SUBHIERARCHIES OF THE SECOND LEVEL IN THE STRAUBING-THÉRIEN HIERARCHY

ONDŘEJ KLÍMA*

Department of Mathematics and Statistics, Masaryk University
Kotlářská 2, 611 37 Brno, Czech Republic
klima@math.muni.cz

LIBOR POLÁK*

Department of Mathematics and Statistics, Masaryk University
Kotlářská 2, 611 37 Brno, Czech Republic
polak@math.muni.cz

Received (Day Month Year)
Revised (Day Month Year)

Communicated by [editor]

In a recent paper we assigned to each positive variety $\mathcal{V}$ and each nonnegative integer k the class of all finite unions of finite intersections or Boolean combinations of the languages of the form $L_0a_1L_1a_2 \dots a_\ell L_\ell$, where $\ell \leq k$, $a_1, \dots, a_\ell$ are letters and $L_0, \dots, L_\ell$ are in the variety $\mathcal{V}$. For these polynomial operators on a wide class of varieties we gave a certain algebraic counterpart in terms of identities satisfied by syntactic (ordered) monoids of languages considered. Here we apply our constructions to particular examples of varieties of languages obtaining four hierarchies of (positive) varieties. Two of them have the 3/2 level of the Straubing-Thérien hierarchy as their limits, and two others tend to the level two of this hierarchy. We concentrate here on the existence of finite bases of identities for corresponding pseudovarieties of (ordered) monoids and we are looking for inclusions among those varieties.

Keywords: Positive varieties of languages; polynomial operators.

1. Introduction

The *positive polynomial operator* PPol assigns to each positive variety of languages $\mathcal{V}$ the class of all finite unions of finite intersections of the languages of the form

$$L_0a_1L_1 \dots a_\ell L_\ell \ , \qquad (*)$$

where A is an alphabet, $a_1, \dots, a_\ell \in A$, $L_0, \dots, L_\ell \in \mathcal{V}(A)$ (i.e. they are over A). Using Boolean combinations, we get *polynomial operator* BPol. Such operators on

*Both authors were supported by the Ministry of Education of the Czech Republic under the project MSM 0021622409 and by the Grant 201/09/1313 of the Grant Agency of the Czech Republic

classes of languages lead to several concatenation hierarchies. Well-known cases are the Straubing-Thérien hierarchy and the group hierarchy. Concatenation hierarchies have been intensively studied by many authors – see Section 8 of Pin's Chapter [8]. The main open problem concerning such hierarchies, which is in fact one of the most interesting open problem in the theory of regular languages, is the membership problem for level 2 in the Straubing-Thérien hierarchy, i.e. the decision problem whether a given regular language can be written as a Boolean combination of languages of the form $(*)$ where L_i's are from level 1 of that hierarchy. It is known that a language is of level 2 if and only if it is a Boolean combination of languages $(*)$ where $L_i = B_i^*$ and $B_i \subseteq A$ $(i = 0, \ldots, \ell)$. Therefore this instance of polynomial operator is the most important case to study.

In the restricted case we fix a nonnegative integer k and we allow only $\ell \leq k$ in $(*)$. This operator was considered mainly in the case that $\mathcal{V}$ is the trivial variety by Simon in [10], in a series of papers by Blanchet-Sadri, see for instance [3], and in a recent paper by the authors [5].

In [6] we considered the restricted case in a general setting and we concentrated on identity problems for corresponding pseudovarieties of ordered monoids and on the question whether those pseudovarieties are generated by a single object.

Here we study four hierarchies of languages which result by considering finite unions of finite intersections or Boolean combinations of languages $(*)$ over the (positive) variety $\mathcal{V}$, such that the set $\mathcal{V}(A)$ is equal either to finite unions of B^*, where $B \subseteq A$, or to finite unions of $\overline{B}$, where $B \subseteq A$ and $\overline{B}$ is the set of all words over A containing exactly the letters from B. Members of all our hierarchies are under level 2 in the Straubing-Thérien hierarchy. Therefore we speak about *subhierarchies*. Our basic questions are to explore inclusions among our varieties and we start to discuss the existence of finite bases for corresponding pseudovarieties of (ordered) monoids. Hopefully our results bring a bit more light into the complexity of the structure of (positive) subvarieties of the second level of the Straubing-Thérien hierarchy.

Section 2 summarizes the background concerning positive varieties of languages and corresponding classes of (finite) ordered monoids. In Section 3 we overview the necessary material from [6] dealing with locally finite varieties and polynomial operators. The next section is devoted to the existence of finite bases of identities for pseudovarieties corresponding to our hierarchies in the case $k = 1$. The last section investigates the inclusions among members of our hierarchies.

2. Preliminaries

For a relation ρ on a set S we define its *dual* relation $\rho^{\mathsf{d}} = \{ (v, u) \in S \times S \mid (u, v) \in \rho \}$. A *quasiorder* ρ on a set S is a reflexive and transitive relation. For such a relation, let $\widehat{\rho} = \rho \cap \rho^{\mathsf{d}}$ (sometimes we write also $(\rho)\widehat{}$) be the corresponding equivalence relation. A relation γ on a monoid $(M, \cdot)$ is *compatible*, if for every

$u, v, w \in M$, we have

$$u \;\gamma\; v \quad \text{implies} \quad uw \;\gamma\; vw, \quad wu \;\gamma\; wv\,.$$

Compatible equivalence relations are called *congruences.*

An *ordered monoid* is a structure $(M, \cdot, 1, \leq)$ where $(M, \cdot, 1)$ is a monoid and $\leq$ is a *compatible* order on $(M, \cdot)$. *Homomorphisms* of ordered monoids are isotone monoid homomorphisms.

Let Y^* be the set of all words over an alphabet Y including the empty word, denoted by λ, endowed by the concatenation product. For a word $u \in Y^*$, let

$$\mathsf{c}(u) = \{\, y \in Y \mid u = u'yu'' \text{ for some } u', u'' \in Y^* \,\}\ .$$

For a set $Z \subseteq Y$, let $\overline{Z} = \{\, u \in Y^* \mid \mathsf{c}(u) = Z \,\}$.

Let us recall now here some basic facts about Eilenberg-type theorems. The Boolean case was invented by Eilenberg [4] and the positive case was introduced by Pin [7].

A *class of languages* $\mathcal{V}$ associates to every finite alphabet A a set $\mathcal{V}(A)$ of regular languages over A. It is called a *positive variety of languages* if

- for each A, $\mathcal{V}(A)$ is closed under finite unions and finite intersections (in particular $\emptyset, A^* \in \mathcal{V}(A)$),
- for each A, $\mathcal{V}(A)$ is closed under derivatives, i.e. $L \in \mathcal{V}(A),\ u, v \in A^*$ implies $u^{-1}Lv^{-1} = \{\, w \in A^* \mid uwv \in L \,\} \in \mathcal{V}(A)$,
- $\mathcal{V}$ is closed under preimages in homomorphisms, i.e. for each A and B: $f : B^* \to A^*$ and $L \in \mathcal{V}(A)$ implies $f^{-1}(L) = \{\, v \in B^* \mid f(v) \in L \,\} \in \mathcal{V}(B)$.

To get the notion of a *Boolean variety* of languages, we use in the first item complements, too. (In literature the authors use usually the name "a variety of languages".)

The meaning of $\mathcal{V} \subseteq \mathcal{W}$ is that $\mathcal{V}(A) \subseteq \mathcal{W}(A)$, for each finite alphabet A. Similarly, $\bigcup_{i\in I}\ \mathcal{V}_i$ is the class of languages defined by $(\bigcup_{i\in I}\ \mathcal{V}_i)(A) = \bigcup_{i\in I}\ \mathcal{V}_i(A)$, for each finite A.

A *pseudovariety* of finite monoids is a class of finite monoids closed under submonoids, homomorphic images and direct products of finite families. Similarly for ordered monoids (see [8]). A *variety* of (ordered) monoids is a class of (ordered) monoids closed under submonoids, homomorphic images and arbitrary direct products. For a variety $\mathbf{V}$ of (ordered) monoids, the class $\mathsf{Fin}\,\mathbf{V}$ of all finite members of $\mathbf{V}$ is a pseudovariety. We call such pseudovarieties *equational.* In fact, in our paper, almost all pseudovarieties are equational.

We fix the set $X = \{x_1, x_2, \dots\}$ of *variables.* An *identity* is a pair (u, v), written as $u = v$, of words from X^*. An identity $u = v$ is *satisfied* in a monoid M if for each homomorphism $\phi : X^* \to M$ we have $\phi(u) = \phi(v)$. In such a case we write $M \models u = v$, and for a set of identities Π, we define

$$\mathsf{Mod}_=(\Pi) = \{\, M \mid (\ \forall\, \pi \in \Pi\ )\ M \models \pi \,\}\,.$$

Let $\mathsf{Id}_=(\mathbf{V})$ be the set of all identities which are satisfied in a class of monoids $\mathbf{V}$.

When considering ordered monoids, an *identity* is a pair (u,v), written as $u \le v$, of words over X. An identity $u \le v$ is *satisfied* in an ordered monoid $(M,\le)$ if for each homomorphism $\phi : X^* \to M$ we have $\phi(u) \le \phi(v)$. In such a case we write $(M,\le) \models u \le v$, and for a set of identities Π, we define

$$\mathsf{Mod}_{\le}(\Pi) = \{\, (M,\le) \mid (\ \forall\, \pi \in \Pi\ )\ (M,\le) \models \pi \,\}\,.$$

Let $\mathsf{Id}_{\le}(\mathbf{P})$ be the set of all identities which are satisfied in a class of ordered monoids $\mathbf{P}$.

A relation γ on X^* is *fully invariant* if, for every homomorphism $\varphi : X^* \to X^*$ and $u, v \in X^*$, we have that $u\ \gamma\ v$ implies $\varphi(u)\ \gamma\ \varphi(v)$.

Result 1 (see [2], [1]). *(i) The operators* $\mathsf{Id}_=$ *and* $\mathsf{Mod}_=$ *are pairwise inverse bijections between varieties of monoids and fully invariant congruences on* X^**.*

(ii) The operators $\mathsf{Id}_{\le}$ *and* $\mathsf{Mod}_{\le}$ *are pairwise inverse bijections between varieties of ordered monoids and fully invariant compatible quasiorders on* X^**.*

For a language $L \subseteq A^*$, we define the relations $\sim_L$ and $\preceq_L$ on A^* as follows: for $u, v \in A^*$ we have

$$u \sim_L v \text{ if and only if } (\ \forall\, p,q \in A^*\ )\ (\ puq \in L \iff pvq \in L\ )\,,$$

$$u \preceq_L v \text{ if and only if } (\ \forall\, p,q \in A^*\ )\ (\ pvq \in L \implies puq \in L\ )\,.$$

The relation $\sim_L$ is the *syntactic congruence* of L on A^*. It is of finite index (i.e. there are only finitely many classes) if and only if L is regular. The quotient structure $\mathsf{M}(L) = A^*/\sim_L$ is called the *syntactic monoid* of L.

The relation $\preceq_L$ is the *syntactic quasiorder* of L and we have $\widehat{\preceq_L} = \sim_L$. Hence $\preceq_L$ induces an order on $\mathsf{M}(L) = A^*/\sim_L$, namely: $[u]_{\sim_L} \le [v]_{\sim_L}$ if and only if $u \preceq_L v$. Then we speak about the *syntactic ordered monoid* of L and we denote this structure by $\mathsf{O}(L)$.

Result 2 (Eilenberg [4], Pin [7]). *(i) Boolean varieties of languages correspond to pseudovarieties of finite monoids. The correspondence, written* $\mathcal{V} \longleftrightarrow \mathbf{V}$*, is given by the following relationships: for a pseudovariety* $\mathbf{V}$ *of monoids, we have, for* $L \subseteq A^*$*,*

$$L \in \mathcal{V}(A) \textit{ if and only if } \mathsf{M}(L) \in \mathbf{V}\,,$$

and, for a Boolean variety $\mathcal{V}$ *of languages,* $\mathbf{V}$ *is the pseudovariety of monoids generated by*

$$\{\, \mathsf{M}(L) \mid A \textit{ finite},\ L \in \mathcal{V}(A) \,\}\,.$$

(ii) Positive varieties of languages correspond to pseudovarieties of finite ordered monoids. The correspondence, written $\mathcal{P} \longleftrightarrow \mathbf{P}$*, is given by the relationships: for a pseudovariety* $\mathbf{P}$ *of ordered monoids, we have, for* $L \subseteq A^*$*,*

$$L \in \mathcal{P}(A) \textit{ if and only if } \mathsf{O}(L) \in \mathbf{P}\,,$$

and, for a positive variety $\mathcal{P}$ of languages, $\mathbf{P}$ *is the pseudovariety of ordered monoids generated by*

$$\{\, \mathsf{O}(L) \mid A \textit{ finite},\ L \in \mathcal{P}(A)\,\}\,.$$

A crucial role in our paper is played by two following classes of languages (see [8]).

For each finite alphabet A, let $\mathcal{J}_1^-(A)$ be the set of all finite unions of the languages of the form B^*, where $B \subseteq A$. Then $\mathcal{J}_1^-$ is a positive variety of languages and the corresponding (equational) pseudovariety of ordered monoids is

$$\mathbf{J}_1^- = \mathsf{Fin\,Mod}_{\leq}(\, x^2 = x,\, xy = yx,\, 1 \leq x\,)\,.$$

We speak about *semilattices with the smallest element 1.*

Let $\mathcal{J}_1(A)$ be the set of all finite unions of the languages of the form $\overline{B}$, where $B \subseteq A$, for each finite set A. Then $\mathcal{J}_1$ is a Boolean variety of languages and the corresponding (equational) pseudovariety of monoids is

$$\mathbf{J}_1 = \mathsf{Fin\,Mod}_{=}(\, x^2 = x,\, xy = yx\,)\,.$$

We speak about *semilattices.*

3. Locally Finite Varieties of Languages and Polynomial Operators of Bounded Length

Here we overview the necessary material from [6]. In that paper and in this contribution we deal with concrete positive varieties of languages which correspond to locally finite pseudovarieties of ordered monoids. Each pseudovariety we consider is formed by finite members of locally finite (i.e. finitely generated ordered monoids are finite) variety of ordered monoids and consequently such a variety of languages can be described by a fully invariant compatible quasiorder on the monoid X^* which has locally finite index – see below.

A relation γ on X^* is *a finite characteristic* if it is a fully invariant compatible quasiorder on the monoid X^* satisfying the condition: for each finite subset Y of the set X, the set Y^* intersects only finitely many classes of $X^*/\,\widehat{\gamma}$.

Given a finite characteristic γ, for each finite or countable infinite alphabet A we define a relation γ_A (or sometimes $\gamma(A)$) called the *natural adaptation* of γ, by an identification of A with a subset of X as follows:

$$\gamma_A = \{\, (u,v) \in A^* \times A^* \mid \varphi(u)\ \gamma\ \varphi(v)\,\}\,,$$

where φ is an injective homomorphism from A^* to X^* such that $\varphi(A) \subseteq X$. Since γ is fully invariant, the definition does not depend on the homomorphism we choose.

We say that γ is a *finite characteristic of a class of languages* $\mathcal{V}$ if γ is a finite characteristic and for every finite alphabet A and for every regular language L over A, we have

$$L \in \mathcal{V}(A) \quad \text{if and only if} \quad \gamma_A \subseteq\, \preceq_L\,.$$

We fix notation for finite characteristics of $\mathcal{J}_1^-$ and $\mathcal{J}_1$:

$$\sigma = \mathsf{Id}_{\leq}\, \mathbf{J}_1^- = \{\, (u,v) \in X^* \times X^* \mid \mathsf{c}(u) \subseteq \mathsf{c}(v) \,\}\,,$$

$$\widehat{\sigma} = \mathsf{Id}_{=}\, \mathbf{J}_1 = \{\, (u,v) \in X^* \times X^* \mid \mathsf{c}(u) = \mathsf{c}(v) \,\}\,.$$

The following result is quite obvious. One can find its proof together with other results from this section in the authors' paper [6].

Result 3. *Let $\mathcal{V}$ be a class of languages and γ be a finite characteristic of $\mathcal{V}$. Then $\mathcal{V}$ is a positive variety of languages, corresponding to the pseudovariety $\mathbf{V} = \mathsf{Fin\,Mod}_{\leq}(\gamma)$. Moreover, if γ is an equivalence relation, then $\mathcal{V}$ is a Boolean variety of languages, corresponding to the pseudovariety $\mathbf{V} = \mathsf{Fin\,Mod}_{=}(\gamma)$.*

Result 4. *Let $\mathcal{V}$ be a positive variety of languages. Then the following conditions are equivalent.*
(i) For each finite alphabet A, the set $\mathcal{V}(A)$ is finite.
(ii) There exists a finite characteristic of $\mathcal{V}$.

A positive variety of languages $\mathcal{V}$ is called *locally finite* if it satisfies conditions from Result 4.

Let $\mathcal{V}$ be a positive variety of languages and let k be a nonnegative integer. We define the class $\mathsf{PPol}_k(\mathcal{V})$ as follows: for a finite alphabet A, $\mathsf{PPol}_k(\mathcal{V})(A)$ consists of finite unions of finite intersections of the languages of the form

$$L_0 a_1 L_1 \dots a_\ell L_\ell, \quad \text{where} \quad 0 \leq \ell \leq k,\ a_1, \dots, a_\ell \in A,\ L_0, \dots, L_\ell \in \mathcal{V}(A)\ . \quad (*)$$

Similarly, using Boolean combinations, we define the classes $\mathsf{BPol}_k(\mathcal{V})$. Clearly $\mathsf{PPol}_k(\mathcal{V}) \subseteq \mathsf{PPol}_{k'}(\mathcal{V})$ for $k \leq k'$. We denote the union of all $\mathsf{PPol}_k(\mathcal{V})$'s by $\mathsf{PPol}(\mathcal{V})$. Similarly for $\mathsf{BPol}_k(\mathcal{V})$'s.

Result 5. *If $\mathcal{V}$ is a positive variety of languages then $\mathsf{PPol}_k(\mathcal{V})$ is a positive variety of languages and $\mathsf{BPol}_k(\mathcal{V})$ is a Boolean variety of languages.*

Let k be a fixed nonnegative integer, let α be a finite characteristic, and let A be a finite or countable infinite alphabet (in particular the set X).

For a word $u \in A^*$, we say that

$$f = (u_0, a_1, u_1, a_2, u_2, \dots, a_\ell, u_\ell)$$

is *a factorization* of u of length $\ell \geq 0$ if $u_0, u_1, \dots, u_\ell \in A^*$, $a_1, a_2, \dots, a_\ell \in A$ and $u_0 a_1 u_1 \dots a_\ell u_\ell = u$. The set of all factorizations of the length at most k of the word u is denoted by $\mathsf{Fact}_k(u)$. For a factorization $f = (u_0, a_1, u_1, \dots, a_\ell, u_\ell)$ of a word $u \in A^*$ and a factorization $g = (v_0, b_1, v_1, \dots, b_m, v_m)$ of a word $v \in A^*$, we write

$$f \leq_\alpha g$$

if $\ell = m$, $a_i = b_i$ for every $i \in \{1, \dots, \ell\}$ and $u_i\ \alpha_A\ v_i$ for every $i \in \{0, 1, \dots, \ell\}$.

We define the relation $(\mathsf{p}_k(\alpha))_A$ on the set A^* as follows: for $u, v \in A^*$, we have

$$u\ (\mathsf{p}_k(\alpha))_A\ v \quad \text{if and only if} \quad (\forall\ g \in \mathsf{Fact}_k(v)\,)\ (\exists\ f \in \mathsf{Fact}_k(u)\,)\ f \leq_\alpha g\ .$$

Note that the relation $(\mathsf{p}_k(\alpha))_X$ is a finite characteristic (see the next result) and therefore the relation $(\mathsf{p}_k(\alpha))_A$ is equal to $((\mathsf{p}_k(\alpha))_X)_A$. We write $\mathsf{p}_k(\alpha)$ instead of $(\mathsf{p}_k(\alpha))_X$.

Result 6. *Let $\mathcal{V}$ be a locally finite positive variety of languages and α be a finite characteristic of $\mathcal{V}$. Then $\mathsf{PPol}_k(\mathcal{V})$ is a locally finite positive variety of languages with the finite characteristic $\mathsf{p}_k(\alpha)$ and $\mathsf{BPol}_k(\mathcal{V})$ is a locally finite Boolean variety of languages with the finite characteristic $(\mathsf{p}_k(\alpha))\widehat{\ }$.*

In this paper we study the hierarchies $\mathsf{PPol}_k(\mathcal{J}_1^-)$, $\mathsf{PPol}_k(\mathcal{J}_1)$, $\mathsf{BPol}_k(\mathcal{J}_1^-)$ and $\mathsf{BPol}_k(\mathcal{J}_1)$. We denote

$$\pi_k^- = \mathsf{p}_k(\sigma) \ \text{ and } \ \pi_k = \mathsf{p}_k(\widehat{\sigma})\,.$$

Next we present finite characteristics for our first two hierarchies explicitly. Let $u, v \in X^*$. Then

- $u\ \pi_k^-\ v$ iff $\forall\ (g_0, a_1, \dots, g_\ell) \in \mathsf{Fact}_k(v)\ \exists\ (f_0, a_1, \dots, f_\ell) \in \mathsf{Fact}_k(u)$ such that $\mathsf{c}(f_0) \subseteq \mathsf{c}(g_0), \dots, \mathsf{c}(f_\ell) \subseteq \mathsf{c}(g_\ell)$,
- $u\ \pi_k\ v$ iff $\forall\ (g_0, a_1, \dots, g_\ell) \in \mathsf{Fact}_k(v)\ \exists\ (f_0, a_1, \dots, f_\ell) \in \mathsf{Fact}_k(u)$ such that $\mathsf{c}(f_0) = \mathsf{c}(g_0), \dots, \mathsf{c}(f_\ell) = \mathsf{c}(g_\ell)$.

For the remaining two hierarchies we can use the intersections with the duals or we can write

- $u\ (\pi_k^-)\widehat{\ }\ v$ iff the sets of minimal elements (with respect to the quasiorder $\leq_\sigma$) of $\mathsf{Fact}_k(u)$ and $\mathsf{Fact}_k(v)$ are equal,
- $u\ \widehat{\pi_k}\ v$ iff $\{\,(\mathsf{c}(f_0), a_1, \mathsf{c}(f_1), \dots, \mathsf{c}(f_\ell)) \mid (f_0, a_1, f_1, \dots, f_\ell) \in \mathsf{Fact}_k(u)\,\} = \{\,(\mathsf{c}(g_0), a_1, \mathsf{c}(g_1), \dots, \mathsf{c}(g_\ell)) \mid (g_0, a_1, g_1, \dots, g_\ell) \in \mathsf{Fact}_k(v)\,\}$.

In the case $k = 1$, we can even write:

- $u\ \pi_1^-\ v$ iff $\mathsf{c}(u) \subseteq \mathsf{c}(v)$ and $(\ \forall\ v_0, v_1 \in A^*,\ a \in A$ such that $v = v_0 a v_1\,)$ $(\ \exists\ u_0, u_1 \in A^*$ such that $u = u_0 a u_1$ and $\mathsf{c}(u_0) \subseteq \mathsf{c}(v_0)$, $\mathsf{c}(u_1) \subseteq \mathsf{c}(v_1)\ )$

and similarly for the three remaining relations.

4. Bases of identities for pseudovarieties corresponding to the first members of our subhierarchies

Our goal now is to try to find finite bases of identities for pseudovarieties from our hierarchies of pseudovarieties of (ordered) monoids. As the results from [3,5] indicate we can not expect that such a finite bases exists for every k. In fact one can use methods by Volkov and Goldberg [11] to show that many pseudovarieties under consideration do not have any finite bases of identities. We will find finite bases in the case $k = 1$.

Note that in this case we consider factorizations of length ℓ with $\ell \leq 1$. But for $\ell = 0$ the condition $f \leq g$ for factorizations of a pair of words is exactly saying that the contents of the considered words are equal or they are in the inclusion. In the case the words are nonempty, this information is contained in the condition $f \leq g$ for factorizations of length $\ell = 1$. So we need not pay attention to the factors of length $\ell = 0$ when considering nonempty words.

4.1. *Identities for pseudovarieties corresponding to* $\mathsf{BPol}_1(\mathcal{J}_1^-)$ *and* $\mathsf{PPol}_1(\mathcal{J}_1^-)$

Let x, y be two different letters from X and let $u \in X^*$ be a word which contains both x and y, i.e. $x, y \in \mathsf{c}(u)$. Then $(uxyx, uyx) \in (\pi_1^-)$. Note that the set of identities

$$uxyx = uyx\ , \quad \text{where} \quad x, y \in \mathsf{c}(u) \tag{1}$$

is equivalent to a pair of identities. Indeed, we distinguish two cases, namely $u = z_1xz_2yz_3$ and $u = z_1yz_2xz_3$ $(z_1, z_2, z_3 \in X^*)$. Those identities are

$$z_1\, x\, z_2\, y\, z_3 \cdot\, x\, y\, x = z_1\, x\, z_2\, y\, z_3 \cdot\, y\, x\ , \tag{1a}$$

$$z_1\, y\, z_2\, x\, z_3 \cdot\, x\, y\, x = z_1\, y\, z_2\, x\, z_3 \cdot\, y\, x\ , \tag{1b}$$

where $z_1, z_2, z_3 \in X$.

We have also the dual version of the set of identities (1), namely

$$xyxu = xyu\, , \quad \text{where} \quad x, y \in \mathsf{c}(u)\, .$$

Similarly as above, this set is equivalent to two identities, which we denote by (1c) and (1d).

When we put $z_1 = z_2 = y = \lambda$ and $z_3 = z$ in (1a) then we obtain the identity

$$xzx^2 = xzx\, , \tag{1e}$$

and similar using (1c) we get

$$x^2zx = xzx\, . \tag{1f}$$

Another set of identities which is satisfied in the pseudovariety corresponding to $\mathsf{BPol}_1(\mathcal{J}_1^-)$ is

$$uxyv = uyxv\ , \quad \text{where} \quad x, y \in \mathsf{c}(u) \cap \mathsf{c}(v) \tag{2}$$

Note that this set is equivalent to the following identities:

$$z_1xz_2yz_3 \cdot xy \cdot t_1xt_2yt_3 = z_1xz_2yz_3 \cdot yx \cdot t_1xt_2yt_3\, , \tag{2a}$$

$$z_1xz_2yz_3 \cdot xy \cdot t_1yt_2xt_3 = z_1xz_2yz_3 \cdot yx \cdot t_1yt_2xt_3\, . \tag{2b}$$

When we work with π_1^- we observe that the identities

$$yzyx \leq yzxyx \quad \text{and} \quad xyzy \leq xyxzy \tag{3}$$

are satisfied in the pseudovariety corresponding to $\mathsf{PPol}_1(\mathcal{J}_1^-)$. Note that the identity

$$x \leq x^2$$

follows from the identities (3). Other identity which is satisfied in the pseudovariety corresponding to $\mathsf{PPol}_1(\mathcal{J}_1^-)$ is

$$xzxtx \leq xztx \ . \tag{4}$$

Note that the set of identities (1) is a consequence of the identities (3) and (4).

Proposition 1. *(i) The six identities* (1a-d) *and* (2a,b) *form a finite basis of identities for the variety of monoids corresponding to* $\mathsf{BPol}_1(\mathcal{J}_1^-)$.

(ii) The four identities (2a,b), (3) *and* (4) *form a finite basis of identities for the variety of ordered monoids corresponding to* $\mathsf{PPol}_1(\mathcal{J}_1^-)$.

Proof. The results follow from the series of lemmas below.

We need a bit more notation. For $u \in X^*$, we denote by:

- $\mathsf{first}(u)$ the sequence of the first occurrences of letters of u (from the left),
- $\mathsf{last}(u)$ the sequence of the last occurrences of letters of u,
- $\mathsf{Sub}_k(u)$ the set of all (scattered) subwords of u of length less or equal to k.

Note that $\mathsf{first}(\lambda) = \mathsf{last}(\lambda) = \lambda$.

Let $u \in X^*$ be a word and $x, y \in \mathsf{c}(u)$, $x \neq y$ be letters such that $xy \notin \mathsf{Sub}_2(u)$ (i.e. the last occurrence of y is before the first occurrence of x in the word u). Then u can be written in the form $u = u_0 y u_1 x u_2$ where $y \notin \mathsf{c}(u_1 x u_2)$ and $x \notin \mathsf{c}(u_0 y u_1)$. We put $\mathsf{int}_{y,x}(u) = \mathsf{c}(u_1)$. Notice that $x, y \notin \mathsf{int}_{y,x}(u)$.

Lemma 2. *Let $u, v \in X^*$ be arbitrary words. Then $(u, v) \in \pi_1^-$ if and only if the following conditions are satisfied*

- $\mathsf{first}(u) = \mathsf{first}(v)$ *and* $\mathsf{last}(u) = \mathsf{last}(v)$*;*
- $\mathsf{Sub}_2(u) \subseteq \mathsf{Sub}_2(v)$*;*
- *for every $x, y \in \mathsf{c}(u)$ such that $x \neq y$ and $xy \notin \mathsf{Sub}_2(v)$ we have* $\mathsf{int}_{y,x}(v) \subseteq \mathsf{int}_{y,x}(u)$.

Proof. Assume $(u, v) \in \pi_1^-$. It is easy to see that $\mathsf{c}(u) = \mathsf{c}(v)$.

Let x and y be two different letters from $\mathsf{c}(v)$. Assume that the first occurrence of the letter x is before the first occurrence of the letter y in v. Let $g = (g_0, x, g_1)$ be a factorization of v such that the central x is the first occurrence of the letter x in v, i.e. $x \notin \mathsf{c}(g_0)$. Under our assumption also $y \notin \mathsf{c}(g_0)$. There is a factorization $f = (f_0, x, f_1)$ of u such that $\mathsf{c}(f_0) \subseteq \mathsf{c}(g_0)$ and $\mathsf{c}(f_1) \subseteq \mathsf{c}(g_1)$. We obtain $x \notin \mathsf{c}(f_0)$

and $y \notin \mathsf{c}(f_0)$. Hence the central x in f is the first occurrence of x in u and the first occurrence of y in u can not be before it. Hence $\mathsf{first}(u) = \mathsf{first}(v)$.

The dual arguments lead to $\mathsf{last}(u) = \mathsf{last}(v)$.

If $xx \notin \mathsf{Sub}_2(v)$ and $x \in \mathsf{c}(v)$ then $v = g_0 x g_1$, where $x \notin \mathsf{c}(g_0) \cup \mathsf{c}(g_1)$. For the factorization (g_0, x, g_1) of v there is a factorization (f_0, x, f_1) of u such that $\mathsf{c}(f_0) \subseteq \mathsf{c}(g_0)$ and $\mathsf{c}(f_1) \subseteq \mathsf{c}(g_1)$. Hence $xx \notin \mathsf{Sub}_2(u)$. In other words $xx \in \mathsf{Sub}_2(u)$ implies $xx \in \mathsf{Sub}_2(v)$.

Observe that for each word w and letters $x, y \in \mathsf{c}(w)$ such that $x \neq y$ we have $xy \in \mathsf{Sub}_2(w)$ if and only if the first occurrence of x in w is before the last occurrence of y in w. Assume, for a moment, that $xy \in \mathsf{Sub}_2(u)$ but $xy \notin \mathsf{Sub}_2(v)$, i.e. the last occurrence of y in v is before the first occurrence of x. Let $g = (g_0, y, g_1)$ be a factorization of v where the central y is the last occurrence of y in v, i.e. $y \notin \mathsf{c}(g_1)$. Under our assumptions also $x \notin \mathsf{c}(g_0)$. There is a factorization $f = (f_0, y, f_1)$ of u such that $\mathsf{c}(f_0) \subseteq \mathsf{c}(g_0)$ and $\mathsf{c}(f_1) \subseteq \mathsf{c}(g_1)$. Hence $x \notin \mathsf{c}(f_0)$ and $y \notin \mathsf{c}(f_1)$. The second condition is saying that the central y in the factorization f is the last occurrence of y in u and hence x does not occur before this y. Consequently, the last occurrence of y in u is before the first occurrence of x in u, which is a contradiction. Hence we proved the second condition.

Now let $x, y \in \mathsf{c}(u)$ be such that $x \neq y$ and $xy \notin \mathsf{Sub}_2(v)$. Then $xy \notin \mathsf{Sub}_2(u)$ is also true. We can write $v = v_0 y v_1 x v_2$ where $y \notin \mathsf{c}(v_1 x v_2)$ and $x \notin \mathsf{c}(v_0 y v_1)$ and also $u = u_0 y u_1 x u_2$ where $y \notin \mathsf{c}(u_1 x u_2)$ and $x \notin \mathsf{c}(u_0 y u_1)$. Let $z \in \mathsf{c}(v_1) = \mathsf{int}_{y,x}(v)$. Then we can write $v_1 = v_1' z v_1''$ and we have a factorization $g = (v_0 y v_1', z, v_1'' x v_2)$ of the word v. Then there is a factorization $f = (f_0, z, f_1)$ of the word u such that $\mathsf{c}(f_0) \subseteq \mathsf{c}(v_0 y v_1') \subseteq \mathsf{c}(v_0 y v_1)$ and $\mathsf{c}(f_1) \subseteq \mathsf{c}(v_1'' x v_2) \subseteq \mathsf{c}(v_1 x v_2)$. Hence $x \notin \mathsf{c}(f_0)$ and $y \notin \mathsf{c}(f_1)$. This means that the central z in f is before the first occurrence of the letter x in u and after the last occurrence of the letter y in u. Thus $z \in \mathsf{c}(u_1)$ and we proved that $\mathsf{int}_{y,x}(v) = \mathsf{c}(v_1) \subseteq \mathsf{c}(u_1) = \mathsf{int}_{y,x}(u)$. The proof of the direct implication is complete.

Conversely, assume that u and v satisfy all three conditions. Note that from the first one we have $\mathsf{c}(u) = \mathsf{c}(v)$. Let $g = (g_0, z, g_1)$ be an arbitrary factorization of v. We distinguish several cases.

Assume that the central z in g is the first occurrence and the last occurrence of this letter in v at the same time. Then it is the unique occurrence of z in v and we have $zz \notin \mathsf{Sub}_2(v)$. Hence $zz \notin \mathsf{Sub}_2(u)$ and there is a unique occurrence of z in u, so we can write $u = u_0 z u_1$ and $z \notin \mathsf{c}(u_0) \cup \mathsf{c}(u_1)$. From the first condition we have $\mathsf{c}(u_0) = \mathsf{c}(g_0)$ and $\mathsf{c}(u_1) = \mathsf{c}(g_1)$.

Assume that the central z in g is the first occurrence of z in v but it is not the last occurrence of z in v. Then $z \notin \mathsf{c}(g_0)$ and $z \in \mathsf{c}(g_1)$. Consider $u = u_0 z u_1$ where the central z is the first occurrence of z in u, i.e. $z \notin \mathsf{c}(u_0)$. From the first condition we have $\mathsf{c}(u_0) = \mathsf{c}(g_0)$ and we would like to show that $\mathsf{c}(u_1) \subseteq \mathsf{c}(g_1)$. So, let $y \in \mathsf{c}(u_1)$. Then $zy \in \mathsf{Sub}_2(u) \subseteq \mathsf{Sub}_2(v)$ and $y \in \mathsf{c}(g_1)$ follows. If the central z in g is the last occurrence of z in v we can use the dual arguments.

Finally, assume that the central z in g is not the first occurrence nor the last occurrence of z in v. If there is no letter with the last occurrence in v before our occurrence of z, then $\mathsf{c}(g_1) = \mathsf{c}(v) = \mathsf{c}(u)$ and we can easily find an appropriate factorization f of u (namely, by the first condition we choose the first occurrence of z in u as a central letter in f) for which the inequality $f \leq_\sigma g$ is true. Dually, in the case when no first occurrence of a letter is after our occurrence of z. Look now at the last occurrence of a letter y before z in v such that there is no last occurrence of some letter between these occurrences of y and z. In the same way we look at the first occurrence of a letter x in v which is after z and there is no first occurrence of some letter between them. Thus $z \in \mathsf{int}_{y,x}(v) \subseteq \mathsf{int}_{y,x}(u)$ by the third condition and we can find the occurrence of z in u between the last occurrence of y and the first occurrence of x, i.e $u = u_1yu_2zu_3xu_4$ where $x \notin \mathsf{c}(u_1yu_2zu_3)$ and $y \notin \mathsf{c}(u_2zu_3xu_4)$. We claim that $f = (u_1yu_2, z, u_3xu_4) \leq_\sigma g = (g_0, z, g_1)$. Indeed, if some letter a occurs in u_1yu_2 then the first occurrence of this letter a is before the first occurrence of x in u and by the first condition the first occurrence of a in v is before the first occurrence of x. By the choice of x, the first occurrence of a in v is before our z in v, i.e. $a \in \mathsf{c}(g_0)$. We proved that $\mathsf{c}(u_1yu_2) \subseteq \mathsf{c}(g_0)$ and one can prove $\mathsf{c}(u_3xu_4) \subseteq \mathsf{c}(g_1)$ in the same manner, i.e. $f \leq_\sigma g$.

In all cases we found such a factorization f of the word u, hence $(u, v) \in \pi_1^-$. □

From the previous lemma we immediately obtain an analogous characterization for the relation $(\pi_1^-)^\frown$.

Lemma 3. *Let $u, v \in X^*$ be arbitrary words. Then $(u, v) \in (\pi_1^-)^\frown$ if and only if the following conditions are satisfied*

- $\mathsf{first}(u) = \mathsf{first}(v)$ *and* $\mathsf{last}(u) = \mathsf{last}(v)$*;*
- $\mathsf{Sub}_2(u) = \mathsf{Sub}_2(v)$*;*
- *for every $x, y \in \mathsf{c}(u)$ such that $x \neq y$ and $xy \notin \mathsf{Sub}_2(v)$, we have* $\mathsf{int}_{y,x}(u) = \mathsf{int}_{y,x}(v)$.

For $w \in X^*$ we define the *skeleton* $\mathsf{skel}(w) \in X^*$ as follows. We remove from w every occurrence of a given letter which is not the first or the last occurrence of this letter in the word w. After deleting of all "interior" occurrences of all letters from w, the resulting word is the skeleton $\mathsf{skel}(w)$ of w. In other words, there is a unique factorization $(\lambda, b_1, w_1, b_2, w_2, \ldots, b_{\ell-1}, w_{\ell-1}, b_\ell, \lambda)$ of the word w, such that

- for every $i = 1, \ldots, \ell - 1$ we have $\mathsf{c}(w_i) \subseteq \{b_1, \ldots, b_i\} \cap \{b_{i+1}, \ldots, b_\ell\}$;
- for every $i = 2, \ldots, \ell - 1$ we have $b_i \notin \{b_1, \ldots, b_{i-1}\}$ or $b_i \notin \{b_{i+1}, \ldots, b_\ell\}$.

We call $b_1w_1b_2w_2 \ldots b_{\ell-1}w_{\ell-1}b_\ell = w$ a *skeleton decomposition* of w and we put $\mathsf{skel}(w) = b_1b_2 \ldots b_{\ell-1}b_\ell$.

Lemma 4. *(i) Let $w \in X^*$. Then* $\mathsf{first}(w) = \mathsf{first}(\mathsf{skel}(w))$, $\mathsf{last}(w) = \mathsf{last}(\mathsf{skel}(w))$ *and* $\mathsf{Sub}_2(w) = \mathsf{Sub}_2(\mathsf{skel}(w))$.

(ii) Let $u, v \in X^*$ satisfy the first two conditions from Lemma 3. Then $\mathsf{skel}(u) = \mathsf{skel}(v)$*. In particular, this is true for* $(u, v) \in (\pi_1^-)^\frown$*.*

Proof. "(i)" : Let $w \in X^*$. By the definition of $\mathsf{skel}(w)$ we have $\mathsf{first}(w) = \mathsf{first}(\mathsf{skel}(w))$, $\mathsf{last}(w) = \mathsf{last}(\mathsf{skel}(w))$ and $\mathsf{Sub}_2(\mathsf{skel}(w)) \subseteq \mathsf{Sub}_2(w)$. On the other hand, if $x, y \in X$ are such that $xy \in \mathsf{Sub}_2(w)$, then one can consider the first occurrence of x and the last occurrence of y in w and see that $xy \in \mathsf{Sub}_2(\mathsf{skel}(w))$. The equality $\mathsf{Sub}_2(\mathsf{skel}(w)) = \mathsf{Sub}_2(w)$ follows.

"(ii)" : For an arbitrary $w \in X^*$ each letter occurs at most twice in $\mathsf{skel}(w)$ and it occurs exactly once if and only if it occurs exactly once in w. Hence the length of the word $\mathsf{skel}(w)$ is equal to the size of the set $\mathsf{c}(w)$ plus the number of letters $x \in X$ such that $xx \in \mathsf{Sub}_2(w)$.

Consider words $u, v \in X^*$ such that $\mathsf{first}(u) = \mathsf{first}(v)$, $\mathsf{last}(u) = \mathsf{last}(v)$, and $\mathsf{Sub}_2(u) = \mathsf{Sub}_2(v)$. Then $\mathsf{skel}(u)$ and $\mathsf{skel}(v)$ have the same length. Let $b_1u_1b_2u_2 \dots b_{\ell-1}u_{\ell-1}b_\ell = u$ and $c_1v_1c_2v_2 \dots c_{\ell-1}v_{\ell-1}c_\ell = v$ be the skeleton decompositions of u and v. Then $\mathsf{skel}(u) = b_1b_2 \dots b_{\ell-1}b_\ell$ and $\mathsf{skel}(v) = c_1c_2 \dots c_{\ell-1}c_\ell$. By (i) we have that $\mathsf{first}(\mathsf{skel}(u)) = \mathsf{first}(\mathsf{skel}(v))$, $\mathsf{last}(\mathsf{skel}(u)) = \mathsf{last}(\mathsf{skel}(v))$ and $\mathsf{Sub}_2(\mathsf{skel}(u)) = \mathsf{Sub}_2(\mathsf{skel}(v))$. Now, assume that $\mathsf{skel}(u) \neq \mathsf{skel}(v)$. Let i be the smallest index such that $b_i \neq c_i$.

If $b_i, c_i \notin \{b_1, \dots, b_{i-1}\}$ then the first occurrence of b_i in $\mathsf{skel}(u)$ is before the first occurrence of c_i in $\mathsf{skel}(u)$ and this is not true for $\mathsf{skel}(v)$. It is a contradiction with $\mathsf{first}(\mathsf{skel}(u)) = \mathsf{first}(\mathsf{skel}(v))$.

If $b_i \in \{b_1, \dots, b_{i-1}\}$ and $c_i \notin \{b_1, \dots, b_{i-1}\} = \{c_1, \dots, c_{i-1}\}$ then $c_ib_i \notin \mathsf{Sub}_2(\mathsf{skel}(u))$ and $b_ib_i \in \mathsf{Sub}_2(\mathsf{skel}(u)) = \mathsf{Sub}_2(\mathsf{skel}(v))$. Hence $b_i = c_j$ for some $j > i$ and $c_ib_i \in \mathsf{Sub}_2(\mathsf{skel}(v)) = \mathsf{Sub}_2(\mathsf{skel}(u))$ which is a contradiction. We can use a similar argument in the case $b_i \notin \{b_1, \dots, b_{i-1}\}$ and $c_i \in \{b_1, \dots, b_{i-1}\}$.

Finally, if $b_i, c_i \in \{b_1, \dots, b_{i-1}\}$ then $b_i \in \{c_{i+1}, \dots, c_\ell\} \setminus \{b_{i+1}, \dots, b_\ell\}$ and $c_i \in \{b_{i+1}, \dots, b_\ell\} \setminus \{c_{i+1}, \dots, c_\ell\}$. Hence the last occurrence of b_i is before the last occurrence of c_i in $\mathsf{skel}(u)$ and this is not true for $\mathsf{skel}(v)$. It is a contradiction with $\mathsf{last}(\mathsf{skel}(u)) = \mathsf{last}(\mathsf{skel}(v))$.

We have obtained a contradiction in all cases, so $\mathsf{skel}(u) = \mathsf{skel}(v)$ holds. □

We say that a word w is a *canonical word* if its skeleton decomposition $w = b_1w_1b_2w_2 \dots b_{\ell-1}w_{\ell-1}b_\ell$ satisfies the following condition:

(C1) for every $i = 1, \dots, \ell - 1$ and positive integers j, j', if $x_jx_{j'}$ is a subword of w_i then $j < j'$.

If the following two conditions are also satisfied we speak about a *balanced canonical word.*

(C2) If $i \in \{1, \dots, \ell - 1\}$ is such that $b_i \in \{b_{i+1}, \dots, b_\ell\}$, then $b_i \in \mathsf{c}(w_i)$ and $\mathsf{c}(w_{i-1}) \subseteq \mathsf{c}(w_i)$ for $i \geq 2$;
(C3) If $i \in \{2, \dots, \ell\}$ is such that $b_i \in \{b_1, \dots b_{i-1}\}$, then $b_i \in \mathsf{c}(w_{i-1})$ and $\mathsf{c}(w_i) \subseteq \mathsf{c}(w_{i-1})$ for $i \leq \ell - 1$.

The role of this notion will be clear from the following lemma which completes the proof of the first statement of the proposition.

Lemma 5. *(i) Let u, v be balanced canonical words and $(u, v) \in (\pi_1^-)^\frown$. Then $u = v$.*

(ii) Let u be an arbitrary word. Then there exists a balanced canonical word w such that $(u, w) \in (\pi_1^-)^\frown$ and the identity $u = w$ is a consequence of the identities (1a-d) *and* (2a,b)*.*

Proof. "(i)" : Let u, v be balanced canonical words such that $(u, v) \in (\pi_1^-)^\frown$. Then $\mathsf{skel}(u) = \mathsf{skel}(v)$ by Lemma 4. Let $u = b_1u_1b_2u_2 \dots b_{\ell-1}u_{\ell-1}b_\ell$ and $v = b_1v_1b_2v_2 \dots b_{\ell-1}v_{\ell-1}b_\ell$ be the skeleton decompositions of u and v.

Since u and v satisfy (C1), the fact $u_i = v_i$ is equivalent to $\mathsf{c}(u_i) = \mathsf{c}(v_i)$. Therefore it is enough to prove that $\mathsf{c}(u_i) = \mathsf{c}(v_i)$ for every $i = 1, \dots, \ell - 1$. To prove this, we consider an arbitrary such i.

If $\{b_1, b_2, \dots, b_i\} \subseteq \{b_{i+1}, \dots, b_\ell\}$ then the assumption $b_j \in \{b_{j+1}, \dots, b_\ell\}$ in (C2) is valid for every $j \in \{1, \dots, i\}$. Hence we have $b_1 \in \mathsf{c}(u_1)$, $b_2 \in \mathsf{c}(u_2), \dots, b_i \in \mathsf{c}(u_i)$ and $\mathsf{c}(u_1) \subseteq \mathsf{c}(u_2) \subseteq \dots \subseteq \mathsf{c}(u_i)$. Thus we have $\{b_1, \dots, b_i\} \subseteq \mathsf{c}(u_i)$ and the opposite inclusion is given by the definition of the skeleton decomposition of a word. The same is true for v, so we obtain $\mathsf{c}(u_i) = \{b_1, \dots, b_i\} = \mathsf{c}(v_i)$.

The dual argument gives the equality $\mathsf{c}(u_i) = \mathsf{c}(v_i)$ if $\{b_{i+1}, \dots, b_\ell\} \subseteq \{b_1, b_2, \dots, b_i\}$. Therefore we can assume there are indices p and q such that $1 \leq p \leq i < q \leq \ell$ and $b_p \notin \{b_{i+1}, \dots, b_\ell\}$, $b_q \notin \{b_1, b_2, \dots, b_i\}$. Moreover, we can consider the largest p and the smallest q satisfying these conditions. In other words, we assume also that for every p' such that $p < p' \leq i$ we have $b_{p'} \in \{b_{i+1}, \dots, b_\ell\}$ and for every q' such that $i < q' < q$ we have $b_{q'} \in \{b_1, b_2, \dots, b_i\}$. These additional assumptions imply that $b_{p'} \in \{b_{p'+1}, \dots, b_i, b_{i+1}, \dots, b_\ell\} = \{b_{i+1}, \dots, b_\ell\}$ for every $p' = p+1, \dots, i$. By (C2) we have $b_{p+1} \in \mathsf{c}(u_{p+1})$, $\dots$, $b_i \in \mathsf{c}(u_i)$ and $\mathsf{c}(u_p) \subseteq \mathsf{c}(u_{p+1}) \subseteq \dots \subseteq \mathsf{c}(u_i)$. Similarly from right $b_{q-1} \in \mathsf{c}(u_{q-2}), \dots, b_{i+1} \in \mathsf{c}(u_i)$ and $\mathsf{c}(u_{q-1}) \subseteq \dots \subseteq \mathsf{c}(u_{i+1}) \subseteq \mathsf{c}(u_i)$. Altogether we observe that

$$\mathsf{int}_{b_p,b_q}(u) = \mathsf{c}(u_pb_{p+1}u_{p+1} \dots u_{i-1}b_iu_ib_{i+1}u_{i+1} \dots u_{q-2}b_{q-1}u_{q-1}) = \mathsf{c}(u_i)\ .$$

The same equality holds for v and we obtain $\mathsf{c}(u_i) = \mathsf{c}(v_i)$, by Lemma 3.

"(ii)" : Let u be an arbitrary word and consider its skeleton decomposition $u = b_1u_1b_2u_2 \dots b_{\ell-1}u_{\ell-1}b_\ell$.

If $b_i \in \{b_{i+1}, \dots, b_\ell\}$, then we can use the identities (1a-d) to add to u_i every letter $x \in \mathsf{c}(u_{i-1}b_i) \setminus \mathsf{c}(u_i)$, one by one (to reach our goal it does not matter the position in the word u_i, or in the word obtained from u_i after some of these steps, we add a letter of $\mathsf{c}(u_{i-1}) \setminus \mathsf{c}(u_i)$, since by the identities (2a,b) we can commute any letters in the word obtained). In this way we transform all u_i's to words which satisfy condition (C2). In the same manner we can transform the final word obtained in this way, which satisfies (C2), to a word satisfying (C3) too. Let v_i be the word that replaced u_i after these steps.

Now we can use the identities (2a,b) to reorder the letters inside v_i's. Finally the identity (1e) can be used on each word v_i to obtain a word w_i that has no more than one occurrence of each letter. This means that we can transform in that way the considered word to a word satisfying condition (C1).

Therefore the identities (1a-d) and (2a,b) can be used to obtain a balanced canonical word which is $(\pi_1^-)^\frown$-related to the given u. $\square$

The first part of Proposition 1 is proved and its second part follows from the following lemma.

Lemma 6. *Let $u, v \in X^*$ be such that $(u, v) \in \pi_1^-$. Then the identity $u \leq v$ is a consequence of the identities* (2a,b)*,* (3) *and* (4)*.*

Proof. The proof will be divided into two parts.

Claim 1. Let $u, v \in X^*$ be such that $(u, v) \in \pi_1^-$. Then there exists a word w such that the identity $u \leq w$ is a consequence of the identities (3) satisfying $(w, v) \in \pi_1^-$ and $\mathsf{Sub}_2(w) = \mathsf{Sub}_2(v)$.

Proof of Claim 1. Recall that $(u, v) \in \pi_1^-$ implies $\mathsf{Sub}_2(u) \subseteq \mathsf{Sub}_2(v)$ by Lemma 2. We prove the claim by an induction with respect to the size of the set $M = \mathsf{Sub}_2(v) \setminus \mathsf{Sub}_2(u)$. If M is the empty set then we can simply put $w = u$.

Let M be of size m and assume that the claim holds for every pair of words (u', v') of $\in \pi_1^-$ which satisfies $|\mathsf{Sub}_2(v') \setminus \mathsf{Sub}_2(u')| < m$.

If $x^2 \in M$ for some $x \in X$, then $u = u_0 x u_1$ where $x \notin \mathsf{c}(u_0) \cup \mathsf{c}(u_1)$. Now we can apply the identity $x \leq x^2$ (consequence of the identities (3)) to obtain $u \leq u_0 x x u_1$. Using Lemma 2 we see that $(u, v) \in \pi_1^-$ implies $(u_0 x x u_1, v) \in \pi_1^-$. We put $w = u_0 x x u_1$ and we have $\mathsf{Sub}_2(w) = \mathsf{Sub}_2(u) \cup \{x^2\}$. Now we can use the induction assumption for the pair (w, v) for which we have $\mathsf{Sub}_2(v) \setminus \mathsf{Sub}_2(w) = M \setminus \{x^2\}$. Thus in the case $x^2 \in M$ we are done, so assume that no x^2 belongs to M.

Now from all pairs $x \neq y$ satisfying $xy \in M$ we choose such one that in the corresponding factorization $u = u_0 y u_1 x u_2$ where $y \notin \mathsf{c}(u_1 x u_2)$ and $x \notin \mathsf{c}(u_0 y u_1)$ the word u_1 is the shortest possible. Since $xy \in M$ we have $v = v_0 x v_1 y v_2$ where $x \notin \mathsf{c}(v_0)$ and $y \notin \mathsf{c}(v_2)$. We know that $\mathsf{first}(v) = \mathsf{first}(u)$, so $y \in \mathsf{c}(v_0)$. In the same manner $x \in \mathsf{c}(v_2)$ follows from $\mathsf{last}(v) = \mathsf{last}(u)$. Hence $x^2, y^2 \in \mathsf{Sub}_2(v)$ and consequently $x^2, y^2 \in \mathsf{Sub}_2(u)$; in particular $x \in \mathsf{c}(u_2)$ and $y \in \mathsf{c}(u_0)$. Now if u_1 contains the first occurrence of a letter z in u then the first occurrence of z is before the first occurrence of x in both u and v and we see that $zy \in M$. This is a contradiction with our choice of the pair x, y. In the same way we can prove that u_1 does not contain the last occurrence of some letter. This means that $\mathsf{c}(u_1) \subseteq \mathsf{c}(u_0)$ and $\mathsf{c}(u_1) \subseteq \mathsf{c}(u_2)$. Let $u_1 = y_1 \dots y_j$ where $y_1, \dots, y_j \in X$. We use one of the identities (3), namely $yzyx \leq yzxyx$, to introduce in u a new occurrence of x immediately before the last letter of u_1. So we have $u = u_0 y y_1 \dots y_{j-1} y_j x u_2 \leq u_0 y y_1 \dots y_{j-1} \cdot x \cdot y_j x u_2$. We use the same identity to introduce in $u_0 y y_1 \dots y_{j-1} x y_j x u_2$ new occurrences of x before all $y_{j-1}, \dots, y_1$ and also before y, step by step. Hence $u = u_0yy_1 \dots y_{j-1}y_jxu_2 \le u_0xyxy_1x \dots xy_{j-1}xy_jxu_2$ is a consequence of the identities (3). Let w be the word on the right hand side. It is not hard to see that $(w, v) \in \pi_1^-$ because $\mathsf{Sub}_2(w) = \mathsf{Sub}_2(u) \cup \{xy\} \subseteq \mathsf{Sub}_2(v)$ and the other invariants from Lemma 2 remain the same. Now we can use the induction assumption for the pair (w, v) and we proved the first claim.

Claim 2. Let $w, v \in X^*$ be such that $(w, v) \in \pi_1^-$ and $\mathsf{Sub}_2(w) = \mathsf{Sub}_2(v)$. Then the identity $w \le v$ is a consequence of the identities (2a,b), (3) and (4).

Proof of Claim 2. By Lemma 5 we can assume that w and v are balanced canonical words. Recall that the identities (1) are consequence of the identities (3) and (4). By Lemmas 2 and 4 we have $\mathsf{skel}(w) = \mathsf{skel}(v)$. Let $w = b_1w_1b_2w_2 \dots b_{\ell-1}w_{\ell-1}b_\ell$ and $v = b_1v_1b_2v_2 \dots b_{\ell-1}v_{\ell-1}b_\ell$ be the skeleton decompositions of words w and v with the same skeleton $b_1 \dots b_\ell$. We have showed some basic properties of words w_i in the skeleton decompositions of canonical balanced canonical words in the proof of item (i) of Lemma 5:

1. If $\{b_1, \dots, b_i\} \subseteq \{b_{i+1}, \dots, b_\ell\}$ then $\mathsf{c}(w_i) = \{b_1, \dots, b_i\}$ and we have $\mathsf{c}(w_i) = \mathsf{c}(v_i)$.

2. Dually in the case $\{b_{i+1}, \dots, b_\ell\} \subseteq \{b_1, \dots, b_i\}$.

3. In the remaining cases there are indices $p < i < q$ such that $b_p \notin \{b_{i+1}, \dots, b_\ell\}$, $b_q \notin \{b_1, \dots, b_i\}$ and for every p' such that $p < p' \le i$ we have $b_{p'} \in \{b_{i+1}, \dots, b_\ell\}$ and for every q' such that $i < q' < q$ we have $b_{q'} \in \{b_1, \dots, b_i\}$. And then $\mathsf{c}(w_i) = \mathsf{int}_{b_p,b_q}(w)$ and $\mathsf{c}(v_i) = \mathsf{int}_{b_p,b_q}(v)$. Hence we can deduce $\mathsf{c}(v_i) \subseteq \mathsf{c}(w_i)$ by Lemma 2.

By condition (C1), v_i is a subword of w_i, which means that either $v_i = w_i$ or v_i can be obtained from w_i by deleting some occurrences of some letters in w_i. Thus for any $s, t \in X^*$ such that $\mathsf{c}(w_i) \subseteq \mathsf{c}(s) \cap \mathsf{c}(t)$, the identity $sw_it \le sv_it$ is a consequence of the identity (4). Then we deduce from (4) the identities $w = b_1w_1b_2w_2b_3 \dots w_{\ell-1}b_\ell \le b_1v_1b_2w_2b_3 \dots w_{\ell-1}b_\ell \le b_1v_1b_2v_2b_3w_3b_4 \dots w_{\ell-1}b_\ell \le \dots \le b_1v_1b_2v_2b_3 \dots v_{\ell-1}b_\ell = v$, and hence also the identity $w \le v$. □

We finished the proof of Proposition 1.

4.2. *Identities for pseudovarieties corresponding to* $\mathsf{BPol}_1(\mathcal{J}_1)$ *and* $\mathsf{PPol}_1(\mathcal{J}_1)$

The proofs in this part are easier because we can use numerous observations from the previous subsection.

We will use the identities (2a,b) again and we introduce a new one

$$xyxxzx = xyxzx\ . \tag{5}$$

It is clear that these three identities are satisfied in the pseudovariety corresponding to $\mathsf{BPol}_1(\mathcal{J}_1)$ and consequently in that for $\mathsf{PPol}_1(\mathcal{J}_1)$. Also it is easy to see that the identity (4) is satisfied in the pseudovariety corresponding to $\mathsf{PPol}_1(\mathcal{J}_1)$.

Proposition 7. *(i) The identities* (2a,b) *and* (5) *form a finite basis of identities for the variety of monoids corresponding to* $\mathsf{BPol}_1(\mathcal{J}_1)$.

(ii) The identities (2a,b), (4), *and* (5) *form a finite basis of identities for the variety of ordered monoids corresponding to* $\mathsf{PPol}_1(\mathcal{J}_1)$.

Proof. We modify Lemma 5 for the relation $\widehat{\pi_1}$ as follows.

Lemma 8. *(i) Let u, v be canonical words such that $(u,v) \in \widehat{\pi_1}$. Then $u = v$.*

(ii) Let u be an arbitrary word. Then there exists a canonical word w such that $(u,w) \in \widehat{\pi_1}$ and the identity $u = w$ is a consequence of the identities (2a,b) *and* (5).

Proof. "(i)" : Since $\widehat{\pi_1} \subseteq (\pi_1^- \widehat{)}$, we have $\mathsf{skel}(u) = \mathsf{skel}(v)$ by Lemma 4 (ii). So, let $u = b_1 u_1 b_2 u_2 \dots b_{\ell-1} u_{\ell-1} b_\ell$ and $v = b_1 v_1 b_2 v_2 \dots b_{\ell-1} v_{\ell-1} b_\ell$ be the skeleton decompositions which satisfy condition (C1). It is enough to prove that $\mathsf{c}(u_i) = \mathsf{c}(v_i)$ for every $i = 1, \dots, \ell-1$. Let i be an arbitrary index and let $x \in \mathsf{c}(v_i)$.

Consider the factorization $g = (g_0, x, g_1)$ of v corresponding to the occurrence of x in v_i. Then we know that $x \in \mathsf{c}(g_0) = \{b_1, \dots, b_i\}$ and $x \in \mathsf{c}(g_1) = \{b_{i+1}, \dots, b_\ell\}$. There is a factorization $f = (f_0, x, f_1)$ of u such that $\mathsf{c}(f_0) = \mathsf{c}(g_0)$ and $\mathsf{c}(f_1) = \mathsf{c}(g_1)$. We claim that for the prefix f_0 of u and the suffix f_1 of u we have $|f_0| \geq |b_1 u_1 b_2 u_2 \dots b_{i-1} u_{i-1} b_i|$ and $|f_1| \geq |b_{i+1} u_{i+1} \dots b_\ell|$. From this claim the statement $x \in \mathsf{c}(u_i)$ follows.

To prove the claim we first assume that $b_i \notin \{b_1, \dots, b_{i-1}\}$. Since f_0 is a prefix of $u = b_1 u_1 b_2 u_2 \dots b_{\ell-1} u_{\ell-1} b_\ell$ such that $\mathsf{c}(f_0) = \mathsf{c}(g_0) = \{b_1, \dots, b_i\}$ we see that $|f_0| \geq |b_1 u_1 b_2 u_2 \dots b_{i-1} u_{i-1} b_i|$. Now assume that $b_i \in \{b_1, \dots, b_{i-1}\}$. Then we have $b_i \notin \{b_{i+1} \dots, b_\ell\} = \mathsf{c}(f_1)$. Hence the suffix f_1 of u is a suffix of $u_i b_{i+1} \dots u_{\ell-1} b_\ell$, in particular $|f_1| \leq |u_i b_{i+1} \dots u_{\ell-1} b_\ell|$. Furthermore $b_i \notin \mathsf{c}(f_1)$ and $x \in \mathsf{c}(g_1) = \mathsf{c}(f_1)$ implies $x \neq b_i$. We can conclude that $|f_1| < |u_i b_{i+1} \dots u_{\ell-1} b_\ell|$, and hence $|f_0| \geq |b_1 u_1 b_2 u_2 \dots b_{i-1} u_{i-1} b_i|$ holds also in this second case. If we consider b_{i+1} we can prove the second half of the claim and we can conclude that $x \in \mathsf{c}(u_i)$.

Analogously we have $\mathsf{c}(u_i) \subseteq \mathsf{c}(v_i)$ for any i.

"(ii)" : Let $u = b_1 u_1 b_2 u_2 \dots b_{\ell-1} u_{\ell-1} b_\ell$ be the skeleton decomposition of u. We use the identities (2a,b) to commute letters inside every u_i and we use the identity (5) to remove redundant occurrences of letters. Thus we can construct w with the required properties. $\square$

Lemma 9. *Let $u, v \in X^*$ be such that $(u,v) \in \pi_1$. Then $\mathsf{skel}(u) = \mathsf{skel}(v)$.*

Proof. If $(u,v) \in \pi_1$ then $(u,v) \in \pi_1^-$ and we have $\mathsf{first}(u) = \mathsf{first}(v)$ and $\mathsf{last}(u) = \mathsf{last}(v)$ by Lemma 2. We claim that $\mathsf{Sub}_2(u) = \mathsf{Sub}_2(v)$. Indeed, the inclusion $\mathsf{Sub}_2(u) \subseteq \mathsf{Sub}_2(v)$ also follows from Lemma 2. Let $xy \in \mathsf{Sub}_2(v)$. If we consider a factorization $g = (g_0, x, g_1)$ of the word v such that $x \notin \mathsf{c}(g_0)$ then $y \in \mathsf{c}(g_1)$. Therefore there is a factorization $f = (f_0, x, f_1)$ of u such that $\mathsf{c}(f_0) = \mathsf{c}(g_0)$ and $\mathsf{c}(f_1) = \mathsf{c}(g_1)$. So the central x in f is the first occurrence of x in u and $y \in \mathsf{c}(f_1)$ and we can conclude that $xy \in \mathsf{Sub}_2(u)$. We proved the claim and the statement trivially follows. $\square$

Lemma 10. *Let u, v be words such that $(u, v) \in \pi_1$. Then the identity $u \leq v$ is a consequence of the identities* (2a,b)*,* (4) *and* (5)*.*

Proof. By Lemma 8 we can assume that both u, v are canonical words. We have $\mathsf{skel}(u) = \mathsf{skel}(v)$ by Lemma 9. So, let $u = b_1 u_1 b_2 u_2 \dots b_{\ell-1} u_{\ell-1} b_\ell$ and $v = b_1 v_1 b_2 v_2 \dots b_{\ell-1} v_{\ell-1} b_\ell$ be the skeleton decompositions. In the proof of the part (i) of Lemma 8 we saw that $\mathsf{c}(v_i) \subseteq \mathsf{c}(u_i)$ for every $i = 1, \dots, \ell - 1$. Then v can be obtained from u by deleting in each u_i every letter of $\mathsf{c}(u_i) \setminus \mathsf{c}(v_i)$, hence the identity $u \leq v$ is a consequence of the identity (4). $\square$

We finished the proof of the proposition.

5. Inclusions between our subhierarchies

Our primary goal is to compare our classes of languages. We are doing that by the usage of the corresponding finite characteristics and our methods use rather combinatorics on words.

Recall that, for any positive variety $\mathcal{V}$ of languages, we have $\mathsf{PPol}_k(\mathcal{V}) \subseteq \mathsf{PPol}_\ell(\mathcal{V})$ and $\mathsf{BPol}_k(\mathcal{V}) \subseteq \mathsf{BPol}_\ell(\mathcal{V})$ for $k < \ell$. Next we show that in the case of varieties $\mathcal{J}_1^-$ and $\mathcal{J}_1$ these inclusions are strict. The positions of our varieties for $k = 0$ is clear – see Figure 1. So let $k \geq 1$ in the rest of this section.

Proposition 11. *The hierarchies* $\mathsf{PPol}_k(\mathcal{J}_1^-)$*,* $\mathsf{PPol}_k(\mathcal{J}_1)$*,* $\mathsf{BPol}_k(\mathcal{J}_1^-)$ *and* $\mathsf{BPol}_k(\mathcal{J}_1)$ *are strict, that is, for $k \neq \ell$, we have* $\mathsf{PPol}_k(\mathcal{J}_1^-) \neq \mathsf{PPol}_\ell(\mathcal{J}_1^-)$ *etc.*

Proof. By Results 2, 3 and 6 it is enough to show $\pi_{k+1}^- \subsetneqq \pi_k^-$ for every nonnegative integer k and similarly for the other relations. The inclusion $\pi_{k+1}^- \subseteq \pi_k^-$ follows directly from the definition (and similarly for the other relations). Let $x \in X$ and let k be a nonnegative integer. Then

$$(x^{k+2}, x^{k+1}) \in \pi_k^- \setminus \pi_{k+1}^-, \quad (x^{2k+1}, x^{2k+2}) \in \pi_k \setminus \pi_{k+1},$$

$$(x^{k+2}, x^{k+1}) \in (\pi_k^-)^\frown \setminus (\pi_{k+1}^-)^\frown, \quad (x^{2k+1}, x^{2k+2}) \in \widehat{\pi_k} \setminus \widehat{\pi_{k+1}}\ . \qquad \square$$

Thus we have two chains of positive varieties of languages $\mathsf{PPol}_k(\mathcal{J}_1^-)$ and $\mathsf{PPol}_k(\mathcal{J}_1)$ and two chains of Boolean varieties of languages $\mathsf{BPol}_k(\mathcal{J}_1^-)$ and $\mathsf{BPol}_k(\mathcal{J}_1)$. The unions $\mathsf{PPol}(\mathcal{J}_1^-)$ and $\mathsf{PPol}(\mathcal{J}_1)$ of the families $(\mathsf{PPol}_k(\mathcal{J}_1^-))_k$ and $(\mathsf{PPol}_k(\mathcal{J}_1))_k$, respectively, coincide and they form the 3/2 level of the Straubing-Thérien hierarchy (consequence of Theorem 8.8 of [8]). Therefore, $\mathsf{BPol}(\mathcal{J}_1^-) = \mathsf{BPol}(\mathcal{J}_1)$ is the second level of this hierarchy. We show a direct argument for these equalities here.

Proposition 12. *It holds* $\mathsf{PPol}(\mathcal{J}_1^-) = \mathsf{PPol}(\mathcal{J}_1)$ *and* $\mathsf{BPol}(\mathcal{J}_1^-) = \mathsf{BPol}(\mathcal{J}_1)$*.*

Proof. Let A be an arbitrary finite alphabet. Since $\mathcal{J}_1^- \subseteq \mathcal{J}_1$ we have $\mathsf{PPol}(\mathcal{J}_1^-)(A) \subseteq \mathsf{PPol}(\mathcal{J}_1)(A)$. To get the opposite inclusion we have to show that an arbitrary language

$$L = \overline{B_0} a_1 \overline{B_1} a_2 \dots a_\ell \overline{B_\ell},\ a_1, \dots, a_\ell \in A,\ B_0, \dots, B_\ell \subseteq A$$

belongs to $\mathsf{PPol}(\mathcal{J}_1^-)(A)$. First observe that for a subset C of A consisting of the letters $c_1, c_2, \dots, c_m$ we can write

$$\overline{C} = \bigcup_{\sigma \in \Sigma} C^* c_{\sigma(1)} C^* c_{\sigma(2)} C^* \dots C^* c_{\sigma(m)} C^* \ , \tag{6}$$

where Σ is the set of all permutations of the set of indices $\{1, \dots, m\}$. When we replace each $\overline{B_i}$ in the expression giving L by the corresponding sum of languages using formula (6) we obtain that $L \in \mathsf{PPol}(\mathcal{J}_1^-)(A)$.

Hence $\mathsf{PPol}(\mathcal{J}_1^-) = \mathsf{PPol}(\mathcal{J}_1)$ and the equality $\mathsf{BPol}(\mathcal{J}_1^-) = \mathsf{BPol}(\mathcal{J}_1)$ follows. □

If we fix a number k, then we see

$$\mathsf{PPol}_k(\mathcal{J}_1^-) \subseteq \mathsf{PPol}_k(\mathcal{J}_1),\ \mathsf{PPol}_k(\mathcal{J}_1^-) \subseteq \mathsf{BPol}_k(\mathcal{J}_1^-)\ ,$$

$$\mathsf{PPol}_k(\mathcal{J}_1) \subseteq \mathsf{BPol}_k(\mathcal{J}_1),\ \mathsf{BPol}_k(\mathcal{J}_1^-) \subseteq \mathsf{BPol}_k(\mathcal{J}_1)\ .$$

Next we show that for a fixed $k \geq 1$, our varieties form (with respect to inclusion) a four-element lattice which is not a chain.

Proposition 13. *For each k, the varieties $\mathsf{PPol}_k(\mathcal{J}_1^-)$, $\mathsf{PPol}_k(\mathcal{J}_1)$, $\mathsf{BPol}_k(\mathcal{J}_1^-)$ and $\mathsf{BPol}_k(\mathcal{J}_1)$ are pairwise different. Moreover, $\mathsf{PPol}_k(\mathcal{J}_1)$ and $\mathsf{BPol}_k(\mathcal{J}_1^-)$ are incomparable, i.e. $\mathsf{PPol}_k(\mathcal{J}_1) \nsubseteq \mathsf{BPol}_k(\mathcal{J}_1^-)$ and $\mathsf{BPol}_k(\mathcal{J}_1^-) \nsubseteq \mathsf{PPol}_k(\mathcal{J}_1)$.*

Proof. Let $x \in X$. Clearly,

$$(x^{k+2}, x^{k+1}) \in (\pi_k \cap (\pi_k^-)^\frown) \setminus \widehat{\pi_k}\ ,$$

$$(x^k, x^{k+1}) \in \pi_k^- \setminus (\pi_k \cup (\pi_k^-)^\frown)\ ,$$

$$(x^{k+1}, x^{k+2}) \in (\pi_k^-)^\frown \setminus \pi_k\ .$$

From the first formula we obtain both $\widehat{\pi_k} \subsetneqq \pi_k$ and $\widehat{\pi_k} \subsetneqq (\pi_k^-)^\frown$ from which $\mathsf{PPol}_k(\mathcal{J}_1) \subsetneqq \mathsf{BPol}_k(\mathcal{J}_1)$ and $\mathsf{BPol}_k(\mathcal{J}_1^-) \subsetneqq \mathsf{BPol}_k(\mathcal{J}_1)$ follow. The second formula yields both $\pi_k \subsetneqq \pi_k^-$ and $(\pi_k^-)^\frown \subsetneqq \pi_k^-$ from which we get $\mathsf{PPol}_k(\mathcal{J}_1^-) \subsetneqq \mathsf{PPol}_k(\mathcal{J}_1)$ and $\mathsf{PPol}_k(\mathcal{J}_1^-) \subsetneqq \mathsf{BPol}_k(\mathcal{J}_1^-)$. Now the third condition says that $(\pi_k^-)^\frown \nsubseteq \pi_k$, hence $\mathsf{PPol}_k(\mathcal{J}_1) \nsubseteq \mathsf{BPol}_k(\mathcal{J}_1^-)$.

To finish the proof we need to show $\mathsf{BPol}_k(\mathcal{J}_1^-) \nsubseteq \mathsf{PPol}_k(\mathcal{J}_1)$, i.e. $\pi_k \nsubseteq (\pi_k^-)^\frown$. Let $x, y \in X$ be a pair of different letters. Then

$$((xy)^{k+1} x (xy)^{k+1}, (xy)^{k+1}) \in \pi_k \setminus (\pi_k^-)^\frown.$$

□

One would expect that the members of our four hierarchies (for $k \geq 1$) form a lattice isomorphic to the product of that four-element lattice with the chain of positive integers. We show that the situation is a bit complicated. Such irregularities are possible just in the case $k \leq 1$. We show that all varieties in the first level are below almost all varieties from the other levels.

Proposition 14. *(i)* $\mathsf{BPol}_1(\mathcal{J}_1) \subseteq \mathsf{BPol}_2(\mathcal{J}_1^-)$.
(ii) $\mathsf{BPol}_1(\mathcal{J}_1) \subseteq \mathsf{PPol}_2(\mathcal{J}_1)$.
(iii) $\mathsf{BPol}_1(\mathcal{J}_1) \subseteq \mathsf{PPol}_3(\mathcal{J}_1^-)$.
(iv) $\mathsf{PPol}_1(\mathcal{J}_1) \not\subseteq \mathsf{PPol}_2(\mathcal{J}_1^-)$.
(v) $\mathsf{BPol}_1(\mathcal{J}_1) \not\subseteq \mathsf{PPol}_2(\mathcal{J}_1^-)$.
(vi) $\mathsf{BPol}_1(\mathcal{J}_1^-) \subseteq \mathsf{PPol}_2(\mathcal{J}_1^-)$.
(vii) $\mathsf{PPol}_2(\mathcal{J}_1^-) \not\subseteq \mathsf{BPol}_1(\mathcal{J}_1)$.

Proof. "(i)" : We want to prove that $\mathsf{BPol}_1(\mathcal{J}_1) \subseteq \mathsf{BPol}_2(\mathcal{J}_1^-)$, that is $(\pi_2^-)^{\frown} \subseteq \widehat{\pi_1}$. Let $u, v \in X^*$ be such that $(u, v) \in (\pi_2^-)^{\frown}$. Since $(\pi_2^-)^{\frown} \subseteq (\pi_1^-)^{\frown}$ we have $\mathsf{skel}(u) = \mathsf{skel}(v)$ by Lemma 4. So, let $u = b_1 u_1 b_2 u_2 \dots u_{\ell-1} b_\ell$ and $v = b_1 v_1 b_2 v_2 \dots v_{\ell-1} b_\ell$ be the skeleton decompositions of u and v with the skeleton $\mathsf{skel}(u) = \mathsf{skel}(v) = b_1 \dots b_\ell$, where $u_1, \dots, u_{\ell-1}, v_1, \dots, v_{\ell-1} \in X^*$. We will prove that $\mathsf{c}(u_i) = \mathsf{c}(v_i)$ for each $i = 1, \dots, \ell - 1$.

For each i we consider a factorization $g = (g_0, b_i, g_1, b_{i+1}, g_2)$ of v, such that $g_0 = b_1 v_1 \dots b_{i-1} v_{i-1}$, $g_1 = v_i$ and $g_2 = v_{i+1} b_{i+2} \dots v_{\ell-1} b_\ell$. Then there is a factorization $f = (f_0, b_i, f_1, b_{i+1}, f_2)$ of u such that $\mathsf{c}(f_0) \subseteq \mathsf{c}(g_0)$, $\mathsf{c}(f_1) \subseteq \mathsf{c}(g_1)$ and $\mathsf{c}(f_2) \subseteq \mathsf{c}(g_2)$. If $b_i \notin \mathsf{c}(g_0)$, then $b_i \notin \{b_1, \dots, b_{i-1}\} = \mathsf{c}(b_1 u_1 \dots b_{i-1} u_{i-1})$ and $b_i \notin \mathsf{c}(f_0)$, which implies that $f_0 = b_1 u_1 \dots b_{i-1} u_{i-1}$. If $b_i \notin \mathsf{c}(g_1 b_{i+1} g_2)$, then $b_i \notin \{b_{i+1}, \dots, b_\ell\} = \mathsf{c}(u_i b_{i+1} \dots u_{\ell-1} b_\ell)$ and $b_i \notin \mathsf{c}(f_1 b_{i+1} f_2)$, which implies that $f_1 b_{i+1} f_2 = u_i b_{i+1} \dots u_{\ell-1} b_\ell$. We can prove the analogue for b_{i+1} and hence $f_1 = u_i$. This means $\mathsf{c}(u_i) = \mathsf{c}(f_1) \subseteq \mathsf{c}(g_1) = \mathsf{c}(v_i)$.

If we exchange the role of v and u we obtain also $\mathsf{c}(v_i) \subseteq \mathsf{c}(u_i)$ and the claim is proved. Now the definition of $\widehat{\pi_1}$ gives immediately $(u, v) \in \widehat{\pi_1}$.

Item (ii) can be proved in the similar way as part (i) taking into account Lemma 9.

"(iii)" : If $(u, v) \in \pi_3^-$ then $\mathsf{Sub}_2(v) \subseteq \mathsf{Sub}_2(u)$, by the definition of π_3^-. Since $\pi_3^- \subseteq \pi_1^-$ we get, by Lemma 2, $\mathsf{first}(u) = \mathsf{first}(v)$, $\mathsf{last}(u) = \mathsf{last}(v)$ and $\mathsf{Sub}_2(u) \subseteq \mathsf{Sub}_2(v)$. Now we can apply Lemma 4 to get $\mathsf{skel}(u) = \mathsf{skel}(v)$. Then one can continue similarly as in part (i).

"(iv)" : Since $(x^2, x^3) \in \pi_2^-$ but $(x^2, x^3) \notin \pi_1$ we see that the inclusion $\pi_2^- \subseteq \pi_1$ does not hold.

Item (v) follows from (iv).

"(vi)" : We want to prove that $\pi_2^- \subseteq (\pi_1^-)^{\frown}$. Let $u, v \in X^*$ be such that $(u, v) \in \pi_2^-$. From the definition of π_2^- we deduce that $\mathsf{Sub}_2(v) \subseteq \mathsf{Sub}_2(u)$ and $\pi_2^- \subseteq \pi_1^-$. Thus by Lemmas 2 and 4 we have $\mathsf{skel}(u) = \mathsf{skel}(v)$. If $|\mathsf{skel}(u)| \leq 1$, then $u = v$. If $|\mathsf{skel}(u)| \geq 2$, with the notation of the proof of (i) we have $\mathsf{c}(u_i) \subseteq \mathsf{c}(v_i)$ for all $i \in \{1, \dots, \ell - 1\}$. Then $(v, u) \in \pi_1 \subseteq \pi_1^-$ and hence $(u, v) \in \pi_1^- \cap (\pi_1^-)^{\mathsf{d}} = (\pi_1^-)^{\frown}$.

"(vii)" : The definitions of $\widehat{\pi_1}$ and π_2^- give $(xyxyxy, xyyxxy) \in \widehat{\pi_1} \setminus \pi_2^-$. □

Final Remark. At present the authors have a rather technical proof of the fact that considered hierarchies for $k \geq 2$ really form the lattice isomorphic to the product of four-element lattice with the chain of positive integers. Thus the poset formed by the members of the four hierarchies is represented by Figure 1. Note that our four hierarchies do not form a lattice with respect to inclusion; for instance, there is no supremum for $\mathsf{PPol}_2(\mathcal{J}_1^-)$ and $\mathsf{BPol}_1(\mathcal{J}_1)$.

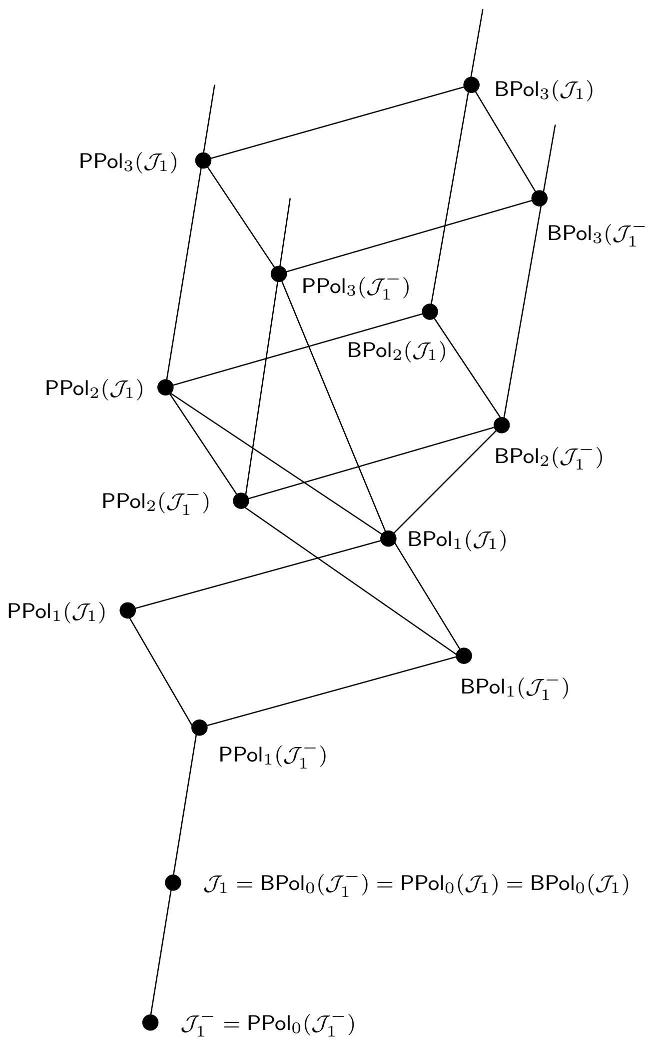

Fig. 1. Hierarchies ordered by inclusion

Acknowledgment

The authors would like to express their deep gratitude to the anonymous referee of the paper. His/her numerous comments and suggestions helped us to avoid several mistakes and increased the understandability of our presentation.

References

[1] S. L. Bloom, Varieties of ordered algebras, *J. Comput. Syst. Sci.* **13** (2) (1976), 200–212

[2] S. Burris and H. P. Sankappanavar, *A Course in Universal Algebra*, Springer, Berlin, 1981

[3] F. Blanchet-Sadri, Equations and monoids varieties of dot-depth one and two, *Theoret. Comput. Sci.* **123** (1994), 239–258

[4] S. Eilenberg, *Automata, Languages and Machines,* Vol. B., Academic Press, New York, 1976

[5] O. Klíma and L. Polák, Hierarchies of piecewise testable languages, in *Developments in Language Theory (DLT 2008)*, Lecture Notes in Computer Science 5257 (2008), 479–490, an extended version: International Journal of Foundations of Computer Science, Vol. **21**, No. 4 (2010), 517–533

[6] O. Klíma and L. Polák, Polynomial operators on classes of regular languages, in *Algebraic Informatics, Third International Conference (CAI 2009)*, Lecture Notes in Computer Science 5725 (2009), 260–277

[7] J.-E. Pin, A variety theorem without complementation, *Russian Mathem. (Iz. VUZ)* **39** (1995), 74–83

[8] J.-E. Pin, Syntactic semigroups, Chapter 10 in *Handbook of Formal Languages*, G. Rozenberg and A. Salomaa eds, Springer, 1997

[9] J.-E. Pin, Algebraic tools for the concatenation product, *Theoretical Computer Science* **292** (2003), 317-342

[10] I. Simon, Piecewise testable events, in *Proc. ICALP 1975*, Lecture Notes in Computer Science 33 (1975), 214–222

[11] M. V. Volkov and I. A. Goldberg, The finite basis problems for monoids of triangular boolean matrices, in Algebraic Systems, Formal Languages, and Conventional and Unconventional Computation Theory, Research Institute for Mathematical Sciences, Kyoto University, 2004, 205–214