International Journal of Algebra and Computation

Rational, Recognizable, and Aperiodic Partially Lossy Queue Languages

CHRIS KÖCHER

Automata and Logics Group, Technische Universität Ilmenau,
P.O. Box 10 05 65, 98684 Ilmenau, Germany
chris.koecher@tu-ilmenau.de

Received (Day Month Year)
Accepted (Day Month Year)

Communicated by [editor]

Partially lossy queue monoids (or plq monoids) model the behavior of queues that can non-deterministically forget specified parts of their content at any time. We call the subsets of this monoid partially lossy queue languages (or plq languages). While many decision problems on recognizable plq languages are decidable, most of them are undecidable if the languages are rational. In particular, in this monoid the classes of rational and recognizable languages do not coincide. This is due to the fact that the class of recognizable plq languages is not closed under multiplication and iteration. However, we can generate the recognizable plq languages using special rational expressions consisting of the Boolean operations and restricted versions of multiplication and iteration. From these special rational expressions we can also obtain an MSO logic describing the recognizable plq languages. Moreover, we provide similar results for the class of aperiodic languages in the plq monoid.

Keywords: Partially Lossy Queues; Transformation Monoid; Rational Sets; Recognizable Sets; Aperiodic Sets; MSO Logic.

Mathematics Subject Classification 2010: 03D05, 03D10, 68P05, 68Q45

1. Introduction

The study of different models of automata along with their expressiveness and algorithmic properties is one of the most important areas in automata theory. Many of these models differ in the mechanism to store their data, e.g., there are finite memories, pushdowns, (blind) counters, and infinite Turing tapes. Another very important mechanism is the so-called fifo queue (or channel), where data can be written to one end and read from the other end of its contents. If we equip these queues with a finite automaton we obtain a Turing powerful computation model (cf. [3]), which results in the undecidability of all non-trivial decision problems on these devices. A surprising result was the decidability of some decision problems like reachability, fair termination, or control-state-maintainability if the fifo queue is allowed to forget any part of its content at any time [1,5,9,22].

A possible approach to verify automata with one of these storage mechanisms is to analyze their monoid of transformations. This is the monoid consisting of all possible sequences of transformations on the storage together with the composition of two such transformation sequences. For example consider a single blind counter. This storage mechanism stores an integer which we can increase or decrease with the help of transformations. Since increasing is the inverse operation of decreasing, we can learn that a blind counter's transformation monoid is $(\mathbb{Z}, +, 0)$. Similarly, we can show that the transformation monoid of a single partially blind counter is the so-called bicyclic semigroup (cf., e.g., [21]) and the one of a stack (or pushdown) is the polycyclic semigroup [26]. A comprehensive survey about the connection between storage mechanisms and monoids can be found in Zetzsche's dissertation [35]. Some basic results on the transformation monoid of reliable queues can be found in [12]. Furthermore, in [16] this author investigated the transformation monoid of lossy queues. When studying the similarities and differences between those two monoids in [18] we found it convenient to join both, the reliable and lossy queues, respectively, into one model: the so-called *partially lossy queues* (or *plq*s for short). Those are given by two disjoint finite sets: a set F of letters that can be forgotten at any time and a set U of letters that are unforgettable. We denote the corresponding transformation monoid by $\mathcal{Q}(F,U)$ and call it the *partially lossy queue monoid* or, for short, *plq monoid*. We call the subsets of this monoid *partially lossy queue languages* or *plq languages* (for short). Hence, with the help of plqs we can argue about reliable and lossy queues at the same time, which results in the unification of some proofs considering these two models.

Another main topic in the theory of automata and formal languages is the study of regular languages. This revealed strong relations to logic, combinatorics, and algebra. For example, we can generalize the notion of regularity from free monoids to arbitrary monoids. This generalization results in two notions: the rational monoid languages, which are a generalization of word languages that are described by regular expressions, and recognizable monoid languages, which are a generalization of word languages accepted by finite automata (see, e.g., [2,28]). Kleene's Theorem [15] states that both notions are equivalent in the free monoid.

In Section 4 we consider some algorithmic properties of the rational plq languages. Such properties encountered increased attention in recent years, e.g., [20] provides a survey on the rational subset membership problem. Since the rational languages in the polycyclic monoid (recall that this is the transformation monoid of a pushdown) are exactly the homomorphic images of some very simple regular languages according to [30], many decision problems like membership, intersection emptiness, universality, inclusion, and recognizability are decidable for rational languages in the polycyclic monoid. In this paper we will see that the rational membership problem of the plq monoid is NL-complete, but the other aforementioned problems are undecidable, which we can prove by reduction from their counterparts in the direct product of $(\mathbb{N}, +, 0)$ and $\{a, b\}^*$ (cf. [10,25]).

If the given plq languages are recognizable, all of the considered problems can be decided using known constructions from automata theory. Hence, the rational plq languages are not effectively recognizable. Especially, we will see that the class of rational languages in the plq monoid is not closed under intersection implying that the classes of rational and recognizable plq languages do not coincide. In contrast, in polycyclic monoids the class of rational languages is closed under Boolean operations. However, the classes of rational and recognizable languages do not coincide in these monoids since there are only two recognizable languages (the empty set and the monoid itself). But since there are even more recognizable languages in the plq monoid and since each recognizable plq language is rational as well due to McKnight's Theorem [23], it is a natural question to ask in which cases a rational plq language is recognizable.

For trace monoids, Ochmański could prove in [27] that it suffices to restrict the usage of the Kleene star in an appropriate way to characterize the recognizable trace languages. In Section 5 of this paper we will use a similar approach to characterize the recognizable plq languages in terms of special rational languages in the plq monoid. Concretely, we will define some special restrictions on the usage of Kleene star and the concatenation and have to add the complement operation to the rational expressions to finally reach this target.

Another famous characterization of the regular languages is the definability in the monadic second-order logic MSO which was proven by Büchi in [4]. This result gave us an even brighter understanding than rational expressions of the formalization of the behavior of finite automata. Similar results about trace monoids can be found in [7]. Hence, this motivates to find another MSO logic describing exactly the recognizable languages in the plq monoid. In this paper we will give such a description.

The last result in this paper regards the connection between the aperiodic languages, star-free languages, and first-order logic. Recall that a language is aperiodic if it is accepted by a counter-free finite automaton and a language is star-free if it can be generated from finite languages by application of Boolean operations and concatenation, only. Schützenberger's Theorem [31] states that both classes coincide in the free monoid. This result gives a procedure to decide whether a given regular language is star-free. Additionally, in [11] it was proven that these classes also coincide in trace monoids. In contrast to these two cases this equality does not hold in the plq monoid. But we can characterize the aperiodic languages in $\mathcal{Q}(F, U)$ with the help of the same restrictions to star-freeness as in our result regarding the rational plq languages. Finally, we prove similar to the results from [7,24] that the aperiodic plq languages in the plq monoid can be described by first-order formulas.

Note that this is the full version of the conference contribution [17].

2. Preliminaries

At first, we need some basic definitions. So, let A be an alphabet. A word $v \in A^*$ is a *prefix* of $w \in A^*$ iff $w \in vA^*$. Similarly, v is a *suffix* of w iff $w \in A^*v$ and v is an *infix* of w iff $w \in A^*vA^*$. If $w = a_1 \dots a_\ell$ with $a_1, \dots, a_\ell \in A$ we denote the *infix of w from positions i to j*, where $1 \leq i \leq \ell$ and $0 \leq j \leq \ell$ by $w[i,j] = a_i \dots a_j$. In particular, we have $w[i,j] = \varepsilon$ if $i > j$. Furthermore, v is a *subword* of w (denoted by $v \preceq w$) iff there are $\ell \in \mathbb{N}$ and $a_1, \dots, a_\ell \in A$ such that $v = a_1 \dots a_\ell$ and $w \in A^*a_1A^*a_2 \dots A^*a_\ell A^*$. Note that $\preceq$ is a partial ordering on A^*. Let $S \subseteq A$. Then we define the *projection* $\pi_S \colon A^* \to S^*$ on S by

$$\pi_S(\varepsilon) = \varepsilon \quad \text{and} \quad \pi_S(aw) = \begin{cases} a\pi_S(w) & \text{if } a \in S \\ \pi_S(w) & \text{otherwise} \end{cases}$$

for each $a \in A$ and $w \in A^*$.

2.1. *Rationality, recognizability, and aperiodicity*

Let $\mathbb{M}$ be a monoid. An $\mathbb{M}$-*language* is a subset $L \subseteq \mathbb{M}$ of $\mathbb{M}$. An $\mathbb{M}$-language $L \subseteq \mathbb{M}$ is called *rational* in $\mathbb{M}$ if it can be constructed from the finite $\mathbb{M}$-languages using union, product, and Kleene iteration. We can see the rational $\mathbb{M}$-languages as a generalization of the languages declared by regular expressions in the free monoid.

An $\mathbb{M}$-language $L \subseteq \mathbb{M}$ is *recognizable* in $\mathbb{M}$ if there are a finite monoid $\mathbb{F}$ and a homomorphism $\phi \colon \mathbb{M} \to \mathbb{F}$ such that $L = \phi^{-1}(\phi(L))$ holds. In this case we say that L is *recognized* by $\mathbb{F}$ via ϕ. We can see the finite monoid $\mathbb{F}$ as a (deterministic) finite automaton accepting L: the set of states is $\mathbb{F}$, the initial state is $\phi(e)$ (where e is the identity of $\mathbb{M}$), and the set of accepting states is $\phi(L)$. There is a transition from p to q labeled with $m \in \mathbb{M}$ iff $p \cdot \phi(m) = q$ holds in $\mathbb{F}$. Then the accepted $\mathbb{M}$-language of this automaton is L.

Now, let $\mathbb{L}$ and $\mathbb{M}$ be two monoids and $\phi \colon \mathbb{L} \to \mathbb{M}$ be a homomorphism. If $L \subseteq \mathbb{L}$ is rational then $\phi(L)$ also is rational. If $M \subseteq \mathbb{M}$ is recognizable then $\phi^{-1}(M)$ is recognizable as well. Note that in general both facts are implications. However, if ϕ is surjective, then recognizability of $\phi^{-1}(M)$ also implies the recognizability of M (cf. e.g., [28]). Additionally, if ϕ is surjective and $M \subseteq \mathbb{M}$ is rational, there also is a rational $\mathbb{L}$-language $L \subseteq \mathbb{L}$ with $\phi(L) = M$. We also know that the class of recognizable $\mathbb{M}$-languages is closed under Boolean operations.

If the monoid $\mathbb{M}$ is finitely generated (i.e., $\mathbb{M} = F^*$ for a finite $\mathbb{M}$-language F), then each recognizable $\mathbb{M}$-language is rational by [23]. For example, this applies to the partially lossy queue monoid $\mathcal{Q}(F, U)$, which we introduce in the succeeding section. The converse direction is not true in general, e.g., in Theorem 4.5 we prove the existence of a rational $\mathcal{Q}(F, U)$-language which is not recognizable. However, in free monoids generated by an alphabet A a language $L \subseteq A^*$ is rational if, and only if, it is recognizable by Kleene's Theorem [15]. In this situation, we call L *regular*.

A recognizable $\mathbb{M}$-language $L \subseteq \mathbb{M}$ is called *aperiodic* in $\mathbb{M}$ if there is $n \in \mathbb{N}$ such that for each $u, v, w \in \mathbb{M}$ we have $uv^nw \in L$ iff $uv^{n+1}w \in L$. By [24] in free monoids

$L \subseteq A^*$ is aperiodic iff it is accepted by a counter-free finite automaton on A. It is an easy exercise to prove that the class of aperiodic $\mathbb{M}$-languages is closed under Boolean operations and homomorphic preimages. By Schützenberger's Theorem [31] a language $L \subseteq A^*$ is aperiodic iff it is star-free. Note that an $\mathbb{M}$-language $L \subseteq \mathbb{M}$ is *star-free* in $\mathbb{M}$ if it can be constructed from finite $\mathbb{M}$-languages using union, product, and complementation.

2.2. *Logic and languages*

In this subsection we recall Büchi's logics on words and their correspondence to languages. Concretely, we model words over an alphabet A as linear orders on the letter positions labeled with letters.

Let A be an alphabet. By FO we denote the set of first-order formulas built up from the atomic formulas of the form

$$x = y\,, \qquad x < y\,, \text{ and } \qquad \Lambda_a(x) \text{ for } a \in A$$

where x and y are variables. To simplify notation we write $\Lambda_S(x)$ instead of $\bigvee_{a\in S} \Lambda_a(x)$ for any non-empty set $S \subseteq A$. Moreover, we write $x \leq y$ instead of $x < y \vee x = y$.

Now let $w = a_1 \dots a_n \in A^*$. The *word model* of w is the relational structure

$$\underline{w} := (\mathrm{dom}(w), <^w, (\Lambda_a^w)_{a\in A})$$

where $\mathrm{dom}(w) := \{1, \dots, n\}$ is the set of letter positions of w, $<^w$ is the natural (strict) order on $\mathrm{dom}(w)$, and $\Lambda_a^w := \{i \in \mathrm{dom}(w) \mid a_i = a\}$ is the set of positions in w labeled with the letter a.

For a formula $\phi \in \mathsf{FO}$ with free variables $x_1, \dots, x_n$ we also write $\phi(x_1, \dots, x_n)$. For a formula $\phi(x_1, \dots, x_n) \in \mathsf{FO}$, a word $w \in A^*$, and positions $p_1, \dots, p_n \in \mathrm{dom}(w)$ we write $(\underline{w}, p_1, \dots, p_n) \models \phi(x_1, \dots, x_n)$ (we say that "ϕ is satisfied in $\underline{w}$") if ϕ evaluates to true on interpretation of $=, <, \Lambda_a$ as equality, $<^w$, and Λ_a^w, respectively, and on interpretation of the free variables as $p_1, \dots, p_n$.

A formula without free variables is called a *sentence*. Then the language defined by the sentence $\phi \in \mathsf{FO}$ is $L(\phi) := \{w \in A^* \mid \underline{w} \models \phi\}$. We say that a language $L \subseteq A^*$ is FO*-definable* if there is $\phi \in \mathsf{FO}$ with $L = L(\phi)$.

By MSO (the *monadic second-order logic*) we denote the second-order extension of FO where all second-order variables are unary. Again, we say that $L \subseteq A^*$ is MSO*-definable* if there is $\phi \in \mathsf{MSO}$ with $L = L(\phi)$.

By [4] a language is regular iff it is MSO-definable. Moreover, a language is star-free (and, hence, aperiodic) iff it is FO-definable according to [24]. For further details on this word logic see, for example, [34].

3. Partially Lossy Queues

A *(fifo-)queue* is an abstract data type which is able to store entries from an alphabet A. The content of such queue is a finite sequence of entries, i.e., a word from

A^*. We can modify the content of a queue with the help of two *actions* for each letter $a \in A$: we can write the letter a (denoted by a) into the queue by appending this letter to the content. So, for a content $w \in A^*$ the result of such write action of a is wa. Additionally, we can read the letter a (denoted by $\overline{a}$) from the queue by removing this letter from the first position of the queue's content. So, reading a from the content $aw \in A^*$ yields the new content w. Note that the queue blocks if w does not start with the letter a (i.e., if $w = \varepsilon$ or $w \in bA^*$ with $b \in A \setminus \{a\}$).

A *partially lossy queue* (or *plq*, for short) is a queue which additionally has an uncontrollable action: it can forget some of its entries at any time of its computation. The entries which can be forgotten are specified by a so-called *lossiness alphabet*:

Definition 3.1. *A* lossiness alphabet *is a tuple* $\mathcal{L} = (F, U)$ *where* F *and* U *are two finite sets with* $F \cap U = \emptyset$. *We call* F *the set of* forgettable *queue entries and* U *the set of* unforgettable *entries.*

From a given lossiness alphabet $\mathcal{L} = (F, U)$ we obtain the alphabet $A_{\mathcal{L}} := F \cup U$ of all possible queue entries. By the alphabet $\Sigma_{\mathcal{L}} := A_{\mathcal{L}} \cup \overline{A_{\mathcal{L}}}$ we denote the set of all controllable queue actions, where $\overline{A_{\mathcal{L}}} := \{\overline{a} \mid a \in A_{\mathcal{L}}\}$ is a disjoint copy of $A_{\mathcal{L}}$ containing the read actions. For a word $w = a_1 a_2 \dots a_n \in A_{\mathcal{L}}^*$ we write $\overline{w} := \overline{a_1}\,\overline{a_2} \dots \overline{a_n}$.

To model the behavior of such partially lossy queue, we require the notion of so-called $\mathcal{L}$-*subwords*: A word $v \in A_{\mathcal{L}}^*$ is an $\mathcal{L}$-*subword* of $w \in A_{\mathcal{L}}^*$ (denoted by $v \preceq_{\mathcal{L}} w$) if we have $\pi_U(w) \preceq v \preceq w$ holds, i.e., if v is a subword of w which still contains all unforgettable letters from w. It is easy to see that $\preceq_{(F,\emptyset)}$ is the subword relation and $\preceq_{(\emptyset,U)}$ is the equality relation.

Next, we model the actions of a plq with the help of a map $\Delta_{\mathcal{L}} \colon A_{\mathcal{L}}^* \times \Sigma_{\mathcal{L}}^* \to 2^{A_{\mathcal{L}}^*}$ where $\Delta_{\mathcal{L}}(w, t)$ contains the set of all possible queue entries after application of the transformation sequence t: for $w \in A_{\mathcal{L}}^*$, $a \in A_{\mathcal{L}}$, and $t \in \Sigma_{\mathcal{L}}^*$ we define

- $\Delta_{\mathcal{L}}(w, \varepsilon) := \{v \in A_{\mathcal{L}}^* \mid v \preceq_{\mathcal{L}} w\}$,
- $\Delta_{\mathcal{L}}(w, at) := \{v \in A_{\mathcal{L}}^* \mid \exists w' \in \Delta_{\mathcal{L}}(w, t) \colon v \preceq_{\mathcal{L}} w'a\}$, and
- $\Delta_{\mathcal{L}}(w, \overline{a}t) := \{v \in A_{\mathcal{L}}^* \mid \exists aw' \in \Delta_{\mathcal{L}}(w, t) \colon v \preceq_{\mathcal{L}} w'\}$.

Example 3.2. Let $\mathcal{L} = (\{a\}, \{b\})$ be a lossiness alphabet. Then we have $\Delta_{\mathcal{L}}(aba, \varepsilon) = \{aba, ab, ba, b\}$, $\Delta_{\mathcal{L}}(aba, \overline{a}) = \{ba, b\}$, and $\Delta_{\mathcal{L}}(aba, \overline{b}) = \{a, \varepsilon\}$.

Now, we want to define the transformation monoid of a partially lossy queue. To this end, we say that two action sequences $s, t \in \Sigma_{\mathcal{L}}^*$ *behave equivalently* (denoted by $s \equiv_{\mathcal{L}} t$) if, and only if, we have $\Delta_{\mathcal{L}}(., s) = \Delta_{\mathcal{L}}(., t)$. In other words, s and t have the same behavior iff they cannot be distinguished by any queue content. This equivalence relation is even a congruence on the free monoid $\Sigma_{\mathcal{L}}^*$. Then the transformation monoid of a plq or the *partially lossy queue monoid* (or *plq monoid*, for short) is the quotient monoid $\mathcal{Q}(\mathcal{L}) := \Sigma_{\mathcal{L}}^* /_{\equiv_{\mathcal{L}}}$. The natural epimorphism on this congruence is $\eta_{\mathcal{L}} \colon \Sigma_{\mathcal{L}}^* \to \mathcal{Q}(\mathcal{L}) \colon t \mapsto [t]_{\equiv_{\mathcal{L}}}$ where $[t]_{\equiv_{\mathcal{L}}} := \{s \in \Sigma_{\mathcal{L}}^* \mid s \equiv_{\mathcal{L}} t\}$ is the equivalence class of t wrt. $\equiv_{\mathcal{L}}$. Since the equivalence classes of $\equiv_{\mathcal{L}}$ identify all action sequences

$t \in \Sigma_{\mathcal{L}}^*$ having the same semantics $\Delta_{\mathcal{L}}(.,t)$, we also call them *transformations*. In the following we consider the rational, recognizable, and aperiodic $\mathcal{Q}(\mathcal{L})$-languages. We also call these languages *partially lossy queue languages* (or *plq languages*, for short).

We know from [18] the following characterization of the behavioral equivalence $\equiv_{\mathcal{L}}$:

Theorem 3.3 ([18, Theorems 3.5 and 3.15]). *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$. The relation $\equiv_{\mathcal{L}}$ is the least congruence on $\Sigma_{\mathcal{L}}^*$ satisfying the following equations for $a, b \in A_{\mathcal{L}}$ and $w \in A_{\mathcal{L}}^*$:*

(i) $b\overline{a} \equiv_{\mathcal{L}} \overline{a}b$ *if* $a \neq b$
(ii) $a\overline{ab} \equiv_{\mathcal{L}} \overline{a}a\overline{b}$
(iii) $bwa\overline{a} \equiv_{\mathcal{L}} bw\overline{a}a$ *if* $b \in U \cup \{a\}$ □

Remark 3.4. Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $A_{\mathcal{L}} = \{a\}$. Then a partially lossy queue on this alphabet acts (independently of $a \in F$ or $a \in U$) like a partially blind counter: writing the letter a increases the number of a's in the queue's content, reading decreases this number if the queue contains at least one a. Hence, the corresponding transformation monoid is the bicyclic semigroup. Since this monoid is already well-known we do not consider this case in this paper.

To handle the equivalence classes of $\equiv_{\mathcal{L}}$ we want to define a normal form on this congruence. We do this by ordering the equations from Theorem 3.3 from left to right, which results in a semi-Thue system called $\mathfrak{R}_{\mathcal{L}}$.

Since the rules of $\mathfrak{R}_{\mathcal{L}}$ are length-preserving and move read actions to the left, it is terminating[a]. Moreover, it is locally confluent[b] by [18, Lemma 3.13] and hence confluent. Therefore, for any word $t \in \Sigma_{\mathcal{L}}^*$ there is a unique, irreducible word $\mathrm{nf}_{\mathcal{L}}(t)$ with $t \to^* \mathrm{nf}_{\mathcal{L}}(t)$, the so-called *normal form* of t.

Example 3.5. Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$, $a, b \in A_{\mathcal{L}}$ with $a \neq b$, and $t = aabb\overline{a}\overline{b}$. If $a \in F$ then we have

$$aabb\overline{a}\overline{b} \xrightarrow{(i)} aab\overline{a}b\overline{b} \xrightarrow{(i)} aa\overline{a}bb\overline{b} \xrightarrow{(iii)} a\overline{a}abb\overline{b} \xrightarrow{(iii)} a\overline{a}ab\overline{b}b$$

and therefore $\mathrm{nf}_{\mathcal{L}}(aabb\overline{a}\overline{b}) = a\overline{a}ab\overline{b}b$. Otherwise, i.e., if $a \in U$, we can extend this derivation as follows:

$$a\overline{a}ab\overline{b}b \xrightarrow{(iii)} a\overline{a}a\overline{b}bb \xrightarrow{(i)} a\overline{a}\overline{b}abb \xrightarrow{(ii)} \overline{a}a\overline{b}abb \xrightarrow{(i)} \overline{a}\overline{b}aabbb$$

and, hence, we obtain $\mathrm{nf}_{\mathcal{L}}(aabb\overline{a}\overline{b}) = \overline{a}\overline{b}aabb$.

From the definition of $\mathfrak{R}_{\mathcal{L}}$ we obtain that a word is in normal form if, and only if, it starts with some read actions followed by a special shuffle of write and read

[a] A semi-Thue system is *terminating* (or *noetherian*) if it has no infinite derivations.
[b] A semi-Thue system over A is *confluent* if for each $u, v, w \in A^*$ with $u \leftarrow^* v \to^* w$ there is $x \in A^*$ with $u \to^* x \leftarrow^* w$. It is *locally confluent* if this conclusion holds at least for the case $u \leftarrow v \to w$.

actions where each read action $\overline{a}$ appears directly right from a, and ends with some write actions. At this juncture, the infixes $a\overline{a}$ in these words are separated by words from $(F \setminus \{a\})^*$, only. Formally, such shuffle of $u \in A_{\mathcal{L}}^*$ and $\overline{v} \in \overline{A_{\mathcal{L}}}^*$ is defined by

$$\langle\!\langle u, \overline{v} \rangle\!\rangle := w_1 a_1 \overline{a_1} w_2 a_2 \overline{a_2} \dots w_\ell a_\ell \overline{a_\ell} \,,$$

where $v = a_1 \dots a_\ell$, $a_1, \dots, a_\ell \in A_{\mathcal{L}}$, $u = w_1 a_1 \dots w_\ell a_\ell$, and $w_i \in (F \setminus \{a_i\})^*$ for each $1 \leq i \leq \ell$. We can see that this special shuffle $\langle\!\langle u, \overline{v} \rangle\!\rangle$ is defined only if v is a special $\mathcal{L}$-subword of u. Concretely, this shuffle is defined if u is a so-called *reduced $\mathcal{L}$-superword* of v: for $v = a_1 a_2 \dots a_n$ the set of all *reduced $\mathcal{L}$-superwords* is

$$\mathrm{redsup}_{\mathcal{L}}(v) := \{ w_1 a_1 w_2 a_2 \dots w_n a_n \mid \forall 1 \leq i \leq n \colon w_i \in (F \setminus \{a_i\})^* \} \,.$$

Note that for each $u \in \mathrm{redsup}_{\mathcal{L}}(v)$ we also have $v \preceq_{\mathcal{L}} u$. However, in the general case $v \preceq_{\mathcal{L}} u$ does not imply $u \in \mathrm{redsup}_{\mathcal{L}}(v)$. For example, we have $\varepsilon \preceq_{\mathcal{L}} a$ if $a \in F$ but $a \notin \mathrm{redsup}_{\mathcal{L}}(\varepsilon) = \{\varepsilon\}$. If $F = \emptyset$ holds, then each word $v \in A_{\mathcal{L}}^*$ has exactly one reduced $\mathcal{L}$-superword $\mathrm{redsup}_{\mathcal{L}}(v) = \{v\}$.

The set of all normal forms is

$$\begin{aligned} \mathrm{NF}_{\mathcal{L}} &:= \{ \overline{u} \langle\!\langle v, \overline{w} \rangle\!\rangle x \mid u, v, w, x \in A_{\mathcal{L}}^*, v \in \mathrm{redsup}_{\mathcal{L}}(w) \} \\ &= \overline{A_{\mathcal{L}}}^* \left(\bigcup_{a \in A_{\mathcal{L}}} (F \setminus \{a\})^* a \overline{a} \right)^* A_{\mathcal{L}}^* \,. \end{aligned}$$

From this equation we can infer that $\mathrm{nf}_{\mathcal{L}}(t) = \overline{w_1} \langle\!\langle w_2, \overline{w_3} \rangle\!\rangle w_4$ is characterized by three components: The first component is the projection to the write actions $\mathrm{wrt}(t) := w_2 w_4 = \pi_A(t)$ (note that the transitions of $\mathfrak{R}_{\mathcal{L}}$ preserve the relative ordering of the write operations). Similarly, the second one is the projection to the read actions $\mathrm{rd}(t) := w_1 w_3$ (note that we suppress the overlines in this projection). Finally, the third component is the *overlap* $\mathrm{rd}_2(t) := w_3$ of t. From these three components we also obtain the words $\mathrm{rd}_1(t) := w_1$ (note that $\mathrm{rd}(t) = \mathrm{rd}_1(t)\mathrm{rd}_2(t)$), $\mathrm{wrt}_1(t) := w_2$, and $\mathrm{wrt}_2(t) := w_4$ (note that $\mathrm{wrt}(t) = \mathrm{wrt}_1(t)\mathrm{wrt}_2(t)$). We also know $\mathrm{wrt}_1(t) \in \mathrm{redsup}_{\mathcal{L}}(\mathrm{rd}_2(t))$ by the argumentation above.

Example 3.6. Recall Example 3.5. There, in case of $a \in F$ we have for $t = aabb\overline{a}\overline{b}$: $\mathrm{wrt}(t) = aabb$, $\mathrm{rd}(t) = ab$, $\mathrm{rd}_1(t) = \varepsilon$, $\mathrm{rd}_2(t) = ab$, $\mathrm{wrt}_1(t) = aab$, and $\mathrm{wrt}_2(t) = b$. Otherwise, if $a \in U$ we have $\mathrm{rd}_1(t) = ab$, $\mathrm{rd}_2(t) = \varepsilon = \mathrm{wrt}_1(t)$, and $\mathrm{wrt}_2(t) = aabb$.

While $\mathrm{rd}_1(t)$ is defined using the semi-Thue system $\mathfrak{R}_{\mathcal{L}}$, it also has a natural meaning: $\mathrm{rd}_1(t)$ is the shortest queue content $w \in A_{\mathcal{L}}^*$ such that $\Delta_{\mathcal{L}}(w, t) \neq \emptyset$, i.e., it is the shortest content where the plq is able to fully apply the actions in t.

The following connections between $\mathfrak{R}_{\mathcal{L}}$ and $\mathrm{nf}_{\mathcal{L}}(w)$ are important for the proofs in this paper:

Proposition 3.7 ([18, Theorem 3.15]). *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet* *with* $|A_\mathcal{L}| \geq 2$ *and let* $v, w \in \Sigma_\mathcal{L}^*$. *Then we have*

$$\begin{aligned} v \equiv_\mathcal{L} w &\iff \mathrm{nf}_\mathcal{L}(v) = \mathrm{nf}_\mathcal{L}(w) \\ &\iff (\mathrm{wrt}(v), \mathrm{rd}(v), \mathrm{rd}_2(v)) = (\mathrm{wrt}(w), \mathrm{rd}(w), \mathrm{rd}_2(w)) . \end{aligned}$$

□

With this main property in mind we can also apply wrt, wrt_i, rd, and rd_i to equivalence classes of $\equiv_\mathcal{L}$ (i.e., elements from $\mathcal{Q}(\mathcal{L})$) instead of words from $\Sigma_\mathcal{L}^*$.

Another question is the description of the word $u\overline{v}$ for any $u, v \in A_\mathcal{L}^*$. We have $\mathrm{wrt}(u\overline{v}) = u$ and $\mathrm{rd}(u\overline{v}) = v$. It remains to describe the overlap $\mathrm{rd}_2(u\overline{v})$:

Lemma 3.8 ([18, Lemma 3.19]). *Let* $\mathcal{L} = (F, U)$ *be a lossiness alphabet with* $|A_\mathcal{L}| \geq 2$ *and* $u, v \in A_\mathcal{L}^*$. *Then* $\mathrm{rd}_2(u\overline{v})$ *is the longest suffix* r_2 *of* v *such that there is a prefix* w_1 *of* u *satisfying* $w_1 \in \mathrm{redsup}_\mathcal{L}(r_2)$. □

Since $\equiv_\mathcal{L}$ is a congruence we can infer $t \equiv_\mathcal{L} \overline{\mathrm{rd}_1(t)}\, \mathrm{wrt}(t)\, \overline{\mathrm{rd}_2(t)}$ for each $t \in \Sigma_\mathcal{L}^*$ from Lemma 3.8.

4. Algorithmic Properties of Rational PLQ Languages

This section studies decision problems concerning the rational partially lossy queue languages. We will see that the classes of rational and recognizable plq languages do not coincide. Especially, we prove that it is undecidable whether a given rational plq language is recognizable. We also show that emptiness of intersections and the unique decipherability in $\mathcal{Q}(\mathcal{L})$ are undecidable. Though, we will see first, that the rational membership problem is NL-complete.

So, let $t \in \Sigma_\mathcal{L}^*$. In our next lemma we want to construct an NFA accepting the equivalence class $[t]_{\equiv_\mathcal{L}}$. In this NFA we map the most of its states to the left-divisors of $[t]_{\equiv_\mathcal{L}}$. The result is, that our NFA has approximately as many states as $[t]_{\equiv_\mathcal{L}}$ has left-divisors. So we should first count the number of left-divisors: to this end, recall that $[t]_{\equiv_\mathcal{L}}$ is fully described by the words $\mathrm{wrt}(t)$, $\mathrm{rd}(t)$, and $\mathrm{rd}_2(t)$ (cf. Proposition 3.7). Since all of these words have length $\mathcal{O}(|t|)$, we also learn that $[t]_{\equiv_\mathcal{L}}$ has at most $\mathcal{O}(|t|^3)$ many left-divisors. Therefore, the constructed NFA also has $\mathcal{O}(|t|^3)$ many states. We will also see in the following that our construction requires only logarithmic temporary space for the construction of this NFA. Note that the proof of this lemma is very close to the proof of [12, Lemma 8.1], which states this result for reliable queues, only.

Lemma 4.1. *Let* $\mathcal{L} = (F, U)$ *be a lossiness alphabet. From* $t \in \Sigma_\mathcal{L}^*$, *one can construct an NFA accepting* $[t]_{\equiv_\mathcal{L}}$ *using logarithmic space, only.*

Proof. Let $t = \alpha_1 \dots \alpha_n$ and let $0 \leq i, j, k \leq n$ be natural numbers. For the triple $p = (i, j, k)$ we define the words $p_1, p_2, p_3 \in A_\mathcal{L}^*$ as follows:

- $p_1 := \mathrm{wrt}(t[1, k])$ as well as
- $p_2 := \mathrm{rd}(t[1, i])$ and $p_3 := \mathrm{rd}(t[i + 1, j])$.

We say that a triple $p = (i, j, k)$ is *good* if, and only if,

(1) there is a prefix p_1' of p_1 with $p_1' \in \mathrm{redsup}_{\mathcal{L}}(p_3)$,
(2) $i = 0$ or $a_i \in \overline{A_{\mathcal{L}}}$ and similarly $j = 0$ or $a_j \in \overline{A_{\mathcal{L}}}$, and
(3) $k = 0$ or $a_k \in A_{\mathcal{L}}$.

We show first, that each good triple p describes an equivalence class $[t_p]_{\equiv_{\mathcal{L}}}$ such that $\mathrm{wrt}(t_p) = p_1$, $\mathrm{rd}(t_p) = p_2p_3$, and $\mathrm{rd}_2(t_p) = p_3$ holds. From the choice of p_1, p_2, and p_3 we learn that p_1 is a prefix of $\mathrm{wrt}(t)$ and p_2p_3 is a prefix of $\mathrm{rd}(t)$. Additionally, since p is good there is a prefix p_1' of p_1 with $p_1' \in \mathrm{redsup}_{\mathcal{L}}(p_3)$. Let p_4 be the complementary suffix of p_1 wrt. p_1'. Then we set $t_p := \overline{p_2}\langle\!\langle p_1', \overline{p_3}\rangle\!\rangle p_4 = \mathrm{nf}_{\mathcal{L}}(t_p)$.

Next, we show that for each $s \in \Sigma_{\mathcal{L}}^*$ where $\mathrm{wrt}(s)$ and $\mathrm{rd}(s)$ are prefixes of $\mathrm{wrt}(t)$ resp. $\mathrm{rd}(t)$ there is a good triple p with $s \equiv t_p$. To this end, recall $\mathrm{rd}(s) = \mathrm{rd}_1(s)\mathrm{rd}_2(s)$, i.e., $\mathrm{rd}(t)$ start with $\mathrm{rd}_1(s)\mathrm{rd}_2(s)$. So, if $\mathrm{rd}_1(s) \neq \varepsilon$ there is $0 < i \leq n$ with $\mathrm{rd}(t[1, i]) = \mathrm{rd}_1(s)$ and $\alpha_i \in \overline{A_{\mathcal{L}}}$. Otherwise we set $i := 0$. If $\mathrm{rd}_2(s) \neq \varepsilon$ there is $i < j \leq n$ with $\mathrm{rd}(t[i + 1, j]) = \mathrm{rd}_2(s)$ and $\alpha_j \in \overline{A_{\mathcal{L}}}$. Otherwise we also set $j := 0$. Finally, if $\mathrm{wrt}(s) \neq \varepsilon$ there is $1 < k \leq n$ with $\mathrm{wrt}(t[1, k]) = \mathrm{wrt}(s)$ and $\alpha_k \in A_{\mathcal{L}}$ since $\mathrm{wrt}(s)$ is a prefix of $\mathrm{wrt}(t)$. Otherwise we set $k := 0$. Then the tuple $p := (i, j, k)$ is good since $p_1' := \mathrm{wrt}_1(s)$ is a prefix of $p_1 = \mathrm{wrt}(s)$ and $p_1' = \mathrm{wrt}_1(s) \in \mathrm{redsup}_{\mathcal{L}}(\mathrm{rd}_2(s)) = \mathrm{redsup}_{\mathcal{L}}(p_3)$. Additionally, we can see $\mathrm{wrt}(s) = p_1 = \mathrm{wrt}(t_p)$, $\mathrm{rd}(s) = p_2p_3 = \mathrm{rd}(t_p)$, and $\mathrm{rd}_2(s) = p_3 = \mathrm{rd}_2(t_p)$ implying $s \equiv_{\mathcal{L}} t_p$.

Now, the set of states of our NFA consists of all good triples $p = (i, j, k)$ and a unique error state $\perp$. The only initial state of the NFA is $\iota := (0, 0, 0)$ (this ensures $t_\iota \equiv_{\mathcal{L}} \varepsilon$). A state $p = (i, j, k)$ is accepting if, and only if, $t_p \equiv t$ holds. Next, we want to define the transitions of the automaton such that, after reading $s \in \Sigma_{\mathcal{L}}^*$, the automaton reaches a state p with $t_p \equiv_{\mathcal{L}} s$, provided that such state exists. Furthermore, we want to make sure that such a state exists whenever $[s]_{\equiv_{\mathcal{L}}}$ is a left-divisor of $[t]_{\equiv_{\mathcal{L}}}$.

So, let $p = (i, j, k)$ be a state (i.e., p is good) and $a \in A_{\mathcal{L}}$. To define the state reached from p after writing a, let $k' > k$ be the minimal write-position in t after k. In other words, we have $k' > k$, $\alpha_{k'} \in A_{\mathcal{L}}$, and $t[k + 1, k' - 1] \in \overline{A_{\mathcal{L}}}^*$. If there is no such k' or if $\alpha_{k'} \neq a$, then the NFA ends up in the error state $\perp$. Otherwise, it moves to $q := (i, j, k')$, which is a good triple and, therefore, a state of the automaton, again. Then we can see:

$$\begin{aligned}
t_p \cdot a &\equiv_{\mathcal{L}} \overline{\mathrm{rd}(t[1, i])}\, \mathrm{wrt}(t[1, k])\, \overline{\mathrm{rd}(t[i + 1, j])} \cdot a && \text{(by Lemma 3.8)} \\
&\equiv_{\mathcal{L}} \overline{\mathrm{rd}(t[1, i])}\, \mathrm{wrt}(t[1, k])a\, \overline{\mathrm{rd}(t[i + 1, j])} && \text{(by Lemma 3.8)} \\
&= \overline{\mathrm{rd}(t[1, i])}\, \mathrm{wrt}(t[1, k'])\, \overline{\mathrm{rd}(t[i + 1, j])} && \\
&\equiv_{\mathcal{L}} t_q\,. && \text{(by Lemma 3.8)}
\end{aligned}$$

Next, we define which state is reached from p after reading $\overline{a}$. Let $j' > j$ be the minimal read-position in t after j. In other words, we have $j' > j$, $\alpha_{j'} \in \overline{A_{\mathcal{L}}}$, and $t[j+1, j'-1] \in A_{\mathcal{L}}^*$. If there is no such j' or if $\alpha_{j'} \neq \overline{a}$, then the NFA ends up in the error state $\perp$. So, assume, that such j' exists and $\alpha_{j'} = \overline{a}$. Additionally let $i' \geq i$

with $i' = 0$ or $\alpha_{i'} \in \overline{A_{\mathcal{L}}}$ be minimal such that there is a prefix w_1 of $\mathrm{wrt}(t[1,k])$ with $w_1 \in \mathrm{redsup}_{\mathcal{L}}(\mathrm{rd}(t[i'+1,j']))$. Then $q := (i', j', k)$ is a good triple and, hence, a valid state of the NFA. We add an $\overline{a}$-edge from p to q in this case. Then we get

$$\begin{aligned} t_p \cdot \overline{a} &\equiv_{\mathcal{L}} \overline{\mathrm{rd}(t[1,i])}\,\mathrm{wrt}(t[1,k])\,\overline{\mathrm{rd}(t[i+1,j])} \cdot \overline{a} && \text{(by Lemma 3.8)} \\ &= \overline{\mathrm{rd}(t[1,i])}\,\mathrm{wrt}(t[1,k])\,\overline{\mathrm{rd}(t[i+1,j'])} \\ &\equiv_{\mathcal{L}} \overline{\mathrm{rd}(t[1,i'])}\,\mathrm{wrt}(t[1,k])\,\overline{\mathrm{rd}(t[i'+1,j'])} && \text{(by Lemma 3.8)} \\ &\equiv_{\mathcal{L}} t_q\,. && \text{(by Lemma 3.8)} \end{aligned}$$

This finishes the construction of the NFA.

Now let $s \in \Sigma_{\mathcal{L}}^*$. If there is an s-labeled path from $\iota = (0,0,0)$ to a non-error state q, we obtain $s \equiv_{\mathcal{L}} t_q$ by induction on $|s|$ from the above calculations. In particular, if q is an accepting state of our NFA, we know $t_q \equiv t$ implying $s \in [t]_{\equiv_{\mathcal{L}}}$.

Now let $[s]_{\equiv_{\mathcal{L}}}$ be a left-divisor of $[t]_{\equiv_{\mathcal{L}}}$. Then $\mathrm{wrt}(s)$ and $\mathrm{rd}(s)$ are prefixes of $\mathrm{wrt}(t)$ and $\mathrm{rd}(t)$, respectively, since wrt and rd are homomorphisms. Then by induction on $|s|$ we obtain an s-labeled path from ι to a state p with $s \equiv_{\mathcal{L}} t_p$. In particular, if $s \in [t]_{\equiv_{\mathcal{L}}}$, then $t_p \equiv_{\mathcal{L}} s \equiv_{\mathcal{L}} t$, i.e., p is accepting. Thus, the NFA accepts $[t]_{\equiv_{\mathcal{L}}}$.

By the construction of the NFA, it is clear that a Turing-machine with t on its input tape can, using logarithmic space on its work tape, write the list of all transitions on its one-way output tape. □

Note that the NFA we have constructed in the proof of Lemma 4.1 is actually deterministic. However, we do not require determinism for the following statement. This theorem states that the rational membership problem is decidable in the plq monoid:

Theorem 4.2. *The following* rational membership problem *for plq monoids is* NL-*complete:*

Input: *A lossiness alphabet* $\mathcal{L} = (F, U)$*, an action sequence* $t \in \Sigma_{\mathcal{L}}^*$*, and an NFA* $\mathfrak{A}$ *over* $\Sigma_{\mathcal{L}}$

Question: *Is there an action sequence* $s \in L(\mathfrak{A})$ *with* $s \equiv_{\mathcal{L}} t$*?*

Proof. Let $t \in \Sigma_{\mathcal{L}}^*$ and let $\mathfrak{A}$ be an NFA over $\Sigma_{\mathcal{L}}$. Let $\mathfrak{B}$ be the NFA accepting $[t]_{\equiv_{\mathcal{L}}}$ from Lemma 4.1 that can be constructed using only logarithmic additional space.

Then there exists $s \in L(\mathfrak{A})$ with $s \equiv_{\mathcal{L}} t$ if, and only if, $L(\mathfrak{A}) \cap [t]_{\equiv_{\mathcal{L}}} \neq \emptyset$ if, and only if, $L(\mathfrak{A}) \cap L(\mathfrak{B}) \neq \emptyset$. Using an on-the-fly construction of $\mathfrak{B}$, this can be decided non-deterministically in logarithmic space. Hence, the problem is in NL.

Since the free monoid $A_{\mathcal{L}}^*$ embeds into $\mathcal{Q}(\mathcal{L})$ and since the rational membership problem for $A_{\mathcal{L}}^*$ is NL-hard [13], we also get NL-hardness for $\mathcal{Q}(\mathcal{L})$. □

Now we will prove some negative algorithmic results on rational plq languages. Concretely, we show that unique decipherability, intersection emptiness, universality, and recognizability are undecidable for rational languages in the plq monoid. Huschenbett et al. inferred in [12, Section 8] similar results for reliable queues from an embedding $\{\alpha,\beta\}^* \times \{\gamma,\delta\}^*$ into $\mathcal{Q}(\emptyset, U)$ for any at least binary alphabet U. Unfortunately, this does not work in arbitrary plq monoids since this direct product does not embed into $\mathcal{Q}(\{a,b\},\emptyset)$ by [18, Theorem 6.14]. Though, we can prove all the undecidability results considered in [12] for any plq monoid.

Some of these results are based on an embedding of the monoid $\{\alpha\}^* \times \{\beta,\gamma\}^*$ into $\mathcal{Q}(\mathcal{L})$. Unfortunately, this does not help for unique decipherability and intersection emptiness since their counterparts in $\{\alpha\}^* \times \{\beta,\gamma\}^*$ are decidable. Hence, we have to prove them directly.

The first considered decision problem is the unique decipherability problem in $\mathcal{Q}(\mathcal{L})$, i.e., the question whether a given finite language T freely generates T^*. In this context, in a monoid $\mathbb{M}$ a set $T \subseteq \mathbb{M}$ *freely generates* T^* if for each $m, n \in \mathbb{N}$ and $s_1, \ldots, s_m, t_1, \ldots, t_n \in T$ with $s_1 \ldots s_m = t_1 \ldots t_n$ we have $m = n$ and $s_i = t_i$ for each $1 \leq i \leq m = n$. We prove that the unique decipherability problem in $\mathcal{Q}(\mathcal{L})$ is undecidable. To this end, we will use the undecidability of this problem in $\{\alpha,\beta\}^* \times \{\gamma,\delta\}^*$. So, let S be a finite language. We obtain a finite plq language $T \subseteq \mathcal{Q}(\mathcal{L})$ with the help of a special encoding of the elements in S and by addition of another transformation. Finally, we can show that S^* is freely generated by S iff this similarly holds for T.

Theorem 4.3. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$. The* unique decipherability problem *in $\mathcal{Q}(\mathcal{L})$ is undecidable:*

Input: *A finite set $T \subseteq \mathcal{Q}(\mathcal{L})$*
Question: *Is T^* freely generated by T?*

Proof. We prove this undecidability by reduction of this question for the monoid $\{\alpha,\beta\}^* \times \{\gamma,\delta\}^*$, which is undecidable by [6, Theorem 3.1]. So, let $a, b \in A_{\mathcal{L}}$ be distinct letters and

$$S := \{(x_1, y_1), \ldots, (x_k, y_k)\} \subseteq \{\alpha,\beta\}^* \times \{\gamma,\delta\}^* .$$

Define the embeddings $f\colon \{\alpha,\beta\}^* \to A_{\mathcal{L}}^*$ and $g\colon \{\gamma,\delta\}^* \to \overline{A_{\mathcal{L}}^*}$ by $f(\alpha) = g(\gamma) = aa$ and $f(\beta) = g(\delta) = ab$. Set $t_0 := [bb\overline{bbb}]_{\equiv_{\mathcal{L}}}$, $t_i := [f(x_i)\overline{g(y_i)}]_{\equiv_{\mathcal{L}}}$ for any $1 \leq i \leq k$, and

$$T := \{t_i \mid 0 \leq i \leq k\} \subseteq \mathcal{Q}(\mathcal{L}) .$$

We show now that S^* is freely generated by S if, and only if, T^* is freely generated by T.

First we assume that S^* is not freely generated by S. Then there are indices $(i_1, \ldots, i_m) \neq (j_1, \ldots, j_n)$ where $m, n > 0$ such that

$$(x_{i_1} \ldots x_{i_m}, y_{i_1} \ldots y_{i_m}) = (x_{j_1} \ldots x_{j_n}, y_{j_1} \ldots y_{j_n}) .$$

Let $\ell := |f(x_{i_1} \dots x_{i_m})|$. Set $p := t_{i_1} \dots t_{i_m} \cdot t_0^\ell$ and $q := t_{j_1} \dots t_{j_n} \cdot t_0^\ell$. It is a simple exercise to prove $\mathrm{nf}_{\mathcal{L}}(t_0^\ell) = \overline{b}^\ell \langle\!\langle b^{2\ell}, \overline{b^{2\ell}} \rangle\!\rangle$. Additionally, by the choice of the two sequences of indices we have

$$\begin{aligned} \mathrm{wrt}(t_{i_1} \dots t_{i_m}) = f(x_{i_1} \dots x_{i_m}) = f(x_{j_1} \dots x_{j_n}) = \mathrm{wrt}(t_{j_1} \dots t_{j_n}) \quad \text{and} \\ \mathrm{rd}(t_{i_1} \dots t_{i_m}) = g(y_{i_1} \dots y_{i_m}) = g(y_{j_1} \dots y_{j_n}) = \mathrm{rd}(t_{j_1} \dots t_{j_n})\,. \end{aligned} \tag{4.1}$$

Now, we consider the overlap rd_2 of the concatenation of $t_{i_1} \dots t_{i_m}$ (resp., $t_{j_1} \dots t_{j_n}$) with $[\overline{b}^\ell]_{\equiv_{\mathcal{L}}}$. To this end, we utilize Lemma 3.8. So, we have to determine $\mathrm{rd}_2(t_{i_1} \dots t_{i_m} \cdot [\overline{b}^\ell]_{\equiv_{\mathcal{L}}})$ and $\mathrm{rd}_2(t_{j_1} \dots t_{j_n} \cdot [\overline{b}^\ell]_{\equiv_{\mathcal{L}}})$. Since $|\mathrm{wrt}(t_{i_1} \dots t_{i_m})|_b < \ell$ holds (by the choice of ℓ and f) we obtain a number $0 \le k \le \ell$ where b^k is the longest suffix of $\mathrm{rd}_2(t_{i_1} \dots t_{i_m}) \cdot b^\ell$ such that there is a prefix w_1 of $\mathrm{wrt}(t_{i_1} \dots t_{i_m})$ with $w_1 \in \mathrm{redsup}_{\mathcal{L}}(b^k)$. By $k \le \ell$ this is even a suffix of b^ℓ and therefore of $\mathrm{rd}_2(t_{j_1} \dots t_{j_n}) \cdot b^\ell$. Hence, we learn

$$\mathrm{rd}_2(t_{i_1} \dots t_{i_m} \cdot [\overline{b}^\ell]_{\equiv_{\mathcal{L}}}) = b^k = \mathrm{rd}_2(t_{j_1} \dots t_{j_n} \cdot [\overline{b}^\ell]_{\equiv_{\mathcal{L}}})\,. \tag{4.2}$$

Finally, we infer the following computation:

$$\begin{aligned} p &= \left[\overline{\mathrm{rd}_1(t_{i_1} \dots t_{i_m})}\mathrm{wrt}(t_{i_1} \dots t_{i_m})\overline{\mathrm{rd}_2(t_{i_1} \dots t_{i_m})} \cdot \overline{b}^\ell \cdot \langle\!\langle b^{2\ell}, \overline{b^{2\ell}} \rangle\!\rangle\right]_{\equiv_{\mathcal{L}}} && \text{(by Lem. 3.8)} \\ &= \left[\overline{\mathrm{rd}_1(t_{i_1} \dots t_{i_m})\mathrm{rd}_2(t_{i_1} \dots t_{i_m})b^{k-\ell}}\mathrm{wrt}(t_{i_1} \dots t_{i_m})\overline{b}^k \cdot \langle\!\langle b^{2\ell}, \overline{b^{2\ell}} \rangle\!\rangle\right]_{\equiv_{\mathcal{L}}} && \text{(by Eq. (4.2))} \\ &= \left[\overline{\mathrm{rd}_1(t_{j_1} \dots t_{j_n})\mathrm{rd}_2(t_{i_1} \dots t_{j_n})b^{k-\ell}}\mathrm{wrt}(t_{i_1} \dots t_{j_n})\overline{b}^k \cdot \langle\!\langle b^{2\ell}, \overline{b^{2\ell}} \rangle\!\rangle\right]_{\equiv_{\mathcal{L}}} && \text{(by Eq. (4.1))} \\ &= \left[\overline{\mathrm{rd}_1(t_{j_1} \dots t_{j_n})}\mathrm{wrt}(t_{j_1} \dots t_{j_n})\overline{\mathrm{rd}_2(t_{j_1} \dots t_{j_n})} \cdot \overline{b}^\ell \cdot \langle\!\langle b^{2\ell}, \overline{b^{2\ell}} \rangle\!\rangle\right]_{\equiv_{\mathcal{L}}} && \text{(by Eq. (4.2))} \\ &= q\,. && \text{(by Lem. 3.8)} \end{aligned}$$

Consequently, T^* is not freely generated by T.

Now let T^* be not freely generated by T. If $[\varepsilon]_{\equiv_{\mathcal{L}}} \in T$ and hence $(\varepsilon, \varepsilon) \in S$ holds, S^* is trivially not freely generated by S. So, we assume that $[\varepsilon]_{\equiv_{\mathcal{L}}} \notin T$ holds and, hence, $(\varepsilon, \varepsilon) \notin S$. Then there are indices $(i_1, \dots, i_m) \neq (j_1, \dots, j_n)$ where $m, n > 0$ such that $t_{i_1} \dots t_{i_m} = t_{j_1} \dots t_{j_n}$. Let $(i'_1, \dots, i'_{m'}), (j'_1, \dots j'_{n'})$ be the above sequences after deletion of all 0's. Since t_0 is the only element in T adding bb into the projections of write and read actions, t_0 does not commute with t_i for any $1 \le i \le k$. Hence, we still have $(i'_1, \dots, i'_{m'}) \neq (j'_1, \dots j'_{n'})$, $\mathrm{wrt}(t_{i'_1} \dots t_{i'_{m'}}) = \mathrm{wrt}(t_{j'_1} \dots t_{j'_{n'}})$, and $\mathrm{rd}(t_{i'_1} \dots t_{i'_{m'}}) = \mathrm{rd}(t_{j'_1} \dots t_{j'_{n'}})$. Then we have:

$$\begin{aligned} f(x_{i'_1} \dots x_{i'_{m'}}) = \mathrm{wrt}(t_{i'_1} \dots t_{i'_{m'}}) = \mathrm{wrt}(t_{j'_1} \dots t_{j'_{n'}}) = f(x_{j'_1} \dots x_{j'_{n'}})\,, \\ g(y_{i'_1} \dots y_{i'_{m'}}) = \mathrm{rd}(t_{i'_1} \dots t_{i'_{m'}}) = \mathrm{rd}(t_{j'_1} \dots t_{j'_{n'}}) = g(y_{j'_1} \dots y_{j'_{n'}})\,. \end{aligned}$$

By injectivity of f and g we infer that

$$(x_{i'_1} \dots x_{i'_{m'}}, y_{i'_1} \dots y_{i'_{m'}}) = (x_{j'_1} \dots x_{j'_{n'}}, y_{j'_1} \dots y_{j'_{n'}})\,,$$

i.e., S^* is not freely generated by S. $\square$

The next problem to consider is the emptiness of intersections of rational plq languages. Given two recognizable plq languages, this problem is decidable since the class of recognizable languages is effectively closed under intersection. However, we will prove that this decidability does not hold for arbitrary rational plq languages. As a corollary we can infer that the class of rational plq languages is not effectively closed under intersection. Afterwards we will prove the existence of two rational plq languages whose intersection is not rational. In other words, the classes of rational and recognizable plq languages do not coincide. Nevertheless, each recognizable language in $\mathcal{Q}(\mathcal{L})$ is rational due to [23] since the plq monoid is finitely generated.

Theorem 4.4. *Let* $\mathcal{L} = (F, U)$ *be a lossiness alphabet with* $|A_{\mathcal{L}}| \geq 2$*. Then the* rational intersection emptiness problem *in* $\mathcal{Q}(\mathcal{L})$ *is undecidable:*

Input: *Two rational plq languages* $S, T \subseteq \mathcal{Q}(\mathcal{L})$
Question: *Does* $S \cap T = \emptyset$ *hold?*

Proof. We prove this by reduction of Post's Correspondence Problem (PCP), which is undecidable by [29]. So, let $a, b \in A_{\mathcal{L}}$ be distinct letters and $I = ((x_1, y_1), \ldots, (x_k, y_k))$ be an instance of the PCP with $x_i, y_i \in A_{\mathcal{L}}^*$. For $1 \leq i \leq k$ we define the transformations $s_i := [a^i b \overline{x_i}]_{\equiv_{\mathcal{L}}}$ and $t_i := [a^i b \overline{y_i}]_{\equiv_{\mathcal{L}}}$. Then we can define rational plq languages as follows:

$$S_I := \{s_i \mid 1 \leq i \leq k\}^+ [\overline{a}\overline{b}^*]_{\equiv_{\mathcal{L}}} \quad \text{and} \quad T_I := \{t_i \mid 1 \leq i \leq k\}^+ [\overline{a}\overline{b}^*]_{\equiv_{\mathcal{L}}} .$$

Now we have to show that $S_I \cap T_I \neq \emptyset$ iff I has a solution. First, let $t \in S_I \cap T_I$. Then by definition of S_I and T_I there are $\ell \in \mathbb{N}$ and indices $i_1, \ldots, i_m$ and $j_1, \ldots, j_n$ with $m, n > 0$ such that $s_{i_1} \ldots s_{i_m} [\overline{a}\overline{b}^\ell]_{\equiv_{\mathcal{L}}} = t = t_{j_1} \ldots t_{j_n} [\overline{a}\overline{b}^\ell]_{\equiv_{\mathcal{L}}}$ holds. Then, we have $x_{i_1} \ldots x_{i_m} ab^\ell = \mathrm{rd}(t) = y_{j_1} \ldots y_{j_n} ab^\ell$ implying $x_{i_1} \ldots x_{i_m} = y_{j_1} \ldots y_{j_n}$. By $a^{i_1} b \ldots a^{i_m} b = \mathrm{wrt}(t) = a^{j_1} b \ldots a^{j_n} b$ we can infer $(i_1, \ldots, i_m) = (j_1, \ldots, j_n)$ which is a solution of I.

For the converse implication we assume that I has a solution $(i_1, \ldots, i_n)$ with $n > 0$. Set $s' := s_{i_1} \ldots s_{i_n}$ and $t' := t_{i_1} \ldots t_{i_n}$ and set $s := s' [\overline{ab^{n+1}}]_{\equiv_{\mathcal{L}}} \in S_I$ and $t := t' [\overline{ab^{n+1}}]_{\equiv_{\mathcal{L}}} \in T_I$. Then we have $\mathrm{wrt}(s) = \mathrm{wrt}(t)$ and $\mathrm{rd}(s) = \mathrm{rd}(t)$. Furthermore, by Proposition 3.7 we have

$$s = \left[\overline{\mathrm{rd}_1(s')} \mathrm{wrt}(s) \overline{\mathrm{rd}_2(s')} \cdot \overline{ab^{n+1}}\right]_{\equiv_{\mathcal{L}}} \quad \text{and} \quad t = \left[\overline{\mathrm{rd}_1(t')} \mathrm{wrt}(t) \overline{\mathrm{rd}_2(t')} \cdot \overline{ab^{n+1}}\right]_{\equiv_{\mathcal{L}}} .$$

Now, we want to utilize Lemma 3.8 to determine $\mathrm{rd}_2(\mathrm{wrt}(s) \overline{\mathrm{rd}_2(s')} \cdot \overline{ab^{n+1}})$ and $\mathrm{rd}_2(\mathrm{wrt}(t) \overline{\mathrm{rd}_2(t')} \cdot \overline{ab^{n+1}})$. Note that we have $|\mathrm{wrt}(s)|_b = |\mathrm{wrt}(t)|_b = n$. Hence, there is a number $0 \leq k \leq n$ such that b^k is the longest suffix of $\mathrm{rd}_2(s') ab^{n+1}$ such that there is a prefix w_1 of $\mathrm{wrt}(s') = \mathrm{wrt}(s)$ with $w_1 \in \mathrm{redsup}_{\mathcal{L}}(b^k)$. By $k < n + 1$ the word b^k is even a proper suffix of ab^{n+1} and, therefore, also of $\mathrm{rd}_2(t') ab^{n+1}$. In other words, we have

$$\mathrm{rd}_2(\mathrm{wrt}(s) \overline{\mathrm{rd}_2(s') ab^{n+1}}) = b^k = \mathrm{rd}_2(\mathrm{wrt}(t) \overline{\mathrm{rd}_2(t') ab^{n+1}}) . \tag{4.3}$$

This finally implies

$$\begin{aligned}
s &= \left[\overline{\mathrm{rd}_1(s')}\mathrm{wrt}(s)\overline{\mathrm{rd}_2(s')}\cdot\overline{ab^{n+1}}\right]_{\equiv_{\mathcal{L}}} && \text{(by Lemma 3.8)}\\
&= \left[\overline{\mathrm{rd}_1(s')\mathrm{rd}_2(s')ab^{n+1-k}}\mathrm{wrt}(s)\overline{b}^k\right]_{\equiv_{\mathcal{L}}} && \text{(by Equation (4.3))}\\
&= \left[\overline{\mathrm{rd}_1(t')\mathrm{rd}_2(t')ab^{n+1-k}}\mathrm{wrt}(t)\overline{b}^k\right]_{\equiv_{\mathcal{L}}} && \text{(by } \mathrm{wrt}(s)=\mathrm{wrt}(t) \text{ and } \mathrm{rd}(s')=\mathrm{rd}(t'))\\
&= \left[\overline{\mathrm{rd}_1(t')}\mathrm{wrt}(t)\overline{\mathrm{rd}_2(t')}\cdot\overline{ab^{n+1}}\right]_{\equiv_{\mathcal{L}}} && \text{(by Equation (4.3))}\\
&= t\,. && \text{(by Lemma 3.8)}
\end{aligned}$$

We learn $s = t \in S_I \cap T_I \neq \emptyset$. Hence, we reduced PCP to the rational intersection emptiness problem of $\mathcal{Q}(\mathcal{L})$ which is therefore undecidable. □

Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$. To prove that the rational plq languages are not closed under intersection and to prove the undecidability of the next problems we use an embedding of $\{\alpha\}^* \times \{\beta,\gamma\}^*$ into the plq monoid $\mathcal{Q}(\mathcal{L})$. Such an embedding is $\psi\colon \{\alpha\}^* \times \{\beta,\gamma\}^* \hookrightarrow \mathcal{Q}(\mathcal{L})$ with $\psi(\alpha,\varepsilon) = [a]_{\equiv_{\mathcal{L}}}$, $\psi(\varepsilon,\beta) = [\overline{a}\overline{b}]_{\equiv_{\mathcal{L}}}$, and $\psi(\varepsilon,\gamma) = [\overline{a}\overline{b}\overline{b}]_{\equiv_{\mathcal{L}}}$ where $a, b \in A_{\mathcal{L}}$ are two distinct letters (cf. [18, Lemma 6.17]). Note that there are some commutations in $\{\alpha\}^* \times \{\beta,\gamma\}^*$:

$$(\alpha,\varepsilon)\cdot(\varepsilon,\beta) = (\alpha,\beta) = (\varepsilon,\beta)\cdot(\alpha,\varepsilon) \text{ and } (\alpha,\varepsilon)\cdot(\varepsilon,\gamma) = (\alpha,\gamma) = (\varepsilon,\gamma)\cdot(\alpha,\varepsilon)\,.$$

The map ψ preserves these commutations:

$$\begin{aligned}
\psi(\alpha,\varepsilon)\cdot\psi(\varepsilon,\beta) = [a\overline{a}\overline{b}]_{\equiv_{\mathcal{L}}} = [\overline{a}\overline{b}a]_{\equiv_{\mathcal{L}}} = \psi(\varepsilon,\beta)\cdot\psi(\alpha,\varepsilon) \quad \text{and}\\
\psi(\alpha,\varepsilon)\cdot\psi(\varepsilon,\gamma) = [a\overline{a}\overline{b}\overline{b}]_{\equiv_{\mathcal{L}}} = [\overline{a}\overline{b}\overline{b}a]_{\equiv_{\mathcal{L}}} = \psi(\varepsilon,\gamma)\cdot\psi(\alpha,\varepsilon)
\end{aligned} \tag{4.4}$$

according to the equations from Theorem 3.3.

Theorem 4.5. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$. The class of rational languages in $\mathcal{Q}(\mathcal{L})$ is not closed under intersection. In particular, there is a rational plq language which is not recognizable.*

Proof. Consider the following rational relations:

$$X := \{(a^m, b^m c^n) \mid m, n \in \mathbb{N}\} \quad \text{and} \quad Y := \{(a^m, b^n c^m) \mid m, n \in \mathbb{N}\}\,.$$

Then $\psi(X)$ and $\psi(Y)$ are rational in $\mathcal{Q}(\mathcal{L})$. Suppose that $\psi(X) \cap \psi(Y)$ is rational. Then there is a regular language $L \subseteq \Sigma_{\mathcal{L}}^*$ with $\eta_{\mathcal{L}}(L) = \psi(X) \cap \psi(Y)$. Since ψ is injective we have

$$\psi(X) \cap \psi(Y) = \psi(X \cap Y) = \psi(\{(a^n, b^n c^n) \mid n \in \mathbb{N}\})\,.$$

Hence, by definition of ψ we have $\mathrm{rd}(L) = \{(ab)^n(abb)^n \mid n \in \mathbb{N}\}$ which would be regular since rd is a homomorphism. But this is a contradiction to the Pumping Lemma.

Towards the second statement, recall that $\mathcal{Q}(\mathcal{L})$ is finitely generated. Then from [23] we know that each recognizable plq language is rational. From the first statement and from the fact that the class of recognizable languages is closed under intersection, we infer that $\psi(X)$ and $\psi(Y)$ are rational but not recognizable. □

From the previous theorem we also infer that the class of rational plq languages is not closed under complement. Otherwise closure under complement and union would also imply closure under intersection, contradicting our theorem above.

By $\mathbb{I}_{\mathcal{L}} \subseteq \mathcal{Q}(\mathcal{L})$ we denote the image of ψ, i.e., ψ is an isomorphism from $\{\alpha\}^* \times \{\beta,\gamma\}^*$ onto $\mathbb{I}_{\mathcal{L}}$. It is easy to see that $\mathbb{I}_{\mathcal{L}} = \{[a]_{\equiv_{\mathcal{L}}}, [\overline{ab}]_{\equiv_{\mathcal{L}}}, [\overline{abb}]_{\equiv_{\mathcal{L}}}\}^*$ holds. In the following lemma we prove that recognizability in this submonoid implies recognizability in the whole plq monoid.

Lemma 4.6. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$ and $T \subseteq \mathbb{I}_{\mathcal{L}}$ be recognizable in $\mathbb{I}_{\mathcal{L}}$. Then T is recognizable in $\mathcal{Q}(\mathcal{L})$.*

Proof. Let $T \subseteq \mathbb{I}_{\mathcal{L}}$ be recognizable in $\mathbb{I}_{\mathcal{L}}$. Then $\psi^{-1}(T) \subseteq \{\alpha\}^* \times \{\beta,\gamma\}^*$ is recognizable in $\{\alpha\}^* \times \{\beta,\gamma\}^*$. Due to Mezei's Theorem [2, Theorem III.1.5] there are regular languages $V_i \subseteq \{\alpha\}^*$ and $W_i \subseteq \{\beta,\gamma\}^*$ with

$$\psi^{-1}(T) = \bigcup_{1\leq i\leq k} V_i \times W_i\,.$$

Now we define homomorphisms $g\colon \{\alpha\}^* \to A_{\mathcal{L}}^*$ by $g(\alpha) = a = \psi(\alpha,\varepsilon)$ and $h\colon \{\beta,\gamma\}^* \to A_{\mathcal{L}}^*$ by $h(\beta) = ab = \psi(\varepsilon,\beta)$ and $h(\gamma) = abb = \psi(\varepsilon,\gamma)$. Then $g(V_i), h(W_i) \subseteq A_{\mathcal{L}}^*$ are regular as well. Hence, $\mathrm{wrt}^{-1}(g(V_i))$ and $\mathrm{rd}^{-1}(h(W_i))$ are recognizable in $\mathcal{Q}(\mathcal{L})$ and therefore

$$\bigcup_{1\leq i\leq k} \mathrm{wrt}^{-1}(g(V_i)) \cap \mathrm{rd}^{-1}(h(W_i))$$

also is recognizable. Finally, we have to prove that this plq language equals T.

First, let $t \in T \subseteq \mathbb{I}_{\mathcal{L}}$. Then there is a tuple $(v,w) \in \{\alpha\}^* \times \{\beta,\gamma\}^*$ with $\psi(v,w) = t$. Since $(v,w) \in \psi^{-1}(t) \subseteq \psi^{-1}(T)$ there is $1 \leq i \leq k$ with $v \in V_i$ and $w \in W_i$. Then we have

$$\begin{aligned} t = \psi(v,w) &= \psi(v,\varepsilon)\psi(\varepsilon,w) = [g(v)\overline{h(w)}]_{\equiv_{\mathcal{L}}} && \text{(by Equation (4.4))} \\ &\in \mathrm{wrt}^{-1}(g(v)) \cap \mathrm{rd}^{-1}(h(w)) \subseteq \mathrm{wrt}^{-1}(g(V_i)) \cap \mathrm{rd}^{-1}(h(W_i))\,. \end{aligned}$$

Conversely, let $1 \leq i \leq k$ and $t \in \mathrm{wrt}^{-1}(g(V_i)) \cap \mathrm{rd}^{-1}(h(W_i))$. Then there are $v \in V_i$ and $w \in W_i$ with $\mathrm{wrt}(t) = g(v)$ and $\mathrm{rd}(t) = h(w)$. From the definition of g and h we infer $\mathrm{wrt}(t) \in a^*$ and $\mathrm{rd}(t) \in \{\varepsilon\} \cup A_{\mathcal{L}}^* b$. Since $a \neq b$ holds, we learn that ε is the only suffix r_2 of $\mathrm{rd}(t)$ having a prefix w_1 of $\mathrm{wrt}(t)$ which is a reduced $\mathcal{L}$-superword of r_2. Hence, we have $\mathrm{rd}_2(t) = \varepsilon$ and, therefore, $t = [\overline{h(w)}g(v)]_{\equiv_{\mathcal{L}}} = \psi(v,w)$. Finally, from $v \in V_i$ and $w \in W_i$ we learn $(v,w) \in \psi^{-1}(T)$ implying $t = \psi(v,w) \in T$.

All in all, we have seen that T is recognizable in $\mathcal{Q}(\mathcal{L})$. □

To prove the undecidability of the rational universality and recognizability problem we use the embedding ψ and the results from Gibbons and Rytter [10] which state that their counterparts in $\{\alpha\}^* \times \{\beta, \gamma\}^*$ are undecidable. Note that rational universality can be reduced to rational inclusion and equality. Hence, these two problems also are undecidable in the plq monoid.

Theorem 4.7. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$. Then the following statements hold:*

(1) The rational universality problem *in $\mathcal{Q}(\mathcal{L})$ is undecidable:*

  Input: *A rational plq language $T \subseteq \mathcal{Q}(\mathcal{L})$*
  Question: *Does $T = \mathcal{Q}(\mathcal{L})$ hold?*

  Consequently, the rational inclusion problem *and the* rational equality problem *in $\mathcal{Q}(\mathcal{L})$ are undecidable.*

(2) The rational recognizability problem *in $\mathcal{Q}(\mathcal{L})$ is undecidable:*

  Input: *A rational plq language $T \subseteq \mathcal{Q}(\mathcal{L})$*
  Question: *Is T recognizable in $\mathcal{Q}(\mathcal{L})$?*

Proof.

(1) Let $T \subseteq \{\alpha\}^* \times \{\beta, \gamma\}^*$ be rational. Then $\psi(T)$ is rational in $\mathcal{Q}(\mathcal{L})$. Due to Lemma 4.6, the language $\mathbb{I}_{\mathcal{L}}$ is recognizable in $\mathcal{Q}(\mathcal{L})$. Therefore, $\mathcal{Q}(\mathcal{L}) \setminus \mathbb{I}_{\mathcal{L}}$ is recognizable and, hence, rational in $\mathcal{Q}(\mathcal{L})$ since this monoid is finitely generated. Consequently, $\psi(T) \cup (\mathcal{Q}(\mathcal{L}) \setminus \mathbb{I}_{\mathcal{L}})$ is rational as well. This plq languages equals $\mathcal{Q}(\mathcal{L})$ iff $\psi(T) = \mathbb{I}_{\mathcal{L}}$, i.e., $T = \{\alpha\}^* \times \{\beta, \gamma\}^*$. But this latter question is undecidable by [10, Theorem 2(Q4)].

(2) Let $T \subseteq \{\alpha\}^* \times \{\beta, \gamma\}^*$ be rational. Then $\psi(T)$ is rational. By Lemma 4.6 $\psi(T)$ is recognizable in $\mathcal{Q}(\mathcal{L})$ iff it is recognizable in $\mathbb{I}_{\mathcal{L}}$. This is the case iff T is recognizable in $\{\alpha\}^* \times \{\beta, \gamma\}^*$. But this latter question is undecidable by [10, Theorem 2(Q6)]. □

5. Characterizations of the Recognizable PLQ Languages

We have already seen that it is undecidable whether a given rational plq language is recognizable (cf. Theorem 4.7(2)). However, we want to characterize in which cases a rational plq language is recognizable. Concretely, we want to give some rational-like expressions which are fully describing the recognizable plq languages. To this end, we have to introduce some special restrictions to the concatenation and iteration of recognizable languages. Additionally, we have to add a non-monotonic operation to our expressions: the complement operation. We call languages constructed via these restricted operations the *q-rational* plq languages (the definition of this notion can be found on page 28). With the help of this definition we will obtain a characterization in the manner of Kleene's Theorem [15].

Additionally, we want to translate Büchi's Theorem [4] to the plq monoid. We do this with the help of special modifications of Büchi's word logic MSO (recall that their definition can be found in subsection 2.2). Concretely, our new logic $\mathsf{MSO_q}$ still identifies an action sequence $w \in \Sigma_{\mathcal{L}}^*$ as a set of positions labeled with letters from $\Sigma_{\mathcal{L}}$. But in contrast to the word logic we have to restrict comparisons between write and read actions to ensure that our logic only describes recognizable plq languages. The full definition of this logic can be found at page 35.

Now, we state our main theorem. As mentioned before we will give the concrete definitions of q-rational languages and the logic $\mathsf{MSO_q}$ later in this section.

Theorem 5.1 (Main Theorem). *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$ and $T \subseteq \mathcal{Q}(\mathcal{L})$. Then the following statements are (effectively) equivalent:*

(A) T is recognizable.
(B) T is q-rational.
(C) T is $\mathsf{MSO_q}$-definable.

We prove these equivalences with the help of three implications in Propositions 5.23, 5.29, and 5.32.

5.1. *Some helpful characterizations*

Before diving into the proof of Theorem 5.1 we prove two further characterizations of the recognizable plq languages which turned out to be convenient for simplification of our proof. We already know these characterizations from Huschenbett et al. [12] which were given there for the transformation monoid $\mathcal{Q}(\emptyset, U)$ of a reliable queue. Here, we generalize these equivalences to plq monoids $\mathcal{Q}(\mathcal{L})$ with arbitrary underlying lossiness alphabet $\mathcal{L} = (F, U)$.

Concretely, we prove the correspondence of recognizability in the plq monoid $\mathcal{Q}(\mathcal{L})$ to regularity in the underlying free monoid $\Sigma_{\mathcal{L}}^*$. We also describe the recognizable plq languages in terms of Boolean combinations of plq languages $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and some special languages $\Omega_\ell \subseteq \mathcal{Q}(\mathcal{L})$ for regular word languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$:

Definition 5.2. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet and $t \in \mathcal{Q}(\mathcal{L})$. Then the* overlap's bounded width *of t is*

$$\mathrm{obw}(t) := \inf\left\{ |\mathrm{rd}_2(s)| \;\middle|\; \begin{array}{c} s \in \mathcal{Q}(\mathcal{L}), \mathrm{wrt}(s) = \mathrm{wrt}(t), \mathrm{rd}(s) = \mathrm{rd}(t), \\ |\mathrm{rd}_2(s)| > |\mathrm{rd}_2(t)| \end{array} \right\} .$$

Similarly, for $w \in \Sigma_{\mathcal{L}}^*$ we may define $\mathrm{obw}(w) := \mathrm{obw}([w]_{\equiv_{\mathcal{L}}})$. Furthermore, for $\ell \in \mathbb{N}$ set

$$\Omega_\ell := \{t \in \mathcal{Q}(\mathcal{L}) \mid \mathrm{obw}(t) > \ell\} .$$

The overlap's bounded width specifies the minimal length of the overlap of a transformation having the same projections and a longer overlap. If no such sequence exists then we set this value to ∞. In particular, we have $t \in \Omega_\ell$ for a transformation $t \in \mathcal{Q}(\mathcal{L})$ and a number $\ell \in \mathbb{N}$ if, and only if, any transformation $s \in \mathcal{Q}(\mathcal{L})$ having the same projections (i.e., $\mathrm{wrt}(s) = \mathrm{wrt}(t)$ and $\mathrm{rd}(s) = \mathrm{rd}(t)$) and a longer overlap (i.e., $|\mathrm{rd}_2(s)| > |\mathrm{rd}_2(t)|$) also satisfies $|\mathrm{rd}_2(s)| > \ell$.

Example 5.3. Let $\mathcal{L} = (\emptyset, \{a, b\})$ and $t = [\overline{ababa}\bar{a}b\bar{b}abab]_{\equiv_\mathcal{L}}$. Then there are two words with the same projections and longer overlaps:

$$s_1 = [\overline{ab}a\bar{a}b\bar{b}a\bar{a}b\bar{b}ab]_{\equiv_\mathcal{L}} \qquad \text{and} \qquad s_2 = [a\bar{a}b\bar{b}a\bar{a}b\bar{b}a\bar{a}b\bar{b}]_{\equiv_\mathcal{L}} .$$

We have $|\mathrm{rd}_2(s_1)| = 4$ and $|\mathrm{rd}_2(s_2)| = 6$. Therefore, we have $\mathrm{obw}(t) = 4$, $\mathrm{obw}(s_1) = 6$, and $\mathrm{obw}(s_2) = \infty$. Hence, $t \in \Omega_3 \setminus \Omega_4$ holds.

From [12, Observation 9.1] we know that the reliable queue languages $\{t \in \mathcal{Q}(\emptyset, U) \mid |\mathrm{rd}_2(t)| > \ell\}$ are not recognizable for any $\ell \in \mathbb{N}$. This also means, there is no non-trivial property of the overlap's width $|\mathrm{rd}_2(t)|$ which is recognizable in the reliable queue monoid $\mathcal{Q}(\emptyset, U)$. An appropriate alternative for the generators of the Boolean algebra of recognizable plq languages has been found in such kind of "over-approximation" of the overlap's width (note that $\mathrm{obw}(t) > |\mathrm{rd}_2(t)|$ holds). Additionally, the following two lemmas provide more motivation of this notion. First, we show that the transformations $t \in \mathcal{Q}(\mathcal{L})$ are fully described by $\mathrm{wrt}(t)$, $\mathrm{rd}(t)$, and $\mathrm{obw}(t)$.

Lemma 5.4. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_\mathcal{L}| \geq 2$ and $s, t \in \mathcal{Q}(\mathcal{L})$. Then we have $\mathrm{wrt}(s) = \mathrm{wrt}(t)$, $\mathrm{rd}(s) = \mathrm{rd}(t)$, and $\mathrm{obw}(s) = \mathrm{obw}(t)$ if, and only if, $s = t$ holds.*

Proof. The implication "$\Leftarrow$" is obvious. So, we only have to prove the converse implication. Let $s, t \in \mathcal{Q}(\mathcal{L})$ be two transformations with $\mathrm{wrt}(s) = \mathrm{wrt}(t)$, $\mathrm{rd}(s) = \mathrm{rd}(t)$, and $\mathrm{obw}(s) = \mathrm{obw}(t)$. Due to Proposition 3.7 it suffices to show $\mathrm{rd}_2(s) = \mathrm{rd}_2(t)$.

By definition of obw(.) we know the following inequation:

$$|\mathrm{rd}_2(s)| < \mathrm{obw}(s) = \mathrm{obw}(t) = \inf\left\{ |\mathrm{rd}_2(u)| \;\middle|\; \begin{array}{c} u \in \mathcal{Q}(\mathcal{L}), \mathrm{wrt}(u) = \mathrm{wrt}(t), \\ \mathrm{rd}(u) = \mathrm{rd}(t), |\mathrm{rd}_2(u)| > |\mathrm{rd}_2(t)| \end{array} \right\} .$$

Recall that $\mathrm{wrt}(s) = \mathrm{wrt}(t)$ and $\mathrm{rd}(s) = \mathrm{rd}(t)$ holds. Then $|\mathrm{rd}_2(s)| > |\mathrm{rd}_2(t)|$ would imply $\mathrm{obw}(t) \leq |\mathrm{rd}_2(s)| < \mathrm{obw}(s) = \mathrm{obw}(t)$, which is impossible. Hence, we have $|\mathrm{rd}_2(s)| \leq |\mathrm{rd}_2(t)|$. By symmetry we also obtain $|\mathrm{rd}_2(t)| \leq |\mathrm{rd}_2(s)|$. In other words, $\mathrm{rd}_2(s)$ and $\mathrm{rd}_2(t)$ are suffixes of $\mathrm{rd}(s) = \mathrm{rd}(t)$ having the same length, i.e., we have $\mathrm{rd}_2(s) = \mathrm{rd}_2(t)$. This finally implies $s = t$ according to Proposition 3.7. $\square$

In other words, we could extend Proposition 3.7 by another equivalence stating that the triple $(\mathrm{wrt}(t), \mathrm{rd}(t), \mathrm{obw}(t))$ is (similar to $(\mathrm{wrt}(t), \mathrm{rd}(t), \mathrm{rd}_2(t))$) a characterization of the transformation $t \in \mathcal{Q}(\mathcal{L})$.

The following statement proves that from a transformation $t \in \mathcal{Q}(\mathcal{L})$ with a small overlap's bounded width we obtain another transformation $s \in \mathcal{Q}(\mathcal{L})$ which has an overlap $\mathrm{rd}_2(s)$ of this bounded width. This knowledge is very helpful to prove Theorem 5.1 and the succeeding Theorem 5.6.

Lemma 5.5. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$, $\ell \in \mathbb{N}$, and $t \in \mathcal{Q}(\mathcal{L})$. Then $\mathrm{obw}(t) \leq \ell$ if, and only if, there is $r_2 \in A_{\mathcal{L}}^{\leq \ell}$ with $\mathrm{rd}(t) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(t)$ such that $|\mathrm{rd}_2(t)| < |r_2|$.*

Proof. We first suppose $\mathrm{obw}(t) > \ell$. Let $r_2 \in A_{\mathcal{L}}^{\leq \ell}$ such that $\mathrm{rd}(t) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(t)$. Furthermore, let $r_1 \in A_{\mathcal{L}}^*$ with $\mathrm{rd}(t) = r_1 r_2$. Set $s := [\overline{r_1}\mathrm{wrt}(t)\overline{r_2}]_{\equiv_{\mathcal{L}}}$. Then we have $\mathrm{wrt}(s) = \mathrm{wrt}(t)$, $\mathrm{rd}(s) = \mathrm{rd}(t)$, and $|\mathrm{rd}_2(s)| \leq |r_2| \leq \ell$. Since $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ and w_1 is a prefix of $\mathrm{wrt}(t) = \mathrm{wrt}(s)$, we have $\mathrm{rd}_2(s) = r_2$ according to Lemma 3.8. Since $|\mathrm{rd}_2(t)| < |\mathrm{rd}_2(s)|$ would imply $\mathrm{obw}(t) \leq |\mathrm{rd}_2(s)| \leq \ell$, we have $|\mathrm{rd}_2(t)| \geq |\mathrm{rd}_2(s)| = |r_2|$.

Now, assume $\mathrm{obw}(t) \leq \ell$. Then there is $s \in \mathcal{Q}(\mathcal{L})$ with $\mathrm{wrt}(s) = \mathrm{wrt}(t)$, $\mathrm{rd}(s) = \mathrm{rd}(t)$, and $|\mathrm{rd}_2(t)| < |\mathrm{rd}_2(s)| \leq \ell$. Consider $r_2 := \mathrm{rd}_2(s)$. Then we have $|r_2| = |\mathrm{rd}_2(s)| \leq \ell$, $\mathrm{rd}(t) = \mathrm{rd}(s) \in A_{\mathcal{L}}^* r_2$, and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix $w_1 \in A_{\mathcal{L}}^*$ of $\mathrm{wrt}(s) = \mathrm{wrt}(t)$. □

With the help of the overlap's bounded width and the knowledge from the previous two statements we are able to state the following equivalences:

Theorem 5.6. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$ and $T \subseteq \mathcal{Q}(\mathcal{L})$. Then the following statements are (effectively) equivalent:*

(1) T is recognizable.
(2) $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is regular.
(3) T is a Boolean combination of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for regular languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$.

Note that $\eta_{\mathcal{L}}^{-1}(T) \subseteq \Sigma_{\mathcal{L}}^*$ is the union of the equivalence classes of the behavioral equivalence $\equiv_{\mathcal{L}}$ in T. From Lemma 3.8 we also know that each such equivalence class contains a word from $\overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$. Consequently, the language $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ contains at least one representative from each equivalence class of the behavioral equivalence $\equiv_{\mathcal{L}}$ in T.

Remark 5.7. Let $\mathcal{L} = (\emptyset, U)$ be a lossiness alphabet with $|U| \geq 2$ (i.e., we consider a reliable queue). Huschenbett et al. proved in [12] that we can extend Theorem 5.6 by the following item:

(4) $\eta_{\mathcal{L}}^{-1}(T) \cap A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^*$ is regular.

However, this theorem cannot be expanded by the statement "$\eta_{\mathcal{L}}^{-1}(T) \cap L$ is regular" where L is either $\overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^*$, $A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$, or the union of both languages.

Moreover, this fourth item is not equivalent for other lossiness alphabet. So, let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $F \neq \emptyset$ and $|A_{\mathcal{L}}| \geq 2$. For two distinct letters

$a, b \in A_{\mathcal{L}}$ with $a \in F$ the language $T := \{[\overline{a}^n a^n b\overline{b}]_{\equiv_{\mathcal{L}}} \mid n \in \mathbb{N}\}$ is not recognizable since

$$\eta_{\mathcal{L}}^{-1}(T) = \{\overline{a}^n a^n b\overline{b} \mid n \in \mathbb{N}\}$$

is no regular language. However, the language $\eta_{\mathcal{L}}^{-1}(T) \cap A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* = \{b\overline{b}\}$ is finite and, therefore, regular.

5.1.1. *The implication "(1)⇒(2)" in Theorem 5.6*

The first implication in Theorem 5.6 that we want to prove is very simple:

Proposition 5.8. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet and $T \subseteq \mathcal{Q}(\mathcal{L})$ be recognizable. Then $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is regular.*

Proof. Since T is recognizable, the set $\eta_{\mathcal{L}}^{-1}(T)$ is recognizable and hence regular. Since the class of regular languages is closed under intersection, $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is regular. □

5.1.2. *The implication "(2)⇒(3)" in Theorem 5.6*

Towards the proof of the second implication in Theorem 5.6 we fix an arbitrary lossiness $\mathcal{L} = (F, U)$ with $|A_{\mathcal{L}}| \geq 2$. Let $T \subseteq \mathcal{Q}(\mathcal{L})$ be a plq language such that $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is regular. We will partition T into two plq languages: $T \cap \Omega_\ell$ and $T \cap \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell$ for an appropriate natural number $\ell \in \mathbb{N}$. We will show then that both plq languages satisfy property (3) of Theorem 5.6. This means, both languages are Boolean combinations of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_k for regular languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $k \in \mathbb{N}$. To prove this result, we first show that any transformation $t \in \Omega_k$ with $k \in \mathbb{N}$ having at least k read actions has the same behavior as one ending with k read actions.

Lemma 5.9. *Let $t \in \mathcal{Q}(\mathcal{L})$ and $r_1, r_2 \in A_{\mathcal{L}}^*$ with $\mathrm{rd}(t) = r_1 r_2$ and $|\mathrm{rd}_2(t)| \leq |r_2| < \mathrm{obw}(t)$. Then we have $t = [\overline{r_1}\mathrm{wrt}(t)\overline{r_2}]_{\equiv_{\mathcal{L}}}$.*

Proof. If $\mathrm{obw}(t) < \infty$ holds, this number is the length of the shortest suffix $s \in A_{\mathcal{L}}^*$ of $\mathrm{rd}(t)$ which is longer than $\mathrm{rd}_2(t)$ and satisfies $p \in \mathrm{redsup}_{\mathcal{L}}(s)$ for a prefix p of $\mathrm{wrt}(t)$. By minimality of $\mathrm{obw}(t)$ we obtain $\mathrm{rd}_2(\mathrm{wrt}(t)\overline{r_2}) = \mathrm{rd}_2(\mathrm{wrt}(t)\overline{\mathrm{rd}_2(t)}) = \mathrm{rd}_2(t)$ for each suffix r_2 of $\mathrm{rd}(t)$ with $|\mathrm{rd}_2(t)| \leq |r_2| < \mathrm{obw}(t)$. Hence, application of Lemma 3.8 yields

$$t = [\overline{\mathrm{rd}_1(t)}\mathrm{wrt}(t)\overline{\mathrm{rd}_2(t)}]_{\equiv_{\mathcal{L}}} = [\overline{r_1}\mathrm{wrt}(t)\overline{r_2}]_{\equiv_{\mathcal{L}}}$$

where r_1 is the complementary prefix of $\mathrm{rd}(t)$ wrt. r_2.

If otherwise $\mathrm{obw}(t) = \infty$, we know $w_1 \notin \mathrm{redsup}_{\mathcal{L}}(r_2)$ for each prefix $w_1 \in A_{\mathcal{L}}^*$ of $\mathrm{wrt}(t)$ and for any suffix $r_2 \in A_{\mathcal{L}}^*$ of $\mathrm{rd}(t)$ which is longer than $\mathrm{rd}_2(t)$. This implies $\mathrm{rd}_2(\mathrm{wrt}(t)\overline{r_2}) = \mathrm{rd}_2(t)$ in this case. Hence, utilization of Lemma 3.8 results in

$$t = [\overline{\mathrm{rd}_1(t)}\mathrm{wrt}(t)\overline{\mathrm{rd}_2(t)}]_{\equiv_{\mathcal{L}}} = [\overline{r_1}\mathrm{wrt}(t)\overline{r_2}]_{\equiv_{\mathcal{L}}}$$

where r_1 is the complementary prefix of $\mathrm{rd}(t)$ wrt. r_2. □

Now, we prove that the aforementioned partitions of $T \subseteq \mathcal{Q}(\mathcal{L})$ coincide with Boolean combinations of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_k for regular languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $k \in \mathbb{N}$. Note that the proof of this result essentially generalizes the proofs of [12, Lemmas 9.9-9.11] which these authors considered only reliable queues.

Lemma 5.10. *Let $\ell \in \mathbb{N}$ and $T \subseteq \mathcal{Q}(\mathcal{L})$ such that $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is recognized by a monoid with ℓ elements. Then the following two plq languages satisfy property (3) of Theorem 5.6:*

(i) $T \cap \Omega_\ell$.
(ii) $T \cap \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell$.

Proof. Let $\mathbb{F}$ be a finite monoid with $|\mathbb{F}| = \ell$ recognizing $L := \eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ via the homomorphism $\phi\colon \Sigma_{\mathcal{L}}^* \to \mathbb{F}$. Furthermore, define $\mu, \overline{\mu}\colon A_{\mathcal{L}}^* \to \mathbb{F}$ such that $\mu(w) = \phi(w)$ and $\overline{\mu}(w) = \phi(\overline{w})$.

First, we show the statement (i) by establishing the following equation

$$\begin{aligned} T \cap \Omega_\ell = & \bigcup_{\substack{r_2 \in A_{\mathcal{L}}^\ell, \alpha,\beta \in \mathbb{F}: \\ \alpha\beta\overline{\mu}(r_2) \in \phi(L)}} \mathrm{rd}^{-1}\big(\overline{\mu}^{-1}(\alpha) r_2\big) \cap \mathrm{wrt}^{-1}\big(\mu^{-1}(\beta)\big) \cap \Omega_\ell \\ & \cup \bigcup_{\substack{r_2 \in A_{\mathcal{L}}^{<\ell}, \beta \in \mathbb{F}: \\ \beta\overline{\mu}(r_2) \in \phi(L)}} \mathrm{rd}^{-1}(r_2) \cap \mathrm{wrt}^{-1}\big(\mu^{-1}(\beta)\big) \cap \Omega_\ell\,. \end{aligned}$$

We denote the left and right hand side of this equation by Y and Z, respectively. Clearly, $Y, Z \subseteq \Omega_\ell$. Hence, it suffices to show, given $t \in \Omega_\ell$, that $t \in Y$ iff $t \in Z$ holds. Towards this equivalence we first have to prove the following statement:

Claim 5.11. *Let $t \in \Omega_\ell$, $r_2 \in A_{\mathcal{L}}^*$ be the longest suffix of $\mathrm{rd}(t)$ with $|r_2| \leq \ell$, and $r_1 \in A_{\mathcal{L}}^*$ be the complementary prefix of $\mathrm{rd}(t)$ wrt. r_2. Then we have $t \in T$ if, and only if, $[\overline{r_1}\mathrm{wrt}(t)\overline{r_2}]_{\equiv_{\mathcal{L}}} \in T$.*

Proof of Claim 5.11. If $|\mathrm{rd}(t)| < \ell$ holds, we have $r_2 = \mathrm{rd}(t)$ and $r_1 = \varepsilon$. Since $\mathrm{obw}(t) > \ell > |\mathrm{rd}(t)|$ holds, we have $\mathrm{obw}(t) = \infty$ in this case. Then from Lemma 5.9 we learn $t = [\mathrm{wrt}(t)\overline{\mathrm{rd}(t)}]_{\equiv_{\mathcal{L}}} = [\overline{r_1}\mathrm{wrt}(t)\overline{r_2}]_{\equiv_{\mathcal{L}}}$ and we are done. So, from now on we assume $|\mathrm{rd}(t)| \geq \ell$. Then we have $r_2 \in A_{\mathcal{L}}^\ell$. From $\mathrm{obw}(t) > \ell$ and Lemma 5.9 we obtain words $x, y, z \in A_{\mathcal{L}}^*$ with $t = [\overline{x}y\overline{zr_2}]_{\equiv_{\mathcal{L}}}$ (note that $r_1 = xz$ and $\mathrm{wrt}(t) = y$ holds). Moreover, due to $|\mathbb{F}| = \ell$ there is $y_0 \in A_{\mathcal{L}}^{\leq \ell}$ with $\phi(y_0) = \phi(y)$. Then from $|y_0| \leq \ell = |r_2|$ we obtain $\overline{x}y_0\overline{zr_2} \equiv_{\mathcal{L}} \overline{xz}y_0\overline{r_0}$ according to Lemma 3.8. Then we see the following:

$$\begin{aligned}
t \in T &\iff \phi(\overline{x}y\overline{z}\overline{r_2}) \in \phi(L) && (\text{since } t = [\overline{x}y\overline{z}\overline{r_2}]_{\equiv_{\mathcal{L}}}) \\
&\iff \phi(\overline{x}y_0\overline{z}\overline{r_2}) \in \phi(L) && (\text{since } \phi(y) = \phi(y_0)) \\
&\iff [\overline{x}y_0\overline{z}\overline{r_2}]_{\equiv_{\mathcal{L}}} \in T \\
&\iff [\overline{x}\overline{z}y_0\overline{r_2}]_{\equiv_{\mathcal{L}}} \in T && (\text{since } \overline{x}y_0\overline{z}\overline{r_2} \equiv_{\mathcal{L}} \overline{x}\overline{z}y_0\overline{r_2}) \\
&\iff \phi(\overline{x}\overline{z}y_0\overline{r_2}) \in \phi(L) \\
&\iff \phi(\overline{x}\overline{z}y\overline{r_2}) \in \phi(L) && (\text{since } \phi(y) = \phi(y_0)) \\
&\iff [\overline{r_1}\,\mathrm{wrt}(t)\,\overline{r_2}]_{\equiv_{\mathcal{L}}} \in T\,.
\end{aligned}$$

This finally finishes the proof of Claim 5.11. □

Now, we resume the proof of the first statement in Lemma 5.10: let $t \in \Omega_\ell$, $r_2 \in A_{\mathcal{L}}^*$ be the longest suffix of $\mathrm{rd}(t)$ with $|r_2| \leq \ell$, and $r_1 \in A_{\mathcal{L}}^*$ be the complementary prefix of $\mathrm{rd}(t)$ wrt. r_2. Then we have:

$$\begin{aligned}
t \in Y &\iff t \in T \\
&\iff [\overline{r_1}\,\mathrm{wrt}(t)\,\overline{r_2}]_{\equiv_{\mathcal{L}}} \in T && (\text{by Claim 5.11}) \\
&\iff \phi(\overline{r_1}\,\mathrm{wrt}(t)\,\overline{r_2}) \in \phi(L) \\
&\iff \overline{\mu}(r_1)\,\mu(\mathrm{wrt}(t))\,\overline{\mu}(r_2) \in \phi(L) \\
&\iff t \in Z\,.
\end{aligned}$$

Now, we prove the second statement by establishing the following equation:

$$T \cap \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell = \bigcup_{0 \leq k < \ell} \; \bigcup_{\substack{r_2 \in A_{\mathcal{L}}^k, \alpha, \beta \in \mathbb{F}: \\ \alpha\beta\overline{\mu}(r_2) \in \phi(L)}} \mathrm{rd}^{-1}\big(\overline{\mu}^{-1}(\alpha) r_2\big) \cap \mathrm{wrt}^{-1}\big(\mu^{-1}(\beta)\big) \cap \Omega_k \setminus \Omega_{k+1}\,.$$

We denote the left- and right-hand side of this equation by Y and Z. From $\mathcal{Q}(\mathcal{L}) = \Omega_0 \supseteq \Omega_1 \supseteq \Omega_2 \supseteq \dots$ we obtain that $\Omega_k \setminus \Omega_{k+1} \subseteq \Omega_0 \setminus \Omega_\ell = \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell$ holds for each $0 \leq k < \ell$. Hence, we have $Y, Z \subseteq \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell$. Then it suffices to show, given $t \in \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell$ that $t \in Y$ holds if, and only if, $t \in Z$ holds.

From $t \notin \Omega_\ell$ we obtain $\mathrm{obw}(t) \leq \ell$. Then there is $0 \leq k < \ell$ with $k+1 = \mathrm{obw}(t)$ (i.e., we have $t \in \Omega_k \setminus \Omega_{k+1}$). By Lemma 5.9 we obtain $t = [\overline{r_1}w\overline{r_2}]_{\equiv_{\mathcal{L}}}$ where $\mathrm{wrt}(t) = w$, $\mathrm{rd}(t) = r_1 r_2$, and $|r_2| = k$. Then we learn:

$$t \in Y \iff \phi(\overline{r_1}w\overline{r_2}) \in \phi(L) \iff \overline{\mu}(r_1)\,\mu(w)\,\overline{\mu}(r_2) \in \phi(L) \iff t \in Z\,.$$

□

Finally, we infer the implication "(2)⇒(3)" in Theorem 5.6:

Proposition 5.12. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$ and $T \subseteq \mathcal{Q}(\mathcal{L})$ such that $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is regular. Then T is a Boolean combination of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for regular languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$.*

Proof. Let $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ be recognized by a finite monoid with ℓ elements. By set theory we have $T = (T \cap \Omega_\ell) \cup (T \cap \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell)$. Due to Lemma 5.10 the right-hand side of this equation is a finite union of plq languages satisfying property (3) of Theorem 5.6. □

5.1.3. *The implication "(3)⇒(1)" in Theorem 5.6*

Again, we fix a lossiness alphabet $\mathcal{L} = (F, U)$ with $|A_{\mathcal{L}}| \geq 2$. We have to prove now that $T \subseteq \mathcal{Q}(\mathcal{L})$ is recognizable if it is a Boolean combination of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for regular langauges $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$. Since the projections wrt and rd are homomorphisms, we obtain recognizability of $\mathrm{wrt}^{-1}(R)$ and $\mathrm{rd}^{-1}(R)$ for any regular word language $R \subseteq A_{\mathcal{L}}^*$. Hence, we only have to show that Ω_ℓ is recognizable for any number $\ell \in \mathbb{N}$. To this end, we prove that $\eta_{\mathcal{L}}^{-1}(\Omega_\ell)$ is FO-definable and therefore - due to McNaughton and Papert's Theorem [24] - even aperiodic.

We first define two FO-formulas embed_ℓ and $\mathsf{overlap}_\ell$ for natural numbers $\ell \in \mathbb{N}$ which describe the following properties: The word model $\underline{w}$ of a queue action sequence $w \in \Sigma_{\mathcal{L}}^*$ satisfies embed_ℓ if, and only if, there is a suffix r_2 of $\mathrm{rd}(w)$ of length ℓ and a prefix w_1 of $\mathrm{wrt}(w)$ such that w_1 is a reduced $\mathcal{L}$-superword of r_2. In other words, there exists an action sequence $v \in \Sigma_{\mathcal{L}}^*$ with the same projections and an overlap of length ℓ.

The formula $\mathsf{overlap}_\ell$ strengthens this as follows: $\underline{w}$ satisfies $\mathsf{overlap}_\ell$ if, and only if, there is such action sequence $v \in \Sigma_{\mathcal{L}}^*$ with the same projections and overlap of length ℓ (i.e., $\underline{w}$ satisfies embed_ℓ) and the overlap of w is at least of length ℓ.

We do this by assigning the last ℓ read actions to variables $x_1, \ldots, x_\ell$ (where x_1 is the position of the right-most read action in w and x_ℓ is the ℓ^{th} last read action) and the corresponding write actions to variables $y_1, \ldots, y_\ell$.

So, let $\ell \in \mathbb{N}$ and $x_1, \ldots, x_\ell, y_1, \ldots, y_\ell$ be variables. Then we define the following formulas:

(1) $\phi_1 := x_\ell < x_{\ell-1} < \cdots < x_1 \wedge y_\ell < y_{\ell-1} < \cdots < y_1$ - This formula guarantees that the x_i's are mutually distinct and in descending order and the same holds for the y_i's.
(2) $\phi_2 := \bigwedge_{i=1}^{\ell} \bigvee_{a \in A_{\mathcal{L}}} (\Lambda_{\overline{a}}(x_i) \wedge \Lambda_a(y_i))$ - This formula ensures that x_i reads the same letter from the queue as y_i writes into it.
(3) $\phi_3 := \forall z\colon \Big((x_\ell \leq z \wedge \Lambda_{\overline{A_{\mathcal{L}}}}(z)) \to \bigvee_{i=1}^{\ell} x_i = z\Big)$ - Satisfaction of this formula requires the x_i's to be the last ℓ read actions in w.
(4) $\phi_4 := \bigwedge_{i=1}^{\ell-1} \bigwedge_{a \in A_{\mathcal{L}}} \forall z\colon ((y_{i+1} < z < y_i \wedge \Lambda_a(y_i)) \to \Lambda_{\overline{A_{\mathcal{L}}} \cup (F \setminus \{a\})}(z))$ and $\phi_5 := \bigwedge_{a \in A_{\mathcal{L}}} \forall z\colon ((z < y_\ell \wedge \Lambda_a(y_\ell)) \to \Lambda_{\overline{A_{\mathcal{L}}} \cup (F \setminus \{a\})}(z))$ - These formulas assure that the infix $w[y_{i+1}+1, y_i-1]$ contains neither the same letter as y_i nor any unforgettable letter. Hence, together with the formulas above, these ones enforce the last ℓ read actions to be an $\mathcal{L}$-subword of a prefix of the write actions.
(5) $\phi_6 := \bigwedge_{i=1}^{\ell} y_i < x_i$ - This formula guarantees that each x_i appears right from y_i.

By conjuction of the formulas above we obtain the announced formulas:

$$\mathsf{embed}_\ell := \exists x_1, \dots, x_\ell, y_1, \dots, y_\ell \colon \phi_1 \wedge \phi_2 \wedge \phi_3 \wedge \phi_4 \wedge \phi_5 \,,$$
$$\mathsf{overlap}_\ell := \exists x_1, \dots, x_\ell, y_1, \dots, y_\ell \colon \phi_1 \wedge \phi_2 \wedge \phi_3 \wedge \phi_4 \wedge \phi_5 \wedge \phi_6 \,.$$

Example 5.13. Let $w = abba\overline{aba}$. If $U = \{a\}$ then we have $w \in L(\mathsf{overlap}_1)$ and $w \in L(\mathsf{overlap}_3)$, but $w \notin L(\mathsf{overlap}_2)$ since each assignment of y_2 to b violates ϕ_5 because of the leading a in w. If $U = \{b\}$ then we have $w \in L(\mathsf{overlap}_1)$, but $w \notin L(\mathsf{overlap}_2)$ and $w \notin L(\mathsf{overlap}_3)$ since each assignment of y_1 to a violates ϕ_4 because of the b at the third position in w.

Now let $w' = \overline{aba}abba$. Then $w' \in L(\mathsf{embed}_\ell)$ if, and only if, $w \in L(\mathsf{embed}_\ell)$ for any $\ell \in \mathbb{N}$. But $w' \in L(\mathsf{overlap}_0)$ and $w' \notin L(\mathsf{overlap}_\ell)$ for any $\ell > 0$.

In the following lemma we describe the words satisfying the formulas embed_ℓ and $\mathsf{overlap}_\ell$ for any $\ell \in \mathbb{N}$. As announced before, in the first case these are the words where the last ℓ read actions are an $\mathcal{L}$-subword of a prefix of their write actions. Furthermore, the words satisfying $\mathsf{overlap}_\ell$ also satisfy embed_ℓ and have an overlap of at least ℓ symbols.

Lemma 5.14. *Let $\ell \in \mathbb{N}$ and $w \in \Sigma_{\mathcal{L}}^*$. Then the following statements hold:*

(1) $w \in L(\mathsf{embed}_\ell)$ if, and only if, there is $r_2 \in A_{\mathcal{L}}^\ell$ with $\mathrm{rd}(w) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(w)$.

(2) $w \in L(\mathsf{overlap}_\ell)$ if, and only if, there is $r_2 \in A_{\mathcal{L}}^\ell$ with $\mathrm{rd}_2(w) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(w)$.

Proof.

(1) First, let $w \in L(\mathsf{embed}_\ell)$. Then there are letters $a_1, \dots, a_\ell, b_1, \dots, b_\ell \in \Sigma_{\mathcal{L}}$ contained in w such that their positions $p_1, \dots, p_\ell, q_1, \dots, q_\ell \in \mathrm{dom}(w)$ satisfy ϕ_1-ϕ_5. By ϕ_2 we have $a_\ell \dots a_1 = \overline{b_\ell \dots b_1}$. Due to ϕ_1 and ϕ_3 we have $\mathrm{rd}(w) \in A_{\mathcal{L}}^* b_\ell \dots b_1$ and $b_\ell \dots b_1 \preceq \mathrm{wrt}(w)$. From $\phi_4 \wedge \phi_5$ we infer that $w_1 \in \mathrm{redsup}_{\mathcal{L}}(b_\ell \dots b_1)$ holds for the prefix $w_1 := \mathrm{wrt}(w[1, q_1])$ of $\mathrm{wrt}(w)$.

For the converse implication, let $r_2 = b_\ell \dots b_1 \in A_{\mathcal{L}}^\ell$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(w)$. Since there is no restriction to the order of x_i and y_i in embed_ℓ, the language $L(\mathsf{embed}_\ell)$ contains a word if, and only if, it contains all words having the same projections. Hence, we can assume that $w \in \mathrm{NF}_{\mathcal{L}}$ and $\mathrm{rd}_2(w) = r_2$ hold. Let $p_1, \dots, p_\ell \in \mathrm{dom}(w)$ be the positions of the last ℓ read actions in w (in descending order) and let $\overline{b_1}, \dots, \overline{b_\ell} \in \overline{A_{\mathcal{L}}}$ be the letters on these positions. Then by definition of $\mathrm{nf}_{\mathcal{L}}(w)$ the positions $p_1 - 1, \dots, p_\ell - 1$ are labeled with $b_1, \dots, b_\ell$. Then it is easy to see, that

$$(\underline{w}, p_1, \dots, p_\ell, p_1 - 1, \dots, p_\ell - 1) \vDash \phi_1 \wedge \phi_2 \wedge \phi_3 \wedge \phi_4 \wedge \phi_5(x_1, \dots, x_\ell, y_1, \dots, y_\ell)$$

holds, i.e., we have $w \in L(\mathsf{embed}_\ell)$.

(2) Let $w \in L(\mathsf{overlap}_\ell)$. Then by (1) there is $r_2 = b_\ell \dots b_1 \in A_{\mathcal{L}}^\ell$ such that $\mathrm{rd}(w) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(w)$. By satisfaction of ϕ_6 we obtain that each b_i is left of $\overline{b_i}$. By Proposition 3.7 we finally obtain $\mathrm{rd}_2(w) \in A_{\mathcal{L}}^* r_2$.

Conversely, let $r_2 = b_\ell \dots b_1 \in A_{\mathcal{L}}^\ell$ such that $\mathrm{rd}_2(w) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(w)$. Then by (1) we have $w \in L(\mathsf{embed}_\ell)$. Hence, we only have to check the satisfaction of ϕ_6. We prove this by induction on the minimal length n of a derivation $w \to_{\mathfrak{R}_{\mathcal{L}}}^n \mathrm{nf}_{\mathcal{L}}(w)$ where $\mathfrak{R}_{\mathcal{L}}$ is the semi-Thue system we obtain by ordering the equations from Theorem 3.3 from left to right.

If $n = 0$ then we have $w = \mathrm{nf}_{\mathcal{L}}(w)$. Set $k := |\mathrm{rd}_2(w)| \geq \ell$. Let $p_1, \dots, p_k \in \mathrm{dom}(w)$ be the positions of the last k read actions (in descending order) in w and let $\overline{b_1}, \dots, \overline{b_k} \in \overline{A_{\mathcal{L}}}$ be the letters on these positions. Due to $\mathrm{wrt}_1(w) \in \mathrm{redsup}_{\mathcal{L}}(\mathrm{rd}_2(w)) = \mathrm{redsup}_{\mathcal{L}}(b_k \dots b_1)$ there is a factorization $v_k b_k v_{k-1} b_{k-1} \dots v_1 b_1 v_0 = \mathrm{wrt}(w)$ with $v_i \in (F \setminus \{b_i\})^*$ for any $1 \leq i \leq k$ and $v_0 = \mathrm{wrt}_2(w) \in A_{\mathcal{L}}^*$. Then, by $w = \mathrm{nf}_{\mathcal{L}}(w)$ the positions $q_i := p_i - 1$ are labeled with b_i and, hence, we have $\mathrm{wrt}(w[q_{i+1}+1, q_i]) = v_i b_i$ for each $1 \leq i \leq k$ (where $q_{k+1} = 0$). Since we also have $w_1 \in \mathrm{redsup}_{\mathcal{L}}(b_\ell \dots b_1)$ for a prefix w_1 of $\mathrm{wrt}(w)$ there is another factorization $v'_\ell b_\ell \dots v'_1 b_1 v'_0 = \mathrm{wrt}(w)$ with $v'_i \in (F \setminus \{b_i\})^*$ for any $1 \leq i \leq \ell$ and $v'_0 \in A_{\mathcal{L}}^*$. Let $q'_\ell, \dots, q'_1$ be the positions of $b_\ell, \dots, b_1$ with $\mathrm{wrt}(w[q'_{i+1}+1, q'_i]) = v'_i b_i$ for each $1 \leq i \leq \ell$ (where $q'_{\ell+1} = 0$). Hence, by $\ell \leq k$ we can infer $q'_i \leq q_i = p_i - 1 < p_i$ for each $1 \leq i \leq \ell$. This finally implies

$$(\underline{w}, p_1, \dots, p_\ell, q'_1, \dots, q'_\ell) \vDash \phi_1 \wedge \phi_2 \wedge \phi_3 \wedge \phi_4 \wedge \phi_5 \wedge \phi_6(x_1, \dots, x_\ell, y_1, \dots, y_\ell),$$

i.e., we learn $w \in L(\mathsf{overlap}_\ell)$.

Now, assume $n > 0$. Then there is $w' \in \Sigma_{\mathcal{L}}^*$ with $w \to_{\mathfrak{R}_{\mathcal{L}}} w' \to_{\mathfrak{R}_{\mathcal{L}}}^{n-1} \mathrm{nf}_{\mathcal{L}}(w)$. By induction hypothesis we know $w' \in L(\mathsf{overlap}_\ell)$ and we have to show $w \in L(\mathsf{overlap}_\ell)$. We know that w' satisfies embed_ℓ. Since the application of any rule of $\mathfrak{R}_{\mathcal{L}}$ transposes only one write action with another read action, we also have $w \in L(\mathsf{embed}_\ell)$.

Let $p_1, \dots, p_\ell, q_1, \dots, q_\ell \in \mathrm{dom}(w') = \mathrm{dom}(w)$ be the positions in w' satisfying ϕ_1-ϕ_6. The transposition of the write and read action has two effects: the position of one read action increases by one, i.e., at most one p_i increases. Additionally, the position of one write action decreases by one, i.e., at most one q_i decreases. This yields new positions $p'_1 \dots, p'_\ell, q'_1, \dots, q'_\ell \in \mathrm{dom}(w)$ satisfying ϕ_1-ϕ_5 in $\underline{w}$ and we have $p'_i \in \{p_i, p_i + 1\}$ and $q'_i \in \{q_i, q_i - 1\}$ for each $1 \leq i \leq \ell$. However, we also have

$$q'_i \leq q_i < p_i \leq p'_i$$

for each $1 \leq i \leq \ell$. In other words, the chosen positions satisfy ϕ_6 and, hence, we have $w \in L(\mathsf{overlap}_\ell)$.

□

From this lemma we also obtain another characterization of the equivalence classes of the behavioral equivalence $\equiv_{\mathcal{L}}$. So, let $w \in \Sigma_{\mathcal{L}}^*$ and $\ell \in \mathbb{N}$ be maximal such that $w \in L(\mathsf{overlap}_\ell)$. Moreover, let $k := \mathrm{obw}(w) < \infty$. Then we know that $w \in L(\mathsf{embed}_k \wedge \neg \mathsf{overlap}_k)$ holds. This implies, that a word $v \in \Sigma_{\mathcal{L}}^*$ satisfies $v \equiv_{\mathcal{L}} w$ whenever the last ℓ read actions of v appear right of their corresponding write actions and at least one of the last k read actions appears left of its corresponding write action. We will use this observation multiple times in the proof of our Main Theorem 5.1. We can also infer the recognizability of the plq languages Ω_ℓ from this observation:

Lemma 5.15. *Let $\ell \in \mathbb{N}$. Then Ω_ℓ is aperiodic and, hence, recognizable.*

Proof. Since $\eta_{\mathcal{L}}$ is surjective, it suffices to show that $\eta_{\mathcal{L}}^{-1}(\Omega_\ell)$ is aperiodic in $\Sigma_{\mathcal{L}}^*$. By Lemmas 5.5 and 5.14 we have

$$\eta_{\mathcal{L}}^{-1}(\Omega_\ell) = L\left(\bigwedge_{k=1}^{\ell} (\mathsf{embed}_k \to \mathsf{overlap}_k)\right) .$$

Then by [24] $\eta_{\mathcal{L}}^{-1}(\Omega_\ell)$ is aperiodic since the given formula is contained in FO. □

Finally, from this lemma we obtain the implication "(3)⇒(1)" in Theorem 5.6 which finishes the proof of Theorem 5.6.

Proposition 5.16. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$ and $T \subseteq \mathcal{Q}(\mathcal{L})$ be a Boolean combination of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for regular languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$. Then T is recognizable.*

Proof. Since the sets $\mathrm{wrt}^{-1}(R)$ and $\mathrm{rd}^{-1}(R)$ are recognizable for any regular language $R \subseteq A_{\mathcal{L}}^*$ and since Ω_ℓ is recognizable for any $\ell \in \mathbb{N}$ due to Lemma 5.15, the plq language T is recognizable by closure properties of the class of recognizable languages in $\mathcal{Q}(\mathcal{L})$. □

5.2. *From recognizability to q-rationality*

Now, we are able to prove the equivalences from our main theorem (Theorem 5.1). In this subsection we prove the first of three implications. Concretely, we show that each recognizable plq language is q-rational. To this end, we first have to define this notion which is a restriction to the classical rational expressions. We need this restriction since the classes of rational and recognizable plq languages do not coincide (cf. Theorem 4.5). Note that this is in contrast to the situation in free monoids, where a word language is rational if, and only if, it is recognizable [15]. Though we can use Ochmański's approach from [27] to generate the recognizable languages. Concretely, we restrict the iteration and the concatenation in the plq monoid in an appropriate way. Unfortunately, we still cannot generate all recognizable plq languages by union, restricted product, and restricted iteration (we will see an example later in this subsection). Hence, we have to add another, non-monotonic[c] operation to our expressions: the complement operation. We call the plq languages generated by those operations *q-rational* and prove that these are exactly the recognizable languages in the plq monoid.

At first, we prove that the class of recognizable plq languages is neither closed under iteration nor it is closed under concatenation:

Lemma 5.17. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_\mathcal{L}| \geq 2$. Then the following statements hold:*

(1) There is a recognizable language $T \subseteq \mathcal{Q}(\mathcal{L})$ such that T^* is not recognizable in $\mathcal{Q}(\mathcal{L})$.
(2) There are recognizable languages $S, T \subseteq \mathcal{Q}(\mathcal{L})$ such that $S \cdot T$ is not recognizable in $\mathcal{Q}(\mathcal{L})$.

Proof. Let $a \in A_\mathcal{L}$ be a letter.

Towards the former statement set $T := \{[a\overline{a}]_{\equiv_\mathcal{L}}\}$ which is recognizable since $\eta_\mathcal{L}^{-1}(T)$ is finite. Using the equation (iii) of Theorem 3.3 we learn

$$\eta_\mathcal{L}^{-1}(T^*) \cap a^*\overline{a}^* = \{a^n\overline{a}^n \mid n \in \mathbb{N}\}\,.$$

This language is not regular and, hence, by closure properties of the class of regular languages $\eta_\mathcal{L}^{-1}(T^*)$ is not regular. Thus, T^* is not recognizable.

Towards the latter statement set $S := \{[a]_{\equiv_\mathcal{L}}\}^*$ and $T := \{[\overline{a}]_{\equiv_\mathcal{L}}\}^*$. Since $\eta_\mathcal{L}^{-1}(S) = a^*$ and $\eta_\mathcal{L}^{-1}(T) = \overline{a}^*$ holds, both plq languages are recognizable in $\mathcal{Q}(\mathcal{L})$. According to equation (iii) of Theorem 3.3 we obtain

$$\eta_\mathcal{L}^{-1}(S \cdot T) \cap \overline{a}a^*\overline{a}^* = \{\overline{a}a^m\overline{a}^n \mid m, n \in \mathbb{N}, m \leq n\}\,.$$

This language is not regular and, hence, $S \cdot T$ is not recognizable. □

Note that Lemma 5.17(1) is a very similar situation as in trace monoids. Here, Ochmański proved in [27] that there are recognizable trace languages such that their iteration is not recognizable anymore. However, the iteration of connected, recognizable trace languages always is recognizable. Hence, this led to the definition of the so-called c-rational expressions.

In the plq monoid, we additionally have to restrict concatenation due to Lemma 5.17(2). According to these restrictions we will define now the so-called *q-rational* languages in the plq monoid. Afterwards we prove that this is a suitable modification of rationality to describe exactly the recognizable plq languages.

Let $\mathcal{L} = (F, U)$ be a lossiness alphabet. At first, we say that a $\mathcal{Q}(\mathcal{L})$-language is *q^+-rational* if it can be obtained by the following rules:

[c]Let $\leq$ be a partial ordering on a set S. A function $f\colon S \to S$ is *monotonic* if for all $x, y \in S$ with $x \leq y$ we also have $f(x) \leq f(y)$.

(1^+) $\mathrm{wrt}^{-1}(\varepsilon)$, $\mathrm{wrt}^{-1}(\emptyset) = \emptyset$, and $\mathrm{wrt}^{-1}(a)$ are q^+-rational for any $a \in A_{\mathcal{L}}$.
(2^+) if $S, T \subseteq \mathcal{Q}(\mathcal{L})$ are q^+-rational then $S \cup T$, $S \cdot T$, and S^* are q^+-rational.

Similarly, we obtain the class of *q^--rational* languages in $\mathcal{Q}(\mathcal{L})$ by replacing wrt^{-1} by rd^{-1} in the rules above.

Observation 5.18. *Let $T \subseteq \mathcal{Q}(\mathcal{L})$. Then the following statements hold:*

(1) T is q^+-rational if, and only if, there is a regular word language $R \subseteq A_{\mathcal{L}}^*$ with $T = \mathrm{wrt}^{-1}(R)$.
(2) T is q^--rational if, and only if, there is a regular word language $R \subseteq A_{\mathcal{L}}^*$ with $T = \mathrm{rd}^{-1}(R)$.

Proof idea. Both equivalences hold since $\mathrm{wrt}^{-1}, \mathrm{rd}^{-1} \colon 2^{A^*} \to 2^{\mathcal{Q}(\mathcal{L})}$ are homomorphisms wrt. union, product, and iteration operations. □

Finally, a $\mathcal{Q}(\mathcal{L})$-language is *q-rational* if it can be constructed by the following rules:

(1) if $T \subseteq \mathcal{Q}(\mathcal{L})$ is q^+- or q^--rational it also is q-rational.
(2) if $S, T \subseteq \mathcal{Q}(\mathcal{L})$ are q-rational then $S \cup T$ and $\mathcal{Q}(\mathcal{L}) \setminus S$ are q-rational.
(3) if $S \subseteq \mathcal{Q}(\mathcal{L})$ is q^+-rational and $T \subseteq \mathcal{Q}(\mathcal{L})$ is q^--rational such that $\mathrm{rd}(T)$ is finite (i.e., T is obtained without usage of the *-operator) then $S \cdot \mathcal{Q}(\mathcal{L}) \cdot T$ is q-rational

Remark 5.19. Recall the classical definition of rational languages. The class of these languages is the closure of the finite languages under union, concatenation, and iteration. All of these operations are monotonic. For q-rationality we additionally need the closure under complement which is not monotonic. Until now it is still an open question whether there is another characterization of the recognizable plq languages using monotonic operations, only.

Example 5.20. Let $T = \{t \in \mathcal{Q}(\mathcal{L}) \mid \mathrm{wrt}(t) \in (ab)^*, \mathrm{rd}(t) = b\}$. Then T is q-rational since we have

$$T = \mathrm{rd}^{-1}(b) \cap (\mathrm{wrt}^{-1}(a) \cdot \mathrm{wrt}^{-1}(b))^* .$$

Note that the class of q-rational languages also is closed under intersection due to Rule (2), i.e., this class is a Boolean algebra.

At first sight, the choice of Rule (3) seems to be some kind of random. The following example shows that we can remove neither the factor "$\mathcal{Q}(\mathcal{L})$", which appears as separator in this product, nor the finiteness of $\mathrm{rd}(T)$. Additionally, we cannot simply remove this rule since the recognizable plq language $\{[a\overline{a}]_{\equiv_{\mathcal{L}}}\}$ cannot be built by application of the Rules (1) and (2), only.

Remark 5.21. Let $a, b \in A_{\mathcal{L}}$ be distinct letters. Then the language $\mathrm{wrt}^{-1}((ab)^*a) \cdot \mathrm{rd}^{-1}(\varepsilon)$ is not recognizable, since $\eta_{\mathcal{L}}^{-1}(\mathrm{wrt}^{-1}((ab)^*a)\,\mathrm{rd}^{-1}(\varepsilon)) \cap (ab)^*\overline{(ab)^*}$ contains exactly those words $(ab)^m\overline{(ab)^n}$ with $m > n$, which is not regular.

Now let $a, b, c \in A_{\mathcal{L}}$ be distinct letters with $a, c \in U$. Then the language

$$\mathrm{wrt}^{-1}(aA_{\mathcal{L}}^*c) \cdot \mathcal{Q}(\mathcal{L}) \cdot \mathrm{rd}^{-1}(aA_{\mathcal{L}}^*c)$$

is not recognizable since

$$L := \eta_{\mathcal{L}}^{-1}(\mathrm{wrt}^{-1}(aA_{\mathcal{L}}^*c) \cdot \mathcal{Q}(\mathcal{L}) \cdot \mathrm{rd}^{-1}(aA_{\mathcal{L}}^*c)) \cap \overline{ab^*c}ab^*c$$

contains exactly those words $\overline{ab^mc}ab^nc$ satisfying $m \neq n$ if $b \in U$ or $m > n$ if $b \in F$ holds. In both cases L is not regular.

Now we can prove the implication "(A)⇒(B)" in Theorem 5.1. To do this, we utilize Theorem 5.6(3). Concretely, we understand a recognizable plq language as a Boolean combination of languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for regular languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$. Then we prove q-rationality by induction on the syntax tree of such expression. The most complicated case in this proof is to show that Ω_ℓ is q-rational. For this proof we need the following lemma:

Lemma 5.22. *Let $\ell \in \mathbb{N}$, $t \in \mathcal{Q}(\mathcal{L})$, and $r_2 = a_1 \dots a_\ell$ with $a_1, \dots, a_\ell \in A_{\mathcal{L}}$. Then we have $\mathrm{rd}_2(t) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(t)$ if, and only if,*

$$t \in \mathrm{wrt}^{-1}\left(\prod_{i=1}^{\ell} F^* a_i\right) \cdot \mathcal{Q}(\mathcal{L}) \cdot \mathrm{rd}^{-1}(r_2) \ .$$

Proof. First, assume $\mathrm{rd}_2(t) \in A_{\mathcal{L}}^* r_2$ and $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(t)$. From Lemma 3.8 we know $t = [\overline{\mathrm{rd}_1(t)}\mathrm{wrt}(t)\overline{\mathrm{rd}_2(t)}]_{\equiv_{\mathcal{L}}}$. Then by assumption we obtain

$$t = \left[\overline{\mathrm{rd}_1(t)}\mathrm{wrt}(t)\overline{\mathrm{rd}_2(t)}\right]_{\equiv_{\mathcal{L}}} \in \mathrm{wrt}^{-1}\left(\prod_{i=1}^{\ell} F^* a_i\right) \cdot \mathcal{Q}(\mathcal{L}) \cdot \mathrm{rd}^{-1}(r_2) \ .$$

Now assume $t \in \mathrm{wrt}^{-1}\left(\prod_{i=1}^{\ell} F^* a_i\right) \cdot \mathcal{Q}(\mathcal{L}) \cdot \mathrm{rd}^{-1}(r_2)$. Then we have $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(t)$ and $\mathrm{rd}(t) \in A_{\mathcal{L}}^* r_2$. Furthermore, there are $w_1, w_2, w_3 \in \Sigma_{\mathcal{L}}^*$ with $t = [w_1w_2w_3]_{\equiv_{\mathcal{L}}}$, $\mathrm{wrt}(w_1) \in \prod_{i=1}^{\ell} F^* a_i$, and $\mathrm{rd}(w_3) = r_2$. Then the letters $\overline{a_1}, \dots, \overline{a_\ell}$ appear to the right of $a_1, \dots, a_\ell$ in $w_1w_2w_3$, i.e., $w_1w_2w_3 \in L(\mathsf{overlap}_\ell)$. Hence, by Lemma 5.14(2), we obtain $\mathrm{rd}_2(w_1w_2w_3) \in A_{\mathcal{L}}^* r_2$, i.e., $\mathrm{rd}_2(t) \in A_{\mathcal{L}}^* r_2$ holds. □

Finally, we can state the following implication:

Proposition 5.23. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$ and $T \subseteq \mathcal{Q}(\mathcal{L})$ be recognizable. Then T is q-rational.*

Proof. By Theorem 5.6(3) T is a Boolean combination of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for regular languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$. We prove the claim by induction on the syntax tree.

At first, assume $T = \mathrm{wrt}^{-1}(R)$ for a regular language $R \subseteq A_{\mathcal{L}}^*$. Then, T is q$^+$-rational by Observation 5.18(1) and, hence, q-rational. Similarly, $T = \mathrm{rd}^{-1}(R)$ is q$^-$-rational and hence q-rational for any regular language $R \subseteq A_{\mathcal{L}}^*$.

Next, suppose $T = \Omega_\ell$ for a number $\ell \in \mathbb{N}$. Then by the Lemmas 5.5 and 5.22 we have

$$\Omega_\ell = \bigcap_{r_2 \in A_{\mathcal{L}}^{\leq \ell}} \Big(\mathcal{Q}(\mathcal{L}) \setminus \big(\mathrm{wrt}^{-1}(W_{r_2} A_{\mathcal{L}}^*) \cap \mathrm{rd}^{-1}(A_{\mathcal{L}}^* r_2) \big) \cup \mathrm{wrt}^{-1}(W_{r_2}) \cdot \mathcal{Q}(\mathcal{L}) \cdot \mathrm{rd}^{-1}(r_2) \Big),$$

where $W_{r_2} = \prod_{i=1}^{k} F^* a_i$ with $r_2 = a_1 \dots a_k$. Since the plq languages $\mathrm{wrt}^{-1}(W_{r_2} A_{\mathcal{L}}^*)$, $\mathrm{rd}^{-1}(A_{\mathcal{L}}^* r_2)$, $\mathrm{wrt}^{-1}(W_{r_2})$, and $\mathrm{rd}^{-1}(r_2)$ are q-rational by the first case of this proof, Ω_ℓ is q-rational due to Rules (2) and (3).

Finally, assume $T = S_1 \cup S_2$, $T = S_1 \cap S_2$, or $T = \mathcal{Q}(\mathcal{L}) \setminus S_1$. Then by induction hypothesis S_1 and S_2 are q-rational. Hence, using Rule (2) T is q-rational as well. □

5.3. *From q-rationality to logical definability*

The second implication from Theorem 5.1 states that each q-rational plq language is definable in a special monadic second-order logic which we call $\mathsf{MSO_q}$. So, the aim of this subsection is to derive a signature from Büchi's word-logic and corresponding structures $\mathcal{S}(w)$ for each word $w \in \Sigma_{\mathcal{L}}^*$ such that we have the following properties:

(1) For two action sequences $v, w \in \Sigma_{\mathcal{L}}^*$ we have $\mathcal{S}(v) \cong \mathcal{S}(w)$ if, and only if, $v \equiv_{\mathcal{L}} w$ holds (i.e., iff v and w induce the same transformation $[v]_{\equiv_{\mathcal{L}}} = [w]_{\equiv_{\mathcal{L}}}$).
(2) The monadic second-order logic on these structures describes exactly the recognizable properties in the plq monoid. For this purpose, we have to exhibit the knowledge from the preceding subsections.

To this end, we have to revisit the rules from the semi-Thue system $\mathfrak{R}_{\mathcal{L}}$ which arises from ordering the equations in Theorem 3.3 from left to right. First, we can observe that the application of any rule of $\mathfrak{R}_{\mathcal{L}}$ to the word w does not change the projections $\mathrm{wrt}(w)$ and $\mathrm{rd}(w)$. For example, if the i^{th} read action in w is right of the j^{th} read action (i.e., $i < j$), then this also holds for any action sequence satisfying $v \equiv_{\mathcal{L}} w$. In particular, the i^{th} read action in w agrees with the i^{th} read action in v in this case.

So, similar to the word models $\underline{w}$, the universe of our new structure $\mathcal{S}(w)$ is the set of positions $\mathrm{dom}(w)$ in w. Additionally, we need a partial ordering $\leq$ satisfying the following property: $p \leq q$ if, and only if: (1) p is the position of the i^{th} write or read action, (2) q is the position of the j^{th} write or read action, and (3) for each $v \equiv_{\mathcal{L}} w$ the i^{th} write / read action in v appears to the left of the j^{th} write / read action in v. Unfortunately, monadic second-order formulas over these structures are able to describe non-recognizable properties as the following remark shows:

Remark 5.24. Consider the first-order formula $\phi := \exists x\colon \Lambda_{A_{\mathcal{L}}}(x) \wedge \forall y\colon \Lambda_{\overline{A_{\mathcal{L}}}}(y) \to x \leq y$, i.e., $[w]_{\equiv_{\mathcal{L}}}$ satisfies ϕ if, and only if, for each $v \equiv_{\mathcal{L}} w$ there is a write action appearing to the left of all read actions. In other words, this is the language of all words $w \in \Sigma_{\mathcal{L}}^*$ with $\mathrm{rd}_1(w) = \varepsilon$. Then we can see

$$\{w \in \Sigma_{\mathcal{L}}^* \mid \mathcal{S}(w) \models \phi\} \cap \overline{a}^* a^* \overline{a}^* = \{a^m \overline{a}^n \mid m \geq n\}$$

which is no regular language. Hence, the language of all transformations $t \in \mathcal{Q}(\mathcal{L})$ satisfying ϕ is not recognizable.

Due to this remark we have to weaken our structures $\mathcal{S}(w)$. A possible approach is to split $\leq$ into the following relations: $\leq_+$ which is the restriction of $\leq$ to the positions of write actions, $\leq_-$ which is the restriction of $\leq$ to the positions of read actions, and $\leq_{\mathrm{rw}}$ which is the restriction of $\leq$ to pairs (p,q) where p is the position of a read action and q is the position of a write action. In other words, we remove those pairs (p,q) from $\leq$ where p is the position of a write action and q is one of a read action. Unfortunately, those weakened structures are still too expressive:

Remark 5.25. Consider the first-order formula

$$\phi := \exists x, y\colon \begin{pmatrix} \Lambda_{A_{\mathcal{L}}}(x) \wedge \forall z\colon (\Lambda_{A_{\mathcal{L}}}(z) \to z \leq_+ x) \wedge \\ \Lambda_{\overline{A_{\mathcal{L}}}}(y) \wedge \forall z\colon (\Lambda_{\overline{A_{\mathcal{L}}}}(z) \to y \leq_- z) \wedge \neg y \leq_{\mathrm{rw}} x \end{pmatrix},$$

i.e., $[w]_{\equiv_{\mathcal{L}}}$ satisfies ϕ if, and only if, there is $v \equiv_{\mathcal{L}} w$ in which the first read action appears to the right of the last write action. Then we have

$$\{w \in \Sigma_{\mathcal{L}}^* \mid \mathcal{S}(w) \models \phi\} \cap \overline{a} a^* \overline{a}^* = \{\overline{a} a^\ell \overline{a}^m \mid m \geq \ell\}.$$

Since the right-hand side of the equation above is not regular, the language of all transformations $t \in \mathcal{Q}(\mathcal{L})$ satisfying ϕ is not recognizable either.

Another way to weaken the expressiveness of these structures is the following: we further split $\leq_{\mathrm{rw}}$ into relations $\leq_{\mathrm{rw},\ell}$ where $\ell > 0$ is a positive integer. This relation $\leq_{\mathrm{rw},\ell}$ contains exactly those tuples (p,q) from $\leq_{\mathrm{rw}}$ where p is the position of the ℓ^{th} last read action. Note that we have

$$\leq_{\mathrm{rw}} = \bigcup_{\ell > 0} \leq_{\mathrm{rw},\ell} .$$

We can see that for a given word $w \in \Sigma_{\mathcal{L}}^*$ and a number $\ell > 0$ all tuples in the relation $\leq_{\mathrm{rw},\ell}$ agree in their first component. Hence, we are able to project these relations to their second component yielding the sets $P_\ell := \{q \in \mathrm{dom}(w) \mid \exists p \in \mathrm{dom}(w)\colon p \leq_{\mathrm{rw},\ell} q\}$.

All in all, we can define the *plq model* of a word $w = \alpha_1\alpha_2\ldots\alpha_n$ with $\alpha_1, \alpha_2, \ldots, \alpha_n \in \Sigma_{\mathcal{L}}$ as the following (infinite) relational structure:

$$\widetilde{w} := (\mathrm{dom}(w), <_+^w, <_-^w, (P_\ell^w)_{\ell>0}, (\Lambda_\alpha^w)_{\alpha\in\Sigma_{\mathcal{L}}})$$

where $\mathrm{dom}(w) := \{1, \ldots, n\}$ is the set of positions in w, $\Lambda_\alpha^w := \{i \in \mathrm{dom}(w) \mid \alpha_i = \alpha\}$ are the labelings of the positions, $<_+^w$ and $<_-^w$ are the natural orderings on $\Lambda_{A_{\mathcal{L}}}^w := \bigcup_{a\in A_{\mathcal{L}}} \Lambda_a^w$ and $\Lambda_{\overline{A_{\mathcal{L}}}}^w$, respectively, and

$$P_\ell^w := \{p \in \Lambda_{A_{\mathcal{L}}}^w \mid \forall v_1, v_2 \in \Sigma_{\mathcal{L}}^*, w \equiv_{\mathcal{L}} v_1v_2, \mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p]) \Rightarrow |\mathrm{rd}(v_2)| < \ell\}.$$

In other words, we have $p \in P_\ell^w$ if, and only if, $\alpha_p \in A_{\mathcal{L}}$ and the ℓ^{th} last read action in w is left of α_p in any equivalently behaving action sequence $v \equiv_{\mathcal{L}} w$. This is conform to the approaches known from [4,7] since the relations $<_+^w$, $<_-^w$, and P_ℓ^w specify which letter have to appear to the left of another one in any word equivalent to w.

Before we show that the monadic second-order formulas on the plq models describe exactly the recognizable plq languages, we first show several basic properties of the plq model $\widetilde{w}$ of an action sequence $w \in \Sigma_{\mathcal{L}}^*$. First, we show that $\widetilde{w}$ identifies the equivalence class $[w]_{\equiv_{\mathcal{L}}}$:

Lemma 5.26. *Let $v, w \in \Sigma_{\mathcal{L}}^*$. Then we have $v \equiv_{\mathcal{L}} w$ if, and only if, $\widetilde{v} \cong \widetilde{w}$.*

Proof. At first, we assume $v \equiv_{\mathcal{L}} w$. Then by Proposition 3.7 there is a permutation σ of $\mathrm{dom}(v) = \mathrm{dom}(w)$ such that σ is compatible with $\leq_+$, $\leq_-$, and Λ_α for any $\alpha \in \Sigma_{\mathcal{L}}$. Additionally, by definition σ is compatible with P_ℓ for any $\ell > 0$. Hence, σ is an isomorphism from $\widetilde{v}$ into $\widetilde{w}$, i.e., $\widetilde{v} \cong \widetilde{w}$.

Now assume $v \not\equiv_{\mathcal{L}} w$. If $\mathrm{wrt}(v) \neq \mathrm{wrt}(w)$ or $\mathrm{rd}(v) \neq \mathrm{rd}(w)$ then we know $\underline{\mathrm{wrt}(v)} \not\cong \underline{\mathrm{wrt}(w)}$ or $\underline{\mathrm{rd}(v)} \not\cong \underline{\mathrm{rd}(w)}$. Note that these word models are sub-structures of the plq models $\widetilde{v}$ resp. $\widetilde{w}$. Hence, any possible bijection from $\mathrm{dom}(v)$ into $\mathrm{dom}(w)$ (if there is any) cannot be compatible with $\leq_+$, $\leq_-$, and Λ_α at the same time. This implies $\widetilde{v} \not\cong \widetilde{w}$ in this case. So, we can assume $\mathrm{wrt}(v) = \mathrm{wrt}(w)$ and $\mathrm{rd}(v) = \mathrm{rd}(w)$ from now on. Let σ be the uniquely determined permutation of $\mathrm{dom}(v) = \mathrm{dom}(w)$ which is compatible with $\leq_+$, $\leq_-$, and Λ_α for any $\alpha \in \Sigma_{\mathcal{L}}$.

By Proposition 3.7 we infer from $v \not\equiv_{\mathcal{L}} w$ that $\mathrm{rd}_2(v) \neq \mathrm{rd}_2(w)$ holds. Since both words are suffixes of $\mathrm{rd}(v) = \mathrm{rd}(w)$ we may assume that $\mathrm{rd}_2(w)$ is a proper suffix of $\mathrm{rd}_2(v)$, i.e., $|\mathrm{rd}_2(v)| > |\mathrm{rd}_2(w)|$. Note that we have $\mathrm{wrt}(v) = \mathrm{wrt}(w) \neq \varepsilon$, since otherwise we would have $\mathrm{rd}_2(v) = \mathrm{rd}_2(w) = \varepsilon$ by definition of $\mathrm{redsup}_{\mathcal{L}}$.

Let $p \in \mathrm{dom}(v)$ be the position of the last write action in v. From Lemma 3.8 we know $v \equiv_{\mathcal{L}} \overline{\mathrm{rd}_1(v)}\mathrm{wrt}(v)\overline{\mathrm{rd}_2(v)}$ and, hence, $p \notin P_{|\mathrm{rd}_2(v)|}^v$. Now, let $w_1, w_2 \in \Sigma_{\mathcal{L}}^*$ with $w \equiv_{\mathcal{L}} w_1 w_2$ and $\mathrm{wrt}(w_1) = \mathrm{wrt}(w[1, \sigma(p)]) = \mathrm{wrt}(w)$ (note that $\sigma(p)$ is the position of the last write action in w). Suppose $|\mathrm{rd}(w_2)| \geq |\mathrm{rd}_2(v)|$. Then we have $w_1 w_2 \in L(\mathsf{overlap}_{|\mathrm{rd}_2(v)|})$ (note that $x \in \mathrm{redsup}_{\mathcal{L}}(\mathrm{rd}_2(v))$ holds for a prefix x of $\mathrm{wrt}(w) = \mathrm{wrt}(v)$) implying $|\mathrm{rd}_2(w)| = |\mathrm{rd}_2(w_1 w_2)| \geq |\mathrm{rd}_2(v)|$ by Lemma 5.14(2) - contradicting to our assumption $|\mathrm{rd}_2(w)| < |\mathrm{rd}_2(v)|$. Hence, we infer $|\mathrm{rd}(w_2)| < |\mathrm{rd}_2(v)|$. Then we have $\sigma(p) \in P_{|\mathrm{rd}_2(v)|}^w$. Thus, σ is no isomorphism from $\widetilde{v}$ into $\widetilde{w}$. Since σ is unique, we have $\widetilde{v} \not\cong \widetilde{w}$. $\square$

Due to Lemma 5.26 we are able to define the *plq model* of a transformation $t \in \mathcal{Q}(\mathcal{L})$ by $\widetilde{t} := \widetilde{\mathrm{nf}_{\mathcal{L}}(t)}$.

By definition the signature of our plq models $\widetilde{w}$ is infinite. However, we can represent these structures finitely, since we can split $\mathbb{N} \setminus \{0\}$ into at most three intervals I_1, I_2, and I_3 such that the sets P_ℓ^w and P_k^w coincide if, and only if, ℓ and k belong to the same interval I_i. These intervals are closely related to $\mathrm{wrt}(w)$,

$\mathrm{rd}(w)$, and $\mathrm{obw}(w)$: if $|\mathrm{wrt}(w)| = 0$ we have $w \in \overline{A_{\mathcal{L}}}^*$ and, hence, $P_\ell^w = \emptyset$ for each $\ell > 0$. The case $|\mathrm{wrt}(w)| > 0$ is considered in the following lemma:

Lemma 5.27. *Let $w \in \Sigma_{\mathcal{L}}^*$ with $|\mathrm{wrt}(w)| > 0$. Then the following statements hold:*

(1) For each $1 \leq \ell < \min\{\mathrm{obw}(w), |\mathrm{rd}(w)| + 1\}$ we have $P_\ell^w = \emptyset$.
(2) For each $\mathrm{obw}(w) \leq \ell \leq |\mathrm{rd}(w)|$ we have $P_\ell^w = P_{\mathrm{obw}(w)}^w \neq \emptyset$.
(3) For each $\ell > |\mathrm{rd}(w)|$ we have $P_\ell^w = \Lambda_{A_{\mathcal{L}}}^w$.

Proof.

(1) Since $\ell < \mathrm{obw}(w)$ holds, we have $[w]_{\equiv_{\mathcal{L}}} \in \Omega_\ell$. Let $r_1, r_2 \in A_{\mathcal{L}}^*$ with $\mathrm{rd}(w) = r_1 r_2$ and $|r_2| = \ell$. Then by Lemma 5.9 there is $v \in \Sigma_{\mathcal{L}}^*$ such that $w \equiv_{\mathcal{L}} v\overline{r_2}$ holds. Consequently, any write action in w can be followed by $\geq \ell$ read actions, i.e., $P_\ell^w = \emptyset$.

(2) First, we show $p := \max \Lambda_{A_{\mathcal{L}}}^w \in P_{\mathrm{obw}(w)}^w \neq \emptyset$ (i.e., the last write position in w is in this set). Let $v_1, v_2 \in \Sigma_{\mathcal{L}}^*$ with $w \equiv_{\mathcal{L}} v_1 v_2$ and $\mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p]) = \mathrm{wrt}(w)$. We show $|\mathrm{rd}(v_2)| < \mathrm{obw}(w)$. Since we have $|\mathrm{rd}_2(w)| < \mathrm{obw}(w) \leq |\mathrm{rd}(w)|$ we obtain $w \in L(\mathsf{embed}_{\mathrm{obw}(w)})$ and $w \notin L(\mathsf{overlap}_{\mathrm{obw}(w)})$. From Lemma 5.14, $w \equiv_{\mathcal{L}} v_1 v_2$, and $w \notin L(\mathsf{overlap}_{\mathrm{obw}(w)})$ we also obtain $v_1 v_2 \notin L(\mathsf{overlap}_{\mathrm{obw}(w)})$. Hence, the satisfaction of ϕ_6 implies that $|\mathrm{rd}_2(v_2)| < \mathrm{obw}(w)$ and, therefore, $p \in P_{\mathrm{obw}(w)}^w$.

Next, we prove $P_\ell^w = P_{\mathrm{obw}(w)}^w$. By definition, the inclusion "$\supseteq$" is trivial. We prove the converse inclusion "$\subseteq$" by induction on ℓ. If $\ell = \mathrm{obw}(w)$, we are done. So, assume $\ell > \mathrm{obw}(w)$. Let $p \in \Lambda_{A_{\mathcal{L}}}^w \setminus P_{\ell-1}^w = \Lambda_{A_{\mathcal{L}}}^w \setminus P_{\mathrm{obw}(w)}^w$. Then there are $v_1, v_2 \in \Sigma_{\mathcal{L}}^*$ with $w \equiv_{\mathcal{L}} v_1 v_2$, $\mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p])$, and $|\mathrm{rd}(v_2)| \geq \ell - 1$. We are done, if $|\mathrm{rd}(v_2)| \geq \ell$. Hence, we assume $|\mathrm{rd}(v_2)| = \ell - 1$ from now on.

Let $r_1, r_2 \in A_{\mathcal{L}}^*$ with $\mathrm{rd}(w) = r_1 r_2$ and $|r_2| = \ell - 1$, i.e., $\mathrm{rd}(v_1) = r_1$ and $\mathrm{rd}(v_2) = r_2$. Additionally, let $r_1 = r_1' a$ with $a \in A_{\mathcal{L}}$ and let $v_1' \in \Sigma_{\mathcal{L}}^*$ be the word which arises from v_1 by removing the rightmost $\overline{a}$.

Suppose $w_1 \notin \mathrm{redsup}_{\mathcal{L}}(a r_2)$ for any prefix w_1 of $\mathrm{wrt}(w)$. Using the rules from the semi-Thue system $\mathfrak{R}_{\mathcal{L}}$, we can move the letter $\overline{a}$ in $v_1 v_2$ to the right-hand side of the write action on position p in w. Then we have $v_1' \overline{a} v_2 \equiv_{\mathcal{L}} v_1 v_2 \equiv_{\mathcal{L}} w$ (since $\mathrm{rd}_2(v_1' \overline{a} v_2) = \mathrm{rd}_2(v_1 v_2)$). Additionally, we know $\mathrm{wrt}(v_1') = \mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p])$, and $|\mathrm{rd}(\overline{a} v_2)| = \ell$. This finally implies $p \in \Lambda_{A_{\mathcal{L}}}^w \setminus P_\ell^w$.

Now, assume $w_1 \in \mathrm{redsup}_{\mathcal{L}}(a r_2)$ for a prefix w_1 of $\mathrm{wrt}(w)$. We know

$$|\mathrm{rd}_2(v_1 v_2)| = |\mathrm{rd}_2(w)| < \mathrm{obw}(w) < \ell\,.$$

By definition of $\mathrm{obw}(w)$, there is $u \in A_{\mathcal{L}}^{\mathrm{obw}(w)}$ with $\mathrm{rd}(v_1 v_2) = \mathrm{rd}(w) \in A_{\mathcal{L}}^* u$, $w_1' \in \mathrm{redsup}_{\mathcal{L}}(u)$ for a prefix w_1' of $\mathrm{wrt}(w)$, and $|\mathrm{rd}_2(v_1 v_2)| < |u| = \mathrm{obw}(w)$. Let $r_2 = b_{\ell-1} \dots b_1$ where $b_1, \dots, b_{\ell-1} \in A_{\mathcal{L}}$ are letters and $p_1, \dots, p_{\mathrm{obw}(w)} \in \mathrm{dom}(w) = \mathrm{dom}(v_1 v_2)$ be the positions of $\overline{b_1}, \dots, \overline{b_{\mathrm{obw}(w)}}$

in v_1v_2. Additionally, let $q_1, \ldots, q_{\text{obw}(w)} \in \text{dom}(w)$ be the positions of the corresponding write actions b_i in v_1v_2, i.e., $\text{wrt}(v_1v_2[q_{i+1}+1, q_i-1]) \in (F \setminus \{b_i\})^*$ holds for each $1 \le i \le \text{obw}(w)$ (where $q_{\text{obw}(w)+1} := 0$). By Lemma 5.14 there is $1 \le i \le \text{obw}(w)$ such that $q_i > p_i$.

Consider the positions $p'_1, \ldots, p'_\ell \in \text{dom}(w)$ of the ℓ right-most read actions (in descending order) and their corresponding write positions $q'_1, \ldots, q'_\ell \in \text{dom}(w)$ in v_1v_2. Then we have $q'_i \ge q_i > p_i = p'_i$ implying, by application of Lemma 5.14, $|\text{rd}_2(v'_1\overline{a}v_2)| < \ell = |ar_2|$. Hence, we can infer $\text{rd}_2(v_1v_2) = \text{rd}_2(v'_1\overline{a}v_2)$. But this finally implies $w \equiv_{\mathcal{L}} v_1v_2 \equiv_{\mathcal{L}} v'_1\overline{a}v_2$ and, therefore, $p \in \Lambda^w_{A_{\mathcal{L}}} \setminus P^w_\ell$.

(3) Since w contains $< \ell$ read actions, no write action can ever be followed by ℓ read actions. Consequently, we have $P^w_\ell = \Lambda^w_{A_{\mathcal{L}}}$. □

Now, we are able to define our logics: by $\mathsf{FO_q}$ we denote the set of all first-order formulas build up from the atomic formulas of the form

$$x = y\,, \quad x <_+ y\,, \quad x <_- y\,, \quad P_\ell(x) \text{ for } \ell > 0\,, \text{ and } \quad \Lambda_\alpha(x) \text{ for } \alpha \in \Sigma_{\mathcal{L}}$$

where x and y are variables. Additionally, by $\mathsf{MSO_q}$ we denote the monadic second-order extension of $\mathsf{FO_q}$.

Let $\phi \in \mathsf{MSO_q}$ be a sentence. The language of transformations defined by ϕ is $T(\phi) := \{t \in \mathcal{Q}(\mathcal{L}) \mid \widetilde{t} \models \phi\}$. We say that $T \subseteq \mathcal{Q}(\mathcal{L})$ is $\mathsf{MSO_q}$*-definable* ($\mathsf{FO_q}$*-definable*) if there is a sentence $\phi \in \mathsf{MSO_q}$ ($\phi \in \mathsf{FO_q}$, respectively) with $T = T(\phi)$.

Remark 5.28. The sets P^w_ℓ are also conform with the special product in the definition of q-rational plq languages. In particular, we have

$$T(\exists x\colon \neg P_\ell(x) \wedge \Lambda_{A_{\mathcal{L}}}(x)) = \text{wrt}^{-1}\big(A^+_{\mathcal{L}}\big) \cdot \mathcal{Q}(\mathcal{L}) \cdot \text{rd}^{-1}\big(A^\ell_{\mathcal{L}}\big)\ .$$

In the proof of implication "(B)⇒(C)" in Theorem 5.1 we need the following notion of restricted quantification: Let $\phi(\vec{x}), \xi(x, \vec{y}) \in \mathsf{MSO_q}$. Then the *restriction* $\phi|_\xi$ *of* ϕ *to* ξ is defined by

$$\phi|_\xi(\vec{x},\vec{y}) := \begin{cases} \phi(\vec{x}) & \text{if } \phi(\vec{x}) \text{ is atomic} \\ \psi|_\xi(\vec{x},\vec{y}) \vee \chi|_\xi(\vec{x},\vec{y}) & \text{if } \phi(\vec{x}) = \psi(\vec{x}) \vee \chi(\vec{x}) \\ \neg\psi|_\xi(\vec{x},\vec{y}) & \text{if } \phi(\vec{x}) = \neg\psi(\vec{x}) \\ \exists x\colon (\psi|_\xi(x,\vec{x},\vec{y}) \wedge \xi(x,\vec{y})) & \text{if } \phi(\vec{x}) = \exists x\colon \psi(x,\vec{x}) \\ \exists X\colon (\psi|_\xi(X,\vec{x},\vec{y}) \wedge \forall x\colon X(x) \to \xi(x,\vec{y})) & \text{if } \phi(\vec{x}) = \exists X\colon \psi(X,\vec{x})\,. \end{cases}$$

In other words, we restrict the quantifiers in ϕ to the values satisfying ξ. Obviously, we have $\phi|_\xi \in \mathsf{FO_q}$ if $\phi, \xi \in \mathsf{FO_q}$ holds.

Finally, we can state the following implication:

Proposition 5.29. *Let* $\mathcal{L} = (F, U)$ *be a lossiness alphabet with* $|A_{\mathcal{L}}| \ge 2$ *and* $T \subseteq \mathcal{Q}(\mathcal{L})$ *be q-rational. Then* T *is* $\mathsf{MSO_q}$*-definable.*

Proof. We prove this by induction on the syntax tree of a q-rational expression generating the plq language T.

If T is q$^+$-rational, then we have $T = \mathrm{wrt}^{-1}(R)$ for a regular language $R \subseteq A_{\mathcal{L}}^*$ by Observation 5.18. By [4] there is an MSO-sentence ϕ with $L(\phi) = R$. Then by replacing of all occurrences of $<$ in ϕ by $<_+$ we obtain an $\mathsf{MSO_q}$-sentence ϕ' with $T(\phi'|_{\Lambda_{A_{\mathcal{L}}}(x)}) = \mathrm{wrt}^{-1}(L(\phi)) = T$.

Similarly, we can prove that T is $\mathsf{MSO_q}$-definable if T is q$^-$-rational (here, we replace $<$ by $<_-$ and restrict the quantifiers to $\Lambda_{\overline{A_{\mathcal{L}}}}$).

If $T = S_1 \cup S_2$ or $T = \mathcal{Q}(\mathcal{L}) \setminus S_1$, where S_1, S_2 are q-rational, there are $\phi_1, \phi_2 \in \mathsf{MSO_q}$ with $T(\phi_1) = S_1$ and $T(\phi_2) = S_2$ by induction hypothesis. Then we have $T = T(\phi_1 \vee \phi_2)$ and $T = T(\neg\phi_1)$, respectively.

Finally, let $T = \mathrm{wrt}^{-1}(R_1) \cdot \mathcal{Q}(\mathcal{L}) \cdot \mathrm{rd}^{-1}(R_2)$ where $R_1 \subseteq A_{\mathcal{L}}^*$ is regular and $R_2 \subseteq A_{\mathcal{L}}^*$ is finite. W.l.o.g. we can assume that $R_2 = \{w\}$ holds. Then there are $\mathsf{MSO_q}$-sentences ϕ_1 and ϕ_2 defining $\mathrm{wrt}^{-1}(R_1)$ and $\mathrm{rd}^{-1}(R_2)$, respectively. Set

$$\phi := \exists x_1, x_2 \colon \phi_1|_{x \leq_+ x_1} \wedge \phi_2|_{x_2 \leq_- x} \wedge \neg P_{|w|}(x_1)\,.$$

Then we have $T = T(\phi)$. □

5.4. *From logical definability to recognizability*

Finally, we have to prove that each $\mathsf{MSO_q}$-definable plq language is recognizable. Concretely, we do this by translation of a formula $\phi \in \mathsf{MSO_q}$ to a formula $\psi \in \mathsf{MSO}$ such that $\eta_{\mathcal{L}}^{-1}(T(\phi)) = L(\psi)$ holds. In this case, the right-hand side of this equation is regular by [4] implying that $T(\phi)$ is recognizable in $\mathcal{Q}(\mathcal{L})$ since $\eta_{\mathcal{L}}$ is surjective.

The most complicated case in our construction is the translation of the atomic formula $P_\ell(x)$ since write and read actions are commutative in certain contexts (cf. Theorem 3.3). For this translation we will utilize the connection between P_ℓ^w and $\mathrm{obw}(w)$ as described in Lemma 5.27.

So, let $w \in \Sigma_{\mathcal{L}}^*$ be a word, $p \in \mathrm{dom}(w)$ be a position in w (which will be represented by x in our formula), and $\ell > 0$. Then we express $p \in P_\ell^w$ as follows:

By definition we have $P_\ell^w \subseteq \Lambda_{A_{\mathcal{L}}}^w$. Consequently, we have $(\underline{w}, p) \models \Lambda_{A_{\mathcal{L}}}(x)$. Additionally, from Lemma 5.27 we can infer two cases: if $|\mathrm{rd}(w)| < \ell$ then we are done. This property can be expressed with the help of an appropriate FO-formula short_ℓ satisfying $L(\mathsf{short}_\ell) = \eta_{\mathcal{L}}^{-1}(\mathrm{rd}^{-1}(A_{\mathcal{L}}^{<\ell}))$. So, we can assume $|\mathrm{rd}(w)| \geq \ell$ from now on. Then, by Lemma 5.27 we have $\mathrm{obw}(w) \leq \ell$, i.e., $[w]_{\equiv_{\mathcal{L}}} \in \mathcal{Q}(\mathcal{L}) \setminus \Omega_\ell$. By Lemma 5.15 the plq language Ω_ℓ is aperiodic implying the existence of an FO-formula Omega_ℓ satisfying $\eta_{\mathcal{L}}^{-1}(\Omega_\ell) = L(\mathsf{Omega}_\ell)$.

Next, we want to determine the values $k := \mathrm{obw}(w)$ and $m := |\mathrm{rd}_2(w)|$ as well as the positions of the last k resp. m read actions and their corresponding write actions in w. To this end, we utilize Lemma 5.14 with some small modifications to the formulas embed_i and $\mathsf{overlap}_i$:

$$\mathsf{embed}'_i(\vec{x}, \vec{y}) := \phi_1 \wedge \phi_2 \wedge \phi_3 \wedge \phi_4 \wedge \phi_5$$
$$\mathsf{overlap}'_i(\vec{x}, \vec{y}) := \phi_1 \wedge \phi_2 \wedge \phi_3 \wedge \phi_4 \wedge \phi_5 \wedge \phi_6\,.$$

Then $(\underline{w}, \vec{p}, \vec{q}) \models \mathsf{embed}'_i(\vec{x}, \vec{y})$ if, and only if, there is $r_2 \in A^i_{\mathcal{L}}$ with $\mathrm{rd}(w) \in A^* r_2$, $w_1 \in \mathrm{redsup}_{\mathcal{L}}(r_2)$ for a prefix w_1 of $\mathrm{wrt}(w)$, where $p_1, \ldots, p_i$ are the positions of the last i read actions in w (in descending order) and $q_1, \ldots, q_i$ are the positions of their corresponding write actions in w. Moreover, we have $(\underline{w}, \vec{p}, \vec{q}) \models \mathsf{overlap}'_i(\vec{x}, \vec{y})$ if, and only if, in addition to the conditions above, we have $\mathrm{rd}_2(w) \in A^*_{\mathcal{L}} r_2$. Hence, we can express "$k = \mathrm{obw}(w)$" as follows:

$$\mathsf{obw}_k(\vec{x}, \vec{y}) := \mathsf{embed}'_k(\vec{x}, \vec{y}) \wedge \neg\mathsf{overlap}'_k(\vec{x}, \vec{y}) \wedge \bigwedge_{i=1}^{k-1} \mathsf{embed}_i \to \mathsf{overlap}_i \,.$$

Additionally, the property "$m = |\mathrm{rd}_2(w)| < k$" can be expressed by the following FO-formula:

$$\mathsf{shuffle}_{m,k}(\vec{x}, \vec{z}) := \mathsf{overlap}'_m(\vec{x}, \vec{z}) \wedge \bigwedge_{i=m+1}^{k} \neg\mathsf{overlap}_i \,.$$

We should note here, that the property "$m = |\mathrm{rd}_2(w)|$" cannot be expressed by an MSO-formula since this property is not recognizable according to [12, Observation 9.1] (at least in the case $F = \emptyset$). In contrast, satisfaction of $\mathsf{shuffle}_{m,k}(\vec{x}, \vec{z})$ by w requires further assumptions on the overlap's bounded width: we require that $k = \mathrm{obw}(w) < \infty$ is a fixed upper bound of $|\mathrm{rd}_2(w)|$.

Now, let $\vec{p}, \vec{q}, \vec{s} \in \mathrm{dom}(w)^\ell$ be positions in w with $(\underline{w}, \vec{p}, \vec{q}) \models \mathsf{shuffle}_{m,k}(\vec{x}, \vec{y})$ and $(\underline{w}, \vec{p}, \vec{s}) \models \mathsf{obw}_k(\vec{x}, \vec{z})$. It is easy to see, that $q_i \le s_i$ holds for each $1 \le i \le m$. In particular, if $q_i = s_i$ holds, then we have $q_j = s_j$ for each $1 \le j < i$. So, let $1 \le i \le m$ be minimal such that $q_{i+1} < s_{i+1}$ holds (where $q_{m+1} := 0$). This value can be determined with the help of the following formula:

$$\mathsf{diff}_i(\vec{y}, \vec{z}) := \bigwedge_{j=1}^{i} y_j = z_j \wedge y_{i+1} < z_{i+1} \,,$$

where "$y_{m+1} < z_{m+1}$" means "true". Then, by utilization of the equations from Theorem 3.3 we can move the read actions on positions $p_{i+1}, \ldots, p_k$ in w to the direct left-hand side of the write action on position s_{i+1}. But it is impossible to transpose the read action on position p_{i+1} with its corresponding write action on position s_{i+1}. In other words, we have $P^w_\ell = P^w_k = \{p' \in \Lambda^w_{A_{\mathcal{L}}} \mid s_{i+1} \le p'\}$.

All in all, we set:

$$\begin{aligned} \mathsf{P}_\ell(x) := (\Lambda_{A_{\mathcal{L}}}(x) \wedge \mathsf{short}_\ell) \vee \Big(&\Lambda_{A_{\mathcal{L}}}(x) \wedge \neg\mathsf{Omega}_\ell \wedge \\ \forall \vec{x}, \vec{y}, \vec{z}\colon \bigwedge_{0 \le i \le m < k \le \ell} &\big(\mathsf{shuffle}_{m,k}(\vec{x}, \vec{y}) \wedge \mathsf{obw}_k(\vec{x}, \vec{z}) \wedge \mathsf{diff}_i(\vec{y}, \vec{z})\big) \to z_{i+1} \le x\Big) \end{aligned}$$

The next two lemmas prove the correctness and completeness of this formula $\mathsf{P}_\ell(x)$:

Lemma 5.30. *Let $\ell > 0$, $w \in \Sigma^*_{\mathcal{L}}$, and $p \in \mathrm{dom}(w)$ with $(\underline{w}, p) \not\models \mathsf{P}_\ell(x)$. Then we have $p \notin P^w_\ell$.*

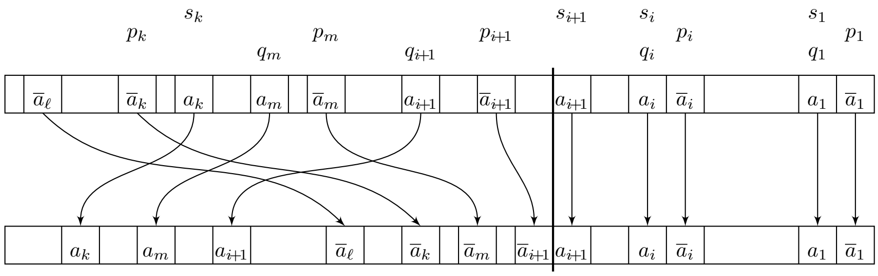

Fig. 1. Visualization of a possible movement of read and write actions. In both words, we have $q_j < p_j$ for all $1 \leq j \leq m$ and $s_{i+1} \geq p_{i+1}$. Hence, both words have an overlap of length m. The write actions on the left-hand side of the bold line are contained in P_ℓ^w, the ones on the right-hand side are not contained in P_ℓ^w.

Proof. If we have $p \notin \Lambda^w_{A_\mathcal{L}}$ we are done since $P^w_\ell \subseteq \Lambda^w_{A_\mathcal{L}}$. So, we can assume $p \in \Lambda^w_{A_\mathcal{L}}$ from now on. Then, we have $\underline{w} \not\models \mathsf{short}_\ell$, i.e., $|\mathrm{rd}(w)| \geq \ell$.

If $\underline{w} \models \mathsf{Omega}_\ell$, we have $\mathrm{obw}(w) > \ell$ and, hence, $P^w_\ell = \emptyset$ by Lemma 5.27. In this case, we are done. Now, we assume $\underline{w} \models \neg\mathsf{Omega}_\ell$, i.e., $\mathrm{obw}(w) \leq \ell$. Then there are $\vec{p}, \vec{q}, \vec{s} \in \mathrm{dom}(w)^\ell$ and $0 \leq i \leq m < k \leq \ell$ such that

- $(\underline{w}, \vec{p}, \vec{q}, \vec{s}) \models \mathsf{shuffle}_{m,k}(\vec{x}, \vec{y}) \wedge \mathsf{obw}_k(\vec{x}, \vec{z}) \wedge \mathsf{diff}_i(\vec{y}, \vec{z})$ and
- $(\underline{w}, p, \vec{s}) \not\models z_{i+1} \leq x$.

As we have argued above we have $m = |\mathrm{rd}_2(w)|$ and $k = \mathrm{obw}(w)$. Additionally, we have:

- $p_1, \ldots, p_k$ are the positions of the last k read actions in w (in descending order),
- $q_1, \ldots, q_m$ are the write positions corresponding to $p_1, \ldots, p_m$, and
- similarly, $s_1, \ldots, s_k$ are the write positions corresponding to $p_1, \ldots, p_k$.

By definition of diff_i the number i is the minimal index such that $q_{i+1} < s_{i+1}$ holds (where $q_{m+1} := 0$). We have to prove that $p < s_{i+1}$ implies $p \notin P^w_\ell$. To this end, let $r_1, r_2, w_1, w_2 \in A^*_\mathcal{L}$ such that $\mathrm{rd}(w) = r_1 r_2$, $|r_2| = i$, $\mathrm{wrt}(w) = w_1 w_2$, and $w_1 = \mathrm{wrt}(w[1, s_{i+1} - 1])$. We first show $w \equiv_\mathcal{L} w_1 \overline{r_1} w_2 \overline{r_2}$. By Proposition 3.7 it suffices to prove $|\mathrm{rd}_2(w)| = |\mathrm{rd}_2(w_1 \overline{r_1} w_2 \overline{r_2})|$. We have:

- the last i read actions are on the right-hand side of each write action and, hence, on the right-hand side of their corresponding write actions and
- the read actions on positions $p_{i+1}, \ldots, p_k$ are on the right-hand side of their corresponding write actions since $q_{i+1} < s_{i+1}$, i.e., the write actions on positions $q_{i+1}, \ldots, q_k$ in w are contained in w_1 which is on the left-hand side of all read actions.

Hence, we have $\underline{w_1 \overline{r_1} w_2 \overline{r_2}} \models \mathsf{overlap}_m$ implying $|\mathrm{rd}_2(w)| = m \leq |\mathrm{rd}_2(w_1 \overline{r_1} w_2 \overline{r_2})|$.

However, we have $\underline{w_1\overline{r_1}w_2\overline{r_2}} \not\models \mathsf{overlap}_k$ since the read action on position p_{i+1} in w (which is the last letter in $\overline{r_1}$) is on the left-hand side of its corresponding write action on position s_{i+1} (which is the first letter in w_2). Hence, we have $|\mathrm{rd}_2(w_1\overline{r_1}w_2\overline{r_2})| < k = \mathrm{obw}(w)$. By minimality of $\mathrm{obw}(w)$ we can infer that $|\mathrm{rd}_2(w)| \geq |\mathrm{rd}_2(w_1\overline{r_1}w_2\overline{r_2})|$ holds. Therefore, we have $w \equiv_{\mathcal{L}} w_1\overline{r_1}w_2\overline{r_2}$.

By $p < s_{i+1}$ we can split $w_1\overline{r_1}w_2\overline{r_2}$ as follows: let $v_1, v_2 \in A_{\mathcal{L}}^*$ with $v_1 = \mathrm{wrt}(w[1,p])$, i.e., v_1 is a prefix of w_1, and v_2 is the complementary suffix of $w_1\overline{r_1}w_2\overline{r_2}$ wrt. v_1. Then, we have $w \equiv_{\mathcal{L}} w_1\overline{r_1}w_2\overline{r_2} = v_1v_2$ and $|\mathrm{rd}(v_2)| = |\mathrm{rd}(w)| \geq \ell$ implying $p \notin P_\ell^w$. □

Lemma 5.31. *Let $\ell > 0$, $w \in \Sigma_{\mathcal{L}}^*$, and $p \in \mathrm{dom}(w) \setminus P_\ell^w$. Then we have $(\underline{w}, p) \not\models \mathsf{P}_\ell(x)$.*

Proof. If we have $p \notin \Lambda_{A_{\mathcal{L}}}^w$ we have obviously $(\underline{w}, p) \not\models \mathsf{P}_\ell(x)$. So, from now on, we assume $p \in \Lambda_{A_{\mathcal{L}}}^w$.

By $p \in \mathrm{dom}(w) \setminus P_\ell^w$ there are $v_1, v_2 \in \Sigma_{\mathcal{L}}^*$ with $w \equiv_{\mathcal{L}} v_1v_2$, $\mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p])$, and $|\mathrm{rd}(v_2)| \geq \ell$. Hence, we have $|\mathrm{rd}(w)| \geq |\mathrm{rd}(v_2)| \geq \ell$ implying $\underline{w} \not\models \mathsf{short}_\ell$.

Next, we consider the value of $\mathrm{obw}(w)$. If $\mathrm{obw}(w) > \ell$, we have $[w]_{\equiv_{\mathcal{L}}} \in \Omega_\ell$ implying $\underline{w} \models \mathsf{Omega}_\ell$. Hence, we are done in this case. Now, assume $\mathrm{obw}(w) \leq \ell$. Then there are $0 \leq m < k \leq \ell$ with $\mathrm{obw}(w) = k$ and $|\mathrm{rd}_2(w)| = m$. Let $\vec{p}, \vec{q}, \vec{s} \in \mathrm{dom}(w)^\ell$ be the following positions in w:

- $p_1, \ldots, p_k$ are the positions of the last k read actions in w (in descending order),
- $q_1, \ldots, q_m$ are the write positions corresponding to $p_1, \ldots, p_m$, and
- similarly, $s_1, \ldots, s_k$ are the write positions corresponding to $p_1, \ldots, p_k$.

Since we have $q_i \leq s_i$ for each $1 \leq i \leq m$, there is a minimal $0 \leq i \leq m$ such that $q_{i+1} < s_{i+1}$ holds (where $q_{m+1} := 0$). Then we have

$$(\underline{w}, \vec{p}, \vec{q}, \vec{s}) \models \mathsf{shuffle}_{m,k}(\vec{x}, \vec{y}) \wedge \mathsf{obw}_k(\vec{x}, \vec{z}) \wedge \mathsf{diff}_i(\vec{y}, \vec{z})\,.$$

Now we prove, that $(\underline{w}, p, \vec{s}) \not\models z_{i+1} \leq x$, i.e., we prove $s_{i+1} > p$. Towards a contradiction, suppose that $s_{i+1} \leq p$ holds.

Recall that $v_1, v_2 \in \Sigma_{\mathcal{L}}^*$ are defined such that $w \equiv_{\mathcal{L}} v_1v_2$, $\mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p])$, and $|\mathrm{rd}(v_2)| \geq \ell$. Consider the positions in v_1v_2 of the letters on positions p, $\vec{p}$, $\vec{q}$, and $\vec{s}$. To this end, let σ be the (uniquely defined) isomorphism mapping the positions in $\underline{w}$ to the ones in $\underline{v_1v_2}$ (note that σ is a permutation of $\mathrm{dom}(w)$). Then we have $\sigma(q_j) < \sigma(p_j)$ for each $1 \leq j \leq m$ since $\mathrm{rd}_2(w) = \mathrm{rd}_2(v_1v_2)$ and, therefore, $\underline{v_1v_2} \models \mathsf{overlap}_m$ holds.

By $\mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p])$ we have $\sigma(p) \leq |v_1|$ and, by $|\mathrm{rd}(v_2)| \geq \ell \geq k$ we have $|v_1| < \sigma(p_j)$ for each $1 \leq j \leq k$. From $w \equiv_{\mathcal{L}} v_1v_2$ we know that $\mathrm{wrt}(w) = \mathrm{wrt}(v_1v_2)$

and $\mathrm{rd}(w) = \mathrm{rd}(v_1v_2)$ hold implying

$$\begin{aligned}
&\sigma(s_k) < \cdots < \sigma(s_{i+1}) && (\text{since } s_k <_+ \cdots <_+ s_{i+1}) \\
&\leq \sigma(p) && (\text{since } s_{i+1} \leq p,\ s_{i+1}, p \in \Lambda^w_{A_\mathcal{L}},\text{ and, hence, } s_{i+1} \leq_+ p) \\
&\leq |v_1| && (\text{since } \mathrm{wrt}(v_1) = \mathrm{wrt}(w[1,p])) \\
&< \sigma(p_k) < \cdots < \sigma(p_1)\,. && (\text{since } |\mathrm{rd}(v_2)| \geq k \text{ and } p_k <_- \cdots <_- p_1)
\end{aligned}$$

Recall that we also have $\sigma(s_j) = \sigma(q_j) < \sigma(p_j)$ for each $1 \leq j \leq i$ by the minimality of i and due to $\underline{v_1v_2} \models \mathsf{overlap}_m$. Then we also obtain $\sigma(s_j) < \sigma(p_j)$ for each $1 \leq j \leq k$, i.e., $\underline{v_1v_2} \models \mathsf{overlap}_k$. From Lemma 5.14 we infer that $|\mathrm{rd}_2(v_1v_2)| \geq k > m = |\mathrm{rd}_2(w)|$ holds. This is a contradiction to $\mathrm{rd}_2(v_1v_2) = \mathrm{rd}_2(w)$ by $w \equiv_\mathcal{L} v_1v_2$ according to Proposition 3.7. Hence, we have $p < s_{i+1}$ implying $(\underline{w}, p) \not\models \mathsf{P}_\ell(x)$. □

Finally, with the help of the formula P_ℓ we can prove the last implication of Theorem 5.1:

Proposition 5.32. *Let $\mathcal{L} = (F,U)$ be a lossiness alphabet with $|A_\mathcal{L}| \geq 2$ and $T \subseteq \mathcal{Q}(\mathcal{L})$ be $\mathsf{MSO_q}$-definable. Then T is recognizable.*

Proof. Let $T \subseteq \mathcal{Q}(\mathcal{L})$ be $\mathsf{MSO_q}$-definable. Then there is $\phi \in \mathsf{MSO_q}$ with $T = T(\phi)$. We construct $\phi' \in \mathsf{MSO}$ such that for each $w \in \Sigma^*_\mathcal{L}$ we have $\widetilde{w} \models \phi$ if, and only if, $\underline{w} \models \phi'$.

The formula ϕ' is inductively defined as follows:

$$\phi' = \begin{cases}
x = y & \text{if } \phi = (x = y) \\
x < y \wedge \Lambda_A(x) \wedge \Lambda_A(y) & \text{if } \phi = x <_+ y \\
x < y \wedge \Lambda_{\overline{A}}(x) \wedge \Lambda_{\overline{A}}(y) & \text{if } \phi = x <_- y \\
\mathsf{P}_\ell(x) & \text{if } \phi = P_\ell(x) \\
\Lambda_a(x) & \text{if } \phi = \Lambda_a(x) \\
X(x) & \text{if } \phi = X(x) \\
\psi' \vee \xi' & \text{if } \phi = \psi \vee \xi \\
\neg\psi' & \text{if } \phi = \neg\psi \\
\exists x\colon \psi' & \text{if } \phi = \exists x\colon \psi \\
\exists X\colon \psi' & \text{if } \phi = \exists X\colon \psi
\end{cases}$$

Then by Lemmas 5.30 and 5.31 we have $\widetilde{w} \models \phi$ if, and only if, $\underline{w} \models \phi'$ for any $w \in \Sigma^*_\mathcal{L}$. Hence, by Büchi's Theorem [4] $\eta_\mathcal{L}^{-1}(T)$ is regular. Since $\eta_\mathcal{L}$ is surjective this implies recognizability of $T = T(\phi)$. □

5.5. *The complexity of the constructions in Theorems 5.1 and 5.6*

Finally, we want to analyze the complexities of the constructions in this section. Recall that we have seen the following circular chain of implications: "(A)/(1) ⇒ (2) ⇒ (3) ⇒ (B) ⇒ (C) ⇒ (A)".

Towards the first implication, let $T \subseteq \mathcal{Q}(\mathcal{L})$ be recognized by a finite monoid $\mathbb{F}$ via the homomorphism $\phi\colon \mathcal{Q}(\mathcal{L}) \to \mathbb{F}$. Then $\eta_{\mathcal{L}}^{-1}(T) \subseteq \Sigma^*$ is regular and recognized via $\eta_{\mathcal{L}} \circ \phi$ (recall that we can understand $\mathbb{F}$ as a DFA accepting $\eta_{\mathcal{L}}^{-1}(T)$). We can obtain another finite monoid $\mathbb{G}$ recognizing $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ in polynomial time by application of the cross-product construction known from automata theory. Note that $|\mathbb{G}| \in \mathcal{O}(|\mathbb{F}|)$ holds in this case.

Now, consider the second implication. So, let $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ be recognized by a finite monoid $\mathbb{F}$. By the proof of Lemma 5.10 the set T is a union of $\mathcal{O}(2^{|\mathbb{F}|})$ many languages which can be expressed by a constant number of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for regular word languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$. In other words, the constructed Boolean combination has exponential size.

Next, let $T \subseteq \mathcal{Q}(\mathcal{L})$ be given by a Boolean combination as in property (3) of Theorem 5.6. If a set $\mathrm{wrt}^{-1}(R)$ or $\mathrm{rd}^{-1}(R)$ is given by a finite automaton accepting $R \subseteq A_{\mathcal{L}}^*$ then a q-rational expression describing such language has exponential size [8]. Otherwise if such plq language is given by a regular expression, an equivalent q-rational expression has linear size. Additionally, the plq language Ω_ℓ is equivalent to a q-rational expression of size exponential in the number ℓ (cf. Lemma 5.15).

Towards the implication "(B)⇒(C)", let $T \subseteq \mathcal{Q}(\mathcal{L})$ be given by a q-rational expression. The translation of q$^+$- and q$^-$-rational expressions into an $\mathsf{MSO_q}$-formula is similar to the translation of a regular expression into an MSO-formula. This is possible in polynomial time when translating the given regular expression into an NFA (this uses the efficient closure properties of regular languages) and then into an MSO-formula as described (e.g.) in [19, Theorem 7.21]. Finally, formulas for unions, complements, and products can be constructed in polynomial time. All in all, we obtain an $\mathsf{MSO_q}$-formula from a q-rational expression in polynomial time.

Finally, let $T \subseteq \mathcal{Q}(\mathcal{L})$ be given by an $\mathsf{MSO_q}$-formula. In Proposition 5.32 we translated this formula into an MSO-formula. The most complicated case is the atomic formula $P_\ell(x)$. Its translation $\mathsf{P}_\ell(x)$ can be obtained from $P_\ell(x)$ in time exponential in the number ℓ. The remaining translations are possible in linear time. Finally, the translation of an MSO-formula into a DFA has non-elementary complexity [32].

6. Characterizations of the Aperiodic PLQ Languages

In the previous section we have seen Kleene- and Büchi-type characterizations of the recognizable languages in the plq monoid. Another more involved task is to describe the aperiodic plq languages. Schützenberger has proven in [31] that the aperiodic word languages are exactly the star-free languages. This result gives us a decision procedure to decide whether a given regular language is star-free. There are also some further characterizations of the aperiodic languages: (1) these are the languages having an aperiodic syntactic monoid[d], (2) these languages are accepted

[d]A finite monoid $\mathbb{F}$ is *aperiodic* iff there is $n \in \mathbb{N}$ with $x^n = x^{n+1}$ for all $x \in \mathbb{F}$.

by so-called counter-free finite automata (these are NFAs which are unable to count modulo any positive integer) [24], (3) these languages are described by the first-order fragment of Büchi's logic on words [24], and (4) these languages are definable in linear temporal logic [14]. Similarly, we know that in the trace monoid a trace language is aperiodic if, and only if, it is star-free [11] if, and only if, it is definable in a special first-order logic [7]. Unfortunately, we cannot translate Schützenberger's Theorem to plq monoids since the class of aperiodic plq languages is not closed under product. For example, the two plq languages $\{[a]_{\equiv_{\mathcal{L}}}\}^*$ and $\{[\overline{a}]_{\equiv_{\mathcal{L}}}\}^*$ are aperiodic, but their product is not even recognizable (cf. Lemma 5.17(2)). Though, we will see that the monoid's product can be restricted such that we can describe exactly the aperiodic languages in the plq monoid. Additionally, we will show that the first-order fragment $\mathsf{FO_q}$ of our special monadic second-order logic $\mathsf{MSO_q}$ describes exactly the aperiodic plq languages.

Before we start to prove the aforementioned equivalences we have to define our restriction of star-freeness: to this end, let $\mathcal{L} = (F, U)$ be a lossiness alphabet. We say that a $\mathcal{Q}(\mathcal{L})$-language is *q^+-star-free* if it can be constructed by the following rules:

(1^+) $\mathrm{wrt}^{-1}(\varepsilon)$ and $\mathrm{wrt}^{-1}(a)$ for any $a \in A_{\mathcal{L}}$ are q^+-star-free.
(2^+) if $S, T \subseteq \mathcal{Q}(\mathcal{L})$ are q^+-star-free then $S \cup T$, $S \cdot T$, and $\mathcal{Q}(\mathcal{L}) \setminus S$ are q^+-star-free.

Similarly, by replacing wrt^{-1} by rd^{-1} in the rules above, we define the class of *q^--star-free* plq languages. Finally, a $\mathcal{Q}(\mathcal{L})$-language is *q-star-free* if it can be obtained by the following rules:

(1) if $T \subseteq \mathcal{Q}(\mathcal{L})$ is q^+- or q^--star-free it also is q-star-free.
(2) if $S, T \subseteq \mathcal{Q}(\mathcal{L})$ are q-star-free then $S \cup T$ and $\mathcal{Q}(\mathcal{L}) \setminus S$ are q-star-free.
(3) if $S \subseteq \mathcal{Q}(\mathcal{L})$ is q^+-star-free and $T \subseteq \mathcal{Q}(\mathcal{L})$ is q^--star-free such that $\mathrm{rd}(T)$ is finite (i.e., T is obtained without usage of the $\setminus$-operator) then $S \cdot \mathcal{Q}(\mathcal{L}) \cdot T$ is q-star-free.

Hence, the only difference of the rules above in comparison to the rules constructing the q-rational plq languages is the missing iteration. Therefore, $T \subseteq \mathcal{Q}(\mathcal{L})$ is q^+-star-free (q^--star-free) if, and only if, there is an aperiodic word language $R \subseteq A_{\mathcal{L}}^*$ with $T = \mathrm{wrt}^{-1}(R)$ ($T = \mathrm{rd}^{-1}(R)$, respectively).

Similarly, to Theorems 5.1 and 5.6 we can state the following result:

Theorem 6.1. *Let $\mathcal{L} = (F, U)$ be a lossiness alphabet with $|A_{\mathcal{L}}| \geq 2$ and $T \subseteq \mathcal{Q}(\mathcal{L})$. Then the following are equivalent:*

(A) T is aperiodic.
(B) $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is aperiodic.
(C) T is a Boolean combination of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for aperiodic languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$.
(D) T is q-star-free.

(E) T is $\mathsf{FO_q}$-definable.

Proof. To prove this theorem we revisit the proofs of Theorems 5.1 and 5.6.

At first we show the implication "(A)⇒(B)". So, let $T \subseteq \mathcal{Q}(\mathcal{L})$ be aperiodic. Then $\eta_{\mathcal{L}}^{-1}(T)$ is aperiodic. Since $\overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ is aperiodic and the class of aperiodic languages is closed under intersection we can infer that $\eta_{\mathcal{L}}^{-1}(T) \cap \overline{A_{\mathcal{L}}}^* A_{\mathcal{L}}^* \overline{A_{\mathcal{L}}}^*$ also is aperiodic.

Next we prove "(B)⇒(C)". We recall the proof of Lemma 5.10 and assume that the recognizing monoid $\mathbb{F}$ is aperiodic. All of the arguments of the inverse projections wrt^{-1} or rd^{-1} are either single words, $\mu^{-1}(\alpha)$, $\overline{\mu}^{-1}(\alpha)$ for any $\alpha \in \mathbb{F}$, or products of those languages. Anyway these plq languages are aperiodic since aperiodicity is preserved under inverse homomorphisms (note that the class of aperiodic word languages is closed under product according to Schützenberger's Theorem [31]). Consequently, T is a Boolean combination of plq languages of the form $\mathrm{wrt}^{-1}(R)$, $\mathrm{rd}^{-1}(R)$, and Ω_ℓ for aperiodic languages $R \subseteq A_{\mathcal{L}}^*$ and numbers $\ell \in \mathbb{N}$.

In the proof of Proposition 5.23 we see that Ω_ℓ is even q-star-free and hence the implication "(C)⇒(D)" holds.

In the construction of the proof of Proposition 5.29 each of the used formulas can be expressed in first-order logic. Therefore, we have "(D)⇒(E)".

Finally we have to prove "(E)⇒(A)". Each translation of an $\mathsf{FO_q}$-formula as seen in Proposition 5.32 results in a formula in FO. Therefore, by [24] each $\mathsf{FO_q}$-definable plq language is aperiodic. □

7. Conclusion and Open Problems

In this paper we investigated some generalizations of famous results from automata theory concerning rational, recognizable, star-free, and aperiodic languages. Concretely, we considered these classes in the so-called partially lossy queue monoid (plq monoid for short). This monoid models the behavior of a partially lossy queue and was first introduced in [18]. Its subsets are called partially lossy queue languages (or plq languages, for short).

First, we proved that, in contrast to Kleene's theorem [15], the classes of rational and recognizable languages in the plq monoid do not coincide. Hence, we defined some restrictions to concatenations and iterations of languages in the plq monoid. Then by closure of the Boolean operations and the aforementioned restricted operations we obtain the so-called q-rational plq languages. We proved in this paper that a plq language is recognizable if, and only if, it is q-rational. Additionally, we defined a modification of Büchi's MSO logic (cf. [4]) called $\mathsf{MSO_q}$ which describes exactly the recognizable languages in the plq monoid.

Similarly, we introduced so-called q-star-free expressions generating the aperiodic plq languages (similar to Schützenberger's theorem [31]) and proved that the first-order fragment of our logic $\mathsf{MSO_q}$ describes the aperiodic languages (similar to McNaughton-Papert's theorem [24]).

There are some open questions relating to the languages in the plq monoid. So, we do not know a rational-like characterization of the recognizable plq languages using monotonic operations, only (note that the complement and intersection operations are not monotonic). Additionally, possibly one finds a characterization of the recognizable or aperiodic plq languages in terms of a temporal logic similar to Kamp's Theorem [14] stating that the logic LTL describes exactly the aperiodic word languages.

One may also consider the aperiodicity problem of rational or recognizable plq languages, i.e., the question whether a given rational (or recognizable, resp.) language in the plq monoid is aperiodic.

Moreover, one could study some sub-classes of the star-free and aperiodic plq languages. For example, maybe there is a simple generalization of the dot-depth hierarchy or the Straubing-Thérien hierarchy such that the n^{th} level of this hierarchy corresponds to the $\mathcal{B}\Sigma_n$-fragment of the $\mathsf{MSO_q}$ logic and the $(n-1/2)^{\text{th}}$ level corresponds to the Σ_n-fragment of $\mathsf{MSO_q}$ (similar to Thomas' result in [33]).

Acknowledgment

The author would like to thank Dietrich Kuske and the anonymous reviewers of the conference version [17] and the submitted version of this full paper for their helpful suggestions to improve this paper.

References

[1] P. A. Abdulla and B. Jonsson, Verifying programs with unreliable channels, *Information and Computation* **127**(2) (1996) 91–101.
[2] J. Berstel, *Transductions and Context-Free Languages* (Teubner Studienbücher, 1979).
[3] D. Brand and P. Zafiropulo, On Communicating Finite-State Machines, *Journal of the ACM* **30**(2) (1983).
[4] J. R. Büchi, Weak Second-Order Arithmetic and Finite Automata, *Mathematical Logic Quarterly* **6**(1-6) (1960) 66–92.
[5] G. Cécé, A. Finkel and S. Purushotaman Iyer, Unreliable channels are easier to verify than perfect channels, *Information and Computation* **124**(1) (1996) 20–31.
[6] M. Chrobak and W. Rytter, Unique decipherability for partially commutative alphabet, in *Mathematical Foundations of Computer Science*, **233** of *Lecture Notes in Computer Science* Springer, (1986) 256–263.
[7] W. Ebinger, Logical Definability of Trace Languages, in V. Diekert and G. Rozenberg (eds.), *The Book of Traces* (World scientific, 1995)
[8] A. Ehrenfeucht and P. Zeiger, Complexity measures for regular expressions, in *Proceedings of the Sixth Annual ACM Symposium on Theory of Computing* Association for Computing Machinery, (1974) 75–79.
[9] A. Finkel, Decidability of the termination problem for completely specified protocols, *Distributed Computing* **7**(3) (1994) 129–135.
[10] A. Gibbons and W. Rytter, On the decidability of some problems about rational subsets of free partially commutative monoids, *Theoretical Computer Science* **48** (1986) 329–337.

[11] G. Guaiana, A. Restivo and S. Salemi, On aperiodic trace languages, in *8th Annual Symposium on Theoretical Aspects of Computer Science*, **480** of *Lecture Notes in Computer Science* Springer, (1991) 76–88.

[12] M. Huschenbett, D. Kuske and G. Zetzsche, The monoid of queue actions, *Semigroup forum* **95**(3) (2017) 475–508.

[13] N. D. Jones, Space-bounded reducibility among combinatorial problems, *Journal of Computer and System Sciences* **11**(1) (1975) 68–85.

[14] J. Kamp, *Tense Logic and the Theory of Linear Order*, dissertation, University of California at Los Angeles, (Los Angeles, 1968).

[15] S. C. Kleene, Representation of events in nerve nets and finite automata, tech. rep., DTIC Document (1951).

[16] C. Köcher, *Einbettungen in das Transformationsmonoid einer vergesslichen Warteschlange*, Master's Thesis, Technische Universität Ilmenau, (Ilmenau, 2016).

[17] C. Köcher, Rational, Recognizable, and Aperiodic Sets in the Partially Lossy Queue Monoid, in *STACS 2018*, **96** of *LIPIcs* Dagstuhl Publishing, (2018) 45:1–45:14.

[18] C. Köcher, D. Kuske and O. Prianychnykova, The Inclusion Structure of Partially Lossy Queue Monoids and their Trace Submonoids, *RAIRO - Theoretical Informatics and Applications* **52**(1) (2018) 55–86.

[19] L. Libkin, *Elements of Finite Model Theory*, Texts in Theoretical Computer Science (Springer, 2004).

[20] M. Lohrey, The rational subset membership problem for groups: A survey, in *Groups St Andrews*, **422** of *London Mathematical Society Lecture Note Series* Cambridge University Press, (2013) 368–389.

[21] M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and Its Applications (Cambridge University Press, 2002).

[22] B. Masson and P. Schnoebelen, On verifying fair lossy channel systems, in *Mathematical Foundations of Computer Science*, **2420** of *Lecture Notes in Computer Science* Springer, (2002) 543–555.

[23] J. D. McKnight, Kleene quotient theorems, *Pacific Journal of Mathematics* **14**(4) (1964) 1343–1352.

[24] R. McNaughton and S. A. Papert, *Counter-Free Automata*, **65** (The MIT Press, 1971).

[25] A. Muscholl and H. Petersen, A Note on the Commutative Closure of Star-Free Languages, *Information Processing Letters* **57**(2) (1996) 71–74.

[26] M. Nivat and J.-F. Perrot, Une généralisation du monoıde bicyclique, *Comptes Rendus de l'Académie des Sciences de Paris* **271** (1970) 824–827.

[27] E. Ochmański, Regular behaviour of concurrent systems, *Bulletin of the EATCS* **27** (1985) 56–67.

[28] J.-É. Pin, Mathematical foundations of automata theory, *Lecture notes LIAFA, Université Paris* **7** (2010).

[29] E. L. Post, A variant of a recursively unsolvable problem, *Bulletin of the American Mathematical Society* **52**(4) (1946) 264–268.

[30] E. Render and M. Kambites, Rational subsets of polycyclic monoids and valence automata, *Information and Computation* **207**(11) (2009) 1329–1339.

[31] M. P. Schützenberger, On finite monoids having only trivial subgroups, *Information and control* **8**(2) (1965) 190–194.

[32] L. J. Stockmeyer and A. R. Meyer, Word problems requiring exponential time, in *Proceedings of the Fifth Annual ACM Symposium on Theory of Computing* Association for Computing Machinery, (1973) 1–9.

[33] W. Thomas, Classifying regular events in symbolic logic, *Journal of Computer and*

System Sciences **25**(3) (1982) 360–376.

[34] W. Thomas, Languages, Automata, and Logic, in G. Rozenberg and A. Salomaa (eds.), *Beyond Words*, *Handbook of Formal Languages* **3** (Springer, 1997) pp. 389–455.

[35] G. Zetzsche, *Monoids as Storage Mechanisms*, dissertation, Technische Universität Kaiserslautern, (Kaiserslautern, 2016).