

Köcher, Chris:

Rational, recognizable, and aperiodic sets in the partially lossy queue monoid

<i>Original published in:</i>	35th Symposium on Theoretical Aspects of Computer Science / STACS 35, 2018 Caen, France. - Saarbrücken/Wadern, Germany : Schloss Dagstuhl - Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, February, 2018. - (2018), art. 45, 14 pp. ISBN 978-3-95977-062-0 (Leibniz International Proceedings in Informatics (LIPIcs) ; 96)
<i>Original published:</i>	February 2018
<i>ISSN:</i>	1868-8969
<i>DOI:</i>	10.4230/LIPIcs.STACS.2018.45
<i>[Visited:</i>	2020-03-02]



This work is licensed under a [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/). To view a copy of this license, visit [http://creativecommons.org/licenses/by/3.0/](https://creativecommons.org/licenses/by/3.0/)

Rational, Recognizable, and Aperiodic Sets in the Partially Lossy Queue Monoid

Chris Köcher

Technische Universität Ilmenau, Automata and Logics Group
chris.koecher@tu-ilmenau.de

Abstract

Partially lossy queue monoids (or plq monoids) model the behavior of queues that can forget arbitrary parts of their content. While many decision problems on recognizable subsets in the plq monoid are decidable, most of them are undecidable if the sets are rational. In particular, in this monoid the classes of rational and recognizable subsets do not coincide. By restricting multiplication and iteration in the construction of rational sets and by allowing complementation we obtain precisely the class of recognizable sets. From these special rational expressions we can obtain an MSO logic describing the recognizable subsets. Moreover, we provide similar results for the class of aperiodic subsets in the plq monoid.

2012 ACM Subject Classification Theory of computation → Algebraic language theory, Theory of computation → Models of computation, Information systems → Data structures

Keywords and phrases Partially Lossy Queues, Transformation Monoid, Rational Sets, Recognizable Sets, Aperiodic Sets, MSO logic

Digital Object Identifier 10.4230/LIPIcs.STACS.2018.45

Funding Supported by the DFG-Project “Speichermechanismen als Monoide”, KU 1107/9-1.

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 state automaton we obtain a Turing complete computation model [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 [8, 5, 1, 17].

To obtain some algebraic results on the behavior of these storage mechanisms we can model them as monoid of transformations. So, a single blind counter induces $(\mathbb{Z}, +)$ and a pushdown induces a polycyclic monoid [12]. Some important results on the transformation monoid of reliable queues can be found in [11]. Furthermore, in [14] we considered the transformation monoid of lossy queues. When studying the similarities and differences between those two monoids in [15] we found it convenient to join both, the reliable and lossy queues, respectively, into one model, the so-called *partially lossy queues* (or *plqs*). Those are given by their underlying alphabet A as well as a subset $U \subseteq A$ of letters that are unforgettable while the letters contained in $A \setminus U$ can be forgotten at any time. We denote the corresponding transformation monoid by $\mathcal{Q}(A, U)$ and call it the *partially lossy queue monoid* or *plq monoid*. Hence, with the help of plqs we can argue about reliable and lossy



© Chris Köcher;

licensed under Creative Commons License CC-BY

35th Symposium on Theoretical Aspects of Computer Science (STACS 2018).

Editors: Rolf Niedermeier and Brigitte Vallée; Article No. 45; pp. 45:1–45:14

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



SYMPOSIUM
ON THEORETICAL
ASPECTS
OF COMPUTER
SCIENCE

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 subsets, which are a generalization of languages that are described by regular expressions, and recognizable subsets, which are a generalization of sets accepted by finite automata (see, e.g., [2, 22]). Kleene's Theorem [13] states that both notions are equivalent in the free monoid.

In Section 3 we consider some algorithmic properties of rational subsets of the plq monoid. Such properties encountered increased attention in recent years, e.g., [16] provides a survey on the membership problem for rational sets. Since the rational sets in the polycyclic monoid (recall that this is the transformation monoid of a pushdown) are exactly the homomorphic images of a special subclass of the regular languages by [24], many decision problems like membership, intersection, universality, inclusion, and recognizability are decidable in this monoid. In this paper we will see that the membership problem of the plq monoid is NL-complete, but the other problems are undecidable, which we can prove by reduction from their counterparts in the direct product of $(\mathbb{N}, +)$ and $\{a, b\}^*$ (cf. [20, 9]).

If the given subsets are recognizable, all of the considered decision problems in plq monoids are decidable by known constructions from automata theory. Hence, the rational subsets are not effectively recognizable. Especially, we will see that the class of rational subsets in the plq monoid is not closed under intersection implying that the classes of rational and recognizable subsets do not coincide. In contrast, in polycyclic monoids the class of rational sets is closed under Boolean operations. However, the classes of rational and recognizable subsets do not coincide in these monoids since there are only two recognizable sets (the empty set and the monoid itself). But since there are even more recognizable sets in the plq monoid and since each recognizable subset is rational as well due to McKnight's Theorem [18], it is a natural question to ask in which cases a rational subset is recognizable.

For trace monoids, Ochmański could prove in [21] that it suffices to restrict the usage of the Kleene star in an appropriate way to characterize the recognizable subsets in the trace monoid. In Section 4 of this paper we will use an approach similar to Ochmański's to characterize the recognizable sets in terms of special rational sets in the plq monoid. Concretely, we will define some special restrictions on the usage of Kleene star and the concatenation to 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, Chapter 10]. Hence, this motivates to find another MSO logic describing exactly the recognizable subsets 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 subsets, star-free subsets, and first-order logic. Recall that a set is aperiodic if it is accepted by a counter-free finite automaton and a set is star-free if it can be generated from finite sets by application of Boolean operations and concatenation, only. Schützenberger's Theorem [25] 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 [10] 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 subsets in $\mathcal{Q}(A, U)$ with the help

of the same restrictions to star-freeness of subsets as in our result regarding the rational subsets. Finally, we prove similar to the results from [19, 7] that the aperiodic subsets in the plq monoid can be described by first-order formulas.

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 $w \in A^*vA^*$. 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 the *projection* $\pi_S: A^* \rightarrow S^*$ to S is the homomorphism induced by $\pi_S(a) = a$ for each $a \in S$ and $\pi_S(a) = \varepsilon$ for each $a \in A \setminus S$. Moreover, v is an *S -prefix* of w (denoted $v \leq_S w$) if there is a prefix w' of w such that $\pi_S(w') \preceq v \preceq w'$. In other words, we have $v \leq_S w$ if v is a subword of a prefix of w and contains all the letters from S in this prefix, e.g., we have $aa \leq_{\{a\}} abaab$ and $aa \not\leq_{\{b\}} abaab$. Note that $v \leq_\emptyset w$ means that v is a subword of w and $v \leq_A w$ means that v is a prefix of w .

2.1 Partially Lossy Queues

The partially lossy queue monoid (or plq monoid) models the behavior of a fifo-queue whose entries come from a finite set A . The unreliability of the queue stems from the fact that it can forget certain letters that we collect in the set $A \setminus U$. In other words, letters from $U \subseteq A$ are *non-forgettable* and those from $A \setminus U$ are *forgettable*.

So, let A be an alphabet of possible queue entries and let $U \subseteq A$ be the set of non-forgettable letters. The states of the queue are the words from A^* . Furthermore, we have some basic controllable actions on these queues: writing of a symbol $a \in A$ (denoted by a) and reading of $a \in A$ (denoted by \bar{a}). Thereby, we assume that the set \bar{A} of all these reading operations \bar{a} is a disjoint copy of A . So, $\Sigma := A \cup \bar{A}$ is the set of all controllable operations on the partially lossy queue. For a word $u = a_1 \dots a_n \in A^*$ we write \bar{u} for the word $\bar{a}_1 \bar{a}_2 \dots \bar{a}_n$.

Formally, the action $a \in A$ appends the letter a to the state of the queue. The action $\bar{a} \in \bar{A}$ tries to cancel the letter a from the beginning of the current state of the queue. If this state does not start with a then the queue ends up in an error state. The lossiness of the queue is modeled by allowing it to forget arbitrary letters from $A \setminus U$ of its content at any moment.

Since a partially lossy queue with an underlying alphabet $A = \{a\}$ (independently of U) acts like a partially blind counter, the corresponding plq monoid is the bicyclic semigroup. On the first sight, the equality of these two transformation monoids seems to be counterintuitive. But it might be explained by the following observation: let \mathcal{A} be an NFA equipped with one reliable counter. Then \mathcal{A} accepts the same language as this NFA equipped with a lossy counter. Hence, from now on, we may exclude this case and assume $|A| \geq 2$.

Before defining the plq monoid we want to identify sequences of operations that have the same effect on any queue. In [15, Proposition 3.21] we proved that $u, v \in \Sigma^*$ act equally (denoted by $u \equiv v$) if, and only if, they can be transformed into each other by applying the equations from the following definition, only.

► **Definition 2.1.** Let $U \subseteq A$ be two finite sets. We define the binary relation $\equiv \subseteq (\Sigma^*)^2$ as the least congruence on Σ^* satisfying the following equations for $a, b \in A$, $c \in U$, and $w \in A^*$:

- (a) $b\bar{a} \equiv \bar{a}b$ if $a \neq b$
- (b) $a\bar{a}b \equiv \bar{a}a\bar{b}$
- (c) $cw\bar{a}\bar{a} \equiv cw\bar{a}a$
- (d) $awa\bar{a} \equiv aw\bar{a}a$

Then the *partially lossy queue monoid* or *plq monoid* induced by (A, U) is the quotient $\mathcal{Q}(A, U) := \Sigma^* / \equiv$. The natural epimorphism of \equiv is $\eta: \Sigma^* \rightarrow \mathcal{Q}(A, U): w \mapsto [w]$.

To handle the equivalence classes of \equiv we want to define a normal form on this congruence. We do this by ordering the equations from Definition 2.1 from left to right, which results in an infinite semi-Thue system called \mathcal{R} .

Since the rules of \mathcal{R} are length-preserving and move read actions to the left, it is terminating. Moreover, it is locally confluent by [15] and hence confluent. Therefore, for any word $u \in \Sigma^*$ there is a unique, irreducible word $\text{nf}(u)$ with $u \rightarrow^* \text{nf}(u)$, the so-called *normal form* of u .

► **Example 2.2.** Let $a, b \in A$ with $a \neq b$ and $q = aabb\bar{a}\bar{b}$. If $a \notin U$ then we have

$$aabb\bar{a}\bar{b} \rightarrow aab\bar{a}\bar{b}\bar{b} \rightarrow aa\bar{a}\bar{b}\bar{b}\bar{b} \rightarrow a\bar{a}abb\bar{b} \rightarrow a\bar{a}abb\bar{b}$$

and therefore $a\bar{a}abb\bar{b} = \text{nf}(aabb\bar{a}\bar{b})$. Otherwise, i.e., if $a \in U$, we can apply Rule c to $a\bar{a}abb\bar{b}$ and hence obtain $\text{nf}(aabb\bar{a}\bar{b}) = \bar{a}baabb$.

From the definition of \mathcal{R} we obtain that a word is in normal form if it starts with some read operations followed by a special shuffle of write and read operations where each read action \bar{a} appears directly right from a . Thereby, the infixes $a\bar{a}$ in these words are divided by words from $(A \setminus (U \cup \{a\}))^*$, only. Formally, such shuffle of $u \in A^*$ and $\bar{v} \in \bar{A}^*$ is defined by $\langle\langle u, \bar{v} \rangle\rangle = w_1 a_1 \bar{a}_1 w_2 a_2 \bar{a}_2 \dots w_\ell a_\ell \bar{a}_\ell w_{\ell+1}$, where $v = a_1 \dots a_\ell$, $a_1, \dots, a_\ell \in A$, $u = a_1 w_1 \dots w_\ell a_\ell w_{\ell+1}$, and $w_i \in (A \setminus (U \cup \{a_i\}))^*$ for each $1 \leq i \leq \ell$. Then the set of all normal forms is

$$\text{NF} = \{\bar{u}\langle\langle v, \bar{w} \rangle\rangle \mid u, v, w \in A^*, v \leq_U w\} = \bar{A}^* \left(\bigcup_{a \in A} (A \setminus (U \cup \{a\}))^* a\bar{a} \right)^* A^*.$$

From this equation we can infer that $\text{nf}(u) = \bar{u}_1 \langle\langle u_2, \bar{u}_3 \rangle\rangle$ is characterized by three components: The first component is the projection to the write actions $\pi(u) := u_2 = \pi_A(u)$ (note that the transitions of \mathcal{R} preserve the relative ordering of the write operations). Similarly, the second is the projection to the read actions $\bar{\pi}(u) := u_1 u_3$ (note that we suppress the overlines in this projection). Finally, the third component is the *overlap* $\bar{\pi}_2(u) := u_3$ of u . Note that the characterization of NF from above implies that $\bar{\pi}_2(u) \leq_U \pi(u)$ holds. Additionally, we can define $\bar{\pi}_1(u) := u_1$.

► **Example 2.3.** Recall Example 2.2. There, in case of $a \notin U$ we have for $u = aabb\bar{a}\bar{b}$: $\pi(u) = aabb$, $\bar{\pi}(u) = ab$, $\bar{\pi}_1(u) = \varepsilon$, and $\bar{\pi}_2(u) = ab$. Otherwise, if $a \in U$ we have $\bar{\pi}_1(u) = ab$ and $\bar{\pi}_2(u) = \varepsilon$.

While $\bar{\pi}_1(u)$ is defined using the semi-Thue system \mathcal{R} , it also has a natural meaning: $\bar{\pi}_1(u)$ is the shortest queue such that there is a run of the plq on execution of u that does not end up in the error state.

By [15, Proposition 3.21] the following holds about \mathcal{R} and $\text{nf}(u)$:

► **Proposition 2.4.** *Let $u, v \in \Sigma^*$. Then we have*

$$u \equiv v \iff \text{nf}(u) = \text{nf}(v) \iff (\pi(u), \bar{\pi}(u), \bar{\pi}_2(u)) = (\pi(v), \bar{\pi}(v), \bar{\pi}_2(v)). \quad \blacktriangleleft$$

With this main property in mind we can also apply π , $\bar{\pi}$, $\bar{\pi}_1$, and $\bar{\pi}_2$ to equivalence classes of \equiv (i.e., elements from $\mathcal{Q}(A, U)$) instead of words from Σ^* .

Another question is the description of the normal form of $u\bar{v}$ for any $u, v \in A^*$. We have $\pi(u\bar{v}) = u$ and $\bar{\pi}(u\bar{v}) = v$. It remains to describe the overlap $\bar{\pi}_2(u\bar{v})$.

► **Lemma 2.5.** *Let $u, v \in A^*$. Then $\bar{\pi}_2(u\bar{v})$ is the longest suffix v' of v that satisfies $v' \leq_U u$.* ◀

Since \equiv is a congruence we can infer $u \equiv \overline{\bar{\pi}_1(u)} \pi(u) \overline{\bar{\pi}_2(u)}$ for each $u \in \Sigma^*$ from Lemma 2.5.

2.2 Rationality, Recognizability, and Aperiodicity

Let \mathcal{M} be a monoid. A subset $L \subseteq \mathcal{M}$ is called *rational* if it can be constructed from the finite subsets of \mathcal{M} using union, concatenation, and Kleene iteration. The subset L is *recognizable* if there are a finite monoid \mathcal{F} and a homomorphism $\phi: \mathcal{M} \rightarrow \mathcal{F}$ such that $L = \phi^{-1}(\phi(L))$, i.e., if L is accepted by an \mathcal{M} -automaton. It is well-known that the image of a rational set under a homomorphism is rational again and that the homomorphic preimage of a recognizable set also is recognizable. Furthermore, the class of recognizable subsets of \mathcal{M} is closed under Boolean operations. Moreover, in a finitely generated monoid each recognizable set is rational by [18]. For example, this applies to $\mathcal{Q}(A, U)$ since this monoid is finitely generated. The converse direction is not true in general, e.g., in Theorem 3.4 we prove the existence of a rational subset of the plq monoid which is not recognizable. However, in free monoids generated by some alphabet Γ a subset $L \subseteq \Gamma^*$ is rational if, and only if, it is recognizable by Kleene's Theorem [13]. In this situation, we call L *regular*.

A recognizable set $L \subseteq \mathcal{M}$ is called *aperiodic* if there is $n \in \mathbb{N}$ such that for each $u, v, w \in \mathcal{M}$ we have $uv^n w \in L$ iff $uv^{n+1}w \in L$. It follows from [19] L is aperiodic if it is accepted by a counter-free \mathcal{M} -automaton. It is an easy exercise to prove that the class of aperiodic subsets is closed under Boolean operations and homomorphic preimages. By Schützenberger's Theorem [25] a language $L \subseteq \Gamma^*$ is aperiodic iff it is star-free. Recall that a set $L \subseteq \mathcal{M}$ is *star-free* if it can be constructed from finite subsets of \mathcal{M} using union, concatenation, and complementation.

2.3 Logic and Languages

In this subsection we recall the logics on words and their correspondence to languages known from [26].

Let Γ be an alphabet. By FO we denote the set of first-order formulas built up from the atomic formulas of the form $x = y$, $x < y$, and $Q_a(x)$ for $a \in \Gamma$ where x and y are variables. To simplify notation we write $Q_S(x)$ instead of $\bigvee_{a \in S} Q_a(x)$ for any $S \subseteq \Gamma$.

Now let $w = a_1 \dots a_n \in \Gamma^*$. The *word model* for w is the relational structure $\underline{w} = (\text{dom}(w), <^w, (Q_a^w)_{a \in \Gamma})$ where $\text{dom}(w) = \{1, \dots, n\}$ is the set of letter positions of w , $<^w$ is the natural order on $\text{dom}(w)$, and $Q_a^w = \{i \in \text{dom}(w) \mid a_i = a\}$ is the set of positions of letters labeled with a . Then we write $\underline{w} \models \phi[p_1, \dots, p_n]$ for $p_1, \dots, p_n \in \text{dom}(w)$ and a formula $\phi \in \text{FO}$ (i.e., ϕ is satisfied in \underline{w}) if ϕ evaluates to true on interpretation of $=$, $<$, Q_a as equality, $<^w$, and Q_a^w , respectively, and on interpretation of the free variables in ϕ as p_i 's. Then the language defined by the sentence ϕ is $L(\phi) = \{w \in \Gamma^* \mid \underline{w} \models \phi\}$. We say that a language $L \subseteq \Gamma^*$ is *FO-definable* if there is $\phi \in \text{FO}$ with $L = L(\phi)$.

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

Büchi's Theorem [4] states that a language is regular iff it is MSO-definable. Moreover, a language is star-free and hence aperiodic iff it is FO-definable by [19].

3 Algorithmic Properties of Rational Subsets

This section studies decision problems concerning the rational subsets of $\mathcal{Q}(A, U)$. We will see that the classes of rational and recognizable subsets do not coincide. Especially, we prove that we cannot decide whether a given rational subset of the plq monoid is recognizable. Additionally, we prove that emptiness of intersection and the unique decipherability in $\mathcal{Q}(A, U)$ are undecidable. Though, we will see first, that the uniform membership problem is NL-complete.

So, let $w \in \Sigma^*$. Then one can show that the number of left-divisors of $[w]$ in $\mathcal{Q}(A, U)$ is at most $|w|^3$. Recall that in a monoid \mathcal{M} u is a *left-divisor* of w if there is v such that $uv = w$. Hence, we can obtain a DFA with only $|w|^3$ many states that accepts $[w]$. In particular, similar to [11, Lemma 8.1] we can prove an even stronger result by using only logarithmic space on construction of this DFA. This implies the following theorem:

► **Theorem 3.1.** *Let A be an at least binary alphabet and $U \subseteq A$. Then the following rational subset membership problem for $\mathcal{Q}(A, U)$ is NL-complete: Given a word $w \in \Sigma^*$ and an NFA \mathcal{A} over Σ . Is there a word $v \in L(\mathcal{A})$ with $w \equiv v$?*

Proof. Let $w \in \Sigma^*$ and let \mathcal{A} be an NFA over Σ . Let \mathcal{B} be the aforementioned DFA that can be constructed using only logarithmic additional space.

Then there exists $v \in L(\mathcal{A})$ with $w \equiv v$ if, and only if, $L(\mathcal{A}) \cap [w] \neq \emptyset$ if, and only if, $L(\mathcal{A}) \cap L(\mathcal{B}) \neq \emptyset$. Using an on-the-fly construction of \mathcal{B} , this can be decided nondeterministically in logarithmic space. Hence, the problem is in NL.

Since the free monoid A^* embeds into $\mathcal{Q}(A, U)$ and since the rational subset membership problem for A^* is NL-hard, we also get NL-hardness for $\mathcal{Q}(A, U)$. ◀

Now we will prove some negative algorithmic results on rational subsets of the plq monoid. In [11, Section 8] these undecidabilities for reliable queues could be inferred from an embedding of $\{a, b\}^* \times \{c, d\}^*$ into $\mathcal{Q}(A, A)$. Unfortunately, this does not work in arbitrary plq monoids since this direct product does not embed into $\mathcal{Q}(\{a, b\}, \emptyset)$ by [15, Theorem 6.14]. Though, we can prove all the undecidability results considered in [11] for any plq monoid.

Some of these results are based on an embedding of the monoid $\{a\}^* \times \{c, d\}^*$ into $\mathcal{Q}(A, U)$. Unfortunately, this does not help for the following two problems since their counterparts in $\{a\}^* \times \{c, d\}^*$ are decidable. Hence, we have to prove them directly.

The first considered decision problem is the unique decipherability problem in $\mathcal{Q}(A, U)$, i.e., the question whether a given finite set S freely generates S^* . To this end, we will use the undecidability of this problem in $\{a, b\}^* \times \{c, d\}^*$ by encoding the elements of the given set and adding another item.

► **Theorem 3.2.** *Let A be an at least binary alphabet and $U \subseteq A$. Then, given a finite set $S \subseteq \mathcal{Q}(A, U)$, it is undecidable whether S^* is freely generated by S .*

Proof. We prove this undecidability by reduction of this question for the monoid $\{a, b\}^* \times \{c, d\}^*$, which is undecidable by [6, Theorem 3.1]. So, let $a, b \in A$ be distinct letters and $S = \{(x_1, y_1), \dots, (x_k, y_k)\}$. Define the embeddings $f: \{a, b\}^* \rightarrow A^*$ and $g: \{c, d\}^* \rightarrow A^*$ by

$f(a) = g(c) = aa$ and $f(b) = g(d) = ab$. Set $q_0 := [\overline{bbbbb}]$, $q_i := [f(x_i)\overline{g(y_i)}]$ for any $1 \leq i \leq k$, and $T := \{q_i \mid 0 \leq i \leq k\} \subseteq \mathcal{Q}(A, U)$. Then we can show that S^* is freely generated by S iff T^* is freely generated by T . ◀

The next problem to consider is the emptiness of intersections of rational subsets in the plq monoid. Given two recognizable sets, this problem is decidable since the class of recognizable subsets is effectively closed under intersection. However, we will prove that this decidability does not hold for arbitrary rational subsets. As a corollary we can infer that the class of rational subsets is not effectively closed under intersection. Afterwards we will prove the existence of two rational subsets whose intersection is not rational. In consequence, the classes of rational and recognizable subsets do not coincide. Nevertheless, each recognizable set in $\mathcal{Q}(A, U)$ is rational due to [18] since the plq monoid is finitely generated.

► **Theorem 3.3.** *Let A be an at least binary alphabet and $U \subseteq A$. Then the emptiness of the intersection of two rational subsets of $\mathcal{Q}(A, U)$ is undecidable.*

Proof. We prove this by reduction of Post's Correspondence Problem (PCP), which is undecidable by [23]. So, let $a, b \in A$ be distinct letters and $I = ((x_1, y_1), \dots, (x_k, y_k))$ be an instance of the PCP with $x_i, y_i \in A^*$. We define the following rational sets

$$X_I := \{p_i = [a^i b \overline{x_i}] \mid 1 \leq i \leq k\}^+ [\overline{a}] [\overline{b}]^* \quad \text{and} \quad Y_I := \{q_i = [a^i b \overline{y_i}] \mid 1 \leq i \leq k\}^+ [\overline{a}] [\overline{b}]^*.$$

We can show then that $X_I \cap Y_I \neq \emptyset$ if, and only if, I has a solution. ◀

To prove that the rational subsets are not closed under intersection and to prove the undecidability of the next problems we use an embedding of $\{a\}^* \times \{b, c\}^*$ into the plq monoid. Let $a, b \in A$ be distinct letters. Such an embedding is $\psi: \{a\}^* \times \{b, c\}^* \rightarrow \mathcal{Q}(A, U)$ with $\psi(a, \varepsilon) = [a]$, $\psi(\varepsilon, b) = [\overline{ab}]$, and $\psi(\varepsilon, c) = [\overline{abb}]$ by [15, Section 6.2].

► **Theorem 3.4.** *Let A be an at least binary alphabet and $U \subseteq A$. Then the set of rational subsets of $\mathcal{Q}(A, U)$ is not closed under intersection. In particular, there is a rational subset of $\mathcal{Q}(A, U)$ which is not recognizable.*

Proof. Consider the following rational relations:

$$R_1 = \{(a^m, b^m c^n) \mid m, n \in \mathbb{N}\} \quad \text{and} \quad R_2 = \{(a^m, b^n c^m) \mid m, n \in \mathbb{N}\}.$$

Then $\psi(R_1)$ and $\psi(R_2)$ are rational in $\mathcal{Q}(A, U)$. Suppose that $\psi(R_1) \cap \psi(R_2)$ is rational. Then there is a regular language $S \subseteq \Sigma^*$ with $\psi(R_1) \cap \psi(R_2) = \eta(S)$. Since ψ is injective we have $\psi(R_1) \cap \psi(R_2) = \psi(R_1 \cap R_2) = \psi(\{(a^n, b^n c^n) \mid n \in \mathbb{N}\})$. Hence, $\overline{\pi}(S) = \{(ab)^n (abb)^n \mid n \in \mathbb{N}\}$ would be regular since $\overline{\pi}$ is a homomorphism. But this is a contradiction to the Pumping Lemma. ◀

Gibbons and Rytter proved in [9] that universality and recognizability are undecidable in $\{a\}^* \times \{b, c\}^*$. Since ψ is an embedding of this monoid into the plq monoid, these undecidabilities imply the undecidability of their counterparts in the plq monoid.

► **Theorem 3.5.** *Let A be an at least binary alphabet and $U \subseteq A$. Then universality, inclusion, equality, and recognizability of rational subsets of $\mathcal{Q}(A, U)$ are undecidable.* ◀

4 Characterizations of the Recognizable Subsets

In Section 3 we have shown many decision problems on rational subsets of the plq monoid to be undecidable. We know that all these problems are decidable if the given subsets are recognizable from the known constructions in automata theory. Here, we want to give characterizations of the recognizable subsets in the manner of Kleene's and Büchi's Theorem [13, 4], i.e., we characterize the recognizable sets as certain rational sets and by logical means. At first, we state the main theorem. Later in this section we give the definitions of q -rational subsets and MSO_q and prove the correctness of this theorem.

► **Theorem 4.1 (Main Theorem).** *Let A be an at least binary alphabet, $U \subseteq A$, and $S \subseteq \mathcal{Q}(A, U)$. Then the following are equivalent:*

- (A) S is recognizable.
- (B) S is q -rational.
- (C) S is MSO_q -definable.

4.1 Some Helping Characterizations

Before we prove Theorem 4.1 we state two further characterizations which turned out to be convenient for simplification of the proof of Theorem 4.1. We know these characterizations from [11] for the recognizable subsets in the reliable queue monoid $\mathcal{Q}(A, A)$ and generalize them to plq monoids $\mathcal{Q}(A, U)$ with arbitrary subsets $U \subseteq A$. On the one hand, we prove the correspondence of recognizability in the plq monoid to regularity in the underlying free monoid. On the other hand, we show that each recognizable subset is a Boolean combination of sets $\pi^{-1}(R)$, $\bar{\pi}^{-1}(R)$ where $R \subseteq A^*$ is regular and some special sets Ω_ℓ for any $\ell \in \mathbb{N}$:

► **Definition 4.2.** Let $q \in \mathcal{Q}(A, U)$. Then the *overlap's bounded width* of q is

$$\omega(q) := \inf\{|\bar{\pi}_2(p)| : p \in \mathcal{Q}(A, U), \pi(p) = \pi(q), \bar{\pi}(p) = \bar{\pi}(q), |\bar{\pi}_2(q)| < |\bar{\pi}_2(p)|\}.$$

Furthermore, for $\ell \in \mathbb{N}$ set $\Omega_\ell := \{q \in \mathcal{Q}(A, U) \mid \omega(q) > \ell\}$.

The overlap's bounded width specifies the minimal length of the overlap of a word with the same projections having a longer overlap. If such word does not exist then we set this value to ∞ .

► **Example 4.3.** Let $A = U = \{a, b\}$ and $q = \overline{abab\bar{a}\bar{a}\bar{b}\bar{b}abab}$. Then there are two words with the same projections and longer overlaps: $q_1 = \overline{aba\bar{a}\bar{b}\bar{b}a\bar{a}\bar{b}\bar{b}ab}$ and $q_2 = \overline{a\bar{a}\bar{b}\bar{b}a\bar{a}\bar{b}\bar{b}a\bar{a}\bar{b}\bar{b}}$. We have $|\bar{\pi}_2(q_1)| = 4$ and $|\bar{\pi}_2(q_2)| = 6$. Therefore, we have $\omega(q) = 4$, $\omega(q_1) = 6$, and $\omega(q_2) = \infty$. Hence, $q \in \Omega_3 \setminus \Omega_4$ holds.

From [11, Observation 9.1] we know that any non-trivial property of the overlap's width $|\bar{\pi}_2(q)|$ is not recognizable in $\mathcal{Q}(A, A)$. An appropriate alternative for the generators of the Boolean algebra of recognizable subsets was found in such kind of "overapproximation" of the overlap's length (note that $\omega(q) > |\bar{\pi}_2(q)|$ holds). Additionally, the following observations provide some more motivation of this notion:

► **Observation 4.4.** *Every $q \in \mathcal{Q}(A, U)$ is completely described by $\pi(q)$, $\bar{\pi}(q)$, and $\omega(q)$. ◀*

► **Observation 4.5.** *Let $\ell \in \mathbb{N}$ and $w \in \Sigma^*$. Then $\omega([w]) \leq \ell$ if, and only if, there is $u \in A^{\leq \ell}$ with $\bar{\pi}(w) \in A^*u$ and $u \subseteq_U \pi(w)$ such that $|\bar{\pi}_2(w)| < |u|$. ◀*

Now we can state the following equivalences which can be proven similar to [11, Theorem 9.4].

► **Theorem 4.6.** *Let A be an at least binary alphabet, $U \subseteq A$, and $S \subseteq \mathcal{Q}(A, U)$. Then the following are equivalent:*

1. S is recognizable.
2. $\eta^{-1}(S) \cap \overline{A^* A^* \overline{A^*}}$ is regular.
3. S is a Boolean combination of sets of the form $\pi^{-1}(R)$ or $\overline{\pi^{-1}(R)}$ for some regular $R \subseteq A^*$ and the sets Ω_ℓ for some $\ell \in \mathbb{N}$. ◀

4.2 From Recognizability to Q-Rational Subsets

In this subsection we prove that each recognizable subset in the plq monoid is q-rational. To this end, we first need to define this notion which is a restriction to the rational expressions. We need this restriction since we cannot translate Kleene's Theorem [13] to plq monoids due to Theorem 3.4. Though, we can use Ochmański's approach from [21] to generate the recognizable subsets. Concretely, we restrict the Kleene star and the concatenation of the plq monoid in an appropriate way. We call the sets generated by those operations *q-rational* and prove that these are exactly the recognizable subsets in the plq monoid.

At first, we prove that the class of recognizable subsets is not closed under iteration:

► **Remark.** Let $S = \{[a\overline{a}]\}$, which is trivially recognizable. Then $\eta^{-1}(S^*) \cap A^* \overline{A^*} \subseteq \Sigma^*$ is the set of all words $a^n \overline{a}^n$ with $n \in \mathbb{N}$ by Rule d in Definition 2.1. This language is not regular. Hence, $\eta^{-1}(S^*)$ is also not regular and therefore S^* is not recognizable.

This is a very similar situation as in trace monoids. Here, Ochmański proved in [21] that it suffices to restrict iteration to obtain some kind of rational expressions that are generating all the recognizable subsets [21]. Unfortunately, the class of recognizable subsets in the plq monoid also is not closed under product.

► **Remark.** Let $S = \{[a]\}^*$ and $T = \{[\overline{a}]\}^*$, which are recognizable. Then $\eta^{-1}(S \cdot T) \cap \overline{A^* A^* \overline{A^*}} \subseteq \Sigma^*$ is the set of all words $\overline{u_1} u_2 \overline{u_3}$ with $u_1, u_2, u_3 \in a^*$ and $u_1 = \varepsilon$ or $|u_2| \leq |u_3|$ by Rule d in Definition 2.1. Since this language is not regular, $S \cdot T$ is not recognizable.

Hence, we have to restrict the use of the monoid's product in the construction of the so-called *q-rational* subsets. Next, we will define these subsets and afterwards we prove that these are a suitable restriction of rationality to describe exactly the recognizable subsets. But at first, we say that a subset of $\mathcal{Q}(A, U)$ is *q⁺-rational* if it can be obtained by the following rules:

- (1⁺) $\pi^{-1}(\varepsilon)$, $\pi^{-1}(\emptyset) = \emptyset$, and $\pi^{-1}(a)$ for any $a \in A$ are q⁺-rational
 (2⁺) if $S_1, S_2 \subseteq \mathcal{Q}(A, U)$ are q⁺-rational then $S_1 \cup S_2$, $S_1 \cdot S_2$, and S_1^* are q⁺-rational
 Similarly, by replacing π^{-1} by $\overline{\pi^{-1}}$ in the rules above, we define the class of *q⁻-rational* subsets of $\mathcal{Q}(A, U)$.

► **Observation 4.7.** *Let $S \subseteq \mathcal{Q}(A, U)$. Then S is q⁺-rational (q⁻-rational) if, and only if, there is some regular $R \subseteq A^*$ with $S = \pi^{-1}(R)$ ($S = \overline{\pi^{-1}(R)}$, resp.).* ◀

Finally, a subset of $\mathcal{Q}(A, U)$ is *q-rational* if it can be constructed from the following rules:

- (1) if $S_1 \subseteq \mathcal{Q}(A, U)$ is q⁺- or q⁻-rational it also is q-rational
- (2) if $S_1, S_2 \subseteq \mathcal{Q}(A, U)$ are q-rational then $S_1 \cup S_2$ and $\mathcal{Q}(A, U) \setminus S_1$ are q-rational
- (3) if $S_1 \subseteq \mathcal{Q}(A, U)$ is q⁺-rational and $S_2 \subseteq \mathcal{Q}(A, U)$ is q⁻-rational such that $\overline{\pi}(S_2)$ is finite (i.e., S_2 is obtained without usage of the *-operator) then $S_1 \cdot \mathcal{Q}(A, U) \cdot S_2$ is q-rational

► **Example 4.8.** Let $S = \{q \in \mathcal{Q}(A, U) \mid \pi(q) \in (ab)^*, \overline{\pi}(q) = b\}$. Then S is q-rational since we have $S = \overline{\pi^{-1}(b)} \cap (\pi^{-1}(a) \cdot \pi^{-1}(b))^*$. Note that the class of q-rational subsets 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. But we can remove neither the factor $\mathcal{Q}(A, U)$, which appears as separator in this product, nor the finiteness of $\bar{\pi}(S_2)$. Additionally, we cannot simply remove this rule since the set $\{\{a\bar{a}\}\}$ cannot be built by application of the Rules 1 and 2, only.

Now we can prove the implication “A \Rightarrow B” in Theorem 4.1. To do this, we utilize Theorem 4.63. Concretely, we do this by induction on the syntax tree of such kind of expression that each recognizable subset is q-rational. 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 4.9.** *Let $\ell \in \mathbb{N}$, $q \in \mathcal{Q}(A, U)$ and $u = a_1 \dots a_\ell \in A^*$. Then we have $u \leq_U \pi(q)$ and $\bar{\pi}_2(q) \in A^*u$ if, and only if, $q \in \pi^{-1}\left(\prod_{i=1}^\ell (A \setminus U)^* a_i\right) \cdot \mathcal{Q}(A, U) \cdot \bar{\pi}^{-1}(u)$. ◀*

Finally, we can state the following implication:

► **Proposition 4.10.** *Let $S \subseteq \mathcal{Q}(A, U)$ be recognizable. Then S is q-rational.*

Proof. We use Theorem 4.63 to prove the claim by induction. At first, if $S = \pi^{-1}(R)$ or $S = \bar{\pi}^{-1}(R)$ where $R \subseteq A^*$ is regular, then S is q-rational by Observation 4.7.

Next, let $\ell \in \mathbb{N}$ and $S = \Omega_\ell$. Then by Observation 4.5 and Lemma 4.9 we have

$$\Omega_\ell = \bigcap_{u \in A^{\leq \ell}} (\mathcal{Q}(A, U) \setminus (\pi^{-1}(W_u A^*) \cap \bar{\pi}^{-1}(A^* u)) \cup \pi^{-1}(W_u) \cdot \mathcal{Q}(A, U) \cdot \bar{\pi}^{-1}(u)),$$

where $W_u = \prod_{i=1}^k (A \setminus U)^* a_i$ with $u = a_1 \dots a_k$. Since the sets $\pi^{-1}(W_u A^*)$, $\bar{\pi}^{-1}(A^* u)$, $\pi^{-1}(W_u)$, and $\bar{\pi}^{-1}(u)$ are q-rational by Observation 4.7, Ω_ℓ is q-rational as well.

Finally, the class of q-rational subsets is closed under Boolean operations. ◀

4.3 From Q-Rational Subsets to Logic

The second implication from Theorem 4.1 states that each q-rational subset is definable in a special monadic second-order logic which we call MSO_q . Here, we try to exhibit the knowledge from the preceding subsection such that this logic defines exactly the recognizable subsets. In fact, we have to add some modifications to Büchi’s MSO -logic from [4]. At first, we should understand $p \leq^w q$ as follows: the letter a on position p in w cannot be moved to the right of the letter b on position q without violating any of the rules from Definition 2.1 (recall that \mathcal{R} only swaps letters). In other words, for any $v \in [w]$ the letter a appears left from b in v . Additionally, we have to restrict comparisons of write and read operations:

► **Remark.** It is not possible to compare arbitrary letters in w without any restrictions. For example, let

$$\phi = \exists x, y: (Q_A(x) \wedge \forall z: (Q_A(z) \rightarrow z \leq x) \wedge Q_{\bar{A}}(y) \wedge \forall z: (Q_{\bar{A}}(z) \rightarrow y \leq z) \wedge \neg x \leq y),$$

i.e., \underline{w} satisfies ϕ iff the first read action can be moved to the right of the last write action. Then we have $L(\phi) \cap \bar{a}^* a^* \bar{a}^* = \{\bar{a}^k a^\ell \bar{a}^m \mid k = 0 \text{ or } m \geq \ell\}$. Since this language is not regular, the subset of $\mathcal{Q}(A, U)$ of the elements satisfying ϕ is not recognizable either.

By FO_q we denote the set of all first-order formulas build up from the atomic formulas of the form $x = y$, $x <_+ y$, $x <_- y$, $P_\ell(x)$ for $\ell \in \mathbb{N}_+$, and $Q_a(x)$ for $a \in A$ where x and y are variables. Additionally, by MSO_q we denote the monadic second-order extension of FO_q .

Now let $w = a_1 \dots a_n \in \Sigma^*$. The *plq model* for w is the relational structure $\tilde{w} := (\text{dom}(w), <_+^w, <_-^w, (P_\ell^w)_{\ell \in \mathbb{N}_+}, (Q_a^w)_{a \in \Sigma})$ where $\text{dom}(w) = \{1, \dots, n\}$, $Q_a^w = \{i \mid a_i = a\}$, $<_+^w$ and $<_-^w$ are the natural orderings on $Q_A^w = \bigcup_{a \in A} Q_a^w$ and $Q_{\bar{A}}^w$, respectively, and

$$P_\ell^w = \{i \in Q_A^w \mid \forall v_1, v_2 \in \Sigma^*: (w \equiv v_1 v_2 \wedge \pi(v_1) = \pi(a_1 \dots a_i)) \rightarrow |\bar{\pi}(v_2)| < \ell\},$$

i.e., we have $i \in P_\ell^w$ iff $a_i \in A$ and the ℓ th last read action in w is left from a_i and cannot be moved to the right of a_i . 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 . Hence, we can infer that \tilde{w} identifies the equivalence class $[w]$:

► **Lemma 4.11.** *Let $v, w \in \Sigma^*$. Then we have $v \equiv w$ if, and only if, $\tilde{v} \cong \tilde{w}$.* ◀

Therefore, we can define the *plq model* $\tilde{q} := \widetilde{\text{nf}(q)}$ for $q \in \mathcal{Q}(A, U)$.

Now let $\phi \in \text{MSO}_q$. The set defined by ϕ is $S(\phi) = \{q \in \mathcal{Q}(A, U) \mid \tilde{q} \models \phi\}$. We say that $S \subseteq \mathcal{Q}(A, U)$ is *MSO_q-definable* (*FO_q-definable*) if there is $\phi \in \text{MSO}_q$ ($\phi \in \text{FO}_q$, respectively) with $S = S(\phi)$.

► **Remark.** The sets P_ℓ^w also are conform to the special product in the definition of q-rational subsets into logics. In particular, we have $S(\exists x: \neg P_\ell(x)) = \pi^{-1}(A^+) \cdot \mathcal{Q}(A, U) \cdot \bar{\pi}^{-1}(A^\ell)$.

Now we prove that each q-rational subset is MSO_q-definable. In the proof of implication “B⇒C” in Theorem 4.1 we need the following notion: Let $\phi, \xi(x) \in \text{MSO}$. Then there is a formula $\phi|_\xi \in \text{MSO}$ which restricts the quantifiers in ϕ to values satisfying $\xi(x)$. Thereby, we have $\phi|_\xi \in \text{FO}$ iff $\phi, \xi \in \text{FO}$.

Finally, we can state:

► **Proposition 4.12.** *Let $S \subseteq \mathcal{Q}(A, U)$ be q-rational. Then S is MSO_q-definable.*

Proof. If S is q⁺-rational then we have $S = \pi^{-1}(R)$ for some regular $R \subseteq A^*$. By [4] there is an MSO-formula ϕ with $L(\phi) = R$. Then by replacing of all occurrences of $<$ in ϕ by $<_+$ we obtain an MSO_q-formula ϕ' with $S(\phi'|_{Q_A(x)}) = \pi^{-1}(L(\phi)) = S$.

Similarly, we can prove that S is MSO_q-definable if S is q⁻-rational (here, we replace $<$ by $<_-$ and restrict to $Q_{\bar{A}}$).

If $S = S_1 \cup S_2$ or $S = \mathcal{Q}(A, U) \setminus S_1$, where S_1, S_2 are q-rational there are $\phi_1, \phi_2 \in \text{MSO}_q$ with $S(\phi_1) = S_1$ and $S(\phi_2) = S_2$. Then we have $S = S(\phi_1 \vee \phi_2)$ and $S = S(\neg \phi_1)$, respectively.

Finally, let $S = \pi^{-1}(R) \cdot \mathcal{Q}(A, U) \cdot \bar{\pi}^{-1}(F)$ where $R \subseteq A^*$ is regular and $F \subseteq A^*$ is finite. W.l.o.g. we can assume that $F = \{w\}$ holds. Then there are MSO_q-formulas ϕ_R and ϕ_F defining $\pi^{-1}(R)$ and $\bar{\pi}^{-1}(F)$, respectively. Set

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

Then we have $S = S(\phi)$. ◀

4.4 From Logic to Recognizability

Finally, we have to prove that each MSO_q-definable subset is recognizable. To do this, we utilize Theorem 4.62. In other words, given $\phi \in \text{MSO}_q$ we construct a formula $\psi \in \text{MSO}$ such that $\eta^{-1}(S(\phi)) \cap \bar{A}^* A^* \bar{A}^* = L(\psi) \cap \bar{A}^* A^* \bar{A}^*$ holds. Since the right-hand side of this equation is regular by [4], we can infer that $S(\phi)$ is recognizable.

The translation of formula $P_\ell(x)$ is the most complicated case in our construction since write and read actions commute in certain contexts given in Definition 2.1. Concretely, we will translate $\neg P_\ell(x)$ since it seems to be easier to understand. Hence, we start with this case. At first, we prove that there is an FO-formula describing the words in which the last ℓ read actions are U -prefixes of the write actions:

► **Lemma 4.13.** *Let $\ell \in \mathbb{N}$. There is a sentence $\text{overlap}_\ell \in \text{FO}$ such that $w \in L(\text{overlap}_\ell)$ if, and only if, there is $u \in A^\ell$ with $u \leq_U \pi(w)$ and $\bar{\pi}_2(w) \in A^* u$ for any $w \in \Sigma^*$.*

Proof. There is an NFA \mathcal{A} which guesses the last ℓ read actions, verifies afterwards whether these are a U -prefix of the write actions, and checks whether each of the last ℓ read actions appear after their corresponding write action. This NFA can be constructed without usage of any counters. Hence, its accepted language is aperiodic. By [19] there is a formula $\text{overlap}_\ell \in \text{FO}$ with $L(\text{overlap}_\ell) = L(\mathcal{A})$. \blacktriangleleft

Now let $w \in \bar{A}^* A^* \bar{A}^*$ and $p \in \text{dom}(w)$. Then we express $p \notin P_\ell^w$ as follows:

If we have $p \in Q_{\bar{A}}^w$ we are ready. So, assume $p \in Q_A^w$ from now on. At first, we choose the last ℓ read actions from w . Let q_1, \dots, q_ℓ be their positions.

If $p < q_1$ then we are ready. So, assume $q_1 < p$ from now on. Then there are two words $u, v \in \Sigma^*$ such that $w \equiv uv$, u ends with the letter on position p in w , and v starts with the letter on position q_1 in w . Since \equiv is a congruence we can assume that $v = \bar{\pi}_1(v)\pi(v)\bar{\pi}_2(v)$ holds. Let $\bar{\pi}(v) = b_1 \dots b_\ell$. Then \bar{b}_i can be moved to the left-hand side of the letter on position p in w if, and only if, uv does not satisfy $\text{overlap}_{\ell-i+1}$.

Finally, there may be some letters \bar{b}_i from $\bar{\pi}_1(v)$ that can be moved to the right in w . This is possible if, and only if, one of the following two cases hold: on the one hand, this is possible if $b_i \dots b_\ell \not\leq_U \pi(w)$. On the other hand, if $b_i \dots b_\ell \leq_U \pi(w)$ and the write action corresponding to \bar{b}_i appears right from position p in w .

All of the above mentioned requirements can be expressed in MSO-formulas. Hence, we can construct $\text{co-P}_\ell(x) \in \text{MSO}$ such that $\underline{w} \models \text{co-P}_\ell[p]$ if, and only if, $p \notin P_\ell^w$. Therefore, we can state the following:

► **Proposition 4.14.** *Let $S \subseteq \mathcal{Q}(A, U)$ be MSO_q -definable. Then S is recognizable.*

Proof. Let $S \subseteq \mathcal{Q}(A, U)$ be MSO_q -definable. Then there is $\phi \in \text{MSO}_q$ with $S = S(\phi)$. We construct $\phi' \in \text{MSO}$ by the following modifications of ϕ :

- replace “ $x <_+ y$ ” by “ $x < y \wedge Q_A(x) \wedge Q_A(y)$ ”
- replace “ $x <_- y$ ” by “ $x < y \wedge Q_{\bar{A}}(x) \wedge Q_{\bar{A}}(y)$ ”
- replace “ $P_\ell(x)$ ” by “ $\neg \text{co-P}_\ell[x]$ ”

Then we can prove that $\tilde{w} \models \phi$ if, and only if, $\underline{w} \models \phi'$ for any $w \in \bar{A}^* A^* \bar{A}^*$. Hence, by Büchi’s Theorem [4] $\eta^{-1}(S) \cap \bar{A}^* A^* \bar{A}^*$ is regular, i.e., S is recognizable due to Theorem 4.6. \blacktriangleleft

5 Characterizations of the Aperiodic Subsets

In the previous section we have seen a Kleene- and Büchi-type characterization of the recognizable subsets in the plq monoid. Another more involved task is to describe the aperiodic subsets in the plq monoid. Schützenberger has proven in [25] that the aperiodic subsets in the free monoid are exactly the star-free languages. This result gives us a decision procedure to decide whether a given regular language is star-free. Another similar result for trace monoids can be found in [10]. These two results cannot be translated to plq monoids since the class of aperiodic subsets is not closed under product. Though, we will see that we can restrict the use of the product to describe exactly the aperiodic subsets of the plq monoid.

Another characterization of the aperiodic languages was proven by [19]: similar to Büchi’s Theorem [4] McNaughton and Papert proved that these are exactly the FO-definable languages. Here, we will see that analogously the aperiodic subsets in the plq monoid are the FO_q -definable subsets.

Before we give these characterizations we have to define the restriction of star-freeness. We say that a subset of $\mathcal{Q}(A, U)$ is q -star-free if it can be constructed by the Rules 1-3 in which we replace “ $S = S_1^*$ ” by “ $S = \mathcal{Q}(A, U) \setminus S_1$ ” in the rules 2 and (2⁻).

Similarly, to Theorem 4.1 we can state and prove the following result:

► **Theorem 5.1.** *Let A be an at least binary alphabet, $U \subseteq A$, and $S \subseteq \mathcal{Q}(A, U)$. Then the following are equivalent:*

- (A) S is aperiodic.
- (B) L is q -star-free.
- (C) L is FO_q -definable. ◀

References

- 1 Parosh Aziz Abdulla and Bengt Jonsson. Verifying programs with unreliable channels. *Inf. Comput.*, 127(2):91–101, 1996. doi:10.1006/inco.1996.0053.
- 2 Jean Berstel. *Transductions and Context-Free Languages*. Teubner Studienbücher, 1979. doi:10.1007/978-3-663-09367-1.
- 3 Daniel Brand and Pitro Zafiropulo. On communicating finite-state machines. *J. ACM*, 30(2):323–342, 1983. doi:10.1145/322374.322380.
- 4 J. Richard Büchi. Weak second-order arithmetic and finite automata. *Mathematical Logic Quarterly*, 6(1-6):66–92, 1960. doi:10.1002/malq.19600060105.
- 5 Gérard Cécé, Alain Finkel, and S. Purushothaman Iyer. Unreliable channels are easier to verify than perfect channels. *Inf. Comput.*, 124(1):20–31, 1996. doi:10.1006/inco.1996.0003.
- 6 Marek Chrobak and Wojciech Rytter. Unique decipherability for partially commutative alphabet (extended abstract). In Jozef Gruska, Branislav Rován, and Juraj Wiedermann, editors, *Mathematical Foundations of Computer Science 1986, Bratislava, Czechoslovakia, August 25-29, 1996, Proceedings*, volume 233 of *Lecture Notes in Computer Science*, pages 256–263. Springer, 1986. doi:10.1007/BFb0016249.
- 7 Volker Diekert and Grzegorz Rozenberg. *The book of traces*. World scientific, 1995. doi:10.1142/9789814261456.
- 8 Alain Finkel. Decidability of the termination problem for completely specified protocols. *Distributed Computing*, 7(3):129–135, 1994. doi:10.1007/BF02277857.
- 9 Alan Gibbons and Wojciech Rytter. On the decidability of some problems about rational subsets of free partially commutative monoids. *Theor. Comput. Sci.*, 48(3):329–337, 1986. doi:10.1016/0304-3975(86)90101-5.
- 10 Giovanna Guaiana, Antonio Restivo, and Sergio Salemi. On aperiodic trace languages. In Christian Choffrut and Matthias Jantzen, editors, *STACS 91, 8th Annual Symposium on Theoretical Aspects of Computer Science, Hamburg, Germany, February 14-16, 1991, Proceedings*, volume 480 of *Lecture Notes in Computer Science*, pages 76–88. Springer, 1991. doi:10.1007/BFb0020789.
- 11 Martin Huschenbett, Dietrich Kuske, and Georg Zetsche. The monoid of queue actions. *Semigroup forum*, 95(3):475–508, 2017. doi:10.1007/s00233-016-9835-4.
- 12 Mark Kambites. Formal languages and groups as memory. *Communications in Algebra*, 37(1):193–208, 2009. doi:10.1080/00927870802243580.
- 13 Stephen C. Kleene. Representation of events in nerve nets and finite automata. *Automata Studies*, 1956.
- 14 Chris Köcher. Einbettungen in das Transformationsmonoid einer vergesslichen Warteschlange. Master’s thesis, TU Ilmenau, 2016. doi:10.13140/RG.2.2.21984.99842.
- 15 Chris Köcher, Dietrich Kuske, and Olena Prianychnykova. The transformation monoid of a partially lossy queue. *RAIRO - Theoretical Informatics and Applications*, 2018. To appear.
- 16 Markus Lohrey. The rational subset membership problem for groups: a survey. In *Groups St Andrews*, volume 422, pages 368–389. Cambridge University Press, 2013. doi:10.1017/CB09781316227343.024.

- 17 Benoît Masson and Philippe Schnoebelen. On verifying fair lossy channel systems. In Krzysztof Diks and Wojciech Rytter, editors, *Mathematical Foundations of Computer Science 2002, 27th International Symposium, MFCS 2002, Warsaw, Poland, August 26-30, 2002, Proceedings*, volume 2420 of *Lecture Notes in Computer Science*, pages 543–555. Springer, 2002. doi:10.1007/3-540-45687-2_45.
- 18 J. D. McKnight. Kleene quotient theorems. *Pacific Journal of Mathematics*, 14(4):1343–1352, 1964.
- 19 Robert McNaughton and Seymour A. Papert. *Counter-Free Automata*, volume 65. The MIT Press, 1971.
- 20 Anca Muscholl and Holger Petersen. A note on the commutative closure of star-free languages. *Inf. Process. Lett.*, 57(2):71–74, 1996. doi:10.1016/0020-0190(95)00187-5.
- 21 Edward Ochmański. Regular behaviour of concurrent systems. *Bulletin of the EATCS*, 27:56–67, 1985.
- 22 Jean-Éric Pin. Mathematical foundations of automata theory. *Lecture notes LIAFA, Université Paris*, 7, 2010.
- 23 Emil L. Post. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52(4):264–268, 1946.
- 24 Elaine Render and Mark Kambites. Rational subsets of polycyclic monoids and valence automata. *Inf. Comput.*, 207(11):1329–1339, 2009. doi:10.1016/j.ic.2009.02.012.
- 25 Marcel Paul Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8(2):190–194, 1965. doi:10.1016/S0019-9958(65)90108-7.
- 26 Wolfgang Thomas. Languages, automata, and logic. In Grzegorz Rozenberg and Arto Salomaa, editors, *Handbook of Formal Languages: Volume 3 Beyond Words*, pages 389–455. Springer, 1997. doi:10.1007/978-3-642-59126-6_7.