau^{d_\tau(v)}(v) = r_\tau$,
- the mapping $\ell_\tau : N_\tau \to 2^{\mathcal{S}(L)^1}$ is a labelling of nodes with sets of elements of $\mathcal{S}(L)^1$ satisfying the condition

$$\forall v \in N_\tau \setminus \{r_\tau\} : \ell_\tau(v) \subseteq \ell_\tau(\pi_\tau(v)). \tag{5}$$

We denote by $\mathcal{T}(L)$ the set of all L-trees.

Now we define a quasi-order $\sqsubseteq$ on $\mathcal{T}(L)$. For $\tau, \vartheta \in \mathcal{T}(L)$ we set $\tau \sqsubseteq \vartheta$ if and only if there exists a mapping $H : N_\tau \to N_\vartheta$ which satisfies

$$\forall v \in N_\tau : \ell_\tau(v) \subseteq \ell_\vartheta(H(v)), \tag{6}$$

$$\forall v \in N_\tau \setminus \{r_\tau\}\ \exists k \in \mathbb{N} : H(\pi_\tau(v)) = \pi_\vartheta^k(H(v)). \tag{7}$$

The relation $\sqsubseteq$ is in fact a well quasi-order on $\mathcal{T}(L)$ due to the famous Kruskal's Tree Theorem [12]. Actually, the usual formulation of the theorem states that labelled trees are well quasi-ordered by a relation which slightly differs from $\sqsubseteq$, since in its defining condition one requires the mapping H to map different edges of τ to disjoint paths of ϑ. But this means that with respect to the quasi-order $\sqsubseteq$ more pairs of trees are comparable than with respect to the standard one and therefore $\sqsubseteq$ is a well quasi-order too by Lemma 6.

Let us now prove that the assumption (5) ensures that if $\tau \sqsubseteq \vartheta$ then there exists a mapping H verifying this which maps the root of τ to the root of ϑ and every immediate successor of any node $v \in N_\tau$ to an immediate successor of $H(v) \in N_\vartheta$.

Lemma 22. *Let $\tau, \vartheta \in \mathcal{T}(L)$. Then $\tau \sqsubseteq \vartheta$ if and only if there exists $H : N_\tau \to N_\vartheta$ satisfying* (6) *and*

$$H(r_\tau) = r_\vartheta, \tag{8}$$

$$\forall v \in N_\tau \setminus \{r_\tau\}: H(\pi_\tau(v)) = \pi_\vartheta(H(v)); \tag{9}$$

in particular, the equality $d_\tau(v) = d_\vartheta(H(v))$ holds for every $v \in N_\tau$.

Proof. Let $H : N_\tau \to N_\vartheta$ be a mapping such that both conditions (6) and (7) hold. We will verify that the mapping $G : N_\tau \to N_\vartheta$ defined for every node $v \in N_\tau$ by the rule $G(v) = \pi_\vartheta^{d_\vartheta(H(v)) - d_\tau(v)}(H(v))$ satisfies all conditions (6), (8) and (9). Clearly we have $G(r_\tau) = \pi_\vartheta^{d_\vartheta(H(r_\tau))}(H(r_\tau)) = r_\vartheta$, hence (8) holds. In order to verify (9), take any $v \in N_\tau \setminus \{r_\tau\}$ and consider $k \in \mathbb{N}$ such that $H(\pi_\tau(v)) = \pi_\vartheta^k(H(v))$. Then

$$\begin{aligned}G(\pi_\tau(v)) &= \pi_\vartheta^{d_\vartheta(H(\pi_\tau(v)))-d_\tau(\pi_\tau(v))}(H(\pi_\tau(v)))\\ &= \pi_\vartheta^{d_\vartheta(\pi_\vartheta^k(H(v)))-d_\tau(v)+1}(\pi_\vartheta^k(H(v)))\\ &= \pi_\vartheta(\pi_\vartheta^{d_\vartheta(H(v))-k-d_\tau(v)+k}(H(v))) = \pi_\vartheta(G(v)).\end{aligned}$$

Finally, we have $\ell_\tau(v) \subseteq \ell_\vartheta(H(v)) \subseteq \ell_\vartheta(G(v))$ due to (5) and therefore G satisfies also (6). $\square$

Theorem 23. *If $K \subseteq A^*$ is an arbitrary language and $L \subseteq A^*$ is a regular language, then the largest solution and the largest ε-free solution of the inequality $XK \subseteq LX$ are regular.*

Proof. First note that we can assume $\varepsilon \notin L$, otherwise the largest solution is clearly equal to A^*.

In order to define a quasi-order $\leqslant_L$ on A^*, we construct a mapping $\varphi: A^* \to \mathcal{T}(L)$ as follows. For $w \in A^*$ let $N_{\varphi(w)}$ be the set of all finite sequences $(w_1, \ldots, w_n)$, where $n \in \mathbb{N}_0$ and $w_1, \ldots, w_n \in L$, such that the word $w_1 \cdots w_n$ is a prefix of w. The root $r_{\varphi(w)}$ of $\varphi(w)$ is the empty sequence and the predecessor mapping is given by the rule $\pi_{\varphi(w)}(w_1, \ldots, w_n) = (w_1, \ldots, w_{n-1})$. Finally, we put an element $s \in \mathcal{S}(L)^1$ into the set $\ell_{\varphi(w)}(w_1, \ldots, w_n)$ if and only if there exist words $\bar{w} \in L^*$ and $\tilde{w} \in A^*$ such that $w_1 \cdots w_n \bar{w}\tilde{w} = w$ and $\sigma_L(\tilde{w}) = s$.

Now we are ready to define the desired quasi-order. For $v, w \in A^*$ we set $v \leqslant_L w$ if and only if $\varphi(v) \sqsubseteq \varphi(w)$. Because $\sqsubseteq$ is a well quasi-order, by Lemma 6 this rule defines a well quasi-order on A^*.

Before proceeding, let us state one simple observation about the mapping φ which is easy to verify and will be of use later in the proof.

Claim 1. *For arbitrary words $w \in A^*$ and $x_1, \ldots, x_k \in L$, where $k \in \mathbb{N}_0$, the L-tree $\varphi(w)$ is isomorphic to the subtree of $\varphi(x_1 \cdots x_k w)$ rooted at the node $(x_1, \ldots, x_k)$ via the mapping $H: N_{\varphi(w)} \to N_{\varphi(x_1\cdots x_k w)}$ sending each node $(w_1, \ldots, w_n) \in N_{\varphi(w)}$ to the node $(x_1, \ldots, x_k, w_1, \ldots, w_n) \in N_{\varphi(x_1\cdots x_k w)}$. In particular, we have $w \leqslant_L x_1 \cdots x_k w$.*

We are going to prove that every solution of the inequality $XK \subseteq LX$ is contained in a regular solution upward closed with respect to $\leqslant_L$, which immediately implies that the largest solution is regular. Let $M \subseteq A^*$ be any solution of $XK \subseteq LX$ and consider the language $P = \{w \in A^* \mid \exists\, v \in M: v \leqslant_L w\}$. First observe that $\varepsilon \in P$ if and only if $\varepsilon \in M$ because from $v \leqslant_L \varepsilon$ it follows that $\sigma_L(v) \in \ell_{\varphi(v)}(r_{\varphi(v)}) \subseteq \ell_{\varphi(\varepsilon)}(r_{\varphi(\varepsilon)}) = \{1\}$, which means $v = \varepsilon$ since σ_L was defined on A^* by adding a new element 1 to $\mathcal{S}(L)$. As we trivially have $M \subseteq P$, to conclude that both statements of the theorem are true it suffices to verify that the language P is regular and satisfies $PK \subseteq LP$, which is the aim of the rest of the proof.

In order to show $PK \subseteq LP$, take any words $w \in P$ and $u \in K$. Then there exists $v \in M$ such that $v \leqslant_L w$. Let $H: N_{\varphi(v)} \to N_{\varphi(w)}$ be a mapping verifying this inequality and let us additionally assume that it satisfies (8) and (9), which is possible due to Lemma 22. Because M is a solution of $XK \subseteq LX$, there exist words $x \in L$ and $y \in M$ such that $vu = xy$. We have to distinguish two situations.

Let us first assume that v is a prefix of x. Then we have $x = v\bar{x}$ and $u = \bar{x}y$ for a certain word $\bar{x} \in A^*$. Because $\sigma_L(v) \in \ell_{\varphi(v)}(r_{\varphi(v)}) \subseteq \ell_{\varphi(w)}(r_{\varphi(w)})$, there exist words $\bar{w} \in L^*$ and $\tilde{w} \in A^*$ which satisfy $\bar{w}\tilde{w} = w$ and $\sigma_L(\tilde{w}) = \sigma_L(v)$. Now we can calculate $\sigma_L(\tilde{w}\bar{x}) = \sigma_L(v\bar{x}) = \sigma_L(x)$, and since $x \in L$, this implies $\tilde{w}\bar{x} \in L$. If $\bar{w} = \varepsilon$ then we get $wu = (\tilde{w}\bar{x})y \in LM \subseteq LP$. Otherwise we have $\bar{w} = \bar{w}_1\bar{w}_2$ for some words $\bar{w}_1 \in L$ and $\bar{w}_2 \in L^*$. By means of Claim 1 we obtain $y \leqslant_L \bar{w}_2(\tilde{w}\bar{x})y$, which demonstrates that $wu = \bar{w}_1(\bar{w}_2\tilde{w}\bar{x}y) \in LP$ as required.

Now let us consider the case when the word x is a proper prefix of v. Then there is a word $\bar{v} \in A^+$ satisfying $v = x\bar{v}$ and $y = \bar{v}u$. In particular, we have $(x) \in N_{\varphi(v)}$ and due to (8) and (9) we obtain $H(x) = (w_0)$ for a certain word $w_0 \in L$. Consequently $w = w_0\hat{w}$, where $\hat{w} \in A^*$. In the following we construct a mapping $G: N_{\varphi(y)} \to N_{\varphi(\hat{w}u)}$ verifying the inequality $y \leqslant_L \hat{w}u$; this inequality immediately gives the desired fact $wu = w_0(\hat{w}u) \in LP$.

The definition of G consists of two parts. First we define G for elements of $N_{\varphi(y)}$ which form prefixes of the word $\bar{v}$. Let $k \in \mathbb{N}_0$ and let $(y_1, \ldots, y_k) \in N_{\varphi(y)}$ be such that $y_1, \ldots, y_k \in L$ and $y_1 \cdots y_k$ is a prefix of $\bar{v}$. Then $xy_1 \cdots y_k$ is a prefix of v and so we can consider the sequence $H(x, y_1, \ldots, y_k)$. Due to (9) and $H(x) = (w_0)$, this sequence is of the form $H(x, y_1, \ldots, y_k) = (w_0, w_1, \ldots, w_k)$, where $w_1, \ldots, w_k \in L$ and $w_1 \cdots w_k$ is a prefix of

$\hat{w}$. Then we define $G(y_1, \ldots, y_k) = (w_1, \ldots, w_k)$. Before dealing with the second part of the definition of G, let us verify that this part of the definition is in accord with conditions (6) and (7).

The verification of (7) is straightforward since (9) holds for H and thus, for $k \geqslant 1$,

$$\begin{aligned} H(x, y_1, \ldots, y_{k-1}) &= H(\pi_{\varphi(v)}(x, y_1, \ldots, y_k)) \\ &= \pi_{\varphi(w)}(H(x, y_1, \ldots, y_k)) = (w_0, w_1, \ldots, w_{k-1}), \end{aligned}$$

which shows

$$G(\pi_{\varphi(y)}(y_1, \ldots, y_k)) = (w_1, \ldots, w_{k-1}) = \pi_{\varphi(\hat{w}u)}(G(y_1, \ldots, y_k)).$$

In order to verify (6), take any element $s \in \ell_{\varphi(y)}(y_1, \ldots, y_k)$. We are going to demonstrate $s \in \ell_{\varphi(\hat{w}u)}(w_1, \ldots, w_k)$. By the definition of the labelling $\ell_{\varphi(y)}$, one can find words $\bar{y} \in L^*$ and $\tilde{y} \in A^*$ satisfying $y_1 \cdots y_k \bar{y}\tilde{y} = y$ and $\sigma_L(\tilde{y}) = s$. Now we have to compare the lengths of $y_1 \cdots y_k \bar{y}$ and $\bar{v}$ and consider two different situations.

First assume that the word $y_1 \cdots y_k \bar{y}$ is a prefix of $\bar{v}$ and let $\hat{y} \in A^*$ be the word which satisfies $\tilde{y} = \hat{y}u$. Then $v = xy_1 \cdots y_k \bar{y}\hat{y}$ and therefore $\sigma_L(\hat{y}) \in \ell_{\varphi(v)}(x, y_1, \ldots, y_k)$. Because (6) is valid for H, this implies $\sigma_L(\hat{y}) \in \ell_{\varphi(w)}(w_0, w_1, \ldots, w_k)$. In other words, we obtain $w = w_0 w_1 \cdots w_k \bar{w}\tilde{w}$ and $\sigma_L(\tilde{w}) = \sigma_L(\hat{y})$ for certain words $\bar{w} \in L^*$ and $\tilde{w} \in A^*$. Consequently $\hat{w}u = w_1 \cdots w_k \bar{w}\tilde{w}u$, which shows

$$s = \sigma_L(\tilde{y}) = \sigma_L(\hat{y}u) = \sigma_L(\tilde{w}u) \in \ell_{\varphi(\hat{w}u)}(w_1, \ldots, w_k).$$

In the case $\bar{v}$ is a proper prefix of $y_1 \cdots y_k \bar{y}$, we can write $\bar{v} = y_1 \cdots y_k \bar{y}_0 \hat{y}_0$ and $u = \hat{y}_1 \bar{y}_1 \tilde{y}$ for certain words $\bar{y}_0, \bar{y}_1 \in L^*$, $\hat{y}_0 \in A^*$ and $\hat{y}_1 \in A^+$ satisfying $\hat{y}_0 \hat{y}_1 \in L$. Then $\sigma_L(\hat{y}_0) \in \ell_{\varphi(v)}(x, y_1, \ldots, y_k)$, and therefore, $\sigma_L(\hat{y}_0) \in \ell_{\varphi(w)}(w_0, w_1, \ldots, w_k)$ since the mapping H satisfies (6). This provides us with $\bar{w} \in L^*$ and $\tilde{w} \in A^*$ such that $w_0 w_1 \cdots w_k \bar{w}\tilde{w} = w$ and $\sigma_L(\tilde{w}) = \sigma_L(\hat{y}_0)$. Altogether, we get $\hat{w}u = w_1 \cdots w_k \bar{w}\tilde{w}\hat{y}_1 \bar{y}_1 \tilde{y}$, where $\bar{w}\tilde{w}\hat{y}_1 \bar{y}_1 \in L^*$ because $\tilde{w}\hat{y}_1 \in L$ is a consequence of $\sigma_L(\tilde{w}\hat{y}_1) = \sigma_L(\hat{y}_0 \hat{y}_1)$. Hence $s = \sigma_L(\tilde{y}) \in \ell_{\varphi(\hat{w}u)}(w_1, \ldots, w_k)$.

Now we proceed to the second part of the definition of G. In order to deal with elements of $N_{\varphi(y)}$ which form prefixes of y longer than $\bar{v}$, we first choose certain sequences of words from L for all nodes of $N_{\varphi(\bar{v})}$. Let $l \in \mathbb{N}$ and let $y_1, \ldots, y_{l-1} \in L$ and $\bar{y}_l \in A^*$ be words satisfying $y_1 \cdots y_{l-1}\bar{y}_l = \bar{v}$, i.e. $(y_1, \ldots, y_{l-1}) \in N_{\varphi(\bar{v})}$ and $\bar{y}_l$ is the corresponding suffix of $\bar{v}$. Using (9) we obtain $H(x, y_1, \ldots, y_{l-1}) = (w_0, w_1, \ldots, w_{l-1})$ for some words $w_1, \ldots, w_{l-1} \in L$ such that $w_1 \cdots w_{l-1}$ is a prefix of $\hat{w}$. Because

$$\sigma_L(\bar{y}_l) \in \ell_{\varphi(v)}(x, y_1, \ldots, y_{l-1}) \subseteq \ell_{\varphi(w)}(w_0, w_1, \ldots, w_{l-1})$$

holds due to (6), there exist an integer $m \in \mathbb{N}_0$ and words $\bar{w}_1, \ldots, \bar{w}_m \in L$ and $\tilde{w} \in A^*$ satisfying $w_0 w_1 \cdots w_{l-1}\bar{w}_1 \cdots \bar{w}_m \tilde{w} = w$ and $\sigma_L(\tilde{w}) = \sigma_L(\bar{y}_l)$. Let us choose such words $\bar{w}_1, \ldots, \bar{w}_m, \tilde{w}$ arbitrarily and define

$$F(y_1, \ldots, y_{l-1}) = (w_1, \ldots, w_{l-1}, \bar{w}_1, \ldots, \bar{w}_m, \tilde{w}). \tag{10}$$

Now let $k \in \mathbb{N}$ and let $(y_1, \ldots, y_k) \in N_{\varphi(y)}$ be such that $y_1, \ldots, y_k \in L$ and the word $\bar{v}$ is a proper prefix of $y_1 \cdots y_k$. Then there exist $l \in \mathbb{N}$, $1 \leqslant l \leqslant k$, and words $\bar{y}_l \in A^*$ and $\tilde{y}_l \in A^+$ such that $y_l = \bar{y}_l \tilde{y}_l$, $\bar{v} = y_1 \cdots y_{l-1}\bar{y}_l$ and the word $\tilde{y}_l y_{l+1} \cdots y_k$ is a prefix of u. Let us consider the sequence (10) previously chosen for the node $(y_1, \ldots, y_{l-1}) \in N_{\varphi(\bar{v})}$ and define

$$G(y_1, \ldots, y_k) = (w_1, \ldots, w_{l-1}, \bar{w}_1, \ldots, \bar{w}_m, \tilde{w}\tilde{y}_l, y_{l+1}, \ldots, y_k).$$

Note that this sequence belongs to $N_{\varphi(\hat{w}u)}$ because $\sigma_L(\tilde{w}\tilde{y}_l) = \sigma_L(\bar{y}_l \tilde{y}_l) = \sigma_L(y_l)$, which implies $\tilde{w}\tilde{y}_l \in L$. It is easy to see that (6) holds for this part of the definition since if we denote by z the product of all words in the sequence $G(y_1, \ldots, y_k)$, we obtain $(y_1 \cdots y_k)^{-1} y = z^{-1}(\hat{w}u)$ and therefore in this case we have even an equality $\ell_{\varphi(y)}(y_1, \ldots, y_k) = \ell_{\varphi(\hat{w}u)}(G(y_1, \ldots, y_k))$.

To prove that G is a required mapping, it remains to verify condition (7) for the second part of the definition of G. If $l < k$ then $\bar{v}$ is a proper prefix of $y_1 \cdots y_{k-1}$, therefore the second part of the definition applies also to the sequence $(y_1, \ldots, y_{k-1})$ and consequently condition (7) holds for $(y_1, \ldots, y_k)$ because the words $\bar{w}_1, \ldots, \bar{w}_m, \tilde{w}$ are fixed for a given sequence $(y_1, \ldots, y_{l-1})$. And if $l = k$ then the first part of the definition applies to the sequence $(y_1, \ldots, y_{k-1})$

and we obtain

$$\begin{aligned} G(\pi_{\varphi(y)}(y_1, \ldots, y_k)) &= G(y_1, \ldots, y_{k-1}) \\ &= (w_1, \ldots, w_{k-1}) \\ &= \pi^{m+1}_{\varphi(\hat{w}u)}(w_1, \ldots, w_{k-1}, \bar{w}_1, \ldots, \bar{w}_m, \tilde{w}\tilde{y}_k) \\ &= \pi^{m+1}_{\varphi(\hat{w}u)}(G(y_1, \ldots, y_k)). \end{aligned}$$

The proof of the theorem will be complete when we demonstrate that P is a regular language. Because P is upward closed with respect to the well quasi-order $\leqslant_L$, it is a finite union of languages of the form $\langle v\rangle = \{w \in A^* \mid v \leqslant_L w\}$ for a word $v \in A^*$. Therefore, it remains to show that each language $\langle v\rangle$ is regular. This is an immediate consequence of the following formula which describes $\langle v\rangle$ inductively using operations preserving regularity.

Claim 2. *Let $v \in A^*$ be an arbitrary word. Then*

$$\langle v\rangle = \bigcap\{L \cdot \langle x^{-1}v\rangle \mid x \in L \text{ is a prefix of } v\} \cap \bigcap\{L^* \cdot \sigma_L^{-1}(s) \mid s \in \ell_{\varphi(v)}(r_{\varphi(v)})\}. \tag{11}$$

First, take any $w \in \langle v\rangle$ and let us prove that w belongs to the language on the right-hand side of (11). Employing Lemma 22, we get a mapping $H : N_{\varphi(v)} \to N_{\varphi(w)}$ satisfying (6), (8) and (9). Let $x \in L$ be a prefix of v. Then $H(x) = (y)$ for a certain prefix $y \in L$ of the word w. Because the labelled tree $\varphi(x^{-1}v)$ is isomorphic to the subtree of $\varphi(v)$ rooted at (x) and similarly $\varphi(y^{-1}w)$ is isomorphic to the subtree of $\varphi(w)$ rooted at (y) due to Claim 1, the restriction of H to the subtree of $\varphi(v)$ rooted at (x) shows that $x^{-1}v \leqslant_L y^{-1}w$. Therefore $w \in L \cdot \langle x^{-1}v\rangle$ holds. Now consider any element $s \in \ell_{\varphi(v)}(r_{\varphi(v)})$. Then $s \in \ell_{\varphi(w)}(r_{\varphi(w)})$ by (6), which can be reformulated as $w \in L^* \cdot \sigma_L^{-1}(s)$, and so the word w lies in all languages on the right-hand side of (11).

In order to prove the reverse inclusion, let a word w belong to the right-hand side of (11). In particular, this means that for every prefix $x \in L$ of v there exist a prefix $y_x \in L$ of w and a mapping $H_x : N_{\varphi(x^{-1}v)} \to N_{\varphi(y_x^{-1}w)}$ satisfying (6), (8) and (9). This enables us to construct a mapping $H : N_{\varphi(v)} \to N_{\varphi(w)}$ by defining $H(r_{\varphi(v)}) = r_{\varphi(w)}$ and $H(x, v_1, \ldots, v_n) = (y_x, w_1, \ldots, w_n)$ whenever $H_x(v_1, \ldots, v_n) = (w_1, \ldots, w_n)$. In other words, the restriction of H to the subtree rooted at (x) is defined as the composition $\kappa_x \circ H_x \circ \iota_x$, where ι_x denotes the isomorphism between the subtree of $\varphi(v)$ rooted at (x) and $\varphi(x^{-1}v)$ and κ_x denotes the isomorphism between $\varphi(y_x^{-1}w)$ and the subtree of $\varphi(w)$ rooted at (y_x). This implies that condition (6) holds for all nodes of $\varphi(v)$ distinct from the root. In order to verify (6) also for the root of $\varphi(v)$, observe that for every $s \in \ell_{\varphi(v)}(r_{\varphi(v)})$ we have $w \in L^* \cdot \sigma_L^{-1}(s)$, which is equivalent to $s \in \ell_{\varphi(w)}(r_{\varphi(w)})$. Finally, condition (9) holds for H since it is valid for all mappings H_x.

This concludes the proof of Claim 2 and the required regularity of P follows. □

Although Theorem 23 does not provide us with an algorithm calculating largest solutions of inequalities of the form $XK \subseteq LX$, where K and L are regular languages, it enables us to algorithmically decide whether a given word belongs to the largest solution of such an inequality. The argument we use to demonstrate this can be found in [5] for the case of commutation equations; it consists of proving that both the largest solution and its complement can be recursively enumerated.

In [16] it is proved that complements of largest solutions of arbitrary finite systems of language inequalities with Boolean operations and concatenation are recursively enumerable. Actually, a uniform procedure for all such equations is described there. This in particular means that there exists a procedure which for given regular languages K and L recursively enumerates all words over A which do not belong to the largest solution of the inequality $XK \subseteq LX$. This result is based on the observation that if the defender has no winning strategy for a given position $w \in A^*$, then the attacker has a finite winning strategy for w. The reason why this is true is that every word has only finitely many prefixes which the defender can choose from. On the other hand, there exists also a procedure which for given regular languages K and L recursively enumerates all words from the largest solution of $XK \subseteq LX$; because the largest solution is regular by Theorem 23, it is sufficient to enumerate for every word w all regular languages containing w and for each of them test whether it is a solution. Altogether, we obtain the following corollary:

Corollary 24. *There exists an algorithm which decides for given regular languages K and L and a word $w \in A^*$ whether w belongs to the largest solution of the inequality $XK \subseteq LX$. Such an algorithm exists also for largest ε-free solutions.*

Now we give an example which in particular shows that it is essential to consider the whole tree structure associated with each word, not only the corresponding elements of the syntactic semigroup and the lengths of paths in the tree.

Example 25. Consider the alphabet $A = \{a, b, c, d, e, f, g, h, i\}$ and let $K = \{e, i\}$ and

$$L = \{a, b, bc, f, fg, gh, hc, e, i\} \cup cdKiK \cup dKeK \cup g^*\{b, h\}cdK^3.$$

One can show that the largest solution of the inequality $XK \subseteq LX$ is the language

$$L^* \cup L^*cdKi \cup L^*dKe \cup L^*g^*\{b, h\}cdK^2 \cup L^+g^*\{b, h\}cdK \cup bcdK \cup LL^+g^*\{b, h\}cd \cup Lbcd.$$

In order to calculate this solution, observe that if u belongs to a solution M of the inequality, then for every $n \in \mathbb{N}$ we have $ue^n \in MK^n \subseteq L^nM$. Because $\varepsilon \notin L$, if we take n sufficiently large, we deduce that the word ue^3 is a prefix of a word from L^*. Since this is possible only if $ue^3 \in L^*$, there remain only few cases to deal with.

We know that the largest solution of the inequality $XK \subseteq LX$ is upward closed with respect to the well quasi-order $\leqslant_L$, therefore one can find finitely many elements of the set $\mathcal{T}(L)$ characterizing the solution, i.e. minimal elements of the image of the solution under the mapping φ. In our case, there are four one-node trees corresponding to the words ε, cdi^2, de^2 and $gbcde^2$, which generate the languages L^*, L^*cdKi, L^*dKe and $L^*g^*\{b, h\}cdK^2$, respectively. Further there are four trees with more than one node corresponding to the words $agbcde$, $bcde$, a^2gbcd and $abcd$, respectively:

$$\begin{tikzcd} 0 \arrow[d] \\ bcde \end{tikzcd} \qquad \begin{tikzcd}[column sep=small] & bcde \arrow[dl] \arrow[dr] & \\ cde & & de \end{tikzcd} \qquad \begin{tikzcd} 0 \arrow[d] \\ 0 \arrow[d] \\ bcd \end{tikzcd} \qquad \begin{tikzcd}[column sep=small] & 0 \arrow[d] & \\ & bcd \arrow[dl] \arrow[dr] & \\ cd & & d \end{tikzcd}$$

where each node is labelled with the set of σ_L-images of all words written in its successor nodes (including the node itself) and by the symbol 0 we mean the zero element of the semigroup $\mathcal{S}(L)$. These trees generate the languages $L^+g^*\{b, h\}cdK$, $L^*\{b, fgh\}cdK$, $LL^+g^*\{b, h\}cd$ and $L^+\{b, fgh\}cd$, respectively.

Finally, let us point out that the word $fghcd$ does not belong to the solution even though the tree corresponding to it is very similar to the one of $abcd$ described above:

$$\begin{tikzcd}[column sep=small] & 0 \arrow[dl] \arrow[dr] & \\ bcd \arrow[d] & & bcd \arrow[d] \\ cd & & d \end{tikzcd}$$

This is a consequence of the facts $abcd \equiv_L fghcd$ and $bcd \equiv_L hcd \equiv_L ghcd$. Moreover, the same equivalences hold and therefore labels of these two trees are equal even if we consider elements of the syntactic semigroup of L^+ instead of L. In fact, when the inequality is viewed as a game, the difference between these two trees is that for the word $fghcd$ the defender has to make his decision immediately after the first turn of the attacker whereas for the word $abcd$ he can decide according to the attacker's second move.

Acknowledgements

I am very grateful to Ondřej Klíma for carefully reading the manuscript and for providing me with numerous helpful comments. I also appreciate Jean-Éric Pin's suggestion to investigate a connection between decomposition quasi-orders and unavoidability of sets of words, which allowed to simplify the proof of Theorem 10.

References

[1] F. Baader, R. Küsters, Unification in a description logic with transitive closure of roles, in: R. Nieuwenhuis, A. Voronkov (Eds.), Proc. eighth Internat. Conf. on Logic for Programming, Artificial Intelligence and Reasoning (LPAR 2001), Lecture Notes in Computer Science, vol. 2250, Springer, Berlin, 2001, pp. 217–232.

[2] W. Bucher, A. Ehrenfeucht, D. Haussler, On total regulators generated by derivation relations, Theoret. Comput. Sci. 40 (1985) 131–148.
[3] C. Choffrut, J. Karhumäki, N. Ollinger, The commutation of finite sets: a challenging problem, Theoret. Comput. Sci. 273 (2002) 69–79.
[4] J.H. Conway, Regular Algebra and Finite Machines, Chapman & Hall, London, 1971.
[5] K. Culik II, J. Karhumäki, P. Salmela, Fixed point approach to commutation of languages, in: N. Jonoska, Gh. Păun, G. Rozenberg (Eds.), Aspects of Molecular Computing, Lecture Notes in Computer Science, vol. 2950, Springer, Berlin, 2004, pp. 119–131.
[6] A. de Luca, S. Varricchio, Finiteness and Regularity in Semigroups and Formal Languages, Springer, Berlin, 1999.
[7] A. Ehrenfeucht, D. Haussler, G. Rozenberg, On regularity of context-free languages, Theoret. Comput. Sci. 27 (1983) 311–332.
[8] J.M. Howie, Fundamentals of Semigroup Theory, Clarendon Press, Oxford, 1995.
[9] J. Karhumäki, I. Petre, Conway's problem for three-word sets, Theoret. Comput. Sci. 289 (2002) 705–725.
[10] J. Karhumäki, I. Petre, Two problems on commutation of languages, in: Gh. Păun, G. Rozenberg, A. Salomaa (Eds.), Current Trends in Theoretical Computer Science, World Scientific, Singapore, 2004, pp. 477–494.
[11] D. Kirsten, The finite power problem revisited, Inform. Process. Lett. 84 (2002) 291–294.
[12] J.B. Kruskal, Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture, Trans. Amer. Math. Soc. 95 (1960) 210–225.
[13] J.B. Kruskal, The theory of well-quasi-ordering: a frequently discovered concept, J. Combin. Theory Ser. A 13 (1972) 297–305.
[14] M. Kunc, The power of commuting with finite sets of words, manuscript (2004), available at http://www.math.muni.cz/∼kunc/, extended abstract in: Proc. 22nd Symp. on Theoretical Aspects of Computer Science (STACS 2005), Lecture Notes in Computer Science, vol. 3404, Springer, Berlin, 2005, pp. 569–580.
[15] E.L. Leiss, Language Equations, Springer, New York, 1999.
[16] A. Okhotin, Decision problems for language equations, 2003, submitted for publication, available at http://www.cs.queensu.ca/home/okhotin/.
[18] G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, vol. 1, Springer, Berlin, 1997.
[19] A. Suschkewitsch, Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit, Math. Ann. 99 (1928) 30–50.
[20] S. Zhang, Efficient simplicity testing of automata, Theoret. Comput. Sci. 99 (1992) 265–278.