HIERARCHIES AND REDUCIBILITIES ON REGULAR LANGUAGES RELATED TO MODULO COUNTING [*]

Victor L. Selivanov [1]

Abstract. We discuss some known and introduce some new hierarchies and reducibilities on regular languages, with the emphasis on the quantifier-alternation and difference hierarchies of the quasi-aperiodic languages. The non-collapse of these hierarchies and decidability of some levels are established. Complete sets in the levels of the hierarchies under the polylogtime and some quantifier-free reducibilities are found. Some facts about the corresponding degree structures are established. As an application, we characterize the regular languages whose balanced leaf-language classes are contained in the polynomial hierarchy. For any discussed reducibility we try to give motivations and open questions, in a hope to convince the reader that the study of these reducibilities is interesting for automata theory and computational complexity.

Mathematics Subject Classification. 03D05, 03C13, 68Q15.

1. Introduction

The notion of hierarchy appeared in descriptive set theory as a classification tool for characterizing complexity of sets studied in analysis. The notion of reducibility appeared in computability theory and plays a central role in the classification of undecidable problems.

[1] A.P. Ershov Institute of Informatics Systems, Siberian Division of the Russian Academy of Sciences; vseliv@nspu.ru

Later, different notions of hierarchies and reducibilities were employed in different branches of computation theory and of definability theory (examples in descriptive set theory are the Borel hierarchy and the Wadge reducibility, in complexity theory – the polynomial-time hierarchy and the polynomial-time m-reducibility, in finite model theory – the logical hierarchies and reducibilities and so on). Some of these hierarchies and reducibilities turned out to be also quite important for the corresponding fields.

Hierarchies in automata theory (*e.g.* the dot-depth hierarchy) were introduced long ago [7]. More recently, people began to consider reducibilities inducing nontrivial degree structures on the regular sets (*i.e.*, on the languages recognized by finite automata) [3,12,13,38,39,43]. In particular, there exists a natural quantifier-free reducibility that fits the dot-depth hierarchy in the sense that every level is downward closed and has a complete set under this reducibility.

In this paper, we continue to discuss some known and introduce some new hierarchies and reducibilities on the regular sets, with the emphasis on the quantifier-alternation and difference hierarchies of the quasi-aperiodic languages (axiomatized by sentences of signatures containing predicates that count positions modulo a given number). The non-collapse of these hierarchies and decidability of some levels are established. Complete sets in the levels of the hierarchies under the polylogtime and quantifier-free reducibilities are found. Several facts on the corresponding degree structures are established. As an application, we characterize the regular languages whose leaf-language classes (in the balanced model of leaf language definability) are contained (uniformly on oracles) in the polynomial hierarchy.

This paper is closely related to [12–14,26,27,35,38,39], and we often refer to the results and proofs there. Reading of our paper would become much easier with these sources at hand.

In Section 2 we recall some notions and known facts and state some new facts related to the logical approach to automata theory. In Section 3 we consider hierarchies of regular languages induced by the quantifier-alternation hierarchies of first-order formulas. Sections 4–7 are devoted to the difference hierarchies over levels of the quantifier-alternation hierarchies. In Section 8 we discuss the important polylogtime reducibility closely related to the so called leaf language approach to complexity classes which is described rather comprehensively in [45]. In Sections 9 and 10 we consider some versions of the quantifier-free reducibility [38,39] which fit the introduced hierarchies. In Section 11 we present some results on the first-order reducibilities. We conclude in Section 12 with mentioning some other reducibilities and open questions.

2. Regular languages and logic

We use (mostly without definitions here) some standard terminology and notation from computability theory, automata theory and complexity theory, say the terminology on reducibilities and degrees, the notation of languages by regular

expressions or the concept of polynomial-time non-deterministic Turing machine. Letters A, B will denote alphabets which are always assumed to contain at least two symbols. By A^+ we denote the set of all non-empty words over A, and by A^* the set of all words (including the empty word ε). For any k, let A^k denote the set of words over A of length k; notations $A^{\leq k}$ and $A^{>k}$ are defined in the same manner. Since usually we work with a fixed alphabet A, we normally do not mention the alphabet explicitly. The length of a word w is denoted $|w|$. For every $i < |w|$, $w(i)$ denotes the ith letter of w (thus, we start the numbering of letters in w with 0). By $\#_a(w)$ we denote the number of entries of the letter a in the word w.

Since we use the logical approach to regular languages [4,20,40,42] and the empty structures are not usual in logic, we work mostly with the languages of non-empty words $L \subseteq A^+$. Correspondingly, the complement $\overline{L}$ of such a language L is defined by $\overline{L} = A^+ \setminus L$. As usual, $P(A^+)$ denotes the power set of A^+. By $P'(A^+)$ we denote the class of all non-trivial (i.e. distinct from $\emptyset$ and A^+) languages over A. By $\mathcal{R}$ ($\mathcal{R}'$) we denote the class of all regular (respectively, regular non-trivial) languages over A. For a class $\mathcal{C}$ of languages, let $\mathrm{BC}(\mathcal{C})$ be the Boolean closure of $\mathcal{C}$, i.e., the closure of $\mathcal{C}$ under union and complement. By co-$\mathcal{C}$ we denote the class of complements of languages in $\mathcal{C}$.

Relate to any alphabet $A = \{a, \ldots\}$ the signature $\sigma = \sigma_A = \{\leq, Q_a, \ldots, \bot, \top, p, s\}$ where $\leq$ is a binary relation symbol, Q_a (for each $a \in A$) is a unary relation symbol, $\bot$ and $\top$ are constant symbols, and p, s are unary function symbols. A word $u = u_0 \ldots u_n \in A^+$ may be considered as a structure $\mathbf{u} = (\{0, \ldots, n\}; \leq, Q_a, \ldots)$ of signature σ, where $\leq$ has its usual meaning, $Q_a (a \in A)$ are unary predicates on $\{0, \ldots, n\}$ defined by $Q_a(i) \leftrightarrow u_i = a$, the symbols $\bot$ and $\top$ denote respectively the least and the greatest elements, while p and s are respectively the predecessor and successor functions on $\{0, \ldots, n\}$ satisfying $p(0) = 0$ and $s(n) = n$. For a sentence ϕ of σ, let $L_\phi = \{u \in A^+ \mid \mathbf{u} \models \phi\}$. Sentences ϕ, ψ are treated as equivalent when $L_\phi = L_\psi$. In [20] it was shown that the class of FO_σ-axiomatizable languages (i.e., languages of the form L_ϕ, where ϕ ranges through the first-order sentences of σ), coincides with the class of regular aperiodic languages (known also as star-free languages).

Remark. We use in this paper the term "axiomatizable" to denote the languages of the form L_ϕ instead of the more popular in the literature on automata theory term "definable" because our term corresponds better to the old tradition in logic. Note that the term "finitely axiomatizable" would be even more appropriate but we can use the abbreviated form safely because consider only finitely axiomatizable languages.

We will consider also some subsignatures and enrichments of the signature σ. In particular, let $\rho = \{<, Q_a, \ldots\}$, and for any positive integer d let τ_d be the signature $\sigma \cup \{P_d^0, \ldots, P_d^{d-1}\}$, where P_d^r is the unary predicate true on the positions of a word which are equivalent to r modulo d. By FO_{τ_d}-*axiomatizable language* we mean any language of the form L_ϕ, where ϕ is a first-order sentence of signature τ_d.

Note that signature τ_1 is essentially the same as σ because P_1^0 is the valid predicate. In contrast, for $d > 1$ FO_{τ_d}-axiomatizable languages need not be aperiodic. *E.g.*, the sentence $P_2^1(\top)$ defines the language L consisting of all words of even length which is known to be non-aperiodic. It is easy to see that the predicates $P_d^1, \ldots, P_d^{d-1}$ may be eliminated from the formulas of signature τ_d (*e.g.* $P_d^1(x)$ is essentially equivalent to $P_d^0(s(x))$). Nevertheless, for technical reasons we will not remove them from τ_d. We are also interested in the infinite signature $\tau = \bigcup_d \tau_d$. Note that the signatures $\rho_d = \rho \cup \{P_d^0, \ldots, P_d^{d-1}\}$ and $\rho_{\mathcal{P}} = \bigcup\{\rho_d \mid d \in \mathcal{P}\}$, for each set $\mathcal{P}$ of positive integers, were discussed in [5,10,35].

Let us formulate a precise characterization of the introduced classes of languages in terms of syntactic monoids and homomorphisms (for details see *e.g.* Chap. 5 of [35]). For a language L, let $M(L)$ be its syntactic monoid and $\eta_L : A^* \to M(L)$ the canonical syntactic homomorphism. We denote the semigroup operation on $M(L)$ by $\cdot$. As is well-known, L is regular iff $M(L)$ is finite. A language L is called *aperiodic* if there is no non-trivial group $(G; \cdot) \subseteq (M(L); \cdot)$ (as usual, $\subseteq$ here means the substructure relation). A language L is called *quasi-aperiodic* [35] if there is no non-trivial group $(G; \cdot) \subseteq (\eta_L(A^d); \cdot)$ for each $d > 0$ where $\eta_L(A^d)$ is the image of A^d under η_L. We call a language L *d-quasi-aperiodic* (for any fixed $d > 0$), if there is no non-trivial group $(G; \cdot) \subseteq (\eta_L((A^d)^+); \cdot)$. Note that L is aperiodic iff L is 1-quasi-aperiodic.

Theorem 2.1.

    (1) *A regular language L is quasi-aperiodic iff it is FO_τ-axiomatizable.*

    (2) *A regular language L is d-quasi-aperiodic iff it is FO_{τ_d}-axiomatizable.*

    (3) *A regular language L is aperiodic iff it is FO_σ-axiomatizable.*

For the proof of (1) see Theorem VI.4.1 in [35]. The proof of (2) is implicitly contained in the proof of that Theorem VI.4.1. For a detailed proof of (2) (even for an arbitrary set $\mathcal{P}$ of moduli) see [10]. The assertion (3) is a particular case of (2) and is a classical result of Schützenberger, McNaughton and Papert [20,25].

The last theorem implies the following important decidability result (for details see *e.g.* [10,35]).

Corollary 2.2. *The classes of languages from the last theorem are decidable.*

Remark. In [10] several additional interesting facts about the first-order axiomatizable languages were established, in particular $\text{FO}_{\tau_a \cup \tau_b} = \text{FO}_{\tau_c}$ where c is the least common multiple of a and b, and $\text{FO}_{\tau_a} \cap \text{FO}_{\tau_b} = \text{FO}_{\tau_d}$ where d is the greatest common divisor of a and b.

From the interpretation of signature symbols in the word structures **u** it follows that the nonempty words correspond bijectively to (the isomorphism types of) the finite models of the theory $\text{CLO}_A^{\tau_d}$ of signature τ_d (CLO stand for "colored linear orders") with the following axioms:

    – $\leq$ is a linear order,

    – any element satisfies exactly one of the predicates $Q_a (a \in A)$,

    – $\forall x(\bot \leq x \leq \top)$,

    – $\forall x(p(x) \leq x \land \neg\exists y(p(x) < y < x))$,

- $\forall x(x \leq s(x) \wedge \neg \exists y(x < y < s(x)))$,
- $\forall x(x > \bot \to p(x) < x)$,
- $\forall x(x < \top \to x < s(x))$,
- $P_d^0(\bot)$,
- any element satisfies exactly one of the predicates $P_d^0, \ldots, P_d^{d-1}$,
- $\forall x < \top(P_d^r(x) \to P_d^{r+1}(s(x)))$ for $0 \leq r < d-1$,
- $\forall x < \top(P_d^{r-1}(x) \to P_d^0(s(x)))$.

Sometimes it is technically more convenient to consider "relational" versions of the signature τ_d (and of the theory CLO^{τ_d}). *E.g.*, one could take the signature $\tau_d' = \{\leq, Q_a, \ldots, P_d^r, \bot, \top, S\}$, where $S(x, y)$ is a binary relation symbol interpreted as "x is an immediate predecessor of y". The signatures τ_d and τ_d' are equivalent for most of our purposes, in particular the quantifier alternation hierarchies over them (discussed in the next section) coincide. So we may use any of the signatures when appropriate.

Remark. Analogs of many results of this paper hold (with similar proofs) also for some signatures not discussed explicitly in what follows, in particular for the signatures ρ_d mentioned above. We present all details for the signatures τ_d and τ because they are better related to the leaf language definability.

For any set $\mathcal{P}$ of positive integers, let $\mathrm{FO} + \mathrm{MOD}(\mathcal{P})$ denote the class of languages axiomatized by σ-sentences using modulo counting quantifiers $\exists^{(q,r)}$ with moduli q in $\mathcal{P}$, along with the usual first-order quantifiers. It is known (see [37] or Chap. 7 of [35]) that the class $\mathrm{FO} + \mathrm{MOD} = \mathrm{FO} + \mathrm{MOD}(\{1, 2, \ldots\})$ consists exactly of languages with solvable syntactic monoid. Any FO_{τ_d}-axiomatizable language is $\mathrm{FO} + \mathrm{MOD}(\{d\})$-axiomatizable. The language $L \subseteq \{a, b\}^+$ consisting of the words with even number of entries of a is $\mathrm{FO} + \mathrm{MOD}(\{2\})$-axiomatizable but not quasi-aperiodic. More information on the logical approach may be found in [11,23,26,35,40,41].

3. Quantifier-alternation hierarchies

In this section we consider hierarchies of regular languages induced by the quantifier-alternation hierarchy of first-order formulas in prenex normal form. For any signature θ we obtain the corresponding hierarchy called θ-hierarchy.

For $n > 0$, let Σ_n^σ be the class of all languages L_ϕ, where ϕ ranges through the Σ_n-sentences of σ. Let $\Pi_n^\sigma = \mathrm{co}\text{-}\Sigma_n^\sigma$ and $\Delta_n^\sigma = \Sigma_n^\sigma \cap \Pi_n^\sigma$. In [40] it was shown that the σ-hierarchy (more exactly, the sequence $\{\mathrm{BC}(\Sigma_n^\sigma)\}_{n>0}$) coincides with the dot-depth hierarchy which is a popular object of automata theory. If we take the smaller signature $\rho = \{<, Q_a, \ldots\}$, we obtain the ρ-hierarchy $\{\Sigma_n^\rho\}$ known as the Straubing-Thérien hierarchy. The mentioned hierarchies of regular languages have natural characterizations in terms of regular expressions [23,40]; we do not recall these characterizations because we work here only with the logical definition.

Levels of the τ_d-hierarchy are denoted $\Sigma_n^{\tau_d}, \Pi_n^{\tau_d}, \Delta_n^{\tau_d}$. Of course, the τ_d-hierarchy exhausts the class of τ_d-axiomatizable languages, and the τ_1-hierarchy coincides

with the σ-hierarchy. To our knowledge, the τ_d-hierarchy for $d > 1$ (as well as the τ-hierarchy discussed below) was not considered in the literature so far.

One could also consider similar hierarchies of languages for signatures like $\tau_{d_1} \cup \cdots \cup \tau_{d_k}$, but in fact we do not obtain new hierarchies in this way. The next result refines a related fact in [10] mentioned in the previous section.

Proposition 3.1. *For all $n, c, e > 0$ there holds $\Sigma_n^{\tau_c \cup \tau_e} = \Sigma_n^{\tau_d}$, where d is the least common multiple of c and e.*

Proof. For the inclusion from left to right, we have to show that the predicates P_c^0 and P_e^0 are easily (say, by the quantifier-free formulas) expressible through the predicates P_d^r. Indeed, let $d = ck$. Then $P_c^0(x)$ is equivalent to $P_d^0(x) \vee \cdots \vee P_d^{k-1}(x)$, and similarly for e.

For the converse inclusion, let a be the greatest common divisor of c, e. Then $d = c'e$, where $c' = c/a$. Since, by the previous paragraph, P_c^0 is easily expressible through P_d^0, it suffices to express P_d^0 through $P_{c'}^0$ and P_e^0. Since c' and e have no common divisors, it holds $P_d^0(x) \leftrightarrow P_{c'}^0(x) \wedge P_e^0(x)$. $\qquad\square$

By $\Sigma_n^\tau, \Pi_n^\tau, \Delta_n^\tau$ we denote the levels of the τ-hierarchy. From the last proposition we immediately obtain the following relation between the introduced hierarchies.

Corollary 3.2. *For any $n > 0$ there hold $\Sigma_n^\tau = \bigcup_d \Sigma_n^{\tau_d}$, $\Pi_n^\tau = \bigcup_d \Pi_n^{\tau_d}$ and $\Delta_n^\tau = \bigcup_d \Delta_n^{\tau_d}$.*

Obviously, any Σ-level of the quantifier-alternation hierarchies is closed under union and intersection and contains $\emptyset$ and L^+ as elements, while any Δ-level is closed under union, intersection and complement.

The main result about the quantifier-alternation hierarchies is that they do not collapse. This can be checked by the standard method of Ehrenfeucht-Fraïssé games developed in [32,33,40,41] for the ρ- and σ-hierarchies and the corresponding difference hierarchies. Those proofs are generalized to the hierarchies of this section in a straightforward way. Following the referees request, we provide some details of the proof.

Theorem 3.3. *For any $n > 0$ there holds $\Sigma_n^\rho \nsubseteq \Pi_n^\tau$. In particular, $\Sigma_n^\tau \nsubseteq \Pi_n^\tau$ and $\Sigma_n^{\tau_d} \nsubseteq \Pi_n^{\tau_d}$ for all $n, d > 0$.*

We start with introducing some notions and establishing some lemmas. Let ν be a finite set of variables. Recall that dealing with the Ehrenfeucht-Fraïssé games is more comfortable for the signatures without function symbols, like the signatures ρ and τ_d' from the previous section.

As we already noted, $\Sigma_k^{\tau_d} = \Sigma_k^{\tau_d'}$ for all $d, k \geq 1$. In this section we work with formulas and structures of the signature τ_d'.

By ν-*structure* we mean a word from A^+ (interpreted as a structure of signature τ_d') together with an assignment of values to the variables from ν. Note [23,35] that the ν-structures may be considered as nonempty words $(a_0, U_0) \cdots (a_n, U_n)$ over the bigger alphabet $A \times P(\nu)$ where U_i is the set of variables assigned to the position i of the "usual" word $a_0 \cdots a_n$; the sets $U_0, \ldots, U_r$ are pairwise disjoint

and exhaust $P(\nu)$. Formulas of signature τ'_d with free variables in ν are interpreted in the ν-structures in the usual way. The $\emptyset$-structures are identified with the nonempty words over A. Since the modulo d is fixed till the proof of Theorem 3.3, we do not mention it explicitly in the technical notions like $S_{k,r}$ or $\leqslant_{k,r}$ we are going to introduce (the exact notation would be $S^d_{k,r}$ or $\leqslant^d_{k,r}$, respectively).

Define for any $k \geq 0$ and $r \geq 1$ the set $S_{k,r}$ of formulas of signature τ'_d by induction on k as follows. Let $S_{0,r}$ be the set of quantifier-free formulas in disjunctive normal form without repetitions of the elementary conjunctions and without repetitions of members of the elementary conjunctions. For $k > 0$, let $S_{k,r}$ be the set of finite disjunctions of pairwise distinct formulas $\exists z^k_1 \exists z^k_2 \ldots \exists z^k_p \neg \phi$ where $p \leq r$ and $\phi \in S_{k-1,r}$. Thus, all bounded variables of the formulas in $S_{k,r}$ are among the variables $z^i_1, z^i_2, \ldots z^i_r$, $i \in \{1, \ldots, k\}$ fixed in advance. The next assertion is obvious.

Lemma 3.4.

(1) *Let ν be a finite set of variables, $k \geq 0$ and $r \geq 1$. Then the set of sentences in $S_{k,r}$ with free variables among ν is finite.*

(2) *For any $k \geq 1$, the class of languages axiomatized by the sentences from $\bigcup_r S_{k,r}$ coincides with $\Sigma^{\tau_d}_k$.*

For all ν, $k \geq 0$ and $r \geq 1$, define a preorder $\leqslant_{k,r}$ on the ν-structures as follows: $u \leqslant_{k,r} v$ iff $u \models \phi$ implies $v \models \phi$, for all formulas $\phi \in S_{k,r}$ with free variables among ν. As usual, let $\equiv_{k,r}$ denote the associated equivalence relation. By the previous lemma, the relation $\equiv_{k,r}$ is of finite index. By *upper set* of a preorder we mean a set of elements closed upwards under the preorder relation.

Lemma 3.5. *For all $L \subseteq A^+$ and $k \geq 1$, $L \in \Sigma^{\tau_d}_k$ iff L is an upper set in $(A^+; \leqslant_{k,r})$ for some $r \geq 1$.*

Proof. Let $L \in \Sigma^{\tau_d}_k$, then $L = L_\phi$ for some Σ_k-sentence ϕ of signature τ'_d. By Lemma 3.4, ϕ is equivalent to a sentence from $S_{k,r}$ for some r. Then L is an upper set in $(A^+; \leqslant_{k,r})$. Conversely, let L be an upper set in $(A^+; \leqslant_{k,r})$. Since L is a finite union of upper "cones" $\check{u} = \{v \mid u \leqslant_{k,r} v\}$ in $(A^+; \leqslant_{k,r})$, it suffices to show that any such a cone $\check{u}$ is in $\Sigma^{\tau_d}_k$. Clearly, $\check{u} = L_\phi \in \Sigma^{\tau_d}_k$ where ϕ is $\bigwedge \{\psi \in S_{k,r} \mid u \models \psi\}$. $\square$

Next we characterize the introduced preorder in terms of a k-round Ehrenfeucht-Fraïssé game $G_{k,r}(u,v)$ defined for any given ν-structures u and v as follows. There are two players denoted 1 and 2. The player 1 wants to show that the structures are distinct while the player 2 wants to show they are similar. Each player has his/her copies of the kr variables $z^j_1, \ldots, z^j_r$, $j = 1, \ldots, k$ (realized *e.g.* as kr pebbles labeled by the variables). For $k = 0$, the players make no moves at all and simply determine the winner by the following rule: the player 2 wins iff $u \equiv_{0,r} v$. Now let $k \geq 1$. At the first round, player 1 chooses a number $p \leq r$ and puts his/her pebbles $z^k_1, \ldots, z^k_p$ at some positions of the ν-structure u, forming thus a $\{y_1, \ldots, y_m, z^k_1, \ldots, z^k_p\}$-structure u', where $\nu = \{y_1, \ldots, y_m\}$. The player 2 answers by putting his/her variables $z^k_1, \ldots, z^k_p$ at some positions of v forming a

$\{y_1, \ldots, y_m, z_1^k, \ldots, z_p^k\}$-structure v'. If $k = 1$, the game is over and the winner is determined by the rule above for u' and v'. Otherwise, the players continue as in the game $G_{k-1,r}(v', u')$. The notion of a winning strategy for each player is defined in the obvious way. Since the game $G_{k,r}(u, v)$ is finite, one of the players has a winning strategy.

Lemma 3.6. *Let u and v be ν-structures and let $k \geqslant 0$, $r \geqslant 1$. Then $u \leqslant_{k,r} v$ iff player 2 has a winning strategy in $G_{k,r}(u, v)$.*

Proof. Let $u \leq_{k,r} v$. We check by induction on k that player 2 has a winning strategy in $G_{k,r}(u, v)$. For $k = 0$ this is clear because u and v satisfy the same quantifier-free formulas. Let $k > 0$. Towards a contradiction, suppose that player 2 does not have a winning strategy, hence player 1 has a winning strategy. Assign the values to $z_1^k, \ldots, z_p^k$ $(p \leq r)$ in u according to the first move of player 1 by this strategy. Then we obtain a structure u' such that for all values of $z_1^k, \ldots, z_p^k$ $(p \leq r)$ in v (resulting in a structure v') player 1 has a winning strategy in $G_{k-1,r}(v', u')$. By induction hypothesis, $v' \not\leqslant_{k-1,r} u'$. Let ϕ be the disjunction of all formulas from $S_{k-1,r}$ that are false in u'. Then $v' \not\models \neg\phi$. Since the values of $z_1^k, \ldots, z_p^k$ in v were arbitrary,

$$v \not\models \exists z_1^k, \ldots, z_p^k \neg\phi \text{ and } u \models \exists z_1^k, \ldots, z_p^k \neg\phi,$$

which is a contradiction because the formula $\exists z_1^k, \ldots, z_p^k \neg\phi$ is equivalent to a formula in $S_{k,r}$.

The opposite implication is also proved by induction on k. If $k = 0$ and player 2 has a winning strategy then $u \equiv_{0,r} v$ and we are done. Now let $k > 0$ and fix a winning strategy for player 2 in $G_{k,r}(u, v)$. Towards a contradiction, suppose that $u \not\leqslant_{k,r} v$, then there is a $\psi \in S_{k,r}$ such that $u \models \psi$ and ¨ $v \not\models \psi$. W.l.o.g. we can assume that ψ is of the form $\exists z_1^k, \ldots, z_r^k \neg\phi$ where $\phi \in S_{k-1,r}$. Since $u \models \psi$, player 1 can put the variables $z_1^k, \ldots, z_r^k$ in u in such a way that the resulting structure u' satisfies $\neg\phi$. Taking the values of $z_1^k, \ldots, z_r^k$ in v according to the winning strategy for player 2 we obtain a structure v' that does not satisfy $\neg\phi$. But player 2 has a winning strategy in $G_{k-1,r}(v', u')$. By induction, $v' \leq_{k-1,r} u'$ and therefore $v' \not\models \phi$. A contradiction.   □

The next corollary of the previous lemma states that the concatenation of ν-structures respects (under a reasonable assumption) the introduced preorder. We omit the obvious proof using the game characterization of $\leqslant_{k,r}$.

Lemma 3.7. *Let ν be a finite set of variables, $k \geq 0$, $r \geq 1$, and u_1, u_2, v_1, v_2, $u_1 u_2$, $v_1 v_2$ be ν-structures. Then $u_1 \leqslant_{k,r} v_1$ and $u_2 \leqslant_{k,r} v_2$ imply $u_1 u_2 \leqslant_{k,r} v_1 v_2$.*

For all $k \geq 0$ and $r \geq 1$, define $c(k, r)$ by induction on k as follows: $c(0, r) = 1$ and $c(k, r) = r + (r+1)c(k-1, r) + 1$. Until the end of this section, an upper index on a word means the corresponding power of this word under concatenation.

Lemma 3.8. *Let $w \in A^+$, $|w| \equiv 0 \pmod{d}$, and $N_1, N_2 \geq c(k, r)$. Then $w^{N_1} \leq_{k,r} w^{N_2}$.*

Proof. We can assume that $N_1 \neq N_2$. It suffices to prove the assertion for $|N_1 - N_2| = 1$ because for $|N_1 - N_2| > 1$ it will then follow by induction on $|N_1 - N_2|$. We have to show that $w^{N_1} \leq_{k,r} w^{N_1+z}$ where $z = N_2 - N_1 \in \{-1, 1\}$. The proof is by induction on k. For $k = 0$ the inequality $w^{N_1} \leq_{k,r} w^{N_1+z}$ is true because w^{N_1} and w^{N_1+z} satisfy the same atomic sentences of signature τ'_d (the sentences are of the form $i = j$, $i < j$, $S(i,j)$, $Q_a(i)$, $P_d^0(i), \ldots, P_d^{d-1}(i)$ for $i, j \in \{\bot, \top\}$).

We assume that the inequality holds for $k-1 \geq 1$ and prove it for k by describing a winning strategy for player 2 in $G_{k,r}(w^{N_1}, w^{N_2})$. Let in the first round player 1 put p his/her pebbles (for some $p \leq r$) in w^{N_1}. Since $N_1 \geq c(k,r)$ the resulting structure will have a factor $w^{c(k-1,r)+1}$ without pebbles, *i.e.* the word w^{N_1} may be factorized as $w_1 w^{c(k-1,r)+1} w_2$ where w_1 and w_2 are some powers of w containing all the p pebbles $z_1^k, \ldots, z_p^k$. Hence, after the first round the left structure looks as $w_1' w^{c(k-1,r)+1} w_2'$. Note that, before the answer of player 2, the right word looks as $w^{N_1+z} = w_1 w^{c(k-1,r)+1+z} w_2$. Player 2 puts his/her pebbles $z_1^k, \ldots, z_p^k$ in this word in a way to obtain $w_1' w^{c(k-1,r)+1+z} w_2'$. By induction hypothesis,

$$w^{c(k-1,r)+1+z} \leq_{k-1,r} w^{c(k-1,r)+1},$$

hence, by Lemma 3.7,

$$w_1' w^{c(k-1,r)+1+z} w_2' \leq_{k-1,r} w_1' w^{c(k-1,r)+1} w_2'.$$

Player 2 continues by the winning strategy in $G_{k-1,r}(w_1'w^{c(k-1,r)+1+z}w_2',$ $w_1' w^{c(k-1,r)+1} w_2')$. $\square$

For any fixed k and r, define operations F and G on A^+ by $F(u) = u^N$ and $G(u,v) = v^M u v^M$ where $N = r + (r+1)(2c(k,r)+1)$ and $M = N + c(k,r)$.

Lemma 3.9. *Let $u, v \in A^+$, $|u| \equiv |v| \equiv 0 \pmod{d}$, $k \geq 0$, $r \geq 1$, and $u \leq_{k,r} v$. Then $F(v) \leq_{k+1,r} G(u,v)$.*

Proof. We describe a winning strategy for player 2 in $G_{k+1,r}(F(v), G(u,v))$. Let in the first round player 1 put his/her p pebbles ($p \leq r$) in the word $w_1 = F(v) = v^N$. By definition of N, the resulting structure has a factor $v^{2c(k,r)+1}$ without pebbles, *i.e.* the word w_1 may be factorized as $u_1 v^{2c(k,r)+1} u_2 = u_1 v^{c(k,r)} v v^{c(k,r)} u_2$ where u_1 and u_2 are some powers of v containing all the p pebbles. The resulting structure looks then as $w_1' = u_1' v^{c(n,r)} v v^{c(n,r)} u_2'$. Note that the right word $w_2 = G(u,v)$ may be factorized as $w_1 v^{M-N} u v^{M-N} w_1 = u_1 v^{m_1} u v^{m_2} u_2$ where $m_1 \geq c(k,r)$ and $m_2 \geq c(k,r)$. Player 2 answers by putting his/her p pebbles in w_2 in a way to obtain $w_2' = u_1' v^{m_1} u v^{m_2} u_2'$. By Lemma 3.8, $v^{m_1} \leq_{k,r} v^{c(n,r)}$. By Lemma 3.7, $w_2' \leq_{k,r} w_1'$ hence player 2 has a winning strategy in $G_{k,r}(w_2', w_1')$. Player 2 proceeds by following this winning strategy. $\square$

Proof of Theorem 3.3. To simplify notation, we prove here the following slightly weaker alphabet-dependent version: for any $k \geq 1$ there is a language H_k over the alphabet $A_k = \{0, 1, \ldots, k\}$ with $H_k \in \Sigma_k^\rho \setminus \Pi_k^\tau$. For the alphabet-independent

version, see remarks in the corresponding proof in Section 5 where we establish a strengthening of Theorem 3.3.

We define the set H_k by induction on k as follows:

$$H_1 = A_1^* \cdot 1 \cdot A_1^* \ \text{ and } \ H_{k+1} = A_{k+1}^* \cdot (k+1) \cdot (A_k^* \setminus H_k) \cdot (k+1) \cdot A_{k+1}^* \ \text{ for } k \geq 1$$

(cf. Lem. 3.1 in [38,39]). Obviously, $H_k \in \Sigma_k^\rho$ for each $k \geq 1$, so it remains to show that $H_k \notin \Pi_k^\tau$. By Corollary 3.2, it suffices to show that $H_k \notin \Pi_k^{\tau d}$ for each $d \geq 1$. By Lemma 3.5, it suffices to find for any $k, r \geq 1$ words $u_k, v_k \in A_k^+$ such that $u_k \leqslant_{k,r} v_k$, $u_k \notin H_k$ and $v_k \in H_k$. It is easy to see (by looking at the game $G_{1,r}(u_1, v_1)$) that the words $u_1 = 0^d$ and $v_1 = 0^d 10^{d-1} 0^d = 0^d 10^{2d-1}$ have the desired properties for $k = 1$ and for all $r \geq 1$.

Fix $r \geq 1$ and suppose by induction on k that we have u_k and v_k with the desired properties. By Lemma 3.7, $x \leq_{k,r} y$ where $x = (k+1)^d u_k (k+1)^d$ and $y = (k+1)^d v_k (k+1)^d$. Set $u_{k+1} = F(y)$ and $v_{k+1} = G(x,y)$. By Lemma 3.9, $u_{k+1} \leqslant_{k+1,r} v_{k+1}$. By definition of H_{k+1}, $u_{k+1} \notin H_{k+1}$ and $v_{k+1} \in H_{k+1}$. $\square$

4. ABSTRACT DIFFERENCE HIERARCHIES

In this section we make a couple of general remarks about the difference hierarchy (DH) known also as the Boolean hierarchy.

Let S be any set. By a *base in* S we mean any class $\mathcal{C}$ of subsets of S which is closed under union and intersection and contains $\emptyset$ and S as elements. For any $k < \omega$, let $\mathcal{C}(k)$ be the class of sets $\bigcup_i (L_{2i} \setminus L_{2i+1})$, where $L_0 \supseteq L_1 \supseteq \cdots$ is a descending sequence of sets from $\mathcal{C}$ and $L_i = \emptyset$ for $i \geq k$. The sequence $\{\mathcal{C}(k)\}_{k<\omega}$ is known as the *difference hierarchy over* $\mathcal{C}$. As is well-known, $\mathcal{C}(k) \cup$co-$\mathcal{C}(k) \subseteq \mathcal{C}(k+1)$ for every k, and $\bigcup_k \mathcal{C}(k) = \mathrm{BC}(\mathcal{C})$.

We will need the following relation between the DH's over different bases.

Proposition 4.1. *Let* $f : T \to S$ *be a surjection,* $\mathcal{C}$ *a base in* S, $\mathcal{M}$ *a base in* T, *and* $f(M) \in \mathcal{C}$ *for all* $M \in \mathcal{M}$. *Then* $f^{-1}(L) \in \mathcal{M}(k)$ *implies* $L \in \mathcal{C}(k)$ *for all* $L \subseteq S$ *and* $k < \omega$.

Proof. Let $f^{-1}(L) \in \mathcal{M}(k)$, then $f^{-1}(L) = \bigcup_i (M_{2i} \setminus M_{2i+1})$ for some $M_0 \supseteq M_1 \supseteq \cdots$, $M_k = \emptyset$, $M_i \in \mathcal{M}$. Then $f(M_0) \supseteq f(M_1) \supseteq \cdots$, $f(M_k) = \emptyset$, and $f(M_i) \in \mathcal{C}$ for all i. It suffices to check that $L = \bigcup_i (f(M_{2i}) \setminus f(M_{2i+1}))$ (then $L \in \mathcal{C}(k)$, as desired).

First we check the inclusion $L \subseteq \bigcup_i (f(M_{2i}) \setminus f(M_{2i+1}))$. Let $y \in L$. Since f is a surjection, $y = f(x)$ for some $x \in T$. Since $x \in f^{-1}(L)$, $x \in M_0$, hence $y \in f(M_0)$. It remains to show that $y \in f(M_{2i+1})$ implies $y \in f(M_{2i+2})$. Let $y = f(x_1)$ for some $x_1 \in M_{2i+1}$. Since $x_1 \in f^{-1}(L)$, $x_1 \in M_{2i+2}$ and therefore $y \in M_{2i+2}$.

The inclusion

$$\overline{L} \subseteq \overline{\bigcup_i (f(M_{2i}) \setminus f(M_{2i+1}))}$$

is checked in the same manner. $\square$

The next notion is well-known from descriptive set theory [18].

Definition 4.2. A base $\mathcal{C}$ is said to have the *separation property*, if for every two disjoint sets $L, K \in \mathcal{C}$ there is a set $M \in \mathcal{C} \cap \text{co-}\mathcal{C}$ with $L \subseteq M \subseteq \overline{K}$. We say that M *separates* L from K (note that it is equivalent to say that $\overline{M}$ separates K from L).

The following easy fact is often of use.

Proposition 4.3. *Let* $\{L_i\}_{i \in I}$ *and* $\{K_j\}_{j \in J}$ *be finite families of subsets of* S *and let* $M_{i,j}$ *separates* L_i *from* K_j *for all* $i \in I, j \in J$. *Then the set* $\bigcup_i \bigcap_j M_{i,j}$ *separates* $L = \bigcup_i L_i$ *from* $K = \bigcup_j K_j$.

Proof. We have $L_i \subseteq M_{i,j} \subseteq \overline{K}_j$ for all $i \in I, j \in J$. Then $L_i \subseteq \bigcap_j M_{i,j} \subseteq \bigcap_j \overline{K}_j$ for all $i \in I$ Therefore, $L = \bigcup_i L_i \subseteq \bigcup_i \bigcap_j M_{i,j} \subseteq \bigcap_j \overline{K}_j = \overline{K}$. $\qquad\square$

We will need the following fact about the DH's over bases with the separation property.

Proposition 4.4. *Let* $\mathcal{C}$ *be a base with the separation property. Then for any* $k < \omega$ *the class* $\mathcal{C}(k+1) \cap \text{co-}\mathcal{C}(k+1)$ *coincides with the class of sets of the form* $(U \cap L) \cup (\overline{U} \cap K)$, *where* $U \in \mathcal{C} \cap \text{co-}\mathcal{C}$, $L \in \mathcal{C}(k)$ *and* $K \in \text{co-}\mathcal{C}(k)$.

Proof. Let $M = (U \cap L) \cup (\overline{U} \cap K)$ where $U \in \mathcal{C} \cap \text{co-}\mathcal{C}$, $L \in \mathcal{C}(k)$ and $K \in \text{co-}\mathcal{C}(k)$. One easily checks (even without using the separation property) that $M \in \mathcal{C}(k+1) \cap \text{co-}\mathcal{C}(k+1)$.

Simplifying notation, we prove the opposite inclusion only for the typical particular case $k = 2$. Let $M \in \mathcal{C}(k+1) \cap \text{co-}\mathcal{C}(k+1)$. Then

$$M = \overline{L}_0 \cup (L_1 \setminus L_2) \text{ and } \overline{M} = \overline{K}_0 \cup (K_1 \setminus K_2), \tag{1}$$

for some sets $L_0 \supseteq L_1 \supseteq L_2$ and $K_0 \supseteq K_1 \supseteq K_2$ from $\mathcal{C}$. Taking complements from (1), we obtain

$$\overline{M} = (L_0 \setminus L_1) \cup L_2 \text{ and } M = (K_0 \setminus K_1) \cup K_2, \tag{2}$$

hence L_2 and K_2 are disjoint. By the separation property, $L_2 \subseteq U \subseteq \overline{K}_2$ for some $U \in \mathcal{C} \cap \text{co-}\mathcal{C}$. Intersecting the first equality in (1) with $\overline{U}$ we obtain

$$M \cap \overline{U} = (\overline{L}_0 \cap \overline{U}) \cup (L_1 \cap \overline{U}) = (\overline{L}_0 \cup L_1) \cap \overline{U}.$$

Intersecting the second equality in (2) with U we obtain $M \cap U = (K_0 \setminus K_1) \cap U$. Therefore, $M = (U \cap L) \cup (\overline{U} \cap K)$ where $L = (K_0 \setminus K_1) \in \mathcal{C}(2)$ and $K = (\overline{L}_0 \cup L_1) \in \text{co-}\mathcal{C}(2)$. $\qquad\square$

Next we formulate an obvious fact showing that the separation property survives under some operation on bases.

Proposition 4.5. *Let* $\{\mathcal{C}_i\}_{i \in I}$ *be a family of bases in* S *such that every* $\mathcal{C}_i$ *has the separation property and for all* $i, j \in I$ *there exists a* $k \in I$ *with* $\mathcal{C}_i \cup \mathcal{C}_j \subseteq \mathcal{C}_k$. *Then* $\mathcal{C} = \bigcup_i \mathcal{C}_i$ *is a base in* S *with the separation property.*

Relate to any partial order (in fact, all the following notions and results in this section apply also to preorders) $P = (P; \leq)$ the base $\mathcal{C}$ in P consisting of all upper sets of P (*i.e.*, the sets $L \subseteq P$ such that $x \in L$ and $x \leq y$ imply $y \in L$), including the empty set. By *alternating chain* of length k for a set $K \subseteq P$ we mean a sequence $(x_0, \ldots, x_k)$ of elements of P such that $x_0 \leq \cdots \leq x_k$ and $x_i \in K \leftrightarrow x_{i+1} \notin K$ for every $i < k$. Such a chain is called an 1-alternating chain if $x_0 \in K$, otherwise it is called a 0-alternating chain. Variants of the following fact frequently appear when dealing with the DH's.

Proposition 4.6. *Let $P = (P; \leq)$ be a partial order and $\mathcal{C}$ the base of upper sets in P. For all $K \subseteq P$ and $k < \omega$, $K \in \mathcal{C}(k)$ iff K has no 1-alternating chain of length k.*

Proof. Let first $K \in \mathcal{C}(k)$, then $K = \bigcup_i (L_{2i} \setminus L_{2i+1})$, where $L_0 \supseteq L_1 \supseteq \cdots$ is a descending sequence of upper sets and $L_k = \emptyset$. Towards a contradiction, suppose that $(x_0, \ldots, x_k)$ is an 1-alternating chain of length k for K. Since $x_0 \in K$, we have $x_0 \in L_0$. Since L_0 is an upper set and $x_0 \leq x_1$, $x_1 \in L_0$. Since $x_1 \notin K$, $x_1 \in L_1$. Continuing in this manner, we obtain $x_k \in L_k = \emptyset$, a contradiction.

In the opposite direction, let K have no 1-alternating chain of length k. For any $i < \omega$, let L_i be the set of all $x \in P$ such that there is an 1-alternating chain $(x_0, \ldots, x_i)$ for K with $x_i \leq x$. Then $L_0 \supseteq L_1 \supseteq \cdots$ is a descending sequence of upper sets and $L_k = \emptyset$. It remains to check that $K = \bigcup_i (L_{2i} \setminus L_{2i+1})$. First we check the inclusion from left to right. Let $x \in K$. Then $x \in L_0$. It suffices to show that $x \in L_{2i+1}$ implies $x \in L_{2i+2}$. Let $x \in L_{2i+1}$, then there is an 1-alternating chain $(x_0, \ldots, x_{2i+1})$ for K with $x_{2i+1} \leq x$. Then $(x_0, \ldots, x_{2i+1}, x)$ is an 1-alternating chain for K, hence $x \in L_{2i+2}$.

It remains to show that if $x \in P \setminus K$ then $x \notin \bigcup_i (L_{2i} \setminus L_{2i+1})$. If $x \notin L_0$, we are done. Now let $x \in L_0$. It suffices to show that $x \in L_{2i}$ implies $x \in L_{2i+1}$, and this is done exactly as in the preceding paragraph. $\square$

We conclude this section by a result about the DH's in well partial orders (wpo) that is closely related to some facts in automata theory, *e.g.* to results like Theorem 3.3 in [34]. Recall that a partial order $(P; \leq)$ is a *wpo* if it has neither infinite descending chains nor infinite antichains. By an ω-*alternating chain* for a set $K \subseteq P$ we mean an ω-sequence $x_0, x_1, \ldots$ of elements of P such that $x_0 \leq x_1 \leq \cdots$ and $x_i \in K \leftrightarrow x_{i+1} \notin K$ for every $i < \omega$.

Proposition 4.7. *Let $P = (P; \leq)$ be a wpo and $\mathcal{C}$ the base of upper sets in P. For all $K \subseteq P$, $K \in \mathrm{BC}(\mathcal{C})$ iff K does not have ω-alternating chains.*

Proof. From left to right, the assertion follows from the last proposition and the equality $\mathrm{BC}(\mathcal{C}) = \bigcup_k \mathcal{C}(k)$. It remains to show that for any $K \in P \setminus BC(\mathcal{C})$ there is an ω-alternating chain. By the last proposition, there are alternating chains for K of arbitrary finite length.

Let ω^* be the set of all finite sequences of natural numbers, including the empty sequence ε. Let $P \cup \{\bot\}$ be the partial order obtained by adjoining a new smallest element $\bot$ to P. We construct a partial function $u : \omega^* \to P \cup \{\bot\}$ as follows. Set $u(\varepsilon) = \bot$ and suppose, by induction on $|\sigma|, \sigma \in \omega^*$, that $u(\sigma)$ is already defined.

If $|\sigma|$ is even then find $m \in \omega$ and the elements $v_0, \ldots, v_{m-1} \in P$ (if any) which enumerate without repetitions the $\leq$-minimal elements in $X = \{v \in K \mid u(\sigma) \leq v\}$ (since P is a wpo, the set X is finite). Then we set $u(\sigma i) = v_i$ for $i < m$ and $u(\sigma i)$ is undefined for $i \geq m$. For $|\sigma|$ odd, the definition is similar but this time we use the set $X = \{v \in P \setminus K \mid u(\sigma) \leq v\}$.

From the construction it follows that $\{\sigma \in \omega^* \mid u(\sigma)$ is defined$\}$ is an infinite finitely branching tree (under the relation of being an initial segment). By König lemma, there is an infinite path through this tree. The image of this path under u provides a desired infinite alternating chain for K. $\qquad\square$

5. DIFFERENCE HIERARCHIES OVER $\Sigma_n^{\tau_{\mathrm{d}}}$ AND Σ_n^{τ}

Here we establish the non-collapse property of the DH's over the bases $\Sigma_k^{\tau_{\mathrm{d}}}$ and Σ_k^{τ}, for each $k \geq 1$. In [32,33] it was shown that the DH's over the bases Σ_k^{ρ} and Σ_k^{σ}, $k \geq 1$, do not collapse. The methods of those papers (using the Ehrenfeucht-Fraïssé games) apply in a straightforward way to obtain the following.

Theorem 5.1. *For all $n, k > 0$ it holds $\Sigma_k^{\rho}(n) \not\subseteq \mathrm{co}\text{-}\Sigma_k^{\tau}(n)$. In particular, $\Sigma_k^{\tau_{\mathrm{d}}}(n) \not\subseteq \mathrm{co}\text{-}\Sigma_k^{\tau_{\mathrm{d}}}(n)$ and $\Sigma_k^{\tau}(n) \not\subseteq \mathrm{co}\text{-}\Sigma_k^{\tau}(n)$ for all $k, n, d > 0$.*

Before proving the theorem, first we note that from Corollary 3.2 it follows that $\Sigma_k^{\tau}(n) = \bigcup_d \Sigma_k^{\tau_{\mathrm{d}}}(n)$, for all $k, n \geq 1$. Next we establish some additional facts on the preorders $\leqslant_{k,r}$ from Section 3. Recall that those preorders actually also depend on a modulo d fixed in advance.

Lemma 5.2. *For all $k, n \geq 1$ and $L \subseteq A^+$, $L \in \Sigma_k^{\tau_{\mathrm{d}}}(n)$ iff there is an $r \geq 1$ such that L does not have 1-alternating chains of length n in $(A^+; \leqslant_{k,r})$.*

Proof. Let $L \in \Sigma_k^{\tau_{\mathrm{d}}}(n)$, so $L = \bigcup_i (L_{2i} \setminus L_{2i+1})$ for some sets $L_0 \supseteq L_1 \supseteq \cdots, L_n = \emptyset$ from $\Sigma_k^{\tau_{\mathrm{d}}}$. By Lemma 3.5, there is an $r \geq 1$ such that all $L_0.L_1, \ldots$ are upper sets in $(A^+; \leqslant_{k,r})$. By Proposition 4.6, L does not have 1-alternating chains of length n in $(A^+; \leqslant_{k,r})$. The opposite implication is checked in a similar fashion. $\qquad\square$

Using the operations F, G from Section 3, let us define an operation $(u_0, \ldots, u_n) \mapsto (\tilde{u}_0, \ldots, \tilde{u}_n)$ on tuples of words in A^+ as follows:

$$\tilde{u}_0 = F^n(u_n^0), \ \tilde{u}_1 = F^{n-1}(u_{n-1}^1), \ldots, \ \tilde{u}_n = u_0^n$$

where $u_i^0 = u_i$ for all $i \leq n$, $u_i^1 = G(u_i^0, u_n^0)$ for all $i \leq n-1$, $u_i^2 = G(u_i^1, u_{n-1}^1)$ for all $i \leq n-2$ and so on. Note that in this definition the upper indices on words are just indices, not the powers of words, while the upper indices on F mean composition.

Lemma 5.3. *Let $u_0 \leqslant_{k,r} \cdots \leqslant_{k,r} u_n$ and $|u_i| \equiv 0 \ (mod \ d)$ for all $i \leq n$. Then $\tilde{u}_0 \leqslant_{k+1,r} \cdots \leqslant_{k+1,r} \tilde{u}_n$.*

Proof. We have $u_0^0 \leqslant_{k,r} \cdots \leqslant_{k,r} u_n^0$. By definition of the operation G and by Lemma 3.7, $u_0^1 \leqslant_{k,r} \cdots \leqslant_{k,r} u_{n-1}^1$. Iterating this argument we obtain that $u_0^i \leqslant_{k,r} \cdots \leqslant_{k,r} u_{n-i}^i$ for all $i \leq n$. By Lemma 3.9,

$$F(u_{n-i}^i) \leqslant_{k+1,r} G(u_{n-i-1}^i, u_{n-i}^i) = u_{n-i-1}^{i+1}$$

for all $i < n$. By Lemma 3.8,

$$\tilde{u}_i = F^{n-i}(u_{n-i}^i) \leqslant_{k+1,r} F^{n-i-1}(u_{n-i-1}^{i+1}) = \tilde{u}_{i+1}$$

for all $i < n$, as desired. $\square$

Proof of Theorem 5.1. As in the proof of Theorem 3.3, for simplicity of notation we prove first the following alphabet-dependent version: for all $k, n \geq 1$ there is a language L_k^n over the alphabet $A_k = \{0, 1, \ldots, k\}$ with $L_k^n \in \Sigma_k^\rho(n) \setminus \text{co-}\Sigma_k^\tau(n)$.

We define the sets H_k^n by induction on k as follows:

$$H_1^n = (A_1^* \cdot 1)^n \cdot A_1^* \text{ and } H_{k+1}^n = A_{k+1}^* \cdot (k+1) \cdot (A_k^* \setminus H_k^n) \cdot (k+1) \cdot A_{k+1}^* \text{ for } k > 1.$$

In particular, H_1^n is the set of words over the binary alphabet $\{0, 1\}$ having at least n entries of 1, and H_2^n is the set of words over $\{0, 1, 2\}$ having a factor $2u2$ where u is a binary word with less than n entries of 1.

Obviously, $H_k^n \in \Sigma_k^\rho$ for each $k, n \geq 1$, $H_k^1 \supseteq H_k^2 \supseteq \cdots$ for any odd k, and $H_k^1 \subseteq H_k^2 \subseteq \cdots$ for any even k. Let L_k^n be the standard Boolean combination of $H_k^1, \ldots, H_k^n$ from the definition of the nth level of the DH, in particular,

$$L_1^4 = (H_1^1 \setminus H_1^2) \cup (H_1^3 \setminus H_1^4) \text{ and } L_2^4 = (H_1^2 \setminus H_2^1) \cup (H_2^4 \setminus H_2^3).$$

Then we have $L_k^n \in \Sigma_k^\rho(n)$ for each $k \geq 1$ and it remains to show that $L_k^n \notin \text{co-}\Sigma_k^\tau(n)$. Since $\Sigma_k^\tau(n) = \bigcup_d \Sigma_k^{\tau_d}(n)$, it suffices to show that $L_k^n \notin \text{co-}\Sigma_k^{\tau_d}(n)$ for each $d \geq 1$. By Lemma 5.2, it suffices to find for any $r \geq 1$ a 0-alternating chain $(u_{k0}, \ldots, u_{kn})$ for L_k^n in $(A^+; \leqslant_{k,r})$.

The chains are defined by induction on k as follows: set

$$u_{10} = 0^d, \ u_{1i} = 0^d(10^{2d-1})^i \ \text{for } i \in \{1, \ldots, n\},$$

and $(u_{k+1,0}, \ldots, u_{k+1,n}) = (\tilde{u}_{k0}, \ldots, \tilde{u}_{kn})$ for $k \geq 1$. Obviously, $|u_{ki}| \equiv 0 \pmod{d}$ for all $k \geq 1$ and $i \leq n$. Since it is easy to find a winning strategy for player 2 in the game $G_{1,r}(u_{1,i}, u_{1,i+1})$ for each $i < n$, we have $u_{10} \leqslant_{1,r} \cdots \leqslant_{1,r} u_{1n}$. By definition of H_1^n and L_1^n, $u_{1i} \in L_1^n$ iff i is even, for all $i \leq n$. Thus, $(u_{10}, \ldots, u_{1n})$ is a 0-alternating chain for L_1^n in $(A^+; \leqslant_{1,r})$. Assume by induction that $(u_{k0}, \ldots, u_{kn})$, $k \geq 1$, is a 0-alternating chain for L_k^n in $(A^+; \leqslant_{k,r})$. By Lemma 5.3, $u_{k+1,0} \leqslant_{k+1,r} \cdots \leqslant_{k+1,r} u_{k+1,n}$. By definition of H_{k+1}^n and L_{k+1}^n, $u_{k+1,i} \in L_{k+1}^n$ iff i is even, for all $i \leq n$. Thus, $(u_{k+1,0}, \ldots, u_{k+1,n})$ is a 0-alternating chain for L_{k+1}^n in $(A^+; \leqslant_{k+1,r})$. This completes the proof of the alphabet-dependent version.

For the alphabet-independent version, a slight modification of the proof above suffices. Namely, for any fixed alphabet A with $0, 1 \in A$ we consider instead of H_k^n above the sets M_k^n defined by

$$M_1^n = (A^* \cdot 1)^n \cdot A^* \text{ and } M_{k+1}^n = A^* \cdot 1^{2k} \cdot (A^* \setminus M_k^n) \cdot 1^{2k} \cdot A^* \text{ for } k \geq 1.$$

A slight modification of the argument above shows that the corresponding sets L_k^n have the desired property. $\qquad \square$

6. Difference hierarchies within $\mathrm{BC}(\Sigma_1^{\tau_d})$

In this section we consider some useful "local" DH's within $\mathrm{BC}(\Sigma_1^{\tau_d})$ and generalize the corresponding theory for signature σ developed independently in [26] and [14]. To this end, we consider some relevant partial orders on words.

Recall that in [14,26] for any $k < \omega$ the following partial order on A^+ was studied: $u \leq_k v$, if $u = v \in A^{\leq k}$ or $u, v \in A^{>k}$, $p_k(u) = p_k(v)$, $s_k(u) = s_k(v)$, and there is a k-embedding $f : u \to v$. Here $p_k(u)$ $(s_k(u))$ is the prefix (respectively, suffix) of u of length k, and the k-embedding f is a monotone injective function from $\{0 \ldots, |u|-1\}$ to $\{0 \ldots, |v|-1\}$ such that $u(i) \cdots u(i+k) = v(f(i)) \cdots v(f(i)+k)$ for all $i < |u| - k$. Note that the relation $\leq_0$ is just the inclusion of words. We say that a k-embedding f is a (k, d)-embedding, if $P_d^r(i)$ implies $P_d^r(f(i))$ for all $i < |u|$ and $r < d$.

Definition 6.1. For all $u, v \in A^+$, $k < \omega$ and $d > 0$, let $u \leq_k^d v$ denote that $u = v \in A^{\leq k}$ or $u, v \in A^{>k}$, $|u| \equiv |v| \pmod{d}$, $p_k(u) = p_k(v)$, $s_k(u) = s_k(v)$, and there is a (k, d)-embedding $f : u \to v$.

Remarks.

1. Note that for $d = 1$ the order $\leq_k^d$ coincides with $\leq_k$, hence our results about $\leq_k^d$ will subsume the corresponding results about $\leq_k$ in [14,26].

2. In the definition of $\leq_k^d$ we could omit the requirement $|u| \equiv |v| \pmod{d}$. The resulting theory would be essentially the same.

Proposition 6.2.
 (1) $(A^+; \leq_k^d)$ *is a well partial order.*
 (2) *If $k \leq l$ and d divides e then $u \leq_l^e v$ implies $u \leq_k^d v$.*

Proof.

(1) We may consider a word $u = u_0 \cdots u_n$ of length $n + 1$ over the alphabet A as a word $g_d(u) = (u_0, r_0) \cdots (u_n, r_n)$ over the alphabet $B = A \times \{0, \ldots, d-1\}$, where r_i (for any $i \leq n$) is the unique number $r < d$ satisfying $P_d^r(i)$. Then g_d is an isomorphic embedding of $(A^+; \leq_k^d)$ into $(B^+; \leq_k)$. Since $(B^+; \leq_k)$ is a wpo [14,26], so is also $(A^+; \leq_k^d)$.

(2) It suffices to show that $u \leq_{k+1}^d v$ implies $u \leq_k^d v$ and $u \leq_k^e v$ implies $u \leq_k^d v$. Let $u \leq_{k+1}^d v$. Then $g_d(u) \leq_{k+1} g_d(v)$, then $g_d(u) \leq_k g_d(v)$ [14,26], hence $u \leq_k^d v$. Now let $u \leq_k^e v$. Since $|u| \equiv |v| \pmod{e}$ implies $|u| \equiv |v| \pmod{d}$ and, as one easily checks, every (k, e)-embedding is also a (k, d)-embedding, we obtain $u \leq_k^d v$. $\qquad \square$

Let $\mathcal{C}_k^d$ be the class of all upper sets in $(A^+; \leq_k^d)$. From Propositions 6.2, 4.6 and 4.7 we immediately obtain the following information about the DH over the base $\mathcal{C}_k^d$.

Theorem 6.3.
  (1) *For all $L \subseteq A^+$ and $n < \omega$, $L \in \mathcal{C}_k^d(n)$ iff L has no 1-alternating chain with respect to $\leq_k^d$ of length n.*
  (2) *For any $L \subseteq A^+$, $L \in \mathrm{BC}(\mathcal{C}_k^d)$ iff L has no ω-alternating chain with respect to $\leq_k^d$.*

The next result generalizes the corresponding fact in [14,26].

Theorem 6.4. *The classes of languages $\mathcal{C}_k^d(n)$ are decidable for all $d > 0$ and $k, n < \omega$ (i.e., given a deterministic finite automaton (dfa) $\mathcal{M}$, it is decidable to check whether $L = L(\mathcal{M})$ is in $\mathcal{C}_k^d(n)$).*

Proof. Our proof here is different from those given in [14,26] for the particular case $d = 1$. Simplifying notation, we consider only a typical particular case $n = 2$, then $\mathcal{C}_k^d(2)$ is the class of differences of sets in $\mathcal{C}_k^d$. It suffices to show that both the relation "$L(\mathcal{M}) \in \mathcal{C}_k^d(2)$" and its negation are computably enumerable (c.e.). By Proposition 6.2(1), $\mathcal{C}_k^d(2)$ coincides with the class of sets $\check{F} = \{y \mid \exists x \in F(x \leq_k^d y)\}$, for finite $F \subseteq A^+$. Therefore, $L(\mathcal{M})$ is in $\mathcal{C}_k^d(2)$ iff there are finite sets $F, G \subseteq A^+$ such that $\mathcal{C}_k^d(2) = \check{F} \setminus \check{G}$. The relation "$\mathcal{C}_k^d(2) = \check{F} \setminus \check{G}$" is computable because one can compute τ_d-sentences ϕ, ψ such that $\check{F} = L_\phi$ and $\check{G} = L_\psi$ (as well as dfa's $\mathcal{A}$ and $\mathcal{B}$ with $L(\mathcal{A}) = \check{F}$) and $L(\mathcal{B}) = \check{G}$). Thus, the relation "$L(\mathcal{M}) \in \mathcal{C}_k^d(2)$" is c.e.

By Theorem 6.3, $L(\mathcal{M}) \notin \mathcal{C}_k^d(2)$ iff there is an 1-alternating chain for $L(\mathcal{M})$ of length 2, iff there are words $x_0 \leq_k^d x_1 \leq_k^d x_2$ such that $x_0, x_2 \in L$ and $x_1 \notin L$. Since the relation "$x \leq_k^d y$" is computable, the relation "$L(\mathcal{M}) \notin \mathcal{C}_k^d(2)$" is c.e. Therefore, the relation "$L(\mathcal{M}) \in \mathcal{C}_k^d(2)$" is computable. $\square$

We conclude this section by a result on the separation property.

Theorem 6.5. *For all $k < \omega$ and $d > 0$, the class $\mathcal{C}_k^d$ has the separation property.*

Proof. Let L, K be disjoint sets in $\mathcal{C}_k^d$. By Proposition 6.2(1), any of L, K is a finite union of "cones" of the form $\{w \in A^+ \mid u \leq_k^d w\}$ (as is well-known and easy to see, a poset P is a wpo iff any upper set in P is a finite union of cones). By Proposition 4.3, w.l.o.g. we may assume that both L and K are such cones, *i.e.* $L = \{w \in A^+ \mid u \leq_k^d w\}$ and $K = \{w \in A^+ \mid v \leq_k^d w\}$ for some $u, v \in A^+$. If $|u| \leq k$ then the set $L = \{u\}$ is in $\mathcal{C}_k^d \cap \mathrm{co\text{-}}\mathcal{C}_k^d$ and separates L from K. The case $|v| \leq k$ is similar hence let us assume that $u, v \in A^{>k}$.

If $p_k(u) \neq p_k(v)$ then the set $\{w \in A^+ \mid p_k(w) = p_k(u)\}$ is in $\mathcal{C}_k^d \cap \mathrm{co\text{-}}\mathcal{C}_k^d$ and separates L from K. The case $s_k(u) \neq s_k(v)$ is considered similarly. If $|u| \not\equiv |v|$ (mod d) then the set $\{w \in A^+ : |w| \equiv |u| \pmod{d}\}$ is in $\mathcal{C}_k^d \cap \mathrm{co\text{-}}\mathcal{C}_k^d$ and separates L from K.

We claim that the remaining case (when $|u|, |v| > k$, $p_k(u) = p_k(v)$, $s_k(u) = s_k(v)$, and $|u| \equiv |v| \pmod{d}$) is not realized because in this case it holds $L \cap K \neq \emptyset$. To see this, let $(u_0, \ldots, u_l)$, $l > 0$, be the sequence of all $(k+1)$-factors of u, *i.e.*

$u_i = u(i) \cdots u(i+k)$ for every $i \leq l$. Let $(v_0, \ldots, v_m)$ be the similar sequence for v. It suffices to show that for suitable words $x_0, \ldots, x_l, y_0, \ldots, y_m \in A^*$ the word $w = u_0 x_0 \cdots u_l x_l v_0 y_0 \cdots v_m y_m v_m$ is in $L \cap K$. Indeed, we only have to care that the lengths of the words $x_0, \ldots, x_l, y_0, \ldots, y_m$ satisfy the following equivalences modulo d:

$$|u_0 x_0| \equiv 0, |u_0 x_0 u_1 x_1| \equiv 1, \ldots, |u_0 x_0 \cdots u_{l-1} x_{l-1}| \equiv l - 1$$

(this guarantees the existence of a (k, d)-embedding of u into w),

$$|u_0 x_0 \cdots u_l x_l| \equiv d - 1, |u_0 x_0 \cdots u_l x_l v_0 y_0| \equiv 0, \ldots, |u_0 x_0$$
$$\cdots u_l x_l v_0 y_0 \cdots v_{m-1} y_{m-1}| \equiv m - 1$$

(this guarantees the existence of a (k, d)-embedding of v into w), and $|w| \equiv |u|$ (this together with the conditions above guarantees $u \leq_k^d w$ and $v \leq_k^d w$). $\square$

From the last proof we easily extract the following characterization of the class $\mathcal{C}_k^d \cap \text{co-}\mathcal{C}_k^d$ in terms of regular expressions.

Corollary 6.6. *The class $\mathcal{C}_k^d \cap \text{co-}\mathcal{C}_k^d$ consists exactly of the finite unions of languages $w_0, w_0(A^d)^* w_1$, where w_0, w_1 are nonempty words of length $\leq k$.*

7. Difference hierarchies over $\Sigma_1^{\tau_d}$ and Σ_1^τ

In this section we provide additional information on the DH's over $\Sigma_1^{\tau_d}$ and Σ_1^τ. First we relate these DH's to those from the preceding section.

Theorem 7.1. *It holds $\Sigma_1^{\tau_d} = \bigcup_k \mathcal{C}_k^d$ and $\Sigma_1^\tau = \bigcup_{k,d} \mathcal{C}_k^d$.*

Proof. The second equality follows from the first one and Corollary 3.2. Towards the first equality, let first $L \in \mathcal{C}_k^d$ for some k, hence L is a finite union of the cones $\{w \in A^+ \mid u \leq_k^d w\}$. It is easy to write down a Σ_1-sentence ϕ of signature τ_d such that $\{w \in A^+ \mid u \leq_k^d w\} = L_\phi$. Therefore, $L \in \Sigma_1^{\tau_d}$.

Conversely, let $L \in \Sigma_1^{\tau_d}$. By Corollary 9.6 (sorry for this reference forwards; though definitely non-elegant, it does not in fact cause any principal problem because there is no logical circle), L is a finite union of sets $w_0(A^d)^* w_1 \cdots (A^d)^* w_n$ where $n \geq 0$ and $w_i \in A^+$. Easy manipulations show that w.l.o.g. we may assume that all the words w_i in any fixed term $w_0(A^d)^* w_1 \cdots (A^d)^* w_n$ have one and the same length k. It is then straightforward to check that $w_0(A^d)^* w_1 \cdots (A^d)^* w_n \in \mathcal{C}_k^d$. Therefore, $L \in \mathcal{C}_k^d$. $\square$

From the proof of Theorem 7.1 and from Corollary 3.2 we obtain the following characterization of $\Delta_1^{\tau_d}$ and Δ_1^τ in terms of regular expressions.

Corollary 7.2.
(1) *For any $d > 0$, the class $\Delta_1^{\tau_d}$ consists exactly of the finite unions of languages $w_0, w_0(A^d)^* w_1$, where $w_0, w_1 \in A^+$.*

(2) *The class Δ_1^τ consists exactly of the finite unions of languages $w_0, w_0(A^d)^* w_1$, where $w_0, w_1 \in A^+$ and $d > 0$.*

The following assertion is an immediate consequence of Theorem 6.5 and Proposition 4.5.

Theorem 7.3. *The classes $\Sigma_1^{\tau_d}$ ($d > 0$) and Σ_1^τ have the separation property.*

Next we generalize some results from [14] about the DH over Σ_1^σ to the DH over $\Sigma_1^{\tau_d}$, for each $d > 0$. We start with recalling (in slightly different notation) some notions from [14].

Let $\mathcal{A} = (Q, A, \delta, i, F)$ be a deterministic finite automaton (dfa) over A and $c_\mathcal{A} = (n+1)^{(n+1)^{n+1}}$, where n is the number of states in $\mathcal{A}$. A word $u \in A^+$ is called $\mathcal{A}$-*idempotent*, if $\delta^{uu} = \delta^u$, where δ^u is the function on the states of $\mathcal{A}$ defined by $\delta^u(s) = \delta(s, u)$. Let $B = B_\mathcal{A}$ be the new (large) alphabet the letters of which being the nonempty words over A of length $\leq c_\mathcal{A}$. Let $U = U_\mathcal{A}$ be the subset of B consisting of all $\mathcal{A}$-idempotents. By *structured word* over $\mathcal{A}$ we mean a word over B of the form $w_0 u_1 w_1 \cdots u_n w_n$, where $n \geq 0$ and all u_i are $\mathcal{A}$-idempotents. Following [14], we denote the structured words as $w_0 \boxed{u_1} w_1 \boxed{u_2} w_2 \cdots \boxed{u_m} w_m$ emphasizing a central role of the idempotents in subsequent considerations. Let $S = S_\mathcal{A}$ be the set of all structured words over B. For any l, let $f_l(w_0 \boxed{u_1} w_1 \boxed{u_2} w_2 \cdots \boxed{u_m} w_m) = w_0 u_1^l x_1 \cdots u_n^l w_n$. We shorten f_1 to f. In [14] it was shown that $f : S_\mathcal{A} \to A^+$ is a surjection.

For $x, y \in S_\mathcal{A}$ we write $x \leq_\mathcal{A} y$ if there exist words $u_i \in A^+$ and $x_i, z_i \in S_\mathcal{A}$ such that

$$
\begin{aligned}
x &= x_0 \boxed{u_1} \quad\quad x_1 \boxed{u_2} \quad\quad x_2 \cdots \boxed{u_m} \quad\quad x_m \quad \text{and}\\
y &= x_0 \boxed{u_1} z_1 \boxed{u_1} x_1 \boxed{u_2} z_2 \boxed{u_2} x_2 \cdots \boxed{u_m} z_m \boxed{u_m} x_m.
\end{aligned}
$$

This partial order was the key in understanding the DH over Σ_1^σ, namely we have the following characterization.

Theorem 7.4 [14]**.** *Let $\mathcal{A}$ be a dfa, $L = L(\mathcal{A}) \subseteq A^+$ and $n \geq 1$. Then $L \in \Sigma_1^\sigma(n)$ if and only if $f^{-1}(L)$ has no 1-alternating chain of length n in $(S_\mathcal{A}; \leq_\mathcal{A})$.*

The cited results from [14] are generalized to the DH over $\Sigma_1^{\tau_d}$ for each $d > 0$ as follows. Let $B^d = B_\mathcal{A}^d$ be the set of words $v \in A^+$ such that $|v| \equiv 0 \pmod{d}$ and $|v| \leq c_\mathcal{A} \cdot d$. Let $U^d = U_\mathcal{A}^d$ be the subset of B^d consisting of all $\mathcal{A}$-idempotents. By d-*structured word* over $\mathcal{A}$ we mean a structured word $w_0 \boxed{u_1} w_1 \boxed{u_2} w_2 \cdots \boxed{u_m} w_m$ such that $w_0, \ldots, w_{m-1} \in B^d$, $u_1, \ldots, u_n \in U^d$, and $w_m \in A^+$, $|w_m| \leq c_\mathcal{A} \cdot d$ (thus, w_n is not necessarily in B^d). Let $S_\mathcal{A}^d$ denote the set of d-structured words over $\mathcal{A}$. The function f_l is defined exactly as above. For $x, y \in S_\mathcal{A}^d$, let $x \leq_\mathcal{A}^d y$ mean that $x \leq_\mathcal{A} y$ and $|f(x)| \equiv |f(y)| \pmod{d}$. We then have the following generalization of some facts in [14].

Proposition 7.5. *Let $\mathcal{A}$ be a dfa and $d > 0$.*
   (1) *The map $f : S_\mathcal{A}^d \to A^+$ is a surjection.*
   (2) *$(S_\mathcal{A}^d; \leq_\mathcal{A}^d)$ is a well partial order.*

(3) If M is an upper set in $(S_{\mathcal{A}}^d; \leq_{\mathcal{A}}^d)$ then $f(M) \in \mathcal{C}^d$.

(4) If $L = L(\mathcal{A}) \in \Sigma_1^{\tau_d}$ then $f^{-1}(L)$ is an upper set in $(S_{\mathcal{A}}^d; \leq_{\mathcal{A}}^d)$.

Proof.

(1) Let $v \in A^+$. Represent v in the form $v = v_0 \cdots v_n$ where $0 < |v_n| \leq d$ and $|v_i| = d$ for all $i < n$. Applying the loop Lemma 6 from [14] (see also a related Th. 2.1 for finite semigroups in [19] attributed to Ramsey) to this factorization of v, we obtain a coarser factorization $v = w_0 u_1 x_1 \cdots u_n w_m$ for some $m \geq 0$ such that $w_0 \boxed{u_1} w_1 \cdots \boxed{u_m} w_m \in S_{\mathcal{A}}^d$, as desired.

(2) is checked exactly as Theorem 7 in [14].

(3) is checked exactly as Theorem 16 (2) in [14].

(4) Let $x \in f^{-1}(L)$, *i.e.* $f(x) \in L$, and $x \leq_{\mathcal{A}} y$. We have to show that $y \in f^{-1}(L)$, *i.e.* $f(y) \in L$. By Theorem 7.1, $L \in \mathcal{C}_k^d$ for some k. It is easy to see that $f_{k+1}(x) \leq_k^d f_{k+1}(y)$. Since $f_{k+1}(x)$ is obtained from $f(x)$ by insertion of idempotents, $f_{k+1}(x) \in L$. Since L is an upper set in $(A^+; \leq_k^d)$, $f_{k+1}(y) \in L$ and hence also $f(y) \in L$. $\qquad\square$

The next generalization of Theorem 7.4 follows immediately from Propositions 7.5 and 4.1.

Theorem 7.6. *Let L be the regular language recognized by a dfa $\mathcal{A}$ and let $n, d > 0$. Then $L \in \Sigma_1^{\tau_d}(n)$ iff $f^{-1}(L)$ has no 1-alternating chain of length n in $(S_{\mathcal{A}}^d; \leq_{\mathcal{A}}^d)$.*

Using the results above we obtain the following decidability result also generalizing the corresponding fact from [14].

Theorem 7.7. *The classes of languages $\Sigma_1^{\tau_d}(n)$ are decidable for all $n, d > 0$ (i.e., given a dfa $\mathcal{A}$, it is decidable to check whether $L = L(\mathcal{A})$ is in $\Sigma_1^{\tau_d}(n)$).*

Proof. One way to show this is to observe that a version of the corresponding argument from Theorem [14] works also for our more general situation. It is also possible to give an alternative shorter proof as follows. It suffices to show that both the relation "$L(\mathcal{A}) \in \Sigma_1^{\tau_d}(n)$" and its negation are c.e. By definition, $L(\mathcal{A})$ is in $\Sigma_1^{\tau_d}(n)$ iff there is a sentence ϕ of signature τ_d such that ϕ is a certain Boolean combination of n existential sentences and $L(\mathcal{A}) = L_\phi$. Since the last relation is computable, the relation "$L(\mathcal{A}) \in \Sigma_1^{\tau_d}(n)$" is c.e. By Theorem 7.6, $L(\mathcal{A}) \notin \Sigma_1^{\tau_d}(n)$ iff there is an 1-alternating chain for $f^{-1}(L)$ of length n in $(S_{\mathcal{A}}^d; \leq_{\mathcal{A}}^d)$ *i.e.* there are $x_0 \leq_{\mathcal{A}}^d \cdots \leq_{\mathcal{A}}^d x_n$ such that $f(x_{2i}) \in L$ and $f(x_{2i+1}) \notin L$. Since f is a computable function and $(S_{\mathcal{A}}^d; \leq_{\mathcal{A}}^d)$ is a computable structure, the relation "$L(\mathcal{M}) \notin \Sigma_1^{\tau_d}(n)$" is c.e. Thus, the relation "$L(\mathcal{M}) \in \Sigma_1^{\tau_d}(n)$" is computable. $\qquad\square$

Remarks.

1. We do not currently know whether the classes of languages $\mathrm{BC}(\Sigma_1^{\tau_d})$, $\Sigma_1^\tau(n)$ (for $n > 1$) and $\mathrm{BC}(\Sigma_1^\tau)$ are decidable. The decidability of Σ_1^τ follows from [12] and a result in the next section.

2. In this paper we discuss only the finite levels of the DH's. As is well-known, in descriptive set theory and computability theory also natural transfinite versions of the DH's (namely the Hausdorff and Ershov hierarchies) are quite useful.

Currently we do not know a useful transfinite version of the DH in the context of automata theory, in particular a transfinite DH over $\Sigma_1^{\tau d}$ that exhausts $\Delta_2^{\tau d}$.

8. PLT-REDUCIBILITY

In this section we start our investigation of reducibilities on regular sets by discussing a reducibility closely related to complexity theory. It was introduced in [3] and independently in [43].

Definition 8.1.
- A language $L \subseteq A^*$ is *polylogtime reducible* to $K \subseteq B^*$, for short $L \leq_{\mathrm{plt}} K$, if there exist functions $f : A^* \times \mathbb{N} \to B$ and $g : A^* \to \mathbb{N}$, computable in polylogarithmic time (on a deterministic Turing machine which treats the input word as an oracle) such that $x \in L \leftrightarrow f(x,1)f(x,2)\ldots f(x,g(x)) \in K$ for every $x \in A^*$.
- By *plt-function* we mean any function of the form $x \mapsto f(x,1)f(x,2)\ldots f(x,g(x))$ where f and g are computable in polylogarithmic time.
- A language L is *plt-decidable*, if its characteristic function is computable in polylogarithmic time.

Examples.
1. The function $u \mapsto pu$ on A^+ which adds a fixed prefix $p \in A^*$ to a word u is a plt-function. The same of course applies to the operation of adding a suffix to a word.

2. Let $h : A^+ \to B^+$ be a semigroup morphism. Such functions are defined by their values $h : A \to B^+$ on the letters of A (*i.e.*, words of length 1) because we have $h(a_0 \cdots a_l) = h(a_0) \cdots h(a_l)$, where $a_i \in A$. It is easy to see that if such a function h is length-multiplying (*i.e.*, it satisfies the property $\forall a, b \in A(|h(a)| = |h(b)|)$) then it is a plt-function.

To explain the relationship of plt-reducibility to complexity theory, let us recall some relevant definitions. Consider a polynomial-time nondeterministic Turing machine M working on an input word x over some alphabet B and printing a letter from another alphabet A after finishing any computation path. We assume w.l.o.g. that the machine M, in every step, splits a computation path into at most two computation paths (hence, a computation path of M on input x can be described by a word from $\{0,1\}^*$). Those printed letters from A are the leaves of the binary tree defined by the nondeterministic choices of M on input x. An ordering of the tuples in the program of M determines a left-to-right ordering of all the leaves. In this way, M may be considered as a deterministic transducer computing a total function $M : B^* \to A^+$. A machine M as above is *balanced* if for every input x there exists an $m \geq 0$ and an $r \in \{0,1\}^m$ such that

$$\{s \mid s \in \{0,1\}^m \wedge s \leq r\} \cdot \{0,1\} \cup \{s \mid s \in \{0,1\}^m \wedge s > r\}$$

is the set of all computation paths of M on x.

Now, relate to any language $L \subseteq A^+$ (called in this situation a *leaf language*) the language $M^{-1}(L) \subseteq B^*$. Denote by $\mathrm{Leaf_b}(L)$ the set of languages $M^{-1}(L)$, for all balanced machines M, and denote by $\mathrm{Leaf_u}(L)$ the set of languages $M^{-1}(L)$, for all machines M (not necessarily balanced) specified above. Obviously, we have $\mathrm{Leaf_b}(L) \subseteq \mathrm{Leaf_u}(L)$ for every language L, and there exist languages L where $\mathrm{Leaf_b}(L) = \mathrm{Leaf_u}(L)$ is unlikely (*e.g.*, $\mathrm{Leaf_b}((11)^*) = \mathrm{P}$, $\mathrm{Leaf_u}((11)^*) = \oplus\mathrm{P}$ and it is widely believed that $\mathrm{P} \neq \oplus\mathrm{P}$, see [45]).

It turns out that many important complexity classes have natural and useful characterizations in terms of leaf languages (see [45] and references therein). The following theorem from [3,43] relates plt-reducibility to the balanced version of leaf language definability.

Theorem 8.2. *For all languages L and K, $L \leq_{\mathrm{plt}} K$ iff $\mathrm{Leaf_b}(L)^{\mathcal{O}} \subseteq \mathrm{Leaf_b}(K)^{\mathcal{O}}$ for every oracle $\mathcal{O}$.*

This result and the fact that the regular languages are most natural to use as leaf languages show that the investigation of plt-reducibility (especially on the regular sets) is important. For some results in this direction see [12,38,39,45]. In particular, there is a greatest element (*i.e.* a complete set) in $(\mathcal{R}; \leq_{\mathrm{plt}})$, and the class of complete sets coincides with the class of regular languages L such that the syntactic monoid $M(L)$ contains a non-solvable subgroup [17]. Here we establish some additional facts on the plt-reducibility.

We start with the following obvious fact in which we, for simplicity of notation, identify preorders with the corresponding quotient partial orders (*i.e.*, degree structures).

Proposition 8.3.
   (1) $\{\emptyset\}$ *and* $\{A^+\}$ *are two distinct minimal elements of the degree structures* $(P(A^+); \leq_{\mathrm{plt}})$ *and* $(\mathcal{R}; \leq_{\mathrm{plt}})$ *which are below any other element.*
   (2) *The structures* $(P(A^+); \leq_{\mathrm{plt}})$ *and* $(\mathcal{R}; \leq_{\mathrm{plt}})$ *are upper semilattices under the supremum operation* $L \oplus K = aL \cup (A \setminus a)K$ *where a is a fixed letter from A.*
   (3) *The structure* $(P'(A^+); \leq_{\mathrm{plt}})$ *is a distributive upper semilattice with the smallest degree consisting exactly of the non-trivial plt-decidable languages.*

Remark. We do not currently know whether the structure $(\mathcal{R}'; \leq_{\mathrm{plt}})$ is a distributive upper semilattice (*i.e.*, whether it satisfies $\forall x, y, z(x \leq y \cup z \rightarrow \exists y' \leq y, z' \leq z(x = y' \cup z')))$.

We need characterizations of some classes of regular languages in terms of the so called forbidden patterns in (the transition graphs of) automata recognizing them.

Definition 8.4. Let $\mathcal{A} = (Q, A, \delta, s_0, F)$ be a dfa.
   (1) A *balanced counting* pattern for $\mathcal{A}$ is formed by states $s_1, \ldots, s_n$ (for some $n \geq 2$) and words x, z, u, v such that: $\delta(s_0, x) = s_1$, $\delta(s_i, u) = s_i$ for all $i \in \{1, \ldots, n\}$, $\delta(s_i, v) = s_{i+1}$ for all $i \in \{1, \ldots, n-1\}$, $\delta(s_n, v) = s_1$, $\delta(s_1, y) \in F$, $\delta(s_2, y) \notin F$, and $|u| = |v|$.
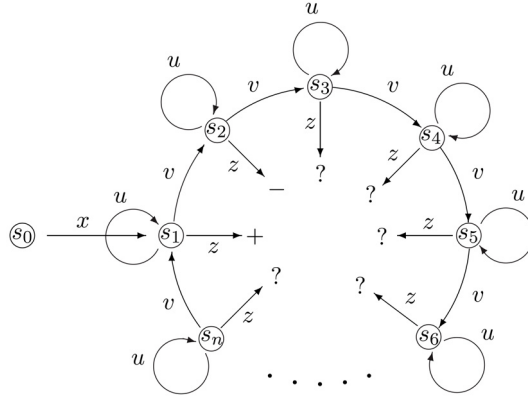
FIGURE 1. The balanced counting pattern with $|u| = |v|$.
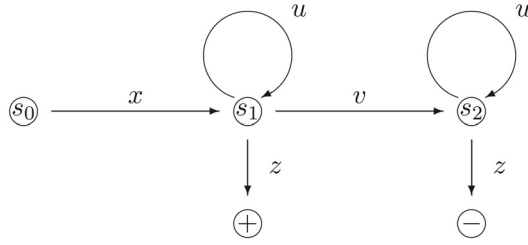

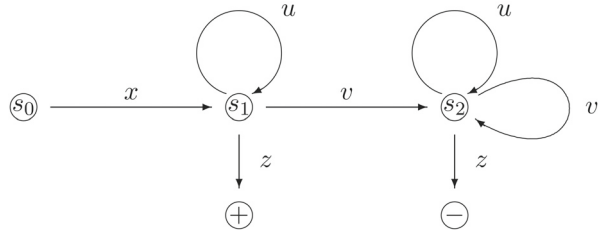FIGURE 2. The balanced coUP pattern with $|u| = |v|$.

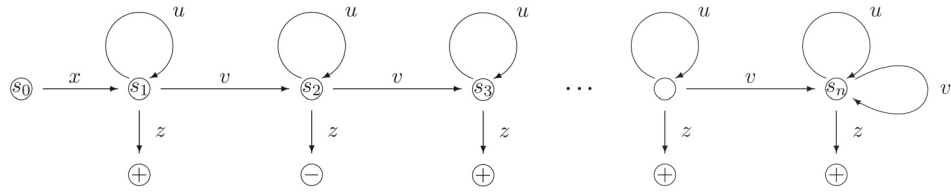FIGURE 3. The balanced coNP pattern with $|u| = |v|$.

FIGURE 4. The balanced co1NP pattern with $|u| = |v|$.

(2) A *balanced coUP*-pattern for $\mathcal{A}$ is formed by states s_1, s_2 and words x, z, u, v such that: $\delta(s_0, x) = s_1$, $\delta(s_i, u) = s_i$ for $i = 1, 2$, $\delta(s_1, v) = s_2$, $\delta(s_1, z) \in F$, $\delta(s_2, z) \notin F$, and $|u| = |v|$.

(3) A *balanced coNP*-pattern for $\mathcal{A}$ is formed by states s_1, s_2 and words x, z, u, v such that: $\delta(s_0, x) = s_1$, $\delta(s_i, u) = s_i$ for $i = 1, 2$, $\delta(s_1, v) = s_2 = \delta(s_2, v)$, $\delta(s_1, z) \in F$, $\delta(s_2, z) \notin F$, and $|u| = |v|$.

(4) A *balanced co1NP*-pattern for $\mathcal{A}$ is formed by states $s_1, s_2, \ldots, s_n$ (for some $n \geq 3$) and words x, z, u, v such that: $\delta(s_0, x) = s_1$, $\delta(s_i, u) = s_i$ for all $i \in \{1, \ldots, n\}$, $\delta(s_i, v) = s_{i+1}$ for all $i \in \{1, \ldots, n-1\}$, $\delta(s_n, v) = s_n$, $\delta(s_1, z), \delta(s_3, z) \ldots, \delta(s_n, z) \in F$, $\delta(s_2, z) \notin F$, and $|u| = |v|$.

The definition is illustrated by Figures 1–4.

Let us formulate a characterization of the relation $L \leq_{\mathrm{plt}} 0^*1(0 \cup 1)^*$ which is a "balanced" version of the corresponding result from [2,24] for the "unbalanced" model. It follows immediately from [12,13].

Theorem 8.5. *Let L be a regular language and $\mathcal{A}$ the minimal automaton recognizing L. The following conditions are equivalent:*

(1) $L \leq_{\mathrm{plt}} 0^*1(0 \cup 1)^*$;

(2) $\mathcal{A}$ *does not have balanced coUP-patterns;*

(3) $\mathcal{A}$ *does not have balanced coNP-, co1NP- and counting patterns;*

(4) L *is a finite union of languages of the form $w_0(A^d)^*w_1 \cdots (A^d)^*w_n$ where $d > 0, n \geq 0$ and $w_i \in A^*$.*

Proof. First note that our versions of the balanced patterns are slightly different from those in [12,13] where they are defined only for the minimal automata. The obvious relation between them is that a dfa $\mathcal{A}$ has a balanced pattern (of any of the four types) in our sense iff the equivalent minimal automaton has a balanced pattern (of the same type) in the sense of [12,13]. Thus the equivalence of (1), (2) and (4) follows from Theorem 8 in [13] while the equivalence of (2) and (3) follows from the proof of Corollary 11 in [13]. $\square$

Our next result states some relationships between the τ-hierarchy and the plt-reducibility.

Theorem 8.6.

(1) *For every $n > 0$ there exists $K \in \Sigma_n^{\sigma}$ such that $L \leq_{\mathrm{plt}} K$ for all $L \in \Sigma_n^{\tau}$.*

(2) *For any regular language L, $L \leq_{\mathrm{plt}} 0^*1(0 \cup 1)^*$ iff $L \in \Sigma_1^{\tau}$.*

Proof.

(1) Let $K = H_n$ be the language over the alphabet $A_n = \{0, 1, \ldots, n\}$ from the proof of Theorem 3.3. Then $H_n \in \Sigma_n^{\rho} \subseteq \Sigma_n^{\sigma}$ for each $n \geq 1$. Let $L \in \Sigma_n^{\tau}$ and $L \subseteq A^+$, then $L \in \Sigma_n^{\tau_d}$ for some $d > 0$. Let $f : A^+ \to A_n^+$ be the function constructed in the proof of Lemma 3.1 in [39] (only this time the formula ϕ from that proof is a quantifier-free formula of signature τ_d). One easily checks that f is a plt-function. Since $L = f^{-1}(H_n)$ by Lemma 3.1 in [39], $L \leq_{\mathrm{plt}} K$.

This proves (1) but only for the alphabet-dependent mode. To obtain the alphabet-independent proof, we slightly modify the encoding used in the proof of

Theorem 3 in [39] (and in [27]). It suffices to find a language K over $A_1 = \{0, 1\}$ with the desired property. For $n = 1$ we can take $K = H_1$. For $n > 1$, consider a length-multiplying injective homomorphism $h : A_n^* \to A_1^*$ such that for any $a \in A_n$ the binary word $h(a)$ is in $0A_1^*0$ and has no factor 00 (such an h obviously exists, *e.g.* for $n = 2$ we can set $h(0) = 011110$, $h(1) = 010110$ and $h(2) = 011010$).

It is easy (similarly to [27] where we used a slightly different encoding with similar properties, so the argument applies also to our encoding here) to relate to any sentence ϕ of signature σ_{A_n} a sentence ϕ' of signature σ_{A_1} such that the following conditions hold:

(i) for any $w \in A^+$, $\mathbf{w} \models \phi$ iff $\mathbf{h}(\mathbf{w}) \models \phi'$;

(ii) if ϕ is a Σ_n-sentence then so is also ϕ'.

Now let ϕ be a Σ_n-sentence of signature σ_{A_n} with $H_n = L_\phi$, and let $C = L_{\phi'}$. Then $C \subseteq A_1^+$, $C \in \Sigma_n^\sigma$ and $H_n = h^{-1}(C)$. But h is a plt-function (see Ex. 2 at the beginning of this section) hence $H_n \leq_{\mathrm{plt}} C$ and C has the desired properties.

(2) Let $L \in \Sigma_1^\tau$. By the proof of (1), $L \leq_{\mathrm{plt}} H_1 = 0^*1(0 \cup 1)^*$. Conversely, let $L \leq_{\mathrm{plt}} 0^*1(0 \cup 1)^*$. By Theorem 8.5, L is a finite union of languages of the form $w_0(A^d)^*w_1 \cdots (A^d)^*w_n$, where $n \geq 0, d > 0$ and $w_i \in A^+$. One easily writes down an existential sentence ϕ of signature τ_d such that $w_0(A^d)^*w_1 \cdots (A^d)^*w_n = L_\phi$. Therefore, $L \in \Sigma_1^\tau$. $\qquad\square$

Corollary 8.7.

(1) *For all $n, d > 0$, there is a plt-complete set in $\Sigma_n^{\tau_\mathrm{d}}$ and Σ_n^τ.*

(2) *The class Σ_1^τ is a principal ideal of $(\mathcal{R}; \leq_{\mathrm{plt}})$.*

(3) *Σ_1^τ coincides with the class of finite unions of languages of the form $w_0(A^d)^*w_1 \cdots (A^d)^*w_n$, where $d > 0, n \geq 0$ and $w_i \in A^+$.*

Proof. Follows from Theorems 8.6(2) and 8.5. $\qquad\square$

Remarks.

1. We do not know whether Σ_n^τ is a principal ideal of $(\mathcal{R}; \leq_{\mathrm{plt}})$ for $n > 1$.

2. A proof similar to the proof of Theorem 8.6(1) (with the use of slight modifications of the proofs from Sect. 3 of [39]) shows that for any alphabet every level of the difference hierarchy over Σ_n^τ has a plt-complete set.

3. By [5], the class of quasi-aperiodic languages coincides with the closure of the class of finite languages and the languages $(A^d)^*$, $d \geq 1$, under the Boolean operations and concatenation. We guess that all the classes $\Sigma_n^\tau, \mathrm{BC}(\Sigma_n^\tau)$ may be characterized in terms of regular expressions in the similar way to the well-known characterization of $\Sigma_n^\sigma, \mathrm{BC}(\Sigma_n^\sigma)$ in [40] and of of $\Sigma_n^\rho, \mathrm{BC}(\Sigma_n^\rho)$ in [23]. The item (3) in Corollary 8.7 provides such a characterization for $n = 1$.

Next we characterize the quasi-aperiodic languages in terms of forbidden patterns. Recall (see *e.g.* Th. V.1.1 in [35]) that the syntactic monoid $M(L)$ may be interpreted as an automaton recognizing L.

Theorem 8.8. *For every regular language L the following conditions are equivalent.*

(1) *L is FO_τ-axiomatizable.*

(2) *Every finite automaton recognizing L has no balanced counting pattern.*
(3) *The automaton $M(L)$ has no balanced counting pattern.*
(4) *The minimal automaton $m(L)$ of L has no balanced counting pattern.*

Proof.

   (1)→(2). Let L be FO$_\tau$-axiomatizable, then $L \in \Sigma_n^\tau$ for some $n > 0$. By Theorem 8.6(1), $L \leq_{\text{plt}} H$ for some $H \in \Sigma_n^\sigma$. By Theorem 8.2, $\text{Leaf}_b(L)^O \subseteq \text{Leaf}_b(H)^O$ for all oracles O. It is known [45] that $\text{Leaf}_b(H)^O \subseteq \text{PH}^O$ for all O, hence $\text{Leaf}_b(L)^O \subseteq \text{PH}^O$ for all O (PH is the class of sets in the polynomial-time hierarchy).

   Toward a contradiction, let $\mathcal{A} = (Q, A, \delta, s_0, F)$ be a dfa which recognizes L and has a balanced counting pattern for some $n > 1$ as in Figure 1, with distinct states $s_1, \ldots, s_n$. Let $K \subseteq \{0,1\}^+$ be the language of all binary words w with $\#_1(w) \equiv i \pmod{n}$ for some $i < n$ such that $\delta(s_i, z)$ is an accepting state of $\mathcal{A}$ (so $0 \in K$ and $1 \notin K$). One easily checks that the plt-function (see examples of the plt-functions at the beginning of this section) $f : \{0,1\}^+ \to A^+$ defined by $f(y) = xh(y)z$, where $h : \{0,1\}^+ \to A^+$ is the length-multiplying homomorphism satisfying $h(0) = u$ and $h(1) = v$, reduces K to L. By the proof of Lemma 6 from [1], w.l.o.g. we may assume that n is prime. Let $M_n \subseteq \{0,1\}^+$ be the set of words with $\#_1(w) \equiv 0 \pmod{n}$. Define a function g on $\{0,1\}^+$ as follows. Let $|g(w)| = |w|^{n-1}$ and for every $i \in \{1, \ldots, |w|^{n-1}\}$ the i-th letter in $g(w)$ is 1 iff $w(i_1) = \cdots = w(i_{n-1}) = 1$, where $(i_1, \ldots, i_{n-1})$ is the i-th tuple in the lexicographic ordering of $\{1, \ldots, |w|\}^{n-1}$. One easily checks that g is a plt-function and $\#_1(g(w)) = (\#_1(w))^{n-1}$. By Fermat's theorem, $\#_1(g(w)) \equiv 0 \pmod{n}$ if $\#_1(w) \equiv 0 \pmod{n}$ and $\#_1(g(w)) \equiv 1 \pmod{n}$ otherwise. Hence, g reduces M_n to K and therefore $M_n \leq_{\text{plt}} L$.

   By Theorem 8.2, $\text{Leaf}_b(M_n)^O \subseteq \text{Leaf}_b(L)^O \subseteq \text{PH}^O$ for all oracles O, which contradicts to a well-known fact of complexity theory (see [45]).

   (2)→(3) and (2)→(4) are obvious.

   (3)→(1). By Theorem 2.1, it suffices to assume that L is not quasi-aperiodic and find a balanced counting pattern in $M(L)$ for some prime n. We have that $\eta_L(A^d)$ contains a non-trivial group for some $d \geq 1$. Then $\eta_L(A^d)$ contains an isomorphic copy $G = (\{[u_0], \ldots, [u_{n-1}]\}; \cdot)$ of the cyclic group $\mathbf{Z}_n = (\{0, \ldots, n-1\}; \oplus)$ for some prime n (so $[u_i] \cdot [u_j] = [u_{i \oplus j}]$, where $\oplus$ is the addition modulo n, $[u_0]$ is the unit of G and $[u_1]$ is a generator of G). Since $[u_0] \neq [u_1]$, there are $y, z \in A^*$ with $yu_0z \in L \leftrightarrow yu_1z \notin L$; let *e.g.* $yu_0z \in L$ and $yu_1z \notin L$. Set $s_i = [yu_i]$ for each $i \in \{1, \ldots, n\}$, $x = yu_0$, $u = u_0$ and $v = u_1$. One easily checks that the states $s_0 = \varepsilon, s_1, \ldots, s_n$ and the words x, z, u, v form a balanced counting pattern for $M(L)$ (see Def. 8.4(1) and Fig. 1).

   (4)→(2). It suffices to show that if a dfa $\mathcal{A} = (Q, A, \delta, s_0, F)$ recognizing L has a balanced counting pattern as in Figure 1 then $m(L)$ has a similar balanced counting pattern. W.l.o.g. we may assume that any state of $\mathcal{A}$ is reachable from the initial state s_0 (otherwise, just remove the unreachable states from $\mathcal{A}$). Define the equivalence relation $\sim$ on Q as follows: $p \sim q$ iff $\forall u \in A^*(\delta(p, u) \in F \leftrightarrow \delta(q, u) \in F)$. As is well-known [35], $m(L)$ is isomorphic to the quotient-automaton
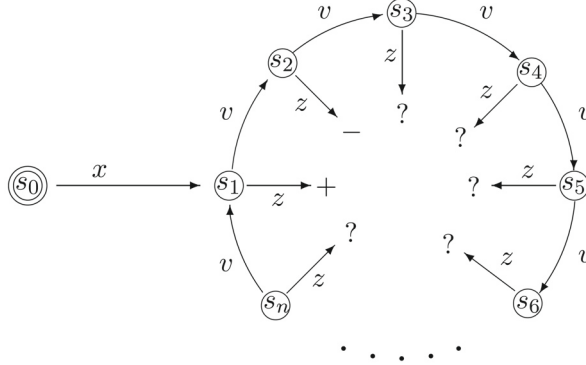
FIGURE 5. The counting pattern with initial state s_0.

$\tilde{\mathcal{A}} = (\tilde{Q}, A, \tilde{\delta}, \tilde{s}_0, \tilde{F})$ of $\mathcal{A}$. The states $\tilde{s}_0, \tilde{s}_1, \ldots \tilde{s}_n$ and the same words x, z, u, v as in Figure 1 form a balanced counting pattern for $\tilde{\mathcal{A}}$. Hence, $m(L)$ has a balanced counting pattern. $\qquad \square$

Remark. As is well-known (see *e.g.* [46] for details), for any dfa $\mathcal{A}$ it holds: $L(\mathcal{A})$ is aperiodic iff $\mathcal{A}$ has no counting pattern (see Fig. 5). Currently we do not know whether the similar characterization holds for the quasi-aperiodic languages and balanced counting patterns (as one might guess from the preceding theorem). Nevertheless, already the characterization from the preceding theorem is useful.

Now we are able to characterize (uniformly on oracles) the regular languages L such that $\mathrm{Leaf}_\mathrm{b}(L)) \subseteq \mathrm{PH}$. The similar characterization for the unbalanced leaf language definability and aperiodic languages is well-known [45].

Theorem 8.9. *A regular language L is quasi-aperiodic iff $\mathrm{Leaf}_\mathrm{b}(L)^O \subseteq \mathrm{PH}^O$ for all oracles O.*

Proof. From left to right the proof is contained in the proof of Theorem 8.8, (1)→(2). Conversely, let L be non-quasi-aperiodic. By the proof of Theorem 8.8, $M_p \leq_\mathrm{plt} L$ for some prime p. By Theorem 8.2, $\mathrm{Leaf}_\mathrm{b}(M_p)^O \subseteq \mathrm{Leaf}_\mathrm{b}(L)^O$ for all O. It is known [45] that $\mathrm{Leaf}_\mathrm{b}(M_p)^O \not\subseteq \mathrm{PH}^O$ for some O. Therefore, $\mathrm{Leaf}_\mathrm{b}(L)^O \not\subseteq \mathrm{PH}^O$ for some O. $\qquad \square$

Corollary 8.10. *The class of regular quasi-aperiodic languages is an ideal of $(\mathcal{R}; \leq_\mathrm{plt})$.*

We conclude this section by a result on an initial segment of the structure $(\mathcal{R}'; \leq_\mathrm{plt})$. This result is implicit in [12,13] (provided we use Th. 8.5 and some known facts on the oracle separations). Nevertheless, we present a more direct proof because its ideas are also used in some proofs below. Let $(P; \leq)$ be an upper semilattice with a least element 0. Recall that an *atom* of P is a minimal non-zero element of P. A semilattice is called *atomic* if below every non-zero element there is an atom.

Theorem 8.11. *The semilattice $(\mathcal{R}'; \leq_\mathrm{plt})$ is atomic with infinitely many atoms.*

Proof. Let $E \subseteq A^+$ be the language of words having at least one letter distinct from a fixed letter $a \in A$. Obviously, $E \equiv_{\mathrm{plt}} 0^*1(0 \cup 1)^*$. For any prime p, let $M_p \subseteq A^+$ consist of all words such that the number of occurrences of letters distinct from a is divided by p. We claim that the languages $E, \overline{E}, M_p$ (p prime) define exactly the atoms of $(\mathcal{R}'; \leq_{\mathrm{plt}})$, *i.e.*:

(1) $E, \overline{E}, M_p$ are pairwise plt-incomparable;
(2) for any $L \in \mathcal{R}'$ of non-smallest plt-degree, at least one of $E, \overline{E}, M_p$ is plt-reducible to L.

The assertion (1) follows from Theorem 8.2 and the well-known oracle separations (alternatively, it may be observed from the definition of plt-reducibility).

(2) Let $L \in \mathcal{R}'$ be of non-smallest degree. By Theorem 8 and Corollary 18 in [13], $L \not\leq_{\mathrm{plt}} E$ or $L \not\leq_{\mathrm{plt}} \overline{E}$. We consider only the first case, the second being dual. By Theorem 8.5, the minimal automaton $\mathcal{A}$ of L contains a balanced coNP-pattern, or a balanced co1NP-pattern, or a balanced counting pattern.

In the case of balanced coNP-pattern as in Figure 3, consider the plt-function $f(y) = xh(y)z$, where $h : A^+ \to A^+$ is the length-multiplying homomorphism satisfying $h(a) = u$ and $h(b) = v$ for all $b \in A \setminus \{a\}$. By the examples of plt-functions at the beginning of this section, h is a plt-function, hence $\overline{E} \leq_{\mathrm{plt}} L$.

In the case of balanced co1NP-pattern, we similarly get $\overline{0^*10^*} = f^{-1}(L)$. Since $E \leq_{\mathrm{plt}} \overline{0^*10^*}$, $E \leq_{\mathrm{plt}} L$.

In the case of balanced counting pattern, let $M'_n \subseteq \{0,1\}^+$ be the set of words with $\#_1(w) \equiv 0 \pmod{n}$. By the proof of Theorem 8.8, $M'_p \leq_{\mathrm{plt}} L$ for some prime p. Obviously, $M_p \equiv_{\mathrm{plt}} M'_p$. Therefore, $M_p \leq_{\mathrm{plt}} L$. $\qquad\square$

Corollary 8.12. Δ_1^τ *coincides with the class of regular plt-decidable languages.*

Proof. According to [12,13], a regular language L is plt-decidable iff $L \leq_{\mathrm{plt}} E$ and $L \leq_{\mathrm{plt}} \overline{E}$. By the proof of Theorem 8.6 this is equivalent to $L \in \Delta_1^\tau$. $\qquad\square$

Since there are non-regular sets plt-reducible to $0^*1(0\cup1)^*$ [12,45], plt-reducibility does not fit the introduced hierarchies. In the next two sections we consider reducibilities which behave better in this respect.

9. QUANTIFIER-FREE REDUCIBILITIES

In [39] the reducibility by quantifier-free formulas of the signature σ was introduced and studied. Here we generalize notions and results of [39] to the signature τ_d for every fixed $d > 0$ and present some new results.

A *$qf\tau_d$-interpretation I over alphabets $A = \{a, \ldots\}$ and $B = \{b, \ldots\}$ is given* by a tuple

$$(\phi_U(\bar{x}), \phi_\leq(\bar{x}, \bar{y}), \phi_\perp(\bar{x}), \phi_\top(\bar{x}), \phi_S(\bar{x}, \bar{y}), \phi_b(\bar{x}), \ldots, \phi_d^0(\bar{x}), \ldots, \phi_d^{d-1}(\bar{x}))$$

where $\bar{x} = (x_1, \ldots, x_n)$ and $\bar{y} = (y_1, \ldots, y_n)$ are sequences of different variables of the same length $n > 0$ (n is fixed in advance) and $\phi_U(\bar{x}), \ldots, \phi_d^0(\bar{x}), \ldots, \phi_d^{d-1}(\bar{x})$

are quantifier-free formulas of $\tau_{d,A}$ with the following properties. Let $u = u_0 \cdots u_l$ be any word over A of length $|u| = l + 1$. Then the set $T = \{\bar{x} \in \{0, \ldots, l\}^n \mid \mathbf{u} \models \phi_U(\bar{x})\}$ should be non-empty and the formulas $\phi_\leq(\bar{x}, \bar{y}), \phi_\perp(\bar{x}), \phi_\top(\bar{x}), \phi_S(\bar{x}, \bar{y})), \phi_b(\bar{x}), ..., \phi_d^0(\bar{x}), \ldots, \phi_d^{d-1}(\bar{x})$ interpreted in $\mathbf{u}$ should define a model of $\mathrm{CLO}^{\tau_d'}{}_B$ with the universe T (the formulas $\phi_\perp(\bar{x}), \phi_\top(\bar{x})$ should be true exactly on the first and the last element, respectively). Since the finite models of $\mathrm{CLO}^{\tau_d'}{}_B$ are in a bijective correspondence with the elements of B^+, any $qf\tau_d$-interpretation I induces a function $u \mapsto u^I$ from A^+ into B^+.

Remark. The definition of the function $u \mapsto u^I$ above is not completely formal, hence one might ask which tuples I really define such a function. One such condition well-known from logic (see *e.g.* [31]) is sufficient for our paper and looks as follows. Given a tuple I as above, relate to any formula ψ of signature $\tau_{d,B}'$ the formula ψ^I of signature $\tau_{d,A}'$ (called *I*-translation of ψ) in the natural way, replacing the signature predicates and constant symbols by the corresponding formulas from I on the universe defined by $\phi_U(\bar{x})$ (*e.g.* the translation of $\forall x(x \leq x)$ and $\forall x(\perp \leq x)$ look like $\forall \bar{x}(\phi_U(\bar{x}) \to \phi_\leq(\bar{x}, \bar{x}))$ and $\forall \bar{x}, \bar{y}(\phi_U(\bar{x}) \wedge \phi_U(\bar{y}) \wedge \phi_\perp(\bar{y}) \to \phi_\leq(\bar{y}, \bar{x}))$, respectively). As is well-known and easy to see, if the sentences $\exists \bar{x}\phi_U(\bar{x}), \exists \bar{x}!\phi_\perp(\bar{x}), \exists \bar{x}!\phi_\top(\bar{x})$ and the *I*-translations of all the axioms of $\mathrm{CLO}_B^{\tau_d'}$ are provable in $\mathrm{CLO}_A^{\tau_d'}$ then $u \mapsto u^I$ is really a function from A^+ into B^+. All interpretations considered in this paper will have this property, hence they will define the functions $u \mapsto u^I$.

Examples.

1. Let $\phi_U(\bar{x})$ be a valid formula, let $\phi_\leq(\bar{x}, \bar{y})$ be a formula that defines the lexicographic ordering between $\bar{x}$ and $\bar{y}$, let $\phi_\perp(\bar{x}), \phi_\top(\bar{x}), \phi_S(\bar{x}, \bar{y}), \phi_d^0(\bar{x}), \ldots, \phi_d^{d-1}$ be defined in the obvious way according to their intended interpretations in the lexicographic ordering, and let $\phi_b(\bar{x}), \ldots$ be chosen arbitrarily but consistent with the axioms $\mathrm{CLO}^{\tau_d}{}_B$ describing the properties of letters. Then u^I is of length $|u|^n$, and the letters of the word u^I are easily computed from the interpretation. Then I is a $qf\tau_d$-interpretation over A and B. Note that $u \mapsto u^I$ is a plt-function.

2. Let $\phi_U(x)$ be a valid formula with one variable x, let $\phi_\leq(x, y)$ be $x \geq y$, let $\phi_a(x)$ be $Q_a(x)$ for any $a \in A$, let $\phi_\perp(\bar{x}), \phi_\top(\bar{x})$ be $x = \top$, $x = \perp$ respectively, and let $\phi_S(x, y)$ be $S(y, x)$. Let $\phi_d^0(\bar{x}), \ldots, \phi_d^{d-1}(\bar{x})$ be defined in the obvious way to satisfy the corresponding axioms of $\mathrm{CLO}_A^{\tau_d}$. Then we obtain a $qf\tau_d$-interpretation I over A and A such that u^I is the reverse of the word $u \in A^+$. Note that $u \mapsto u^I$ is a plt-function.

3. Let $u \mapsto pu$ be the function on A^+ which adds a fixed prefix $p \in A^*$ to a word u. Is there a $qf\tau_d$-interpretation I over A and A such that $u^I = pu$ for any u? For $p = \varepsilon$ the answer is of course positive, otherwise it is negative (since any $qf\tau_d$-interpretation sends words of length 1 to words of length 1). But it is easy to see that there is a $qf\tau_d$-interpretation I over A and A such that $u^I = pu$ for each $u \in A^{\geq 2}$. The same of course applies to the operation of adding a suffix to a word.

4. For any d-length-multiplying semigroup morphism $h : A^+ \to B^+$ (which means that h has the property $\forall a, b \in A(|h(a)| \equiv |h(b)| \pmod{d})$) there is a $qf\tau_d$-interpretation I over A and B such that $u^I = h(u)$ for almost all $u \in A^+$ (i.e. for all but finitely many words). In general, $u \mapsto u^I$ is not a plt-function.

Definition 9.1.

- A function $f : A^+ \to B^+$ is called a $qf\tau_d$-function if there is a $qf\tau_d$-interpretation I over A and B such that $u^I = f(u)$ for almost all $u \in A^+$.
- We say that $L \subseteq A^+$ is $qf\tau_d$-reducible to $K \subseteq B^+$ (in symbols $L \leq_{\mathrm{qf}\tau_\mathrm{d}} K$) if $L = f^{-1}(K)$ for some $qf\tau_d$-function $f : A^+ \to B^+$.

The next theorem generalizes in a straightforward way the corresponding facts from [38,39] obtained there for the case of $qf\sigma$-reducibility.

Theorem 9.2.

(1) *The relation $\leq_{\mathrm{qf}\tau_\mathrm{d}}$ is reflexive and transitive.*
(2) *$\{\emptyset\}$ and $\{A^+\}$ are two distinct minimal elements of the degree structure $(P(A^+); \leq_{\mathrm{qf}\tau_\mathrm{d}})$ which are below any other element.*
(3) *The structure $(P(A^+); \leq_{\mathrm{qf}\tau_\mathrm{d}})$ is an upper semilattice.*
(4) *The structure $(P'(A^+); \leq_{\mathrm{qf}\tau_\mathrm{d}})$ is a distributive upper semilattice with a least element which consists exactly of the non-trivial $\Delta_1^{\tau_\mathrm{d}}$-languages.*
(5) *For all $n \geq 1$ the classes $\bigcup_n \Sigma_n^{\tau_\mathrm{d}}$, $\mathcal{R}$, $FO + MOD(\mathcal{P})$ ($\mathcal{P}$ is any set of positive integers containing d) are ideals of $(P(A^+); \leq_{\mathrm{qf}\tau_\mathrm{d}})$.*
(6) *The classes $\Sigma_n^{\tau_\mathrm{d}}, \Pi_n^{\tau_\mathrm{d}}$, as well as all levels of the difference hierarchy over $\Sigma_n^{\tau_\mathrm{d}}$, are principal ideals of $(P(A^+); \leq_{\mathrm{qf}\tau_\mathrm{d}})$.*
(7) *Let D_n be a $qf\tau_d$-complete set in the n-th level of the difference hierarchy over $\Sigma_1^{\tau_\mathrm{d}}$. Then $D_n \oplus \overline{D}_n$ is the infimum of sets $D_{n+1}, \overline{D}_{n+1}$ under $\leq_{\mathrm{qf}\tau_\mathrm{d}}$.*

Proof. (sketch). The items (1)–(5) are checked in the same way as the corresponding assertions in [38,39].

In the proof of (6) for $\Sigma_n^{\tau_\mathrm{d}}$ in the alphabet-dependent mode, we take again the set H_n from the proof of Theorem 8.6(1) and note that the function f from the proof of Lemma 3.1 in [39] is a $qf\tau_d$-function. For the alphabet-independent version we use an encoding h like that in the proof of Theorem 8.6(1) but with the additional requirement that h is d-length-multiplying (this requirement is also satisfied in an obvious way). For the levels of the DH, similar tricks apply.

The proof of the item (7) makes use of Theorem 7.3 and Proposition 4.4 and is similar to the corresponding proof in [39]. $\square$

Next we state a "parameterized" version of Theorem 8.8. For any fixed $d > 0$, by a *d-balanced counting pattern* for a dfa $\mathcal{A}$ we mean a counting pattern as in Figure 5 but with the additional requirement $|v| \equiv 0 \pmod{d}$. It is easy to see that if $\mathcal{A}$ has such a d-balanced counting pattern then it also has a pattern obtained from Definition 8.4(1) by replacing the equality $|u| = |v|$ on the equivalence $|u| \equiv |v| \equiv 0 \pmod{d}$ (to observe this, take $u = v^n$). Note that the 1-counting pattern coincides with the counting pattern.

Theorem 9.3. *For every regular language L the following conditions are equivalent:*

(1) *L is FO_{τ_d}-axiomatizable.*
(2) *Every finite automaton recognizing L has no d-balanced counting patterns.*
(3) *The automaton $M(L)$ has no d-balanced counting patterns.*
(4) *The minimal automaton $m(L)$ of L has no d-balanced counting patterns.*

Proof. is almost the same as that of Theorem 8.8, only now, in the proof of the implication (3) $\to$ (1), we have $|u_i| \equiv 0 \pmod{d}$ for all $i \leq n$. This guarantees that the resulting counting pattern will be d-balanced. $\qquad\square$

Remark. Results in [10,46] imply the following improvement of the previous result: for any dfa $\mathcal{A}$, $L(\mathcal{A})$ is FO_{τ_d}-axiomatizable iff $\mathcal{A}$ does not have d-balanced counting patterns.

In [38] some relationships between plt- and $qf\sigma$-reducibilities were established. These results also generalize to the signatures τ_d. To see this, we recall definition of the so called languages of finite counting type (*cf.* [6,15,16]). For all $k \geq 1$ and $V \subseteq \omega^k$ we define the language

$$L(V) = \{x \in A_k^+ \mid (\#_1(x), \dots, \#_k(x)) \in V\}, \ A_k = \{0, 1, \dots, k\},$$

where $\#_i(x)$ is the number of occurrences of i in x. Such languages are called languages of counting type. The set V and the language $L(V)$ are said to be *of finite counting type* iff there exists an $m \geq 0$ such that

$$(n_1, \dots, n_k) \in V \leftrightarrow (\min(n_1, m), \dots, \min(n_k, m)) \in V.$$

It is obvious that the languages of finite counting type are aperiodic, in fact they are in $\mathrm{BC}(\Sigma_1^\rho)$.

Proposition 9.4.

(1) *If $L \leq_{\mathrm{qf}\tau_\mathrm{d}} M$ and M is of counting type then $L \leq_{\mathrm{plt}} M$.*
(2) *If $L \leq_{\mathrm{qf}\sigma} M$, $L \subseteq A^+$, and $M = L(V) \subseteq A_k^+$ is of counting type then $L \leq_{\mathrm{qf}\tau_\mathrm{d}} M$.*
(3) *If both L and M are of finite counting type then $L \leq_{\mathrm{qf}\sigma} M$ iff $L \leq_{\mathrm{qf}\tau_\mathrm{d}} M$ iff $L \leq_{\mathrm{plt}} M$.*
(4) *Within the class $\Sigma_1^{\tau_\mathrm{d}}(2)$ of differences of $\Sigma_1^{\tau_\mathrm{d}}$-sets there are infinitely many languages modulo $\equiv_{\mathrm{qf}\tau_\mathrm{d}}$.*

Proof.

(1) Is proved exactly as Theorem 5.3 in [39], only now the interpretation I from that proof is a $qf\tau_d$-interpretation.

(2) Let $L \leq_{\mathrm{qf}\sigma} M$ via the $qf\sigma$-interpretation

$$J = (\psi_U(\bar{x}), \psi_\leq(\bar{x}, \bar{y}), \psi_0(\bar{x}), \dots, \psi_k(\bar{x}), \psi_\perp(\bar{x}), \psi_\top(\bar{x}), \psi_S(\bar{x}, \bar{y})), \ \bar{x} = (x_1, \dots, x_n),$$

over A and A_k. Define a $qf\tau_d$-interpretation

$$I = (\phi_U(\bar{x}), \phi_\le(\bar{x}, \bar{y}), \phi_\perp(\bar{x}), \phi_\top(\bar{x}), \phi_S(\bar{x}, \bar{y}), \phi_0(\bar{x}), \ldots, \phi_k(\bar{x}), \phi_d^0(\bar{x}), \ldots, \phi_d^{d-1}(\bar{x}))$$

over A and A_k (with the same n) exactly as in Example 1 at the beginning of this section where $\phi_0(\bar{x})$ is $\neg\psi_U(\bar{x}) \vee \psi_0(\bar{x})$ and $\phi_i(\bar{x})$ is $\psi_U(\bar{x}) \wedge \psi_i(\bar{x})$ for each $i \in \{1, \ldots, k\}$. Then $\#(u^I) = \#(u^J)$ for all $u \in A^+$ where $\#(x) = (\#_1(x), \ldots, \#_k(x))$. Therefore, for almost all $x \in A^+$ we have

$$x \in L \leftrightarrow x^J \in M \leftrightarrow \#(x^J) \in V \leftrightarrow \#(x^I) \in V \leftrightarrow x^I \in M$$

which means $L \le_{\mathrm{qf}\tau_\mathrm{d}} M$.

(3) Since $L \le_{\mathrm{plt}} M$ implies $L \le_{\mathrm{qf}\sigma} M$ by Theorem 5.4 in [39], the assertion follows from (1) and (2).

(4) Follows from (2) and Corollary 6.2 in [39]. $\qquad\square$

Next we want to establish "parameterized" versions of Theorems 8.5 and 8.11. The notions of d-balanced coUP-, coNP- and co1NP-patterns are obtained from the corresponding notions in Definition 8.4 by replacing the equality $|u| = |v|$ on the equivalence $|u| \equiv |v| \equiv 0 \pmod{d}$.

Theorem 9.5. *Let L be a regular language and $\mathcal{A}$ the minimal automaton recognizing L. The following conditions are equivalent:*
  (1) $L \le_{\mathrm{qf}\tau_\mathrm{d}} 0^*1(0 \cup 1)^*$;
  (2) $\mathcal{A}$ *does not have d-balanced coUP-patterns;*
  (3) $\mathcal{A}$ *does not have d-balanced coNP-, co1NP- and counting patterns;*
  (4) L *is a finite union of languages of the form $w_0(A^d)^*w_1 \cdots (A^d)^*w_n$, where $n \ge 0$ and $w_i \in A^*$.*

Proof. is obtained by repeating (with some obvious modifications) of the corresponding proofs of Theorem 8 and Corollary 11 in [13] for the balanced case. Note that for $d = 1$ we obtain the corresponding result from [2]. $\qquad\square$

Corollary 9.6. $\Sigma_1^{\tau_\mathrm{d}}$ *is the class of finite unions of languages $w_0(A^d)^*w_1 \cdots (A^d)^*w_n$, where $n \ge 0$ and $w_i \in A^+$.*

Proof. The inclusion from right to left is obvious. The opposite inclusion follows from the last theorem and the fact that $0^*1(0 \cup 1)^*$ is $qf\tau_d$-complete in $\Sigma_1^{\tau_\mathrm{d}}$ by the proof of Theorem 9.2(6). $\qquad\square$

Remark. By [10], for any $d \ge 1$ the class of d-quasi-aperiodic languages coincides with the closure of the class of finite languages and the language $(A^d)^*$ under the Boolean operations and concatenation. We guess that the classes $\Sigma_n^{\tau_\mathrm{d}}, \mathrm{BC}(\Sigma_n^{\tau_\mathrm{d}})$ may be characterized in terms of regular expressions in the similar way to the well-known characterization of $\Sigma_n^\sigma, \mathrm{BC}(\Sigma_n^\sigma)$ in [40] and of of $\Sigma_n^\rho, \mathrm{BC}(\Sigma_n^\rho)$ in [23]. Corollary 9.6 provides such a characterization for $n = 1$.
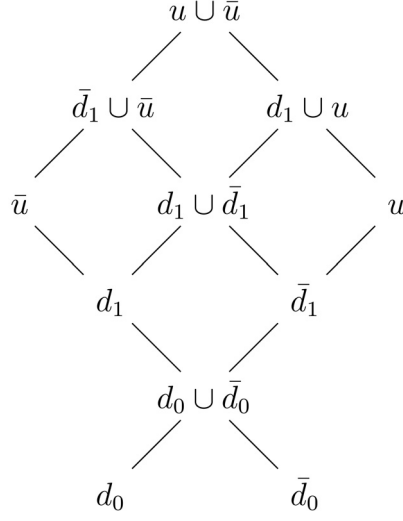
Next we prove an analog of Theorem 8.11.

$$u \cup \bar{u}$$

$$\bar{d}_1 \cup \bar{u} \qquad d_1 \cup u$$

$$\bar{u} \qquad d_1 \cup \bar{d}_1 \qquad u$$

$$d_1 \qquad \bar{d}_1$$

$$d_0 \cup \bar{d}_0$$

$$d_0 \qquad \bar{d}_0$$

FIGURE 6. The principal ideal of $(P(A^+); \leq_{\mathrm{qf}\sigma})$ generated by $U \oplus \overline{U}$.

Theorem 9.7. *The semilattice $(\mathcal{R}'; \leq_{\mathrm{qf}\tau_\mathrm{d}})$ is atomic with infinitely many atoms.*

Proof. We claim that the languages $E, \overline{E}, M_p$ (p prime) from the proof of Theorem 8.11 define exactly the atoms of $(\mathcal{R}'; \leq_{\mathrm{qf}\tau_\mathrm{d}})$, *i.e.*:

(1) $E, \overline{E}, M_p$ are pairwise $qf\tau_d$-incomparable,

(2) for any $L \in \mathcal{R} \setminus \Delta_1^{\tau_\mathrm{d}}$, at least one of the sets $E, \overline{E}, M_p$ is $qf\tau_d$-reducible to L.

(1) The incomparability of E and $\overline{E}$ follows from the fact that E is $qf\tau_d$-complete in $\Sigma_1^{\tau_\mathrm{d}}$ and $\Sigma_1^{\tau_d} \neq \Pi_1^{\tau_d}$. Since M_p is not a quasi-aperiodic language (see [35]), we have $M_p \not\leq_{\mathrm{qf}\tau_\mathrm{d}} E$. Suppose that $E \leq_{\mathrm{qf}\tau_\mathrm{d}} M_p$. Since M_p is of counting type, $E \leq_{\mathrm{plt}} M_p$, contradicting to the proof of Theorem 8.11. Therefore, E and M_p are incomparable. A similar proof shows that $\overline{E}$ and M_p are incomparable. It remains to check that $M_p \not\leq_{\mathrm{qf}\tau_\mathrm{d}} M_q$ for $p \neq q$. Suppose $M_p \leq_{\mathrm{qf}\tau_\mathrm{d}} M_q$. Since the language M_q is of counting type, $M_p \leq_{\mathrm{plt}} M_q$ by Proposition 9.4(1), contradicting to Theorem 8.11.

With Theorem 9.5 at hand, the assertion (2) is proved in the same way as the corresponding assertion in the proof of Theorem 8.11 but with the d-balanced patterns instead of the balanced patterns. One has only to observe that the function g from the proof of Theorem 8.8 is now a $qf\tau_d$-function because the homomorphism h is d-length-multiplying (see Exs. 3 and 4 at the beginning of this section). $\qquad\square$

From distributivity of the semilattice $(\mathcal{R}'; \leq_{\mathrm{qf}\tau_\mathrm{d}})$ we immediately obtain the following result about an initial segment.

Corollary 9.8. *The lattice $(Fin; \subseteq)$ of all finite subsets of ω is isomorphic to an ideal of $(\mathcal{R}'; \leq_{\mathrm{qf}\tau_\mathrm{d}})$.*

We note that the result from [38,39] on the principal ideal of $(P(A^+); \leq_{\mathrm{qf}\sigma})$ generated by the set $U \oplus \overline{U}$ where $U = 0^*10^*$, is true also for the structure $(P(A^+); \leq_{\mathrm{qf}\tau_d})$ for each $d \geq 1$ (this follows from Prop. 9.4 because all the languages defining the principal ideal are of finite counting type). This principal ideal is shown in Figure 6 where u is the degree of U, d_0 is the degree of $\emptyset$, d_1 is the degree of $0^*1(0 \cup 1)^*$ and $\bar{c}$ denotes the dual of a degree c.

Remark. Note that all results about the structure of $qf\tau_a$-degrees above are obtained by using some known technique for the plt-reducibility and the languages of finite counting type. We would like to see methods specially designed for the investigation of the $qf\tau_a$-reducibility.

The reader might wonder about the relationships between the $qf\tau_d$-reducibilities for different d. The first impression could be that there are some dependencies (*e.g.* one could expect that if d is a multiple of e (in symbols, $e \mid d$) then $\leq_{\mathrm{qf}\tau_e}$ implies $\leq_{\mathrm{qf}\tau_d}$). But actually this is not the case.

Theorem 9.9. *For all distinct $d, e > 0$, $qf\tau_d$-reducibility is incomparable with $qf\tau_e$-reducibility.*

Proof. First we observe that $(A^d)^+ \in \Sigma_1^{\tau_e}$ implies $d \mid e$. Indeed, by Corollary 9.6, $(A^d)^+$ is a finite union of languages $L = w_0(A^e)^*w_1 \cdots (A^e)^*w_n$, where $n \geq 0$ and $w_i \in A^+$. Since $(A^d)^+$ is infinite, it contains at least one such a language L with $n > 0$. Since $w = w_0 \cdots w_n$ and $w_0a^ew_1 \cdots w_n$, $a \in A$, are in L, d divides both $|w|$ and $|w| + e$. Therefore, $d \mid e$.

Now consider the following three cases.

Case 1. $d \nmid e$ and $e \nmid d$.

By the observation above, $(A^d)^+ \notin \Sigma_1^{\tau_e}$ and $(A^e)^+ \notin \Sigma_1^{\tau_d}$. Hence, $(A^d)^+ \leq_{\mathrm{qf}\tau_d} 0^*1(0\cup1)^*$, $(A^d)^+ \not\leq_{\mathrm{qf}\tau_e} 0^*1(0\cup1)^*$ and similarly with d, e interchanged. Therefore, the $qf\tau_d$- and $qf\tau_e$-reducibilities are incomparable.

Case 2. $e \mid d$ and $d \neq e$.

In this case $d \nmid e$ hence, by case 1, $qf\tau_d$-reducibility does not imply $qf\tau_e$-reducibility. It remains to check that $qf\tau_e$-reducibility does not imply $qf\tau_d$-reducibility. It suffices to show that $M_c \leq_{\mathrm{qf}\tau_e} (A^d)^+$ but $M_c \not\leq_{\mathrm{qf}\tau_d} (A^d)^+$ where $c = d/e$ and $A = \{0, 1\}$. Suppose $M_c \leq_{qf\tau_d} (A^d)^+$. Since $(A^d)^+ \in \Delta_1^{\tau_d}$, $M_c \in \Delta_1^{\tau_d}$ and hence M_c is quasi-aperiodic. A contradiction.

It remains to show that $M_c \leq_{\mathrm{qf}\tau_e} (A^d)^+$. Let $h : \{0,1\}^+ \to \{0,1\}^+$ be the homomorphism satisfying $h(0) = 0^d$ and $h(1) = 1^e$. Then

$$\#_1(h(w)) = e \cdot \#_1(w) \text{ and } |h(w)| \equiv \#_1h(w) \equiv e \cdot \#_1(w) \ (\mathrm{mod} \ d).$$

Thus,

$$w \in M_c \leftrightarrow \#_1(w) \equiv 0 \ (\mathrm{mod} \ c) \leftrightarrow |h(w)| \equiv 0 \ (\mathrm{mod} \ d).$$

Since $e \mid h(w)$ for each $w \in A^+$, h is a $qf\tau_e$-function (see Ex. 4 at the beginning of this section) and therefore $M_c \leq_{\mathrm{qf}\tau_e} (A^d)^+$.

The remaining case ($d \mid e$ and $d \neq e$) is symmetric with case 2. $\qquad\square$

10. A reducibility for the τ-hierarchy

We have seen that the $qf\tau_d$-reducibility fits the τ_d-hierarchy. Is there a reducibility that fits the τ-hierarchy? In this section we find such a reducibility.

Definition 10.1. We say that a language L is τ-reducible to a language K (in symbols $L \leq_\tau K$), if there exists $c > 0$ such that $L \leq_{qf\tau_d} K$ for all multiples d of c.

Analogs of some results of the previous sections hold true for the τ-reducibility. One of its nice properties is the following.

Theorem 10.2. *The τ-reducibility fits the τ-hierarchy, i.e. every class Σ_n^τ, $n > 0$, is a principal ideal of $(P(A^+); \leq_\tau)$.*

Proof. Let $L \leq_\tau K \in \Sigma_n^\tau$. Then there exist $a, c > 0$ such that $K \in \Sigma_n^{\tau_a}$ and $L \leq_{qf\tau_d} K$ for all d with $c \mid d$. For the number $e = a \cdot d$ we have $L \leq_{qf\tau_e} K$ hence, by Proposition 3.1, $L \in \Sigma_n^{\tau_e} \subseteq \Sigma_n^\tau$. Therefore, Σ_n^τ is an ideal of $(P(A^+); \leq_\tau)$.

It remains to find a τ-complete set in Σ_n^τ. We again take the set $H_n \in \Sigma_n^\sigma$ from the proof of Theorem 8.6(1). As noticed in the proof of Theorem 9.2, H_n is $qf\tau_d$-complete in $\Sigma_n^{\tau_d}$ for all $d > 0$. Let $L \in \Sigma_n^\tau$, then $L \in \Sigma_n^{\tau_c}$ for some $c > 0$. By Proposition 3.1, $L \in \Sigma_n^{\tau_d}$ for all d with $c \mid d$. Thus, $L \leq_{qf\tau_d} H_n$ for all d with $c \mid d$. Therefore, $L \leq_\tau H_n$, as desired. $\qquad\square$

Remark. A proof similar to the proof of Theorem 8.6(1) (with the use of slight modifications of proofs from Sect. 3 of [39]) shows that for any alphabet every level of the difference hierarchy over Σ_n^τ is a principal ideal under the τ-reducibility.

We conclude this section with an analog of Theorem 9.7.

Theorem 10.3.

  (1) *The degree structure $(\mathcal{R}'; \leq_\tau)$ is a semilattice with a smallest element consisting exactly of the nontrivial languages in Δ_1^τ.*
  (2) *The degree structure $(\mathcal{R}'; \leq_\tau)$ is atomic with infinitely many atoms.*

Proof. The assertion (1) easily follows from the previous theorem and the fact that $0^*1(0 \cup 1)^*$ is τ-complete in Σ_n^τ. To prove (2), we show that the sets $E, \overline{E}$ and M_p, p prime, from the proof of Theorem 8.11 again define exactly the atoms of $(\mathcal{R}'; \leq_\tau)$. By the proof of Theorem 9.7, the sets $E, \overline{E}$ and M_p are pairwise $qf\tau_d$-incomparable for all $d > 0$. Thus, they are also τ-incomparable and are not in Δ_1^τ.

It remains to check that any $L \in \mathcal{R}' \setminus \Delta_1^\tau$ is above at least one of $E, \overline{E}, M_p$ under τ-reducibility. We have $L \notin \Delta_1^{\tau_d}$ for all $d > 0$.

Case 1. L is quasi-aperiodic.
By the proof of Theorem 9.7, $E \leq_{qf\tau_d} L$ for all d or $\overline{E} \leq_{qf\tau_d} L$ for all d. Thus, $E \leq_\tau L$ or $\overline{E} \leq_\tau L$.

Case 2. L is not quasi-aperiodic.
By the proof of Theorem 9.7, there is a prime p such that $M_p \leq_{qf\tau_d} L$ for all d. Thus, $M_p \leq_\tau L$. $\qquad\square$

Remark. Though the τ-reducibility has some nice properties, it does not seem natural because it is apparently less constructive than the $qf\tau_d$-reducibilities. Namely, the $qf\tau_d$-reducibility on the regular sets is computably enumerable while the τ-reducibility is, as an obvious computation shows, only in the level Σ_3^0 of the arithmetical hierarchy. We know neither the exact estimations of the relations $\leq_{qf\tau_d}$ and $\leq_\tau$ on $\mathcal{R}$ in the arithmetical hierarchy nor a more constructive reducibility that fits the τ-hierarchy.

11. First order reducibilities

Here we briefly discuss some weaker logical reducibilities, namely the reducibilities $\leq_{fo\tau_d}$ by first-order formulas of signature τ_d. Definition of the FO_{τ_d}-reducibility $\leq_{fo\tau_d}$ is the same as that of $\leq_{qf\tau_d}$, only now the interpretation I consists of first-order formulas of τ_d. The following result is straightforward.

Theorem 11.1.
 (1) *The relation $\leq_{fo\tau_d}$ is reflexive and transitive.*
 (2) *$\{\emptyset\}$ and $\{A^+\}$ are two distinct minimal elements of $(P(A^+); \leq_{fo\tau_d})$ which are below any other element.*
 (3) *The structure $(P(A^+); \leq_{fo\tau_d})$ is an upper semilattice.*
 (4) *The classes $\mathcal{R}$, $FO+MOD(\mathcal{P})$ ($\mathcal{P}$ is any set of positive integers containing d) are ideals of $(P(A^+); \leq_{fo\tau_d})$.*
 (5) *The structures $(P'(A^+); \leq_{fo\tau_d})$ and $(\mathcal{R}'; \leq_{fo\tau_d})$ are distributive upper semilattices with the least element consisting exactly of the non-trivial FO_{τ_d}-axiomatizable languages.*

The next theorem is an analog of Theorems 9.7 and 8.11.

Theorem 11.2. *The semilattice $(\mathcal{R}'; \leq_{fo\tau_d})$ is atomic with infinitely many atoms.*

Proof. We claim that the sets M_p (p prime) from the proof of Theorem 8.11 define exactly the atoms of $(\mathcal{R}'; \leq_{fo\tau_d})$, *i.e.*:
 (1) M_p are pairwise $fo\tau_d$-incomparable;
 (2) for any regular non-FO_{τ_d}-axiomatizable language L, at least one of M_p is $fo\tau_d$-reducible to L.

To prove (1), suppose that $M_q \leq_{fo\tau_d} M_p$ for some $q \neq p$. Since $M_p \in (FO + MOD)(p)$ and $(FO + MOD)(p)$ is an ideal of $(P(A^+); \leq_{fo\tau_d})$, $M_q \in (FO + MOD)(p)$. This contradicts to Theorem VII.2.1 in [35].

(2) Let L be regular and non-FO_{τ_d}-axiomatizable. By Theorem 9.5 and the proof of Theorem 9.7, $M_p \leq_{qf\tau_d} L$ for some prime p. Therefore, $M_p \leq_{fo\tau_d} L$. $\square$

Corollary 11.3. *The lattice $(Fin; \subseteq)$ is isomorphic to an ideal of $(\mathcal{R}'; \leq_{fo\tau_d})$.*

Remarks.
 1. It would be interesting to have more information on structures like $(\mathcal{R}; \leq_{fo_\sigma})$ or $(FO_\tau; \leq_{fo_\sigma})$ which are the structures of degrees of non-aperiodicity of the corresponding classes of languages.

2. Another natural reducibility notion that might deserve attention is the $fo\tau$-reducibility defined by: $L \leq_{\mathrm{fo}\tau} M$ iff $L \leq_{\mathrm{fo}\tau_\mathrm{d}} M$ for some $d \geq 1$. Obviously, $\leq_{\mathrm{fo}\tau}$ is a preorder, hence $(\mathcal{R}; \leq_{\mathrm{fo}\tau})$ is the structure of degrees of non-quasi-aperiodicity of regular languages.

12. Other reducibilities and open questions

There are also other natural reducibilities on the regular sets. *E.g.*, let $\leq_{\mathrm{fom}}$ be defined in the same way as $\leq_{\mathrm{fo}\tau_\mathrm{d}}$ but this time the interpretation I consists of $(FO+MOD)$-formulas. One can easily establish for this reducibility the analog of Theorem 11.1. In particular, the smallest degree in $(\mathcal{R}'; \leq_{\mathrm{fom}})$ consists exactly of the non-trivial regular languages with solvable syntactic monoid. Therefore, there exist at least two distinct (modulo $\equiv_{\mathrm{fom}}$) non-trivial regular languages. We do not know whether there exist three non-trivial regular languages which are pairwise distinct modulo $\equiv_{\mathrm{fom}}$.

We do not also currently know whether there exist regular languages which are complete under the "logical" reducibilities considered above.

Analogs of the open questions from [38] for the $qf\tau_d$-reducibilities seem also natural.

One could consider also the reducibilities by functions computable by natural classes of finite transducers. Such reducibilities were successfully applied for the classification of some classes of regular ω-languages [30,44].

In this paper we used mainly the logical approach to regular languages. As is well-known, there is an equally important and popular algebraic approach (see *e.g.* [9,21,24]) based on the close relationships of the so called varieties of languages to the varieties of finite semigroups. In [36] this approach was extended to the so called $\mathcal{C}$-varieties where $\mathcal{C}$ is a category of word morphisms. The concept of $\mathcal{C}$-variety (and its possible versions like positive $\mathcal{C}$-varieties in the spirit of [22] or $\mathcal{C}$-families in the spirit of [27]) seem to be relevant to the classes of languages discussed above. *E.g.*, for any $n \geq 1$ the class $\mathrm{BC}(\Sigma_n^\tau)$ (over different alphabets) forms a $\mathcal{C}$-variety where $\mathcal{C}$ is the category of length-multiplying homomorphisms and the class $\mathrm{BC}(\Sigma_n^{\tau_\mathrm{d}})$ (over different alphabets) forms a $\mathcal{C}$-variety where $\mathcal{C}$ is the category of d-length-multiplying homomorphisms. May be, it makes sense to look at these relationships in a systematic way.

References

[1] B. Borchert, On the acceptance power of regular languages. *Theor. Comput. Sci.* **148** (1995) 207–225.

[2] B. Borchert, D. Kuske and F. Stephan, On existentially first-order definable languages and their relation to NP. *RAIRO-Theor. Inf. Appl.* **33** (1999) 259–269.

[3] D.P. Bovet, P. Crescenzi and R. Silvestri, A uniform approach to define complexity classes. *Theor. Comput. Sci.* **104** (1992) 263–283.

[4] J.R. Büchi, Weak second-order arithmetic and finite automata. *Z. Math. Logic Grundl. Math.* **6** (1960) 66–92.

[5] D.A.M. Barrington, K. Compton, H. Straubing and D. Thérien, Regular languages in NC^1. *J. Comput. System Sci.* **44** (1992) 478–499.

[6] K. Cronauer, U. Hertrampf, H. Vollmer and K.W. Wagner, The chain method to separate counting classes. *Theor. Comput. Syst.* **31** (1998) 93–108.

[7] R.S. Cohen and J.A. Brzozowski, Dot-depth of star-free events. *J. Comput. System Sci.* **5** (1971) 1–16.

[8] L. Chaubard, J.-E. Pin and H. Straubing, Actions, wreath products of C-varieties and concatenation product. *Theor. Comput. Sci.* **356** (2006) 73–89.

[9] S. Eilenberg, *Automata, Languages and Machines* v. A and B. Academic Press (1974 and 1976).

[10] Z. Esik and M. Ito, Temporal logic with cyclic counting and the degree of aperiodicity of finite automata. *Acta Cybern.* **16** (2003) 1–28.

[11] Z. Esik and K.G. Larsen, Regular languages definable by Lindström quantifiers. *RAIRO-Theor. Inf. Appl.* **37** (2003) 179–241.

[12] C. Glaßer, Polyloglime-reductions decrease dot-depth, in *Proc. of STACS-2005. Lect. Notes Comput. Sci.* **3404** (2005).

[13] C. Glaßer, Languages Polylog-Time Reducible to Dot-Depth 1/2. *J. Comput. System Sci.* **73** (2007) 36–56.

[14] C. Glaßer and H. Schmitz, The Boolean Structure of Dot-Depth One. *J. Autom. Lang. Comb.* **6** (2001) 437–452.

[15] T. Gundermann and G. Wechsung, Counting classes of finite accepting types. *Computers and Artificial Intelligence* **6** (1987) 395–409.

[16] T. Gundermann, N.A. Nasser and G. Wechsung, A survey on counting classes, in *Proc. of Structures in Complexity Theory* (1990) 140–153.

[17] U. Hertrampf, C. Lautemann, T. Schwentick, H. Vollmer and K.W. Wagner, On the power of polynomial time bit-reductions, *Proc. 8th Structure in Complexity Theory* (1993) 200–207.

[18] A.S. Kechris, *Classical Descriptive Set Theory.* Springer, New York (1994).

[19] R. McNaughton, Algebraic decision procedures for local testability. *Math. Syst. Theor.* **8** (1974) 60–76.

[20] R. McNaughton and S. Papert, *Counter-Free Automata.* MIT Press, Cambridge, Massachussets (1971).

[21] J.-E. Pin, *Varieties of Formal Languages.* North Oxford Academic (1986).

[22] J.-E. Pin, Syntactic semigroups, Chap. 10 in *Handbook of language theory*, Vol. I, edited by G. Rozenberg and A. Salomaa. Springer Verlag (1997) 679–746.

[23] D. Perrin and J.-E. Pin, First-order logic and star-free sets. *J. Comput. System Sci.* **32** (1986) 393–496.

[24] J.-E. Pin and P. Weil, Polynomial closure and unambiguous product. *Theor. Comput. Syst.* **30** (1997) 383–422.

[25] M.P. Schützenberger, On finite monoids having only trivial subgroups. *Inform. Control* **8** (1965) 190–194.

[26] V.L. Selivanov, A logical approach to decidability of hierarchies of regular star-free languages, in *Proc. of STACS-2001. Lect. Notes Comput. Sci.* **2010** (2001) 539–550.

[27] V.L. Selivanov, Relating automata-theoretic hierarchies to complexity-theoretic hierarchies. *RAIRO-Theor. Inf. Appl.* **36** (2002) 29–42.

[28] V.L. Selivanov, Some hierarchies and reducibilities on regular languages. University of Würzburg, Technical Report 349 (2004).

[29] V.L. Selivanov, Some reducibilities on regular sets, in *Proc. of CIE-2005. Lect. Notes Comput. Sci.* **3526** (2005) 430–440.

[30] V.L. Selivanov, Fine hierarchy of regular aperiodic ω-languages, in *Proc. of DLT-2007*, edited by T. Harju, J. Karhumäki and A. Lepistö. *Lect. Notes Comput. Sci.* **4588** (2007) 399–410.

[31] J. Shoenfield, *Mathematical Logic.* Addison Wesley, Massachussets (1967).

[32] A.G. Shukin, Difference hierarchies of regular languages. *Comput. Systems, Novosibirsk* **161** (1998) 141–155 (in Russian).

[33] V.L. Selivanov and A.G. Shukin, On hierarchies of regular star-free languages (in Russian). Preprint 69 of A.P. Ershov Institute of Informatics Systems (2000) 28.

[34] J. Stern, Characterizations of some classes of regular events. *Theor. Comput. Sci.* **35** (1985) 163–176.

[35] H. Straubing, *Finite automata, formal logic and circuit complexity.* Birkhäuser, Boston (1994).

[36] H. Straubing, On logical description of regular languages, in *Proc. of LATIN-2002. Lect. Notes Comput. Sci.* **2286** (2002) 528–538.

[37] H. Straubing, D. Thérien and W. Thomas, Regular languages defined with generalized quantifiers. *Inform. Comput.* **118** (1995) 289–301.

[38] V.L. Selivanov and K.W. Wagner, A reducibility for the dot-depth hierarchy. *Proc. 29th Int. Symp. on Mathematical Foundations of Computer Science. Lect. Notes Comput. Sci.* **3153** (2004) 783–793.

[39] V.L. Selivanov and K.W. Wagner, A reducibility for the dot-depth hierarchy. *Theor. Comput. Sci.* **345** (2005) 448–472.

[40] W. Thomas, Classifying regular events in symbolic logic. *J. Comput. System Sci.* **25** (1982) 360–376.

[41] W. Thomas, An application of the Ehrenteucht-Fraïssé game in formal language theory. *Mém. Soc. Math. France Ser. 2* **16** (1984) 11–21.

[42] B.A. Trakhtenbrot, Synthesis of logic networks whose operators are described by means of single-placed predicate calculus. *Doklady Akad. Nauk SSSR* **118** (1958) 646–649.

[43] N.K. Vereshchagin, Relativizable and non-relativizable theorems in the polynomial theory of algorithms. *Izvestiya Rossiiskoi Akademii Nauk* **57** (1993) 51–90 (in Russian).

[44] K.W. Wagner, On ω-regular sets. *Inform. Control* **43** (1979) 123–177.

[45] K.W. Wagner, Leaf language classes. *MCU-2004. Lect. Notes Comput. Sci.* **3354** (2005) 60–81.

[46] T. Wilke, Classifying discrete temporal properties, in Proc. *STACS-99. Lect. Notes Comput. Sci.* **1563** (1999) 32–46.