Semi-linear Lattices and Right One-Way Jumping Finite Automata (Extended Abstract)

Simon Beier and Markus Holzer(✉)

Institut für Informatik, Universität Giessen,
Arndtstr. 2, 35392 Giessen, Germany
{simon.beier,holzer}@informatik.uni-giessen.de

Abstract. Right one-way jumping automata (ROWJFAs) are an automaton model that was recently introduced for processing the input in a discontinuous way. In [S. Beier, M. Holzer: Properties of right one-way jumping finite automata. In Proc. 20th DCFS, number 10952 in LNCS, 2018] it was shown that the permutation closed languages accepted by ROWJFAs are exactly those with a finite number of positive Myhill-Nerode classes. Here a Myhill-Nerode equivalence class $[w]_L$ of a language L is said to be positive if w belongs to L. Obviously, this notion of positive Myhill-Nerode classes generalizes to sets of vectors of natural numbers. We give a characterization of the linear sets of vectors with a finite number of positive Myhill-Nerode classes, which uses rational cones. Furthermore, we investigate when a set of vectors can be decomposed as a finite union of sets of vectors with a finite number of positive Myhill-Nerode classes. A crucial role is played by lattices, which are special semi-linear sets that are defined as a natural way to extend "the pattern" of a linear set to the whole set of vectors of natural numbers in a given dimension. We show connections of lattices to the Myhill-Nerode relation and to rational cones. Some of these results will be used to give characterization results about ROWJFAs with multiple initial states. For binary alphabets we show connections of these and related automata to counter automata.

1 Introduction

Semi-linear sets, Presburger arithmetic, and context-free languages are closely related to each other by the results of Ginsburg and Spanier [10] and Parikh [14]. More precisely, a set is semi-linear if and only it is expressible in Presburger arithmetic, which is the first order theory of addition. These sets coincide with the Parikh images of regular languages, which are exactly the same as the Parikh images of context-free languages by Parikh's theorem that states that the Parikh image of any context-free language is semi-linear. Since then semi-linear sets and results thereof are well known in computer science. Recently, the interest on semi-linear sets has increased significantly. On the one hand, there was renewed

M. Hospodár and G. Jirásková (Eds.): CIAA 2019, LNCS 11601, pp. 70–82, 2019.
https://doi.org/10.1007/978-3-030-23679-3_6

interest in equivalence problems on permutation closed languages [12] which obviously correspond to their Parikh-image, and on the other hand, it turned out that semi-linearity is the key to understand the accepting power of jumping finite automata, an automaton model that was introduced in [13] for discontinuous information processing. Roughly speaking, a jumping finite automaton is an ordinary finite automaton, which is allowed to read letters from anywhere in the input string, not necessarily only from the left of the remaining input. Moreover, semi-linear sets were also subject to descriptional complexity considerations in [3] and [5].

The tight relation between semi-linear sets and jumping automata is not limited to this automaton model, but also turns over to right one-way jumping automata (ROWJFAs), which were introduced in [4], as shown in [1,2]. This device moves its head from left-to-right starting from the leftmost letter in the input, reads and erases some symbols, while it jumps over others, and when it reaches the end of the input word, it returns to the beginning and continues the computation, which is executed deterministically. Most questions on formal language related problems such as inclusion problems, closure properties, and decidability of standard problems concerning ROWJFAs were answered recently in one of the papers [1,2,4]. One of the main results on these devices was a characterization of the induced language family that reads as follows: a permutation closed language L belongs to **ROWJ**, the family of all languages accepted by ROWJFAs, if and only if L can be written as the *finite union* of Myhill-Nerode equivalence classes. Observe, that the overall number of equivalence classes can be infinite. This result nicely contrasts the characterization of regular languages, which requires that the overall number of equivalence classes is finite.

In this paper we try to improve the understanding of the Myhill-Nerode equivalence relation given by a subset of $\mathbb{N}^k$ as defined in [9]. For a subset $S \subseteq \mathbb{N}$ and the induced Myhill-Nerode relation, an equivalence class is called positive if the vectors of the class lie in S. We characterize in which cases linear sets have only a finite number of positive equivalence classes in terms of rational cones, which are a special type of convex cones that are important objects in different areas of mathematics and computer science like combinatorial commutative algebra, geometric combinatorics, and integer programming. A special type of semi-linear sets called lattices is introduced. Their definition is inspired by the mathematical object of a lattice which is of great importance in geometry and group theory, see [6]. These lattices are subgroups of $\mathbb{R}^k$ that are isomorphic to $\mathbb{Z}^k$ and span the real vector space $\mathbb{R}^k$. Our semi-linear lattices are defined like linear sets, but allowing integer coefficients for the period vectors, instead of only natural numbers. However our lattices are still, per definition, subsets of $\mathbb{N}^k$. Lattices have only one positive Myhill-Nerode class and can be decomposed as a finite union of linear sets with only one positive Myhill-Nerode class. We give a characterization of the lattices that can even be decomposed as a finite union of linear sets with linearly independent period sets and only one positive Myhill-Nerode class and again get a connection to rational cones. That is why we study these objects in more detail and show that the set of vectors with only non-negative components in a linear subspace of dimension n of $\mathbb{R}^k$ spanned by a subset of $\mathbb{N}^k$ always forms a rational cone spanned by a linearly independent subset of $\mathbb{N}^k$ if and only if $n \in \{0, 1, 2, k\}$. This result has consequences for the mentioned decompositions of lattices. We show when a subset of $\mathbb{N}^k$ can be decomposed as a finite union of those subsets that have only a finite number of positive Myhill-Nerode classes. That result heavily depends on the theory of lattices.

The obtained results on lattices are applied to ROWJFAs generalized to devices with multiple initial states (MROWJFAs). This slight generalization is in the same spirit as the one for ordinary finite automata that leads to multiple entry deterministic finite automata [7]. We show basic properties of MROWJFAs and inclusion relations to families of the Chomsky hierarchy and related families. A connection between the family of permutation closed languages accepted by MROWJFAs (the corresponding language family is referred to **pMROWJ**) and lattices is shown. This connection allows us to deduce a characterization of languages in **pMROWJ** from our results about lattices and decompositions of subsets of $\mathbb{N}^k$. We also investigate the languages accepted by MROWJFAs and related languages families for the special case of a binary input alphabet and get in some cases different or stronger results than for arbitrary alphabets. We can show that each permutation closed semi-linear language (these are exactly the languages accepted by jumping finite automata) over a binary alphabet is accepted by a counter automaton. Furthermore, each language over a binary alphabet accepted by a ROWJFA is also accepted by a realtime deterministic counter automaton. Our results for lattices lead to a characterization, which is stronger than the one for arbitrary alphabets, of the languages over binary alphabets in **pMROWJ**: these are exactly the languages that are a finite union of permutation closed languages accepted by ROWJFAs, which are characterized by positive Myhill-Nerode classes as stated above.

2 Preliminaries

We use $\subseteq$ for inclusion and $\subset$ for proper inclusion of sets. For a binary relation $\sim$ let $\sim^+$ and $\sim^*$ denote the transitive closure of $\sim$ and the transitive-reflexive closure of $\sim$, respectively. In the standard manner, $\sim$ is extended to $\sim^n$, where $n \geq 0$. Let $\mathbb{Z}$ be the set of integers, $\mathbb{R}$ be the set of real numbers, and $\mathbb{N}$ ($\mathbb{R}_{\geq 0}$, respectively) be the set of integers (real numbers, respectively) which are non-negative. Let $k \geq 0$. For the set $T \subseteq \{1, 2, \dots, k\}$ with $T = \{t_1, t_2, \dots, t_\ell\}$ and $t_1 < t_2 < \cdots < t_\ell$ we define $\pi_{k,T} : \mathbb{N}^k \to \mathbb{N}^{|T|}$ as $\pi_{k,T}(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k) = (\boldsymbol{x}_{t_1}, \boldsymbol{x}_{t_2}, \dots, \boldsymbol{x}_{t_{|T|}})$. The elements of $\mathbb{R}^k$ can be partially ordered by the $\leq$-relation on vectors. For vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ with $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^k$ we write $\boldsymbol{x} \leq \boldsymbol{y}$ if all components of $\boldsymbol{x}$ are less or equal to the corresponding components of $\boldsymbol{y}$. For a set $S \subseteq \mathbb{R}^k$ let $\mathsf{span}(S)$ be the intersection of all linear subspaces of $\mathbb{R}^k$ that are supersets of S. This vector space is also called the *linear subspace of* $\mathbb{R}^k$ *spanned by* S. For a linear subspace V of $\mathbb{R}^k$ let $\dim(V)$ be the dimension of V. For a finite $S \subseteq \mathbb{Z}^k$ the *rational cone spanned by* S is

$\mathsf{cone}(S) = \{ \sum_{\boldsymbol{x_i} \in S} \lambda_i \cdot \boldsymbol{x_i} \mid \lambda_i \in \mathbb{R}_{\geq 0} \} \subseteq \mathbb{R}^k$. A *linearly independent rational cone* in $\mathbb{R}^k$ is a set of the form $\mathsf{cone}(S)$ for a linearly independent $S \subseteq \mathbb{Z}^k$. Each rational cone is a finite union of linearly independent rational cones, see for example [15].

For a $\boldsymbol{c} \in \mathbb{N}^k$ and a finite $P \subseteq \mathbb{N}^k$ let $\mathsf{L}(\boldsymbol{c}, P) = \{ \boldsymbol{c} + \sum_{\boldsymbol{x_i} \in P} \lambda_i \cdot \boldsymbol{x_i} \mid \lambda_i \in \mathbb{N} \}$ and $\mathsf{La}(\boldsymbol{c}, P) = \{ \boldsymbol{c} + \sum_{\boldsymbol{x_i} \in P} \lambda_i \cdot \boldsymbol{x_i} \mid \lambda_i \in \mathbb{Z} \} \cap \mathbb{N}^k$. By definition, $\mathsf{L}(\boldsymbol{c}, P) \subseteq \mathsf{La}(\boldsymbol{c}, P)$. The vector c is called the *constant vector* whereas the set P is called the set of *periods* of $\mathsf{L}(\boldsymbol{c}, P)$ and of $\mathsf{La}(\boldsymbol{c}, P)$. Sets of the form $\mathsf{L}(\boldsymbol{c}, P)$, for a $\boldsymbol{c} \in \mathbb{N}^k$ and a finite $P \subseteq \mathbb{N}^k$, are called *linear* subsets of $\mathbb{N}^k$, while sets of the form $\mathsf{La}(\boldsymbol{c}, P)$ are called *lattices*. A subset of $\mathbb{N}^k$ is said to be *semi-linear* if it is a finite union of linear subsets. For a $\boldsymbol{c} \in \mathbb{N}^k$, $n \geq 0$, and $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n} \in \mathbb{N}^k$ we have that $\mathsf{La}(\boldsymbol{c}, \{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}\})$ is equal to the set of all $\boldsymbol{y} \in \mathbb{N}^k$ such that there exists $\lambda_1, \mu_1, \lambda_2, \mu_2, \ldots, \lambda_n, \mu_n \in \mathbb{N}$ with $\boldsymbol{c} + \sum_{i=1}^{n} \lambda_i \boldsymbol{x_i} = \boldsymbol{y} + \sum_{i=1}^{n} \mu_i \boldsymbol{x_i}$, which is a Presburger set. Since the Presburger sets are exactly the semi-linear sets by [10], every lattice is semi-linear. In order to explain our definitions we give an example.

Example 1. Consider the vector $\boldsymbol{c} = (4, 4)$ and the period vectors $\boldsymbol{p_1} = (1, 2)$ and $\boldsymbol{p_2} = (2, 0)$. A graphical presentation of the linear set $\mathsf{L}(\boldsymbol{c}, P)$ with $P = \{\boldsymbol{p_1}, \boldsymbol{p_2}\}$ is given on the left of Fig. 1. The constant vector $\boldsymbol{c}$ is drawn as a dashed arrow and both periods $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ are depicted as solid arrows. The dots indicate the elements that belong to $\mathsf{L}(\boldsymbol{c}, P)$. The lattice $\mathsf{La}(\boldsymbol{c}, P)$ is drawn in the middle of Fig. 1. Again, the constant vector is dashed, while both periods are solid arrows. Since now integer coefficients are allowed, there are new elements compared to $\mathsf{L}(\boldsymbol{c}, P)$ that belong to $\mathsf{La}(\boldsymbol{c}, P)$. On the right of Fig. 1 it is shown that $\mathsf{La}(\boldsymbol{c}, P)$ can be written as a linear set by using the constant vector $\boldsymbol{0}$ and the three period vectors drawn as solid arrows, that is, $\mathsf{La}(\boldsymbol{c}, P) = \mathsf{L}(\boldsymbol{0}, \{\boldsymbol{p_1}, \boldsymbol{p_2}, \boldsymbol{p_3}\})$, where $\boldsymbol{p_3} = (0, 2)$. $\square$

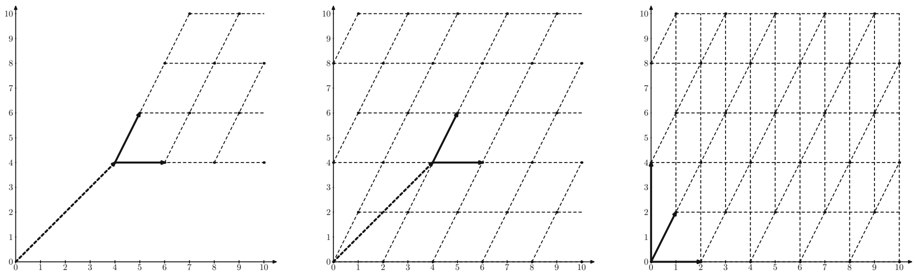

Fig. 1. The linear set $\mathsf{L}(\boldsymbol{c}, P)$ with $\boldsymbol{c} = (4, 4)$ and $P = \{\boldsymbol{p_1}, \boldsymbol{p_2}\}$, where $\boldsymbol{p_1} = (1, 2)$ and $\boldsymbol{p_2} = (2, 0)$ drawn on the left. The black dots indicate membership in $\mathsf{L}(\boldsymbol{c}, P)$. The lattice set $\mathsf{La}(\boldsymbol{c}, P)$ is depicted in the middle. Here the black dots refer to membership in $\mathsf{La}(\boldsymbol{c}, P)$. On the right a representation of $\mathsf{La}(\boldsymbol{c}, P)$ as a linear set is shown. The constant vector $\boldsymbol{0}$ is not shown and the period vectors are drawn as solid arrows.

An important result about semi-linear sets is that each semi-linear set can be written as a finite union of linear sets with linearly independent period sets [8]:

Theorem 2. *Let $k \geq 0$ and $S \subseteq \mathbb{N}^k$ be a semi-linear set. Then, there is $m \geq 0$, vectors $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^k$, and linearly independent sets $P_1, P_2, \ldots, P_m \subseteq \mathbb{N}^k$ such that $S = \bigcup_{i=1}^{m} \mathsf{L}(\boldsymbol{c_i}, P_i)$.*

Now, we recall some basic definitions from formal language theory. Let Σ be an alphabet. Then Σ^* is the set of all words over Σ, including the empty word λ. For a language $L \subseteq \Sigma^*$ define the set $\mathsf{perm}(L) = \cup_{w \in L} \mathsf{perm}(w)$, where $\mathsf{perm}(w) = \{ v \in \Sigma^* \mid v \text{ is a permutation of } w \}$. A language L is called *permutation closed* if $L = \mathsf{perm}(L)$. The length of a word $w \in \Sigma^*$ is denoted by $|w|$. For the number of occurrences of a symbol a in w we use the notation $|w|_a$. If Σ is the ordered alphabet $\Sigma = \{a_1, a_2, \ldots, a_k\}$, the *Parikh-mapping* $\psi : \Sigma^* \to \mathbb{N}^k$ is the function defined by $w \mapsto (|w|_{a_1}, |w|_{a_2}, \ldots, |w|_{a_k})$. The set $\psi(L)$ is called the *Parikh-image* of L. A language $L \subseteq \Sigma^*$ is called *semi-linear* if its Parikh-image $\psi(L)$ is a semi-linear set.

Let M be a monoid, i.e., a set with an associative binary operation and an identity element. For a subset $L \subseteq M$ let $\sim_L$ be the *Myhill-Nerode equivalence relation* on M. So, for two elements $v, w \in M$, we have $v \sim_L w$ if, for all $u \in M$, the equivalence $vu \in L \Leftrightarrow wu \in L$ holds. For $w \in M$, we call the equivalence class $[w]_{\sim_L}$ *positive* if $w \in L$. For $k \geq 0$ and $M = \mathbb{N}^k$ the equivalence relation $\sim_L$ will be written as $\equiv_L$, because that is the notation of this relation on $\mathbb{N}^k$ in [9]. If $L \subseteq \Sigma^*$ is a permutation closed language and $v, w \in L$ we have $v \sim_L w$ if and only if $\psi(v) \equiv_{\psi(L)} \psi(w)$. So, the language L is regular if and only if $\mathbb{N}^{|\Sigma|}/\equiv_{\psi(L)}$ is finite.

Let **REG**, **DCF**, **CF**, and **CS** be the families of regular, deterministic context-free, context-free, and context-sensitive languages. Moreover, we are interested in families of permutation closed languages. These language families are referred to by a prefix **p**. E.g., **pREG** denotes the language family of all permutation closed regular languages. Let **JFA** be the family of all languages accepted by jumping finite automata, see [13]. These are exactly the permutation closed semi-linear languages.

A *right one-way jumping finite automaton with multiple initial states* (MROWJFA) is a tuple $A = (Q, \Sigma, R, S, F)$, where Q is the *finite set of states*, Σ is the *finite input alphabet*, $\Sigma \cap Q = \emptyset$, R is a *partial function* from $Q \times \Sigma$ to Q, $S \subseteq Q$ is the *set of initial or start states*, and $F \subseteq Q$ is the *set of final states*. A *configuration* of A is a string in $Q\Sigma^*$. The *right one-way jumping relation*, symbolically denoted by $\circlearrowright_A$ or just $\circlearrowright$ if it is clear which MROWJFA we are referring to, over $Q\Sigma^*$ is defined as follows. Let $p, q \in Q$, $a \in \Sigma$, $w \in \Sigma^*$. If $R(p, a) = q$, then we have $paw \circlearrowright qw$. In case $R(p, a)$ is undefined, we get $paw \circlearrowright pwa$. So, the automaton jumps over a symbol, when it cannot be read. The *language accepted* by A is

$$L_R(A) = \{ w \in \Sigma^* \mid \exists s \in S, f \in F : sw \circlearrowright^* f \}.$$

We say that A *accepts* $w \in \Sigma^*$ if $w \in L_R(A)$ and that A *rejects* w otherwise. Let **MROWJ** be the family of all languages that are accepted by MROWJFAs.

Furthermore, in case the MROWJFA has a single initial state, i.e., $|S| = 1$, then we simply speak of a right one-way jumping automaton (ROWJFA) and refer to the family of languages accepted by ROWJFAs by **ROWJ**. Obviously, by definition we have $\mathbf{ROWJ} \subseteq \mathbf{MROWJ}$. We give an example of a ROWJFA:

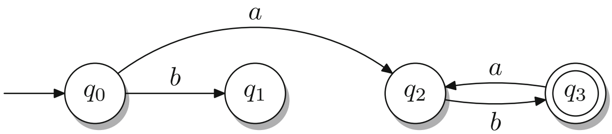

Fig. 2. The ROWJFA A.

Example 3. Let A be the ROWJFA $A = (\{q_0, q_1, q_2, q_3\}, \{a, b\}, R, q_0, \{q_3\})$, where the set R consists of the rules $q_0 b \to q_1$, $q_0 a \to q_2$, $q_2 b \to q_3$, and $q_3 a \to q_2$. The automaton A is depicted in Fig. 2. To show how ROWJFAs work, we give an example computation of A on the input $aabbba$:

$$q_0 aabbba \circlearrowright q_2 abbba \circlearrowright^2 q_3 bbaa \circlearrowright^3 q_2 abb \circlearrowright^2 q_3 ba \circlearrowright^2 q_2 b \circlearrowright q_3$$

That shows $aabbba \in L_R(A)$. Analogously, one can see that every word that contains the same number of a's and b's and that begins with an a is in $L_R(A)$. On the other hand, no other word can be accepted by A, interpreted as an ROWJFA. So, we get $L_R(A) = \{ w \in a\{a,b\}^* \mid |w|_a = |w|_b \}$. Notice that this language is non-regular and not closed under permutation. □

The following characterization of permutation closed languages accepted by ROWJFAs is known from [2].

Theorem 4. *Let L be a permutation closed language. Then, the language L is in* **pROWJ** *if and only if the Myhill-Nerode relation $\sim_L$ has only a finite number of positive equivalence classes.*

3 Lattices, Linear Sets, and Myhill-Nerode Classes

Because of Theorem 4 a permutation closed language is in **ROWJ** if and only if the Parikh-image has only a finite number of positive Myhill-Nerode equivalence classes. In this section we will study these kind of subsets of $\mathbb{N}^k$. Linear sets and lattices will play a key role in our theory. We will investigate decompositions of subsets of $\mathbb{N}^k$ as finite unions of such subsets that have only a finite number of positive equivalence classes. This will lead to characterization results about the language class **pMROWJ** in the next section.

3.1 Connections Between Linear Sets and Rational Cones

It was pointed out in [11] that "rational cones in $\mathbb{R}^d$ are important objects in toric algebraic geometry, combinatorial commutative algebra, geometric combinatorics, integer programming." In the following we will see how rational cones are related to the property of linear sets to have only a finite number of positive Myhill-Nerode equivalence classes. The following property is straightforward.

Lemma 5. *For $k \geq 0$, vectors $\boldsymbol{c}, \boldsymbol{d} \in \mathbb{N}^k$, and a finite set $P \subseteq \mathbb{N}^k$ the map $\mathsf{L}(\boldsymbol{c}, P) \rightarrow \mathsf{L}(\boldsymbol{d}, P)$ given by $\boldsymbol{x} \mapsto \boldsymbol{x} - \boldsymbol{c} + \boldsymbol{d}$ induces a bijection from $\mathsf{L}(\boldsymbol{c}, P)/\equiv_{\mathsf{L}(\boldsymbol{c},P)}$ to $\mathsf{L}(\boldsymbol{d}, P)/\equiv_{\mathsf{L}(\boldsymbol{d},P)}$.* □

Next, we define two properties of subsets of $\mathbb{N}^k$ which involve rational cones. Let $k \geq 0$ and $S \subseteq \mathbb{N}^k$. Then, the set S has the *linearly independent rational cone property* if $\mathsf{span}(S) \cap (\mathbb{R}_{\geq 0})^k = \mathsf{cone}(T)$, for some a linearly independent $T \subseteq \mathbb{N}^k$. The set S has the *own rational cone property* if S is finite and it holds $\mathsf{span}(S) \cap (\mathbb{R}_{\geq 0})^k = \mathsf{cone}(S)$.

A linear set has only a finite number of positive Myhill-Nerode equivalence classes if and only if the period set has the own rational cone property:

Theorem 6. *Let $k \geq 0$ and $P \subseteq \mathbb{N}^k$ be finite. Then, $\left|\mathsf{L}(\mathbf{0}, P)/\equiv_{\mathsf{L}(\mathbf{0},P)}\right| < \infty$ if and only if P has the own rational cone property.* □

For linear sets with linearly independent periods we even get a stronger equivalence than in Theorem 6:

Corollary 7. *For $k \geq 0$ and a linearly independent $P \subseteq \mathbb{N}^k$ the following three conditions are equivalent:*

1. $\left|\mathsf{L}(\mathbf{0}, P)/\equiv_{\mathsf{L}(\mathbf{0},P)}\right| < \infty$
2. $\left|\mathsf{L}(\mathbf{0}, P)/\equiv_{\mathsf{L}(\mathbf{0},P)}\right| = 1$
3. *The set P has the own rational cone property.* □

3.2 Decompositions of Lattices

Lattices defined as subsets of $\mathbb{R}^k$ play an important rule in geometry, group theory, and cryptography, see [6]. Our lattices defined as subsets of $\mathbb{N}^k$ are a natural way to extend "the pattern" of a linear set to $\mathbb{N}^k$. Using lattices we can give a characterization in which cases arbitrary subsets of $\mathbb{N}^k$ can be decomposed as a finite union of subsets with only a finite number of positive Myhill-Nerode classes in the next subsection. This result, in turn, will enable us to prove a characterization result about MROWJFAs in the next section. In this subsection, we will show some decomposition results about lattices: it will turn out that lattices can be decomposed as a finite union of linear sets which have only one positive Myhill-Nerode equivalence class. Since each semi-linear set is the finite union of linear sets with linearly independent period sets by Theorem 2, we will investigate in which cases lattices can even be decomposed as a finite union of linear sets that have linearly independent period sets and only one positive

Myhill-Nerode equivalence class (or only a finite number of positive Myhill-Nerode equivalence classes).

For $k \geq 0$, $\boldsymbol{c}, \boldsymbol{y} \in \mathbb{N}^k$, a finite $P \subseteq \mathbb{N}^k$, and $\boldsymbol{x} \in \mathsf{La}(\boldsymbol{c}, P)$ the vector $\boldsymbol{x}+\boldsymbol{y}$ is in $\mathsf{La}(\boldsymbol{c}, P)$ if and only if $\boldsymbol{y} \in \mathsf{La}(\boldsymbol{0}, P)$. This gives us that each lattice has only one positive Myhill-Nerode equivalence class. On the other hand, each lattice is a finite union of linear sets that have only one positive Myhill-Nerode equivalence class:

Proposition 8. *Let $k \geq 0$, $\boldsymbol{c} \in \mathbb{N}^k$, and $P \subseteq \mathbb{N}^k$ be finite. Then, there is a natural number $m > 0$, $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^k$, and a finite $Q \subseteq \mathbb{N}^k$ such that $\mathsf{La}(\boldsymbol{c}, P) = \bigcup_{i=1}^m \mathsf{L}(\boldsymbol{c_i}, Q)$ and $\left|\mathsf{L}(\boldsymbol{0}, Q)/\equiv_{\mathsf{L}(\boldsymbol{0},Q)}\right| = 1$.* □

The linearly independent rational cone property is connected to the property of lattices to be a finite union of linear sets that have linearly independent period sets and only finitely many positive Myhill-Nerode equivalence classes:

Theorem 9. *For $k \geq 0$, $\boldsymbol{c} \in \mathbb{N}^k$, and a finite $P \subseteq \mathbb{N}^k$ the following three conditions are equivalent:*

1. *There is an $m > 0$, vectors $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^k$, and linearly independent $Q_1, Q_2, \ldots, Q_m \subseteq \mathbb{N}^k$ such that $\mathsf{La}(\boldsymbol{c}, P) = \bigcup_{i=1}^m \mathsf{L}(\boldsymbol{c_i}, Q_i)$ and for all $i \in \{1, 2, \ldots, m\}$ it holds $\left|\mathsf{L}(\boldsymbol{0}, Q_i)/\equiv_{\mathsf{L}(\boldsymbol{0},Q_i)}\right| < \infty$.*
2. *There is an $m > 0$, vectors $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^k$, and a linearly independent $Q \subseteq \mathbb{N}^k$ such that $\mathsf{La}(\boldsymbol{c}, P) = \bigcup_{i=1}^m \mathsf{L}(\boldsymbol{c_i}, Q)$ and $\left|\mathsf{L}(\boldsymbol{0}, Q)/\equiv_{\mathsf{L}(\boldsymbol{0},Q)}\right| = 1$.*
3. *The set P has the linearly independent rational cone property.* □

Because of Theorem 9 it is worthwhile to investigate the linearly independent rational cone property more. Intuitively one might think that this property always holds, but it turns out that this is only the case in dimension $k \leq 3$:

Theorem 10. *Let $k \geq 0$ and $n \in \{0, 1, \ldots, k\}$. Then, the condition that each $S \subseteq \mathbb{N}^k$ with $\dim(\mathsf{span}(S)) = n$ has the linearly independent rational cone property holds if and only if $n \in \{0, 1, 2, k\}$.* □

Thus, for $k \geq 0$ and $n \in \{0, 1, \ldots, k\}$, the condition that for all vectors $\boldsymbol{c} \in \mathbb{N}^k$ and finite sets $P \subseteq \mathbb{N}^k$ with $\dim(\mathsf{span}(P)) = n$ we get a decomposition of the set $\mathsf{La}(\boldsymbol{c}, P)$ as in Theorem 9 is equivalent to the condition $n \in \{0, 1, 2, k\}$.

3.3 A Decomposition Result About Subsets of $\mathbb{N}^k$

Having the decompositions of lattices from Subsect. 3.2, we now turn to a decomposition result about arbitrary subsets of $\mathbb{N}^k$. To state the result, we will work with quasi lattices: let $k \geq 0$ and $S \subseteq \mathbb{N}^k$. The set S is a *quasi lattice* if there is a $\boldsymbol{y} \in \mathbb{N}^k$, an $m \geq 0$, vectors $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^k$, and finite subsets $P_1, P_2, \ldots, P_m \subseteq \mathbb{N}^k$ such that the set $\{\, \boldsymbol{z} \in S \mid \boldsymbol{z} \geq \boldsymbol{y} \,\}$ is equal to $\{\, \boldsymbol{z} \in \bigcup_{j=1}^m \mathsf{La}(\boldsymbol{c_j}, P_j) \mid \boldsymbol{z} \geq \boldsymbol{y} \,\}$.

We can identify a pattern of two linear sets formed by three vectors that gives a sufficient condition for the property of a subset of $\mathbb{N}^k$ to not be a quasi lattice:

Lemma 11. *Let $k \geq 0$ and $S \subseteq \mathbb{N}^k$ such that there are vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{N}^k$ with $\pi_{k,\{j\}}(\boldsymbol{v}) > 0$, for all $j \in \{1, 2, \ldots, k\}$ with $\mathsf{L}(\boldsymbol{u}, \{\boldsymbol{v}\}) \cap S = \emptyset$ and moreover $\mathsf{L}(\boldsymbol{u} + \boldsymbol{w}, \{\boldsymbol{v}, \boldsymbol{w}\}) \subseteq S$. Then, the set S is* not *a quasi lattice.* □

We call subsets of $\mathbb{N}^k$ that allow a pattern as in the above lemma anti-lattices: let $k \geq 0$ and $S \subseteq \mathbb{N}^k$. If there are vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{N}^k$ with $\pi_{k,\{j\}}(\boldsymbol{v}) > 0$ for all $j \in \{1, 2, \ldots, k\}$ so that $\mathsf{L}(\boldsymbol{u}, \{\boldsymbol{v}\}) \cap S = \emptyset$ and $\mathsf{L}(\boldsymbol{u} + \boldsymbol{w}, \{\boldsymbol{v}, \boldsymbol{w}\}) \subseteq S$, the set S is called an *anti-lattice.* semi-linear set is a quasi lattice if and only if it is not an anti-lattice:

Proposition 12. *Let $k \geq 0$ and $S \subseteq \mathbb{N}^k$ be a semi-linear set. Then, the set S is a quasi lattice if and only if S is not an anti-lattice.* □

It follows that each subset of $\mathbb{N}^k$ which has only a finite number of positive Myhill-Nerode equivalence classes is a quasi lattice:

Corollary 13. *For a $k \geq 0$ and a subset $S \subseteq \mathbb{N}^k$ with $|S/\equiv_S| < \infty$ the set S is a quasi lattice.* □

Quasi lattices are related to the property of a subset $S \subseteq \mathbb{N}^k$ to be a finite union of subsets of $\mathbb{N}^k$ that have only a finite number of positive Myhill-Nerode equivalence classes, which holds exactly if S is a finite union of linear sets that have only one positive Myhill-Nerode equivalence class:

Theorem 14. *For a $k \geq 0$ and a subset $S \subseteq \mathbb{N}^k$ the following three conditions are equivalent:*

1. *There is an $m \geq 0$ and subsets $S_1, S_2, \ldots, S_m \subseteq \mathbb{N}^k$ such that $S = \bigcup_{j=1}^{m} S_j$ and for each $j \in \{1, 2, \ldots, m\}$ we have $\left|S_j/\equiv_{S_j}\right| < \infty$.*
2. *There is an $m \geq 0$ and linear sets $L_1, L_2, \ldots, L_m \subseteq \mathbb{N}^k$ such that $S = \bigcup_{j=1}^{m} L_j$ and for each $j \in \{1, 2, \ldots, m\}$ we have $\left|L_j/\equiv_{L_j}\right| = 1$.*
3. *For all subsets $T \subseteq \{1, 2, \ldots, k\}$ and vectors $\boldsymbol{x} \in \mathbb{N}^{|T|}$ it holds that the set $\pi_{k,\{1,2,\ldots,k\}\setminus T}(\{\, \boldsymbol{z} \in S \mid \pi_{k,T}(\boldsymbol{z}) = \boldsymbol{x} \,\})$ is a quasi lattice.* □

In dimension $k \leq 3$ we can strengthen the second condition of Theorem 14, while we can weaken the third condition in dimension $k \leq 2$:

Corollary 15. *For a $k \in \{0, 1, 2, 3\}$ and a subset $S \subseteq \mathbb{N}^k$ the conditions from Theorem 14 are equivalent to the following condition. There is a number $m \geq 0$, vectors $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^k$, and linearly independent sets $P_1, P_2, \ldots, P_m \subseteq \mathbb{N}^k$ such that it holds $S = \bigcup_{j=1}^{m} \mathsf{L}(\boldsymbol{c_j}, P_j)$ and for each $j \in \{1, 2, \ldots, m\}$ we have $\left|\mathsf{L}(\boldsymbol{0}, P_j)/\equiv_{\mathsf{L}(\boldsymbol{0}, P_j)}\right| = 1$. For a $k \in \{0, 1, 2\}$ and a subset $S \subseteq \mathbb{N}^k$ the conditions from Theorem 14 are equivalent to the condition that S is a semi-linear set and a quasi lattice.* □

4 Right One-Way Jumping Finite Automata with Multiple Initial States

In this section we investigate MROWJFAs. To get results about these devices we use results from Subsect. 3.3.

4.1 Results for Arbitrary Alphabets

First, some basic properties are given. Directly from the definition of MROWJFAs we get that the unary languages in **MROWJ** are exactly the unary regular languages and that **MROWJ** consists exactly of the finite unions of languages from **ROWJ**. However, it is not clear that every language from **pMROWJ** is a finite union of languages from **pROWJ**. From [2] we know that a^* and the language $\{ w \in \{a, b\}^* \mid |w|_a = |w|_b \}$ are in **ROWJ**, but the union of these two sets is not in **ROWJ**. Together with the properties of **ROWJ** shown in [2] and [4], this gives us: we have **REG** $\subset$ **ROWJ** $\subset$ **MROWJ** and also **pREG** $\subset$ **pROWJ** $\subset$ **pMROWJ**. The family **MROWJ** is incomparable to **DCF** and to **CF**. Each language in **MROWJ** is semi-linear and contained in the complexity classes **DTIME**(n^2) and **DSPACE**(n). We get **pMROWJ** $\not\subseteq$ **CF** and **pMROWJ** $\subseteq$ **JFA** $\subset$ **pCS**. The letter-bounded languages contained in **MROWJ** are exactly the regular letter-bounded languages.

Now, we will study the language class **pMROWJ** in more detail. The foundation for this will be the next result.

Theorem 16. *The Parikh-image of each language in* **pMROWJ** *is a quasi lattice.* □

Because the Parikh-image of $\{ w \in \{a, b\}^* \mid |w|_a \neq |w|_b \}$ is an anti-lattice, this language is not in **MROWJ**. Thus, we have **pMROWJ** $\subset$ **JFA** and that the family **pMROWJ** is incomparable to **pDCF** and to **pCF**.

To get more detailed results about **pMROWJ**, we define the language operation of *disjoint quotient* of a language $L \subseteq \Sigma^*$ with a word $w \in \Sigma^*$ as follows:

$$\begin{aligned} L/^d w &= \{ v \in \Sigma^* \mid vw \in L, \forall a \in \Sigma : (|v|_a = 0 \vee |w|_a = 0) \} \\ &= (L/w) \cap \{ a \in \Sigma \mid |w|_a = 0 \}^*. \end{aligned}$$

From Theorem 4 we get that the family **pROWJ** is closed under the operations of quotient with a word and disjoint quotient with a word. Let Σ be an alphabet, $\Pi \subseteq \Sigma$, and $L \subseteq \Sigma^*$ be in **MROWJ**. Then, it is easy to see that $L \cap \Pi^*$ is also in **MROWJ**. Thus, we get that if **pMROWJ** is closed under the operation quotient with a word, then **pMROWJ** is also closed under disjoint quotient with a word.

Theorem 4 gives a characterization of the language class **pROWJ** in terms of the Myhill-Nerode relation. The next Corollary is a result in the same spirit for the language class **pMROWJ**. Theorems 4, 14, and 16 give us a characterization of all languages L for which each disjoint quotient of L with a word is contained in **pMROWJ**:

Corollary 17. *For an alphabet Σ and a permutation closed language $L \subseteq \Sigma^*$ the following conditions are equivalent:*

1. *For all $w \in \Sigma^*$ the language $L/^d w$ is in* **pMROWJ**.
2. *There is an $n \geq 0$ and $L_1, L_2, \ldots, L_n \subseteq \Sigma^*$ with $L_1, L_2, \ldots, L_n \in$* **pROWJ** *and $L = \bigcup_{i=1}^{n} L_i$.*

3. *There is an* $n \geq 0$ *and permutation closed languages* $L_1, L_2, \ldots, L_n \subseteq \Sigma^*$ *such that* $L = \bigcup_{i=1}^{n} L_i$ *and for all* $i \in \{1, 2, \ldots, n\}$ *the language* L_i *has only a finite number of positive Myhill-Nerode equivalence classes.*
4. *There is an* $m \geq 0$ *and linear sets* $L_1, L_2, \ldots, L_m \subseteq \mathbb{N}^{|\Sigma|}$ *such that* $\psi(L) = \bigcup_{j=1}^{m} L_j$ *and for each* $j \in \{1, 2, \ldots, m\}$ *we have* $\left|L_j / \equiv_{L_j}\right| = 1$.
5. *For all subsets* $T \subseteq \{1, 2, \ldots, |\Sigma|\}$ *and vectors* $\boldsymbol{x} \in \mathbb{N}^{|T|}$ *it holds that the set* $\pi_{|\Sigma|, \{1,2,\ldots,|\Sigma|\} \setminus T}\left(\{\, \boldsymbol{z} \in \psi(L) \mid \pi_{|\Sigma|, T}(\boldsymbol{z}) = \boldsymbol{x} \,\}\right)$ *is a quasi lattice.* □

For ternary alphabets we can weaken the first condition of the previous corollary, by the fact that the family **JFA** is closed under the operation of disjoint quotient, and strengthen its fourth condition by Corollary 15:

Corollary 18. *For an alphabet* Σ *with* $|\Sigma| = 3$ *and a permutation closed language* $L \subseteq \Sigma^*$ *the following two conditions are equivalent:*

1. *For all unary* $w \in \Sigma^*$ *the language* $L/^d w$ *is in* **pMROWJ**.
2. *There is a number* $m \geq 0$*, vectors* $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^3$*, and linearly independent sets* $P_1, P_2, \ldots, P_m \subseteq \mathbb{N}^3$ *such that* $\psi(L) = \bigcup_{j=1}^{m} \mathsf{L}(\boldsymbol{c_j}, P_j)$ *and for each* $j \in \{1, 2, \ldots, m\}$ *we have* $\left|\mathsf{L}(\boldsymbol{0}, P_j) / \equiv_{\mathsf{L}(\boldsymbol{0}, P_j)}\right| = 1$. □

From Theorem 4 and Corollary 17 we get that the condition that each language from **pMROWJ** is a finite union of languages from **pROWJ** is equivalent to the condition that the family **pMROWJ** is closed under the operation of quotient with a word and to the condition that the family **pMROWJ** is closed under the operation of disjoint quotient with a word.

Consider an alphabet Σ and a language $L \subseteq \Sigma^*$. If for all $w \in \Sigma^*$ the language $L/^d w$ is in **pMROWJ**, then the language L is contained in the complexity class **DTIME**(n), as the next result shows:

Lemma 19. *Let* Σ *be an alphabet,* $n > 0$*, and* $L_1, L_2, \ldots, L_n \subseteq \Sigma^*$ *be in* **pROWJ**. *Then, there is a one-way* $(n \cdot |\Sigma|)$*-head DFA with endmarker accepting* $\bigcup_{j=1}^{n} L_i$. □

4.2 Results for Binary Alphabets

Now, we will investigate **MROWJ** and related language families for binary alphabets. It turns out that for some problems we get different or stronger results than for arbitrary alphabets. From the next theorem it follows that for binary alphabets **pCF** = **JFA**, whereas for arbitrary alphabets it holds **pCF** ⊂ **JFA**.

Theorem 20. *Each permutation closed semi-linear language over a binary alphabet is accepted by a counter automaton.* □

For binary alphabets we have **pROWJ** ⊂ **pDCF**, while for arbitrary alphabets **pROWJ** is incomparable to **pDCF** and to **pCF**:

Proposition 21. *Each language over a binary alphabet in* **pROWJ** *is accepted by a realtime deterministic counter automaton.* □

For the family **pMROWJ** we get the following results. Notice that for arbitrary alphabets **pMROWJ** and **pCF** are incomparable.

Corollary 22. *For binary alphabets it holds that* **pROWJ** $\subset$ **pMROWJ***, that* **pMROWJ** *is incomparable to* **pDCF***, and that* **pMROWJ** $\subset$ **JFA** $=$ **pCF***.* □

If the languages do not need to be closed under permutation, we get for binary alphabets the same inclusion relations between **ROWJ**, **MROWJ**, and **DCF** as for arbitrary alphabets:

Lemma 23. *For binary alphabets* **ROWJ** $\subset$ **MROWJ***. The families* **ROWJ** *and* **MROWJ** *are both incomparable to* **DCF** *over binary alphabets.* □

Theorems 4 and 16, Proposition 12, and Corollary 15 imply a characterization of the languages in **pMROWJ** over a binary alphabet, which is stronger than the statement for arbitrary alphabets in Corollary 17, because we do not need to consider disjoint quotients of a language with a word here.

Corollary 24. *Let Σ be an alphabet with $|\Sigma| = 2$ and $L \subseteq \Sigma^*$ be a permutation closed language. Then, the following conditions are equivalent:*

1. *Language L is in* **pMROWJ***.*
2. *There is an $n \geq 0$ and $L_1, L_2, \ldots, L_n \subseteq \Sigma^*$ with $L_1, L_2, \ldots, L_n \in$* **pROWJ** *and $L = \bigcup_{i=1}^{n} L_i$.*
3. *There is an $n \geq 0$ and permutation closed languages $L_1, L_2, \ldots, L_n \subseteq \Sigma^*$ such that $L = \bigcup_{i=1}^{n} L_i$ and for all $i \in \{1, 2, \ldots, n\}$ the language L_i has only a finite number of positive Myhill-Nerode equivalence classes.*
4. *There is a number $m \geq 0$, vectors $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_m} \in \mathbb{N}^2$, and linearly independent sets $P_1, P_2, \ldots, P_m \subseteq \mathbb{N}^2$ such that $\psi(L) = \bigcup_{j=1}^{m} \mathsf{L}(\boldsymbol{c_j}, P_j)$ and for each $j \in \{1, 2, \ldots, m\}$ we have $\left|\mathsf{L}(\boldsymbol{0}, P_j)/\equiv_{\mathsf{L}(\boldsymbol{0}, P_j)}\right| = 1$.*
5. *The Parikh-image of L is a semi-linear set and a quasi lattice.*
6. *The Parikh-image of L is a semi-linear set and not an anti-lattice.* □

From Corollary 24 it follows that each language from **pMROWJ** over a binary alphabet is a finite union of permutation closed languages accepted by a realtime deterministic counter automaton.

5 Conclusions

We have investigated ROWJFAs with multiple initial states and showed inclusion and incomparability results of the induced language family by using results on semi-linear sets and generalizations thereof. In order to complete the picture of these new language family it remains to study closure properties and decision problems for these devices and moreover to investigate nondeterministic variants of ROWJFAs in general.

References

1. Beier, S., Holzer, M.: Decidability of right one-way jumping finite automata. In: Hoshi, M., Seki, S. (eds.) DLT 2018. LNCS, vol. 11088, pp. 109–120. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98654-8_9
2. Beier, S., Holzer, M.: Properties of right one-way jumping finite automata. In: Konstantinidis, S., Pighizzini, G. (eds.) DCFS 2018. LNCS, vol. 10952, pp. 11–23. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94631-3_2
3. Beier, S., Holzer, M., Kutrib, M.: On the descriptional complexity of operations on semilinear sets. In: Csuhaj-Varjú, E., Dömösi, P., Vaszil, G. (eds.) AFL 2017. EPTCS, vol. 252, pp. 41–55. Debrecen, Hungary (2017). https://doi.org/10.4204/EPTCS.252.8
4. Chigahara, H., Fazekas, S., Yamamura, A.: One-way jumping finite automata. Int. J. Found. Comput. Sci. **27**(3), 391–405 (2016). https://doi.org/10.1142/S0129054116400165
5. Chistikov, D., Haase, C.: The taming of the semi-linear set. In: Chatzigiannakis, I., Mitzenmacher, M., Rabani, Y., Sangiorgi, D. (eds.) ICALP 2016. LIPIcs, vol. 55, pp. 128:1–128:13 (2016). https://doi.org/10.4230/LIPIcs.ICALP.2016.128
6. Conway, J., Horton, S., Neil, J.A.: Sphere Packings, Lattices and Groups, Grundlehren der Mathematischen Wissenschaften, vol. 290, 3rd edn. Springer, New York (1999). https://doi.org/10.1007/978-1-4757-6568-7
7. Gill, A., Kou, L.T.: Multiple-entry finite automata. J. Comput. System Sci. **9**(1), 1–19 (1974). https://doi.org/10.1016/S0022-0000(74)80034-6
8. Ginsburg, S., Spanier, E.H.: Bounded ALGOL-like languages. Trans. Am. Math. Soc. **113**, 333–368 (1964). https://doi.org/10.2307/1994067
9. Ginsburg, S., Spanier, E.H.: Bounded regular sets. Proc. Am. Math. Soc. **17**(5), 1043–1049 (1966). https://doi.org/10.1090/S0002-9939-1966-0201310-3
10. Ginsburg, S., Spanier, E.H.: Semigroups, Presburger formulas, and languages. Pac. J. Math. **16**(2), 285–296 (1966). https://doi.org/10.2140/pjm.1966.16.285
11. Gubeladze, J., Michalek, M.: The poset of rational cones. Pac. J. Math. **292**(1), 103–115 (2018). https://doi.org/10.2140/pjm.2018.292.103
12. Haase, C., Hofman, P.: Tightening the complexity of equivalence problems for commutative grammars. In: Ollinger, N., Vollmer, H. (eds.) STACS 2016. LIPIcs, vol. 47, pp. 41:1–41:14 (2016). https://doi.org/10.4230/LIPIcs.STACS.2016.41
13. Meduna, A., Zemek, P.: Jumping finite automata. Int. J. Found. Comput. Sci. **23**(7), 1555–1578 (2012). https://doi.org/10.1142/S0129054112500244
14. Parikh, R.J.: On context-free languages. J. ACM **13**(4), 570–581 (1966). https://doi.org/10.1145/321356.321364
15. Studený, M.: Convex cones in finite-dimensional real vector spaces. Kybernetika **29**(2), 180–200 (1993). http://www.kybernetika.cz/content/1993/2/180