THE GENERAL VECTOR ADDITION SYSTEM REACHABILITY PROBLEM BY PRESBURGER INDUCTIVE INVARIANTS [*]

JÉRÔME LEROUX

Laboratoire Bordelais de Recherche en Informatique, CNRS, Talence, France
e-mail address: leroux@labri.fr

ABSTRACT. The reachability problem for Vector Addition Systems (VASs) is a central problem of net theory. The general problem is known to be decidable by algorithms exclusively based on the classical Kosaraju-Lambert-Mayr-Sacerdote-Tenney decomposition. This decomposition is used in this paper to prove that the Parikh images of languages recognized by VASs are *semi-pseudo-linear*; a class that extends the semi-linear sets, a.k.a. the sets definable in Presburger arithmetic. We provide an application of this result; we prove that a final configuration is not reachable from an initial one if and only if there exists a semi-linear inductive invariant that contains the initial configuration but not the final one. Since we can decide if a Presburger formula denotes an inductive invariant, we deduce that there exist checkable certificates of non-reachability. In particular, there exists a simple algorithm for deciding the general VAS reachability problem based on two semi-algorithms. A first one that tries to prove the reachability by enumerating finite sequences of actions and a second one that tries to prove the non-reachability by enumerating Presburger formulas.

1. INTRODUCTION

Vector Addition Systems (VASs) or equivalently Petri Nets are one of the most popular formal methods for the representation and the analysis of parallel processes [2]. The reachability problem is central since many computational problems (even outside the parallel processes) reduce to the reachability problem. Sacerdote and Tenney provided in [10] a partial proof of decidability of this problem. The proof was completed in 1981 by Mayr [7] and simplified by Kosaraju [5] from [10, 7]. Ten years later, Lambert[6] provided a more simplified version based on [5]. This last proof still remains difficult and the upper bound complexity of the corresponding algorithm is just known to be non-primitive recursive. Nowadays, it is an open problem wether an elementary upper complexity bound for this problem exists. In fact, the known general reachability algorithms are exclusively based on the Kosaraju-Lambert-Mayr-Sacerdote-Tenney (KLMST) decomposition.

In this paper, by using the KLMST decomposition we prove that the Parikh images of languages recognized by VASs are semi-pseudo-linear, a class that extends the semi-linear

sets, a.k.a. the sets definable in Presburger arithmetic [3]. We provide an application of this result; we prove that a final configuration is not reachable from an initial one if and only if there exists a forward inductive invariant definable in Presburger arithmetic that contains the initial configuration but not the final one. Since we can decide if a Presburger formula denotes a forward inductive invariant, we deduce that there exist checkable certificates of non-reachability. In particular, there exists a simple algorithm for deciding the general VAS reachability problem based on two semi-algorithms. A first one that proves the reachability by enumerating finite sequences of actions denoting a path from the initial configuration to the final one and a second one that proves the non-reachability by enumerating Presburger formulas denoting inductive invariant containing the initial configuration but not the final one.

Outline of the paper: Section 2 introduces the class of *Vector Addition Systems (VASs)*. Section 3 recalls the class of *Marked Reachability Graph Sequences (MRGSs)* and the KLMST decomposition of languages recognized by VASs into finite unions of languages recognized by *perfect MRGSs*. Semi-pseudo-linear sets are introduced in Section 4. In Section 5, Parikh images of languages recognized by perfect MRGSs are proved to be pseudo-linear. In Section 6 we introduce the class of *Petri sets* a subclass of the semi-pseudo-linear sets stable by intersection with every semi-linear set. Reachability sets of VASs from semi-linear sets are proved to be Petri sets in this section. In Section 7 we study approximations of two pseudo-linear sets with an empty intersection. Finally in Section 8 we deduce that if a final configuration is not reachable from an initial one, there exists a forward inductive invariant definable in Presburger arithmetic that contains the initial configuration but not the final one.

2. Vector Addition Systems

We denote by $\mathbb{Q}, \mathbb{Q}_+, \mathbb{Z}, \mathbb{N}$, respectively, the set of rational values, non-negative rational values, the set of integers and the set of non-negative integers. The *components of a vector* $\mathbf{x} \in \mathbb{Q}^n$ are denoted by $(\mathbf{x}[1], \ldots, \mathbf{x}[n])$. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x} \in \mathbb{Q}^n$ and $r \in \mathbb{Q}$. The sum $\mathbf{x}_1 + \mathbf{x}_2$ and the product $r\mathbf{x}$ are naturally defined component wise. Given a function $f : E \to F$ where E, F are sets, we denote by $f(X) = \{f(x) \mid x \in X\}$ for every subset $X \subseteq E$. This definition naturally defines sets $X_1 + X_2$ and RX where $X_1, X_2, X \subseteq \mathbb{Q}^n$ and $R \subseteq \mathbb{Q}$. With slight abuse of notation, $\{\mathbf{x}_1\} + X_2$, $X_1 + \{\mathbf{x}_2\}$, $\{r\}X$ and $R\{\mathbf{x}\}$ are simply denoted by $\mathbf{x}_1 + X_2$, $X_1 + \mathbf{x}_2$, rX and $R\mathbf{x}$.

The lattice $(\mathbb{N}, \leq)$ is completed with an additional element $\top$ such that $k \leq \top$ for every $k \in \mathbb{N} \cup \{\top\}$. The set $\mathbb{N} \cup \{\top\}$ is denoted by $\mathbb{N}_\top$. Given a non-decreasing sequence $(x_i)_{i \geq 0}$ in $(\mathbb{N}_\top, \leq)$ we denote by $\lim_{i \to +\infty}(x_i)$ the *least upper bound* in $\mathbb{N}_\top$. The $\top$ element is interpreted as a "don't care value" by introducing the partial order $\trianglelefteq$ over $\mathbb{N}_\top$ defined by $x_1 \trianglelefteq x_2$ if and only if $x_1 = x_2$ or $x_2 = \top$. Orders $\leq$ and $\trianglelefteq$ are extended component-wise over $\mathbb{N}_\top^n$. The set of minimal elements for $\leq$ of a set $X \subseteq \mathbb{N}^n$ is denoted by $\min(X)$. As $(\mathbb{N}^n, \leq)$ is a *well partially ordered set*, the set $\min(X)$ is finite and $X \subseteq \min(X) + \mathbb{N}^n$ for every $X \subseteq \mathbb{N}^n$.

An *alphabet* is a non-empty finite set Σ. The set of words over Σ is denoted by Σ^*. The empty word is denoted by ϵ. The concatenation of two words σ_1 and σ_2 is simply denoted by $\sigma_1 \sigma_2$. The concatenation of $r \geq 1$ times a word σ is denoted by σ^r. By definition $\sigma^0 = \epsilon$. The number of occurrences of an element $a \in \Sigma$ in a word $\sigma \in \Sigma^*$ is denoted by $|\sigma|_a$. The *Parikh image* of a word σ over Σ is the function $\|\sigma\|_\Sigma : \Sigma \to \mathbb{N}$ defined by $\|\sigma\|_\Sigma(a) = |\sigma|_a$

for every $a \in \Sigma$. This function is simply denoted by $||\sigma||$ when Σ is known without every ambiguity. The *Parikh image* $||\mathcal{L}||$ of a language $\mathcal{L} \subseteq \Sigma^*$ is defined as the set of functions $||\sigma||$ over the words $\sigma \in \mathcal{L}$.

A *Vector Addition System (VAS)* is a tuple $\mathcal{V} = (\Sigma, n, \delta)$ where Σ is an alphabet, $n \in \mathbb{N}$ is the *dimension*, and $\delta : \Sigma \to \mathbb{Z}^n$ is a *displacement function*. In the sequel, such a functions is naturally extended to a function $\delta : \Sigma^* \to \mathbb{Z}^n$ satisfying $\delta(\epsilon) = \mathbf{0}$ and $\delta(\sigma) = \sum_{i=1}^{k} \delta(a_i)$ for every word $\sigma = a_1 \ldots a_k$ of $k \geq 1$ elements $a_i \in \Sigma$. A *configuration* is a vector in $\mathbb{N}^n$ and an *extended configuration* is a vector in $\mathbb{N}_\top^n$. For $a \in \Sigma$, the binary relation $\xrightarrow{a}_\mathcal{V}$ is defined over the set of extended configurations by $\mathbf{x} \xrightarrow{a}_\mathcal{V} \mathbf{x}'$ if and only if $\mathbf{x}' = \mathbf{x} + \delta(a)$ with $\top + z = \top$ by definition for every $z \in \mathbb{Z}$. Let $k \geq 1$. Given a word $\sigma = a_1 \ldots a_k$ of elements $a_i \in \Sigma$, we denote by $\xrightarrow{\sigma}_\mathcal{V}$ the concatenation $\xrightarrow{a_1}_\mathcal{V} \cdots \xrightarrow{a_k}_\mathcal{V}$. By definition $\xrightarrow{\epsilon}_\mathcal{V}$ is the identity binary relation over the set of extended configurations. We denote by $\xrightarrow{*}_\mathcal{V}$ the *reachability binary relation* over the set of extended configurations defined by $\mathbf{x} \xrightarrow{*}_\mathcal{V} \mathbf{x}'$ if and only if there exists $\sigma \in \Sigma^*$ such that $\mathbf{x} \xrightarrow{\sigma}_\mathcal{V} \mathbf{x}'$. Observe that in this case $\mathbf{x}[i] = \top$ if and only if $\mathbf{x}'[i] = \top$. Intuitively the $\top$ element provides a simple way to get rid of some components of a VAS since these components remain equal to $\top$.

Definition 2.1. *The* reachability problem *for a tuple* $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ *where* $(\mathbf{s}, \mathbf{s}')$ *are two configurations of a VAS* $\mathcal{V}$ *consists in deciding if* $\mathbf{s} \xrightarrow{*}_\mathcal{V} \mathbf{s}'$.

Let $\mathbf{m}, \mathbf{m}'$ be two extended configurations. The *language recognized* by $(\mathbf{m}, \mathcal{V}, \mathbf{m}')$ is the set $\mathcal{L}(\mathbf{m}, \mathcal{V}, \mathbf{m}') = \{\sigma \in \Sigma^* \mid \exists \mathbf{s}, \mathbf{s}' \in \mathbb{N}^n \quad \mathbf{s} \trianglelefteq \mathbf{m} \wedge \mathbf{s} \xrightarrow{\sigma}_\mathcal{V} \mathbf{s}' \wedge \mathbf{s}' \trianglelefteq \mathbf{m}'\}$. Given two sets S, S' of configurations, the set $\mathrm{post}_\mathcal{V}^*(S)$ of *reachable configurations from* S and the set $\mathrm{pre}_\mathcal{V}^*(S')$ of *co-reachable configurations from* S' are formally defined by:

$$\mathrm{post}_\mathcal{V}^*(S) = \{\mathbf{s}' \in \mathbb{N}^n \mid \exists \mathbf{s} \in S \quad \mathbf{s} \xrightarrow{*}_\mathcal{V} \mathbf{s}'\}$$

$$\mathrm{pre}_\mathcal{V}^*(S') = \{\mathbf{s} \in \mathbb{N}^n \mid \exists \mathbf{s}' \in S' \quad \mathbf{s} \xrightarrow{*}_\mathcal{V} \mathbf{s}'\}$$

Example 2.2. *A VAS* $\mathcal{V} = (\Sigma, n, \delta)$ *with* $\Sigma = \{a, b\}$, $n = 2$, $\delta(a) = (1, 1)$ *and* $\delta(b) = (-1, -2)$ *is depicted in Figure 1. Observe that* $\mathbf{s} \xrightarrow{a^4 b^3}_\mathcal{V} \mathbf{s}'$ *with* $\mathbf{s} = (0, 2)$ *and* $\mathbf{s}' = (1, 0)$. *Note that* $\mathrm{post}_\mathcal{V}^*(\{\mathbf{s}\}) = \{\mathbf{x} \in \mathbb{N}^2 \mid \mathbf{x}[2] \leq \mathbf{x}[1] + 2\}$ *and* $\mathrm{pre}_\mathcal{V}^*(\{\mathbf{s}'\}) = \{\mathbf{x} \in \mathbb{N}^2 \mid \mathbf{x}[2] \geq 2(\mathbf{x}[1] - 1)\}$.
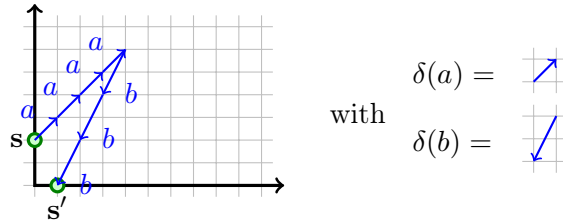

Figure 1: A Vector Addition System.

A *graph* is a tuple $G = (Q, \Sigma, T)$ where Q is a finite set of *states*, Σ is an alphabet, $T \subseteq Q \times \Sigma \times Q$ is a finite set of *transitions*. A *path* π is a word $\pi = t_1 \ldots t_k$ of $k \in \mathbb{N}$ transitions $t_i \in T$ such that there exists $q_0, \ldots, q_k \in Q$ and there exists $a_1, \ldots, a_k \in \Sigma$ such that $t_i = (q_{j-1}, a_j, q_j)$ for every $1 \leq j \leq k$. In this case we say that π *is a path labeled by* $\sigma = a_1 \ldots a_k$ *from* q_0 *to* q_k. In the sequel we denote by $q_0 \xrightarrow{\sigma}_G q_k$ such a path π. If the states

q_0 and q_k are equal, the path π is called a *cycle on this state*. As usual a graph is said to be *strongly connected* if for every pair of states $(q, q') \in Q \times Q$, there exists a path from q to q'.

Remark 2.3. *A* Vector Addition System with States (VASS) *is a tuple* $(Q, \Sigma, T, n, \delta)$ *where* $G = (Q, \Sigma, T)$ *is a graph and* $\mathcal{V} = (\Sigma, n, \delta)$ *is a VAS. A pair in* $Q \times \mathbb{N}^n$ *is called a* VASS configuration. *Let* $\sigma \in \Sigma^*$. *The VASS semantics is defined over the VASS configurations by* $(q, \mathbf{s}) \xrightarrow{\sigma} (q', \mathbf{s}')$ *if and only if* $q \xrightarrow{\sigma}_G q'$ *and* $\mathbf{s} \xrightarrow{\sigma}_\mathcal{V} \mathbf{s}'$. *Note* [4] *that* n*-dim VASSs can be simulated by* $(n + 3)$*-dim VASs.*

Example 2.4. *Recall* [4] *that sets* $\mathrm{post}^*_\mathcal{V}(S)$ *and* $\mathrm{pre}^*_\mathcal{V}(S')$ *are definable in Presburger arithmetic* $\mathrm{FO}\,(\mathbb{N}, +, \leq)$ *if* S *and* S' *are definable in this logic and* $n \leq 5$. *Moreover from* [4] *we deduce an example of a* 6*-dim VAS* $\mathcal{V}$ *and a pair of configurations* $(\mathbf{s}, \mathbf{s}') \notin \xrightarrow{*}_\mathcal{V}$ *such that neither* $\mathrm{post}^*_\mathcal{V}(\{\mathbf{s}\})$ *nor* $\mathrm{pre}^*_\mathcal{V}(\{\mathbf{s}'\})$ *are definable in Presburger arithmetic. This example is obtained by considering the VASS depicted in Figure 2. This VASS has a loop on state* p *and another loop on state* q. *Intuitively, iterating the loop on state* p *transfers the content of the first counter to the second counter whereas iterating the loop on state* q *transfers and multiplies by two the content of the second counter to the first counter. The third counter is incremented each time we come back to state* p *from* q. *In* [4] *the set of reachable configurations from* $(p, (1, 0, 0))$ *is proved equal to* $(\{p\} \times \{\mathbf{x} \in \mathbb{N}^3 \mid \mathbf{x}[1] + \mathbf{x}[2] \leq 2^{\mathbf{x}[3]}\}) \cup (\{q\} \times \{\mathbf{x} \in \mathbb{N}^3 \mid \mathbf{x}[1] + 2\mathbf{x}[2] \leq 2^{\mathbf{x}[3]+1}\})$. *This set is not definable in Presburger arithmetic.*

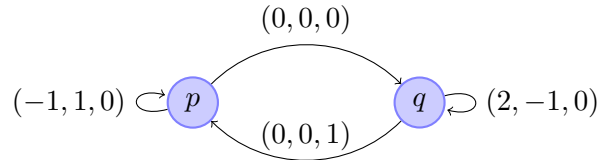

Figure 2: A VASS taken from [4].

3. THE KLMST DECOMPOSITION

The emptiness of $\mathcal{L}(\mathbf{m}, \mathcal{V}, \mathbf{m}')$ can be decided with the *Kosaraju-Lambert-Mayr-Sacerdote-Tenney (KLMST)* decomposition. This decomposition shows that $\mathcal{L}(\mathbf{m}, \mathcal{V}, \mathbf{m}')$ is effectively decomposable as a finite union $\bigcup_{\mathcal{U} \in F} \mathcal{L}(\mathcal{U})$ where $\mathcal{L}(\mathcal{U})$ is the language recognized by a *perfect Marked Reachability Graph Sequence (MRGS)* $\mathcal{U}$. We provide in Section 3.1 a new definition of *perfect MRGS* that does not require complex constructions. This definition is proved equivalent to the original one [6] in Section 3.2. Finally in Section 3.3 we recall the KLMST decomposition.

3.1. **The Perfect MRGSs.** In this section we introduce the class of *Marked Reachability Graph Sequences (MRGSs)* by following notations introduced by Lambert [6]. We also provide a new definition for the class of MRGSs said to be *perfect* [6].

A *reachability graph* for a VAS $\mathcal{V} = (\Sigma, n, \delta)$ is a graph $G = (Q, \Sigma, T)$ with $Q \subseteq \mathbb{N}^n_\top$ and $T \subseteq \{(q, a, q') \in Q \times \Sigma \times Q \mid q \xrightarrow{a}_\mathcal{V} q'\}$. A *marked reachability graph* $\mathcal{M} = (\mathbf{m}, \mathbf{x}, G, \mathbf{x}', \mathbf{m}')$ for $\mathcal{V}$ is a strongly connected reachability graph G for $\mathcal{V}$ equipped with two extended configurations $\mathbf{x}, \mathbf{x}' \in Q$ respectively called the *input state* and the *output*

state, and equipped with two extended configurations $\mathbf{m}, \mathbf{m}'$ satisfying $\mathbf{m} \trianglelefteq \mathbf{x}$ and $\mathbf{m}' \trianglelefteq \mathbf{x}'$ respectively called the *input constraint* and the *output constraint*. An *accepted tuple* for $\mathcal{M}$ is a tuple $(\mathbf{s}, \pi, \mathbf{s}')$ where $\pi = (\mathbf{x} \xrightarrow{\sigma}_G \mathbf{x}')$ is a path in G labeled by σ from the input state $\mathbf{x}$ to the output state $\mathbf{x}'$ and where $\mathbf{s}, \mathbf{s}' \in \mathbb{N}^n$ are two configurations such that $\mathbf{s} \trianglelefteq \mathbf{m}$, $\mathbf{s} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{s}'$ and $\mathbf{s}' \trianglelefteq \mathbf{m}'$. Intuitively the graph G and the input/output states enforce σ to label a path in G from $\mathbf{x}$ to $\mathbf{x}'$. The *input/output constraints* enforce $\mathbf{s}[i]$ and $\mathbf{s}'[i]$ to be equal to $\mathbf{m}[i]$ and $\mathbf{m}'[i]$, respectively, if $\mathbf{m}[i]$ and $\mathbf{m}'[i]$ are not equal to the "don't care value" $\top$.

A *marked reachability graph sequence (MRGS)* for $(\mathbf{m}, \mathcal{V}, \mathbf{m}')$ is a sequence

$$\mathcal{U} = \mathcal{M}_0 a_1 \mathcal{M}_1 \ldots a_k \mathcal{M}_k$$

that alternates elements $a_j \in \Sigma$ and marked reachability graphs $\mathcal{M}_j = (\mathbf{m}_j, \mathbf{x}_j, G_j, \mathbf{x}'_j, \mathbf{m}'_j)$ with $G_j = (Q_j, \Sigma, T_j)$ such that $\mathbf{m}_0 \trianglelefteq \mathbf{m}$ and $\mathbf{m}'_k \trianglelefteq \mathbf{m}'$. An *accepted sequence* for $\mathcal{U}$ is a sequence $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_{0 \leq j \leq k}$ such that $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)$ is an accepted tuple for $\mathcal{M}_j$ for every $0 \leq j \leq k$ and such that $\mathbf{s}'_{j-1} \xrightarrow{a_j}_{\mathcal{V}} \mathbf{s}_j$ for every $1 \leq j \leq k$. The *language recognized* by $\mathcal{U}$ is the set of words of the form $\sigma = \sigma_0 a_1 \sigma_1 \ldots a_k \sigma_k$ such that there exists an accepted sequence $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_{0 \leq j \leq k}$ where π_j is labeled by σ_j. This set is denoted by $\mathcal{L}(\mathcal{U})$. Since $\mathbf{m}_0 \trianglelefteq \mathbf{m}$ and $\mathbf{m}'_k \trianglelefteq \mathbf{m}'$, relations $\mathbf{s}_0 \trianglelefteq \mathbf{m}_0$ and $\mathbf{s}'_k \trianglelefteq \mathbf{m}'_k$ imply $\mathbf{s}_0 \trianglelefteq \mathbf{m}$ and $\mathbf{s}'_k \trianglelefteq \mathbf{m}'$. In particular the inclusion $\mathcal{L}(\mathcal{U}) \subseteq \mathcal{L}(\mathbf{m}, \mathcal{V}, \mathbf{m}')$ holds.

Example 3.1. *Let $\mathcal{V} = (\Sigma, n, \delta)$ be a VAS and let $(\mathbf{s}, \mathbf{s}')$ be a pair of configurations of $\mathcal{V}$. Let us introduce an MRGS $\mathcal{U}$ such that $\mathcal{L}(\mathcal{U}) = \mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$. We consider the graph $G = (Q, \Sigma, T)$ where $Q = \{(\top, \ldots, \top)\}$ and $T = Q \times \Sigma \times Q$, and the marked reachability graph $\mathcal{M} = (\mathbf{s}, (\top, \ldots, \top), G, (\top, \ldots, \top), \mathbf{s}')$. Now just observe that the MRGS $\mathcal{U} = \mathcal{M}$ satisfies $\mathcal{L}(\mathcal{U}) = \mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$.*

Definition 3.2. *An MRGS $\mathcal{U}$ is said to be* perfect *if for every $c \in \mathbb{N}$, there exists an accepted sequence $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_{0 \leq j \leq k}$ for $\mathcal{U}$ such that for every $0 \leq j \leq k$:*

- $\mathbf{s}_j[i] \geq c$ *for every i such that $\mathbf{m}_j[i] = \top$,*
- *there exists a prefix $\mathbf{x}_j \xrightarrow{w_j}_{G_j} \mathbf{x}_j$ of π_j and a configuration $\mathbf{r}_j$ such that $\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}} \mathbf{r}_j$ and such that $\mathbf{r}_j[i] \geq c$ for every i such that $\mathbf{x}_j[i] = \top$, and*
- $|\pi_j|_t \geq c$ *for every $t \in T_j$,*
- *there exists a suffix $\mathbf{x}'_j \xrightarrow{w_j}{}'_{G_j} \mathbf{x}'_j$ of π_j and a configuration $\mathbf{r}'_j$ such that $\mathbf{r}'_j \xrightarrow{w_j}{}'_{\mathcal{V}} \mathbf{s}'_j$ and such that $\mathbf{r}'_j[i] \geq c$ for every i such that $\mathbf{x}'_j[i] = \top$,*
- $\mathbf{s}'_j[i] \geq c$ *for every i such that $\mathbf{m}'_j[i] = \top$.*

3.2. **Original perfect condition.** The perfect condition given in Definition 3.2 is proved equivalent to the original one [6]. The original definition requires additional notions recalled in this section. These results are also used in Section 5 to establish the pseudo-linearity of Parikh images of language recognized by perfect MRGSs.

Let $\mathcal{M} = (\mathbf{m}, \mathbf{x}, G, \mathbf{x}', \mathbf{m}')$ be a marked reachability graph. We say that $\mathcal{M}$ satisfies the *input loop condition* if there exists a sequence $(\mathbf{x} \xrightarrow{w_c}_G \mathbf{x})_c$ of cycles and a non-decreasing sequence $(\mathbf{m}_c)_c$ of extended configurations such that $\mathbf{m} \xrightarrow{w_c}_{\mathcal{V}} \mathbf{m}_c$ for every c and $\lim_{c \to +\infty} \mathbf{m}_c = \mathbf{x}$. Symmetrically, we say that $\mathcal{M}$ satisfies the *output loop condition* if there exists a sequence $(\mathbf{x}' \xrightarrow{w_c}{}'_G \mathbf{x}')_c$ of cycles and a non-decreasing sequence $(\mathbf{m}'_c)_c$ of extended configurations such that $\mathbf{m}'_c \xrightarrow{w_c}{}'_{\mathcal{V}} \mathbf{m}'$ for every c and $\lim_{c \to +\infty} \mathbf{m}'_c = \mathbf{x}'$. The following

Lemma 3.3 and Lemma 3.4 show that these conditions are in EXPSPACE since they reduce to *covering problems* [9].

Lemma 3.3. *The input loop condition is satisfied by $\mathcal{M}$ iff there exist a cycle $\mathbf{x} \xrightarrow{w}_G \mathbf{x}$ and an extended configuration $\mathbf{y}$ satisfying $\mathbf{m} \xrightarrow{w}_{\mathcal{V}} \mathbf{y}$ and satisfying $\mathbf{y}[i] > \mathbf{m}[i]$ for every i such that $\mathbf{m}[i] < \mathbf{x}[i]$.*

Proof. Assume first that $\mathcal{M}$ satisfies the input loop condition. There exist a sequence $(\mathbf{x} \xrightarrow{w_c}_G \mathbf{x})_c$ of cycles and a non-decreasing sequence $(\mathbf{m}_c)_c$ of extended configurations such that $\mathbf{m} \xrightarrow{w_c}_{\mathcal{V}} \mathbf{m}_c$ for every c and $\lim_{c \to +\infty} \mathbf{m}_c = \mathbf{x}$. Let us consider the set I of integers i such that $\mathbf{m}[i] < \mathbf{x}[i]$. Let us prove that for every $i \in I$ there exists an integer c_i such that $\mathbf{m}_c[i] > \mathbf{m}[i]$ for every $c \geq c_i$. Let $i \in I$. Since $\mathbf{m}[i] \trianglelefteq \mathbf{x}[i]$ we deduce that $\mathbf{m}[i] \in \mathbb{N}$ and $\mathbf{x}[i] = \top$. From $\lim_{c \to +\infty} \mathbf{m}_c[i] = \mathbf{x}[i]$ we deduce that there exists an integer $c_i \geq 0$ such that $\mathbf{m}_c[i] > \mathbf{m}[i]$ for every $c \geq c_i$. Now let us consider an integer c such that $c \geq c_i$ for every $i \in I$. Observe that $\mathbf{m}_c[i] > \mathbf{m}[i]$ for every $i \in I$. We have proved that there exist a cycle $\mathbf{x} \xrightarrow{w}_G \mathbf{x}$ with $w = w_c$ and an extended configuration $\mathbf{y} = \mathbf{m}_c$ satisfying $\mathbf{m} \xrightarrow{w}_{\mathcal{V}} \mathbf{y}$ and satisfying $\mathbf{y}[i] > \mathbf{m}[i]$ for every i such that $\mathbf{m}[i] < \mathbf{x}[i]$.

Next, assume that there exist a cycle $\mathbf{x} \xrightarrow{w}_G \mathbf{x}$ and an extended configuration $\mathbf{y}$ satisfying $\mathbf{m} \xrightarrow{w}_{\mathcal{V}} \mathbf{y}$ and satisfying $\mathbf{y}[i] > \mathbf{m}[i]$ for every i such that $\mathbf{m}[i] < \mathbf{x}[i]$.

Let us prove that for every i such that $\mathbf{m}[i] \geq \mathbf{x}[i]$ we have $\mathbf{y}[i] = \mathbf{m}[i]$. The relation $\mathbf{m}[i] \trianglelefteq \mathbf{x}[i]$ implies $\mathbf{m}[i] \leq \mathbf{x}[i]$. Thus $\mathbf{m}[i] = \mathbf{x}[i]$. The paths $\mathbf{m} \xrightarrow{w}_{\mathcal{V}} \mathbf{y}$ and $\mathbf{x} \xrightarrow{w}_{\mathcal{V}} \mathbf{x}$ with $\mathbf{m}[i] = \mathbf{x}[i]$ provides $\mathbf{y}[i] = \mathbf{x}[i]$. We have proved that $\mathbf{y}[i] = \mathbf{m}[i]$.

Therefore $\mathbf{y} \geq \mathbf{m}$ and an immediate induction shows that there exists a non-decreasing sequence $(\mathbf{m}_c)_c$ of extended configurations such that $\mathbf{m} \xrightarrow{w^c}_{\mathcal{V}} \mathbf{m}_c$. Finally, just observe that $(\mathbf{x} \xrightarrow{w^c}_G \mathbf{x})$ is a cycle and $\lim_{c \to +\infty} \mathbf{m}_c = \mathbf{x}$. $\square$

Symetrically, we prove the following lemma.

Lemma 3.4. *The output loop condition is satisfied by $\mathcal{M}$ iff there exist a cycle $\mathbf{x}' \xrightarrow{w'}_G \mathbf{x}'$ and an extended configuration $\mathbf{y}'$ satisfying $\mathbf{y}' \xrightarrow{w'}_{\mathcal{V}} \mathbf{m}'$ and satisfying $\mathbf{y}'[i] > \mathbf{m}'[i]$ for every i such that $\mathbf{m}'[i] < \mathbf{x}'[i]$.*

Let (q, q') a pair of states of a graph $G = (Q, \Sigma, T)$. We say that a function $\mu : Q \to \mathbb{Q}$ satisfies the *Kirchhoff's laws* of (q, G, q') if the following system $\chi_{q,G,q'}(\mu)$ holds where $e : Q \times Q \to \{0, 1\}$ denotes the function that takes the value one iff its two arguments are equal:

$$\chi_{q,G,q'}(\mu) := \bigwedge_{p \in Q} \left(\sum_{t=(p_0,a,p) \in T} \mu(t) + e(q,p) = \sum_{t=(p,a,p_1) \in T} \mu(t) + e(p,q') \right)$$

The Parikh image $||\pi||$ of a path π from a state q to a state q' in a graph $G = (Q, \Sigma, T)$ provides a function $||\pi||$ that satisfies $\chi_{q,G,q'}$. *Euler's Lemma* shows that if G is strongly connected then every function $\mu : T \to \mathbb{N}\backslash\{0\}$ satisfying the Kirchhoff's laws of (q, G, q') is the Parikh image of a path from q to q'. Since $\chi_{q,G,q}$ does not depend on $q \in Q$, this linear system is simply denoted by χ_G in the sequel. Naturally, the Parikh image of a cycle satisfies this linear system.

Let $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_{0 \le j \le k}$ be an accepted sequence of an MRGS $\mathcal{U}$. Observe that $\xi = (\mathbf{s}_j, \mu_j, \mathbf{s}'_j)_{0 \le j \le k}$ with $\mu_j = ||\pi_j||$ is a solution of the linear system given in Figure 3 where $\delta(t)$ denotes $\delta(a)$ for every transition $t = (q, a, q')$. This linear system is called the *characteristic system* of $\mathcal{U}$. A solution ξ of the characteristic system is called *concretizable* if there exists an accepted sequence $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_{0 \le j \le k}$ such that $\xi = (\mathbf{s}_j, ||\pi_j||, \mathbf{s}'_j)_j$. The homogeneous form of the characteristic system, obtained by replacing constant terms by zero is called the *homogeneous characteristic system of* $\mathcal{U}$. This system is given in Figure 3. In the sequel, a solution of the homogeneous characteristic system is denoted by $\xi_0 = (\mathbf{s}_{0,j}, \mu_{0,j}, \mathbf{s}'_{0,j})_j$.

$$
\begin{cases}
\dfrac{\text{for all } 1 \le j \le k}{\mathbf{s}'_{j-1} + \delta(a_j) = \mathbf{s}_j} \\[2em]
\dfrac{\text{for all } 0 \le j \le k}{\mathbf{s}_j + \sum\limits_{t \in T_j} \mu_j(t)\delta(t) = \mathbf{s}'_j} \\[2em]
\dfrac{\text{for all } 0 \le j \le k,\ 1 \le i \le n}{\begin{array}{l}\mathbf{s}_j[i] = \mathbf{m_j}[i] \text{ if } \mathbf{m}_j[i] \in \mathbb{N} \\ \mathbf{s}'_j[i] = \mathbf{m_j}'[i] \text{ if } \mathbf{m}'_j[i] \in \mathbb{N}\end{array}} \\[2.5em]
\dfrac{\text{for all } 0 \le j \le k}{\chi_{\mathbf{x}_j, G_j, \mathbf{x}'_j}(\mu_j)}
\end{cases}
\qquad
\begin{cases}
\dfrac{\text{for all } 1 \le j \le k}{\mathbf{s}'_{0,j-1} = \mathbf{s}_{0,j}} \\[2em]
\dfrac{\text{for all } 0 \le j \le k}{\mathbf{s}_{0,j} + \sum\limits_{t \in T_j} \mu_{0,j}(t)\delta(t) = \mathbf{s}'_{0,j}} \\[2em]
\dfrac{\text{for all } 0 \le j \le k,\ 1 \le i \le n}{\begin{array}{l}\mathbf{s}_{0,j}[i] = 0 \text{ if } \mathbf{m}_j[i] \in \mathbb{N} \\ \mathbf{s}'_{0,j}[i] = 0 \text{ if } \mathbf{m}'_j[i] \in \mathbb{N}\end{array}} \\[2.5em]
\dfrac{\text{for all } 0 \le j \le k}{\chi_{G_j}(\mu_{0,j})}
\end{cases}
$$

Figure 3: On the left the characteristic system. On the right the homogeneous characteristic system.

We say that $\mathcal{U}$ satisfies the *large solution condition* if there exists a non-decreasing sequence $(\xi_c)_{c \in \mathbb{N}}$ of solutions $\xi_c = (\mathbf{s}_{j,c}, \mu_{j,c}, \mathbf{s}'_{j,c})_j$ with components in $\mathbb{N}$ of the characteristic system such that:

- $\lim_{c \to +\infty} \mathbf{s}_{j,c} = \mathbf{m}_j$ for every j,
- $\lim_{c \to +\infty} \mu_{j,c}(t) = \top$ for every j and $t \in T_j$, and
- $\lim_{c \to +\infty} \mathbf{s}_{j,c} = \mathbf{m}'_j$ for every j.

The following lemma shows that the large solution condition is decidable in polynomial time since the condition (i) of this lemma is in PTIME with the *Hermite decomposition* and the condition (ii) is in PTIME with the *interior points method*.

Lemma 3.5. *The large solution condition is satisfied by* $\mathcal{U}$ *iff the following conditions* (i) *and* (ii) *hold:*

(i) *Its characteristic system has a solution* ξ *with components in* $\mathbb{Z}$,

(ii) *Its homogeneous characteristic system has a solution* $\xi_0 = (\mathbf{s}_{0,j}, \mu_{0,j}, \mathbf{s}'_{0,j})_j$ *with components in* $\mathbb{Q}$ *satisfying for every* j:
  - $\star$ $\mathbf{s}_{0,j}[i] > 0$ *for every* i *such that* $\mathbf{m}_j[i] = \top$,
  - $\star$ $\mu_{0,j}(t) > 0$ *for every* $t \in T_j$, *and*
  - $\star$ $\mathbf{s}'_{0,j}[i] > 0$ *for every* i *such that* $\mathbf{m}'_j[i] = \top$.

Proof. Let us consider ξ and ξ_0 satisfying condition (i) and (ii). Since ξ_0 is the solution of a linear system, by multiplying ξ_0 by a positive integer, its components can be assumed in $\mathbb{Z}$. Note that in this case the components are in fact in $\mathbb{N}$. Since there exists an integer $c \geq 0$ such that $\xi + c\xi_0$ has its components in $\mathbb{N}$, by replacing ξ by $\xi + c\xi_0$ we can assume that the components of ξ are in $\mathbb{N}$. Now, just observe that $\xi_c = \xi + c\xi_0$ provides a sequence $(\xi_c)_c$ that proves that $\mathcal{U}$ satisfies the large solution condition.

Next assume that $\mathcal{U}$ satisfies the large solution condition. There exists a sequence $(\xi_c)_c$ proving the large solution condition of $\mathcal{U}$. Let us denote by $\xi = (\mathbf{s}_j, \mu_j, \mathbf{s}'_j)_j$ the first solution of this sequence. This solution naturally satisfies (i). Observe that there exists an integer $c \geq 0$ such that for every j:

- $\mathbf{s}_{j,c}[i] > \mathbf{s}_j[i]$ for every i such that $\mathbf{m}_j[i] = \top$,
- $\mu_{j,c}(t) > \mu_j(t)$ for every $t \in T_j$, and
- $\mathbf{s}'_{j,c}[i] > \mathbf{s}'_j[i]$ for every i such that $\mathbf{m}'_j[i] = \top$.

Notice that $\xi_0 = \xi_c - \xi$ provides a solution of the homogeneous characteristic system satisfying condition (ii). □

By adapting [6], we deduce that the perfect condition given in Definition 3.2 is equivalent to the original one [6] (also equivalent to the θ-condition [5]). More formally, we prove the following Proposition 3.6 (the proof is given in Appendix A).

Proposition 3.6. *An MRGS $\mathcal{U}$ is* perfect *if and only if it satisfies the large solution condition and if its marked reachability graphs satisfy the input and output loop conditions.*

3.3. **The KLMST decomposition.** We provide an informal presentation of the algorithm deciding the emptiness of $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$. This algorithm is based on a well-founded order $\sqsubseteq$ over the MRGSs. During its execution, a finite set F of MRGSs is computed. This set satisfies the invariant $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}') = \bigcup_{\mathcal{U} \in F} \mathcal{L}(\mathcal{U})$. Initially, the algorithm starts with the set $F = \{\mathcal{U}\}$ where $\mathcal{U}$ is an MRGS such that $\mathcal{L}(\mathcal{U}) = \mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ (see Example 3.1). Recursively, while the set F is non empty and it only contains MRGSs that do not satisfy the *perfect condition*, such an MRGS $\mathcal{U}$ is picked up from F. Since $\mathcal{U}$ is not perfect, Proposition 3.6 shows that either it does not satisfy the large solution condition or one of its marked reachability graphs does not satisfy the input or the output loop condition. Considering separately these cases, the algorithm computes a finite set F' of MRGSs satisfying $\mathcal{U}' \sqsubset \mathcal{U}$ for every $\mathcal{U}' \in F'$ and $\mathcal{L}(\mathcal{U}) = \bigcup_{\mathcal{U}' \in F}' \mathcal{L}(\mathcal{U}')$. Then, the algorithm replaces F by $F\backslash\{\mathcal{U}\} \cup F'$ and it restarts the while loop. Since $\sqsubseteq$ is well-founded, the loop termination is guaranteed. When the loop terminates, the set F is either empty or it contains at least one perfect MRGS. If F is non empty the algorithm decides that $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ is non empty, otherwise it decides that $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ is empty. The correctness of the algorithm is obtained by observing that the language recognized by a perfect MRGS is always non empty.

Now, let us assume that the while loop is continuing still there exists in F at least one MRGS that does not satisfy the perfect condition. The loop termination is still guaranty since $\sqsubseteq$ is well-founded and when the while loop terminates we get an eventually empty set F of perfect MRGSs such that $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}') = \bigcup_{\mathcal{U} \in F} \mathcal{L}(\mathcal{U})$. This algorithm provides the following Theorem 3.7.

Theorem 3.7 (Fundamental Decomposition [5, 6])**.** *For every tuple* $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$, *we can effectively compute a finite set* F *of perfect MRGSs for* $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ *such that:*

$$\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}') = \bigcup_{\mathcal{U} \in F} \mathcal{L}(\mathcal{U})$$

4. SEMI-PSEUDO-LINEAR SETS

We introduce the class of semi-pseudo-linear sets.

We first introduce the class of monoids. A *monoid* of $\mathbb{Q}^n$ is a set $M \subseteq \mathbb{Q}^n$ such that $\mathbf{0} \in M$ and $M + M \subseteq M$. Observe that for every $X \subseteq \mathbb{Q}^n$, the set $M = \{\mathbf{0}\} \cup \{\sum_{i=1}^{k} \mathbf{x}_i \mid k \geq 1 \wedge \mathbf{x}_i \in X\}$ is the minimal monoid that contains X with respect to the inclusion. This monoid is called the *monoid generated* by X and denoted X^*. A monoid is said to be *finitely generated* if it can be generated by a finite set.

Let M be a monoid. A vector $\mathbf{a} \in M$ is said to be *interior* to M if for every $\mathbf{x} \in M$ there exists an integer $N \geq 1$ satisfying $N\mathbf{a} \in \mathbf{x} + M$. The *interior of a monoid* M is the set of interior vectors to M. It is denoted by $\mathcal{I}(M)$.

Example 4.1. *Let* $P = \{(1, 1), (-1, 1)\}$. *The monoid* $M = P^*$ *and its interior are depicted in Figure 4.*
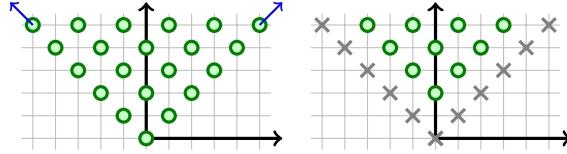

Figure 4: On the left a monoid M. On the right its interior $\mathcal{I}(M)$.

The following Lemma 4.2 characterizes the set $\mathcal{I}(P^*)$ where P is a finite set.

Lemma 4.2. *Let* $P = \{\mathbf{p}_1, \ldots, \mathbf{p}_k\} \subseteq \mathbb{Q}^n$ *with* $k \in \mathbb{N}$. *We have* $\mathcal{I}(P^*) = \{\mathbf{0}\}$ *if* $k = 0$ *and* $\mathcal{I}(P^*) = P^* \cap ((\mathbb{Q}_+ \backslash \{0\})\mathbf{p}_1 + \cdots + (\mathbb{Q}_+ \backslash \{0\})\mathbf{p}_k)$ *if* $k \geq 1$.

Proof. Since the case $k = 0$ is immediate, we assume that $k \geq 1$. Let us first consider an interior vector $\mathbf{a} \in \mathcal{I}(P^*)$. As $\sum_{j=1}^{k} \mathbf{p}_j \in P^*$ and $\mathbf{a} \in \mathcal{I}(P^*)$, there exists $N \geq 1$ such that $N\mathbf{a} \in (\sum_{j=1}^{k} \mathbf{p}_j) + P^*$. Let $\mathbf{p} \in P^*$ such that $N\mathbf{a} = \sum_{j=1}^{k} \mathbf{p}_j + \mathbf{p}$. As $\mathbf{p} \in P^*$, there exists a sequence $(N_j)_{1 \leq j \leq k}$ of elements in $\mathbb{N}$ such that $\mathbf{p} = \sum_{j=1}^{k} N_j \mathbf{p}_j$. Combining this equality with the previous one provides $\mathbf{a} = \sum_{j=1}^{k} \frac{1+N_j}{N} \mathbf{p}_j$. Thus $\mathbf{a} \in (\mathbb{Q}_+ \backslash \{0\})\mathbf{p}_1 + \cdots + (\mathbb{Q}_+ \backslash \{0\})\mathbf{p}_k$. Conversely, let us consider $\mathbf{a} \in P^* \cap ((\mathbb{Q}_+ \backslash \{0\})\mathbf{p}_1 + \cdots + (\mathbb{Q}_+ \backslash \{0\})\mathbf{p}_k)$. Observe that there exists an integer $d \geq 1$ large enough such that $d\mathbf{a} \in (\mathbb{N} \backslash \{0\})\mathbf{p}_1 + \cdots + (\mathbb{N} \backslash \{0\})\mathbf{p}_k$. In particular for every $\mathbf{x} \in P^*$ there exists $N \geq 1$ such that $Nd\mathbf{a} \in \mathbf{x} + P^*$. $\square$

A set $L \subseteq \mathbb{Z}^n$ is said to be *linear* [3] if there exists a vector $\mathbf{b} \in \mathbb{Z}^n$ and a finitely generated monoid $M \subseteq \mathbb{Z}^n$ such that $L = \mathbf{b} + M$. A *semi-linear set* $S \subseteq \mathbb{Z}^n$ is a finite union of linear sets $L_i \subseteq \mathbb{Z}^n$. Recall [3] that sets definable in FO $(\mathbb{N}, +, \leq)$, also called *Presburger sets*, are exactly the non-negative semi-linear sets. By observing that integers are differences of two non-negative integers, we deduce that sets definable in FO $(\mathbb{Z}, +, \leq)$ are exactly the semi-linear sets.

Let us now introduce the class of *pseudo-linear sets* and *semi-pseudo-linear sets*. Intuitively, the pseudo-linear sets extend the *linear sets*, and the semi-pseudo-linear sets extend the semi-linear sets. More formally, a set $X \subseteq \mathbb{Z}^n$ is said to be *pseudo-linear* if there exists $\mathbf{b} \in \mathbb{Z}^n$ and a finitely generated monoid $M \subseteq \mathbb{Z}^n$ such that $X \subseteq \mathbf{b} + M$ and such that for every finite set R of interior vectors to M, there exists $\mathbf{x} \in X$ such that $\mathbf{x} + R^* \subseteq X$. In this case, M is called a *linearizator* for X and the linear set $L = \mathbf{b} + M$ is called a *linearization* of X. A *semi-pseudo-linear set* is a finite union of *pseudo-linear sets*.

Example 4.3. *The set* $X = \{\mathbf{x} \in \mathbb{Z}^2 \mid 0 \leq \mathbf{x}[2] \leq \mathbf{x}[1] \leq 2^{\mathbf{x}[2]}\}$ *is depicted in Figure 5. Observe that* X *is pseudo-linear and* $L = \{\mathbf{x} \in \mathbb{Z}^2 \mid 0 \leq \mathbf{x}[2] \leq \mathbf{x}[1]\}$ *is a linearization of* X. *The set* $Y = \{(2^k, 0) \mid k \in \mathbb{N}\}$ *is not semi-pseudo-linear. However* $Z = X \cup Y$ *is pseudo-linear since* L *is still a linearization of* Z.
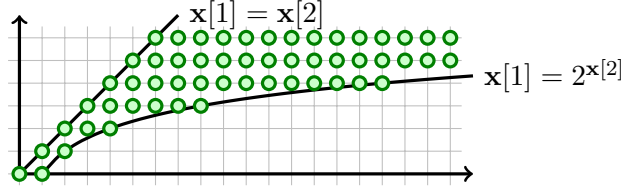

Figure 5: A pseudo-linear set.

Remark 4.4. *Every linear set* $L = \mathbf{b} + M$ *is pseudo-linear.* M *is a linearizator for* L *and* L *is a linearization of* L. *Every semi-linear set is semi-pseudo-linear.*

Remark 4.5. *Semi-pseudo-linear sets can be empty whereas pseudo-linear sets cannot be empty.*

As expected, the class of pseudo-linear sets is stable by linear function images. A function $f : \mathbb{Z}^n \to \mathbb{Z}^{n'}$ is said *linear* if there exists a matrix $A \in \mathbb{Z}^{n \times n'}$ and a vector $\mathbf{v} \in \mathbb{Z}^{n'}$ such that $f(\mathbf{x}) = A\mathbf{x} + \mathbf{v}$ for every $\mathbf{x} \in \mathbb{Z}^n$.

Proposition 4.6. *Images* $X' = f(X)$ *of pseudo-linear sets* X *by a linear function* f *are pseudo-linear. Moreover* $L' = f(L)$ *is a linearization of* X' *for every linearization* L *of* X.

Proof. Let us consider a linear function $f : \mathbb{Z}^n \to \mathbb{Z}^{n'}$ defined by a matrix $A \in \mathbb{Z}^{n \times n'}$ and a vector $\mathbf{v} \in \mathbb{Z}^{n'}$. Let us consider a pseudo-linear set $X \subseteq \mathbb{Z}^n$. As X is pseudo-linear, there exists a linearizator M of X and a vector $\mathbf{b} \in \mathbb{Z}^n$ such that $X \subseteq \mathbf{b} + M$. Let $L = \mathbf{b} + M$. As M is finitely generated there exists a finite set P such that $M = P^*$. We are going to prove that $L' = f(L)$ is a linearization of $X' = f(X)$. Let us consider $\mathbf{b}' = f(\mathbf{b})$ and $P' = \{A\mathbf{p} \mid \mathbf{p} \in P\}$ and observe that $L' = \mathbf{b}' + (P')^*$. In particular L' is a linear set. Since $X \subseteq L$ we deduce that $X' \subseteq L'$. Let us consider a set $R' = \{\mathbf{r}'_1, \ldots, \mathbf{r}'_d\}$ included in the interior of $(P')^*$. As $\mathbf{r}'_i \in (P')^*$ there exists $\mathbf{p}_i \in P^*$ such that $\mathbf{r}'_i = A\mathbf{p}_i$. Lemma 4.2

shows that $\mathbf{r}'_i$ is a sum of vectors of the form $\lambda_{i,\mathbf{p}} A\mathbf{p}$ over all $\mathbf{p} \in P$ where $\lambda_{i,\mathbf{p}} > 0$ is a value in $\mathbb{Q}$. There exists an integer $n_i \geq 1$ large enough such that $n_i \lambda_{i,\mathbf{p}} \in \mathbb{N}\backslash\{0\}$ for every $\mathbf{p} \in P$. We deduce that $\mathbf{r}_i = \sum_{\mathbf{p}\in P} n_i\lambda_{i,\mathbf{p}}\mathbf{p}$ is a vector in P^*. Moreover, from Lemma 4.2 we deduce that $\mathbf{r}_i$ is in the interior of P^*. Let us consider the set R of vectors $\mathbf{r}_i + k_i\mathbf{p}_i$ where k_i is an integer such that $0 \leq k_i < n_i$. As $\mathbf{r}_i \in \mathcal{I}(P^*)$ and $\mathbf{p}_i \in P^*$ we deduce that $\mathbf{r}_i + k_i\mathbf{p}_i \in \mathcal{I}(P^*)$. We have proved that $R \subseteq \mathcal{I}(P^*)$. As L is a linearization of X, there exists $\mathbf{x} \in X$ such that $\mathbf{x} + R^* \subseteq X$. We deduce that $f(\mathbf{x}) + AR^* \subseteq X'$. Let us consider $\mathbf{x}' = f(\mathbf{x}) + A(\sum_{i=1}^{d} \mathbf{r}_i)$ and let us prove that $\mathbf{x}' + (R')^* \subseteq X'$. Consider $\mathbf{r}' \in (R')^*$. There exists a sequence $(\mu'_i)_{1\leq i\leq d}$ of integers in $\mathbb{N}$ such that $\mathbf{r}' = \sum_{i=1}^{d} \mu'_i\mathbf{r}'_i$. The Euclid division of μ'_i by n_i shows that $\mu'_i = k_i + n_i\mu_i$ where $\mu_i \in \mathbb{N}$ and $0 \leq k_i < n_i$. From $n_i\mathbf{r}'_i = A\mathbf{r}_i$ we deduce that $\mathbf{x}' + \mathbf{r}' = f(\mathbf{x}) + A(\sum_{i=1}^{d}(\mathbf{r}_i + k_i\mathbf{p}_i) + \sum_{i=1}^{d} \mu_i\mathbf{r}_i)$. Observe that $\mathbf{r}_i + k_i\mathbf{p}_i$ and $\mathbf{r}_i$ are both in R. We have proved that $\mathbf{x}' + \mathbf{r}' \in f(\mathbf{x}) + AR^*$. Thus $\mathbf{x}' + (R')^* \subseteq X'$. We have proved that L' is a linearization of X'. $\square$

5. The Parikh Images of Perfect MRGSs

The Parikh images of languages recognized by perfect MRGSs are proved to be pseudo-linear in this section. From the KLMST decomposition, we deduce the semi-pseudo-linearity of the Parikh image of $\mathcal{L}(\mathbf{m}, \mathcal{V}, \mathbf{m}')$.

Let us consider a perfect MRGS $\mathcal{U}$ for $(\mathbf{m}, \mathcal{V}, \mathbf{m}')$. We denote by H the solutions with components in $\mathbb{N}$ of the characteristic system of $\mathcal{U}$. We consider the set of concretizable solutions H'. Since the Parikh image of $\mathcal{L}(\mathcal{U})$ is the image by a linear function of H', by Proposition 4.6 it is sufficient to prove that H' is pseudo-linear. Let us introduce the set H_0 of solutions with components in $\mathbb{N}$ of the homogeneous characteristic system. We prove in the sequel that H_0 is a linearizator for H'. First of all observe that H_0 is a finitely generated monoid since $H_0 = P_0^*$ where $P_0 = \min(H_0\backslash\{\mathbf{0}\})$, and P_0 is finite since $\leq$ is a well-order over H_0.

Since $H' \subseteq H$, the following Lemma 5.1 shows that H' is included in $(\xi - \xi_0) + H_0$. We follow notations introduced in Definition 3.2.

Lemma 5.1. *There exists $\xi \in H$ and $\xi_0 \in H_0$ such that $H \subseteq (\xi - \xi_0) + H_0$.*

Proof. As $\mathcal{U}$ satisfies the large solution condition there exists $\xi \in H$. Moreover, Lemma 3.5 shows that there exists a solution ξ_0 with components in $\mathbb{Q}$ of the homogeneous characteristic system satisfying the additional conditions $\mathbf{s}_{0,j}[i] > 0$ if $\mathbf{m}_j[i] = \top$, $\mathbf{s}'_{0,j}[i] > 0$ if $\mathbf{m}'_j[i] = \top$, and $\mu_{0,j}(t) > 0$ for every $t \in T_j$. By multiplying ξ_0 by a positive integer, we can assume that the components of ξ_0 are in $\mathbb{N}$. Note that for every $\xi' \in H$, there exists $c \in \mathbb{N}$ such that $\xi' + c\xi_0 \geq \xi$. As $\min(H)$ is finite, by multiplying ξ_0 by a positive integer we can assume that $\xi' + \xi_0 \geq \xi$ for every $\xi' \in H$. That means $H \subseteq (\xi - \xi_0) + H_0$. $\square$

Now, let us consider a finite set $R_0 = \{\xi_1, \ldots, \xi_d\}$ included in the interior of H_0. We are going to prove that there exists $\xi \in H$ such that $\xi + R_0^* \subseteq H'$. We first prove the following lemma.

Lemma 5.2. *For every $\xi_l = (\mathbf{s}_{l,j}, \mu_{l,j}, \mathbf{s}'_{l,j})_j$ interior vector of H_0, the function $\mu_{l,j}$ is the Parikh image of a cycle $\pi_{l,j} = (\mathbf{x}_j \xrightarrow{\sigma_{l,j}}_{G_j} \mathbf{x}_j)$.*

Proof. Since $\mathcal{U}$ satisfies the large solution condition, Lemma 3.5 shows for every $t \in T_j$, there exists a solution $\xi_0 = (\mathbf{s}_{0,j}, \mu_{0,j}, \mathbf{s}'_{0,j})_j$ in H_0 such that $\mu_{0,j}(t) > 0$. As $H_0 = P_0^*$, for every $t \in T_j$ there exists $\xi_0 \in P_0$ satisfying the same property. As ξ_l is in the interior of H_0, Lemma 4.2 shows that there exists a sequence $(\lambda_{\xi_0})_{\xi_0 \in P_0}$ of positive rational values $\lambda_{\xi_0} \in \mathbb{Q}_{>0}$ such that $\xi_l = \sum_{\xi_0} \lambda_{\xi_0} \xi_0$. In particular, we deduce that $\mu_{l,j}(t) > 0$ for every $t \in T_j$ and for every $0 \leq j \leq k$. As ξ_l satisfies χ_{G_j} we deduce that $\mu_{l,j}$ satisfies the Kirchhoff's laws. As G_j is strongly connected and $\mu_{l,j}(t) \geq 1$ for every $t \in T_j$, Euler's Lemma shows that $\mu_{l,j}$ is the Parikh image of a cycle $\pi_{l,j} = (\mathbf{x}_j \xrightarrow{\sigma_{l,j}}_{G_j} \mathbf{x}_j)$. $\qquad\square$

Since $\mathbf{x}_j \xrightarrow{\sigma_{l,j}}_{\mathcal{V}}$, there exists an integer $c \geq 0$ such that for every $0 \leq j \leq k$ and for every configuration $\mathbf{r}_j$ satisfying $\mathbf{r}_j[i] \geq c$ if $\mathbf{x}_j[i] = \top$ and $\mathbf{r}_j[i] = \mathbf{x}_j[i]$ otherwise, we have $\mathbf{r}_j \xrightarrow{\sigma_{l,j}}_{\mathcal{V}}$.

As $\mathcal{U}$ is perfect, there exists an accepted tuple $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_{0 \leq j \leq k}$ such that for every j, π_j can be decomposed into:

$$\pi_j = (\mathbf{x}_j \xrightarrow{w_j}_{G_j} \mathbf{x}_j \xrightarrow{\sigma_j}_{G_j} \mathbf{x}'_j \xrightarrow{w'_j}_{G_j} \mathbf{x}'_j)$$

and such that the pair of configurations $(\mathbf{r}_j, \mathbf{r}'_j)$ satisfying the following relations:

$$\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}} \mathbf{r}_j \xrightarrow{\sigma_j}_{\mathcal{V}} \mathbf{r}'_j \xrightarrow{w'_j}_{\mathcal{V}} \mathbf{s}'_j$$

also satisfy:
- $\mathbf{r}_j[i] \geq c$ if $\mathbf{x}_j[i] = \top$ and $\mathbf{r}_j[i] = \mathbf{x}_j[i]$ otherwise,
- $\mathbf{r}'_j[i] \geq c$ if $\mathbf{x}'_j[i] = \top$ and $\mathbf{r}'_j[i] = \mathbf{x}'_j[i]$ otherwise.

In particular we have $\mathbf{r}_j \xrightarrow{\sigma_{l,j}}_{\mathcal{V}}$ for every $0 \leq j \leq k$ and for every $1 \leq l \leq d$.

As $\mathbf{s}_{l,j} \geq \mathbf{0}$ and $\mathbf{r}_j \xrightarrow{\sigma_{l,j}}_{\mathcal{V}}$ we deduce that $\mathbf{r}_j + \mathbf{s}_{l,j} \xrightarrow{\sigma_{l,j}}_{\mathcal{V}}$. Moreover, from $\mathbf{s}_{l,j} + \delta(\sigma_{l,j}) = \mathbf{s}'_{l,j}$ we get:

$$\mathbf{r}_j + \mathbf{s}_{l,j} \xrightarrow{\sigma_{l,j}}_{\mathcal{V}} \mathbf{r}_j + \mathbf{s}'_{l,j}$$

As $\mathbf{s}_{l,j}, \mathbf{s}'_{l,j} \geq \mathbf{0}$, an immediate induction shows that for every sequence $n_1, \ldots, n_d \in \mathbb{N}$ we have the following relation:

$$\mathbf{r}_j + \sum_{l=1}^{d} n_l \mathbf{s}_{l,j} \xrightarrow{\sigma_{1,j}^{n_1} \ldots \sigma_{d,j}^{n_d}}_{\mathcal{V}} \mathbf{r}_j + \sum_{l=1}^{d} n_l \mathbf{s}'_{l,j}$$

Let $\xi = (\mathbf{s}_j, \|\pi_j\|, \mathbf{s}'_j)_{0 \leq j \leq k}$. We have proved that $\xi + \sum_{l=1}^{d} n_l \xi_l$ is concretizable. Thus $\xi + R_0^* \subseteq H'$. Therefore H' is pseudo-linear and H_0 is a linearizator for H'. We have proved the following Theorem 5.3.

Theorem 5.3. *The Parikh image of $\mathcal{L}(\mathcal{U})$ is pseudo-linear for every perfect MRGS $\mathcal{U}$.*

From Theorem 3.7 and Theorem 5.3 we deduce the following Corollary 5.4.

Corollary 5.4. *The Parikh image of $\mathcal{L}(\mathbf{m}, \mathcal{V}, \mathbf{m}')$ is semi-pseudo-linear.*

6. Petri Sets

A set $X \subseteq \mathbb{Z}^n$ is said to be a *Petri set* if $X \cap S$ is semi-pseudo-linear for every semi-linear set $S \subseteq \mathbb{Z}^n$. Since $\mathbb{Z}^n$ is a linear set, Petri sets are semi-pseudo-linear. However the converse is not true in general (see Example 6.1). In this section, $\mathrm{post}^*_{\mathcal{V}}(S)$ and $\mathrm{pre}^*_{\mathcal{V}}(S')$ are proved to be Petri sets for every semi-linear sets $S, S' \subseteq \mathbb{N}^n$. This result is used in Section 8 to get a *local analysis* of $\mathrm{post}^*_{\mathcal{V}}(S)$ and $\mathrm{pre}^*_{\mathcal{V}}(S')$ with respect to some semi-linear sets.

Example 6.1. *Let us consider the pseudo-linear set $Z = X \cup Y$ introduced in Example 4.3 and observe that Z is not a Petri set since $Y = Z \cap S$ is not semi-pseudo-linear with $S = (1,0) + \{(1,0)\}^*$.*

Let us prove that $\mathrm{post}^*_{\mathcal{V}}(S) \cap S'$ and $S \cap \mathrm{pre}^*_{\mathcal{V}}(S')$ are semi-pseudo-linear for every semi-linear sets $S, S' \subseteq \mathbb{N}^n$. Since semi-linear sets are finite unions of linear sets we only prove this result for the special case of two linear sets $S = \mathbf{s} + P^*$ and $S' = \mathbf{s}' + (P')^*$ where $\mathbf{s}, \mathbf{s}' \in \mathbb{N}^n$ and $P, P' \subseteq \mathbb{N}^n$ are two finite sets. We consider two alphabets $\Sigma_P, \Sigma_{P'}$ disjoint of Σ and a displacement function $\bar{\delta}$ defined over $\bar{\Sigma} = \Sigma_P \cup \Sigma \cup \Sigma_{P'}$ that extends δ such that:
$$P = \{\bar{\delta}(a) \mid a \in \Sigma_P\} \qquad P' = \{-\bar{\delta}(a) \mid a \in \Sigma_{P'}\}$$
We consider the VAS $\bar{\mathcal{V}} = (\bar{\Sigma}, n, \bar{\delta})$. Intuitively, since $\bar{\delta}(\Sigma_P) \subseteq \mathbb{N}^n$ and $\bar{\delta}(\Sigma_{P'}) \subseteq -\mathbb{N}^n$, words in $\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$ can be reordered into words in $(\Sigma_P^* \Sigma^* \Sigma_{P'}^*) \cap \mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$. More formally, we prove the following lemma.

Lemma 6.2. *Assume that $\mathbf{s} \xrightarrow{\sigma a \sigma'}_{\mathcal{V}} \mathbf{s}'$ holds with $\sigma, \sigma' \in \Sigma^*$, and $a \in \Sigma$. We have:*

- $\mathbf{s} \xrightarrow{a \sigma \sigma'}_{\mathcal{V}} \mathbf{s}'$ *if $\delta(a) \geq \mathbf{0}$.*
- $\mathbf{s} \xrightarrow{\sigma \sigma' a}_{\mathcal{V}} \mathbf{s}'$ *if $\delta(a) \leq \mathbf{0}$.*

Proof. We only consider the case $\delta(a) \geq \mathbf{0}$ since the other case is symmetrical by replacing $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ by $(\mathbf{s}', -\mathcal{V}, \mathbf{s})$ where $-\mathcal{V} = (\Sigma, n, -\delta)$. Let us consider the pair of configurations $(\mathbf{r}, \mathbf{r}')$ such that $\mathbf{s} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{r} \xrightarrow{a}_{\mathcal{V}} \mathbf{r}' \xrightarrow{\sigma'}_{\mathcal{V}} \mathbf{s}'$. Since $\delta(a) \geq \mathbf{0}$ we have $\mathbf{s} \xrightarrow{a}_{\mathcal{V}} \mathbf{s} + \delta(a)$. As $\mathbf{s} + \delta(a) \geq \mathbf{s}$ and $\mathbf{s} \xrightarrow{\sigma}_{\mathcal{V}}$ we deduce that $\mathbf{s} + \delta(a) \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{s} + \delta(a) + \delta(\sigma)$. From $\mathbf{r}' = \mathbf{s} + \delta(\sigma) + \delta(a)$ we deduce the lemma. $\square$

Let us consider the displacement functions f and f' defined over $\bar{\Sigma}$ by:
$$f(a) = \begin{cases} \bar{\delta}(a) & \text{if } a \in \Sigma_P \\ \mathbf{0} & \text{otherwise} \end{cases}$$
$$f'(a) = \begin{cases} -\bar{\delta}(a) & \text{if } a \in \Sigma_{P'} \\ \mathbf{0} & \text{otherwise} \end{cases}$$

Lemma 6.3. *We have $\mathrm{post}^*_{\mathcal{V}}(S) \cap S' = \mathbf{s}' + f'(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$ and $S \cap \mathrm{pre}^*_{\mathcal{V}}(S') = \mathbf{s} + f(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$.*

Proof. Let us consider $\mathbf{c}' \in \mathrm{post}^*_{\mathcal{V}}(S) \cap S'$ and let us prove that $\mathbf{c}' \in \mathbf{s}' + f'(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$. There exists $\mathbf{c} \in S$ and a word $v \in \Sigma^*$ such that $\mathbf{c} \xrightarrow{v}_{\mathcal{V}} \mathbf{c}'$. In particular $\mathbf{c} \xrightarrow{v}_{\bar{\mathcal{V}}} \mathbf{c}'$. Since $S = \mathbf{s} + P^*$ we observe that there exists a word $u \in \Sigma_P^*$ such that $\mathbf{s} \xrightarrow{u}_{\bar{\mathcal{V}}} \mathbf{c}$. Symmetrically since $S' = \mathbf{s}' + (P')^*$ there exists $u' \in \Sigma_{P'}^*$ such that $\mathbf{c}' \xrightarrow{u'}_{\bar{\mathcal{V}}} \mathbf{s}'$. We have proved that $uvu' \in \mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$. Note that $f'(uvu') = -\bar{\delta}(u')$. From $\mathbf{s}' = \mathbf{c}' + \bar{\delta}(u')$ we have proved that $\mathbf{c}' \in \mathbf{s}' + f'(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$.

Conversely, let us consider a vector $\mathbf{c}' \in \mathbf{s}' + f'(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$ and let us prove that $\mathbf{c}' \in \text{post}_{\mathcal{V}}^*(S) \cap S'$. There exists a word $\sigma \in \mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$ such that $\mathbf{c}' = \mathbf{s}' + f'(\sigma)$. Since $\delta(\Sigma_P) \subseteq \mathbb{N}^n$ and $\delta(\Sigma_{P'}) \subseteq -\mathbb{N}^n$, Lemma 6.2 shows that σ can be reordered into a word $\sigma_0 \in \mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}') \cap (\Sigma_P^* \Sigma^* \Sigma_{P'}^*)$. As σ_0 and σ have the same Parikh image we deduce that $f'(\sigma) = f'(\sigma_0)$. In particular, we can assume without loss of generality that $\sigma = uvu'$ with $u \in \Sigma_P^*$, $v \in \Sigma^*$ and $u' \in \Sigma_{P'}^*$. Let us consider the two configurations $\mathbf{c}, \mathbf{c}''$ such that $\mathbf{s} \xrightarrow{u}_{\bar{\mathcal{V}}} \mathbf{c} \xrightarrow{v}_{\bar{\mathcal{V}}} \mathbf{c}'' \xrightarrow{u'}_{\bar{\mathcal{V}}} \mathbf{s}'$. As $\mathbf{s}' = \mathbf{c}'' + \bar{\delta}(u')$ and $f'(\sigma) = -\bar{\delta}(u')$ we deduce that $\mathbf{c}'' = \mathbf{c}'$. Moreover, since $u \in \Sigma_P^*$ we deduce that $\mathbf{c} \in S$ and since $u' \in \Sigma_{P'}^*$ we get $\mathbf{c}' \in S'$. From $v \in \Sigma^*$ we deduce $\mathbf{c} \xrightarrow{v}_{\mathcal{V}} \mathbf{c}'$. We have proved that $\mathbf{c}' \in \text{post}_{\mathcal{V}}^*(S) \cap S'$.

Thus $\text{post}_{\mathcal{V}}^*(S) \cap S' = \mathbf{s}' + f'(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$. Symmetrically we get $S \cap \text{pre}_{\mathcal{V}}^*(S') = \mathbf{s} + f(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$. $\qquad\square$

Observe that sets $\mathbf{s}' + f'(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$ and $\mathbf{s} + f(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$ are images by linear functions of the Parikh image of $\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$. Corollary 5.4 shows that the Parikh image of $\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$ is semi-pseudo-linear. From Proposition 4.6 we deduce the following Theorem 6.4.

Theorem 6.4. $\text{post}_{\mathcal{V}}^*(S)$ and $\text{pre}_{\mathcal{V}}^*(S')$ are Petri sets for every semi-linear sets $S, S' \subseteq \mathbb{N}^n$.

7. PSEUDO-LINEAR SETS INTERSECTIONS

Let X_1, X_2 be two pseudo-linear sets with an empty intersection $X_1 \cap X_2$ and let L_1, L_2 be linearizations of X_1, X_2. Since L_1, L_2 over-approximate X_1, X_2, the intersection $L_1 \cap L_2$ is not empty in general. In this section we introduce a dimension function that satisfies $\dim(L_1 \cap L_2) < \dim(X_1 \cup X_2)$. This dimension function is defined in Section 7.1 and the strict inequality is proved in Section 7.2.

7.1. **Dimension.** A *vector space* V of $\mathbb{Q}^n$ is a set $V \subseteq \mathbb{Q}^n$ such that $\mathbf{0} \in V$, $V + V \subseteq V$ and $\mathbb{Q}V \subseteq V$. Observe that for every set $X \subseteq \mathbb{Q}^n$ the set $V = \{\mathbf{0}\} \cup \{\sum_{i=1}^k \lambda_i \mathbf{x}_i \mid k \geq 1 \wedge \lambda_i \in \mathbb{Q} \wedge \mathbf{x}_i \in X\}$ is the minimal vector space that contains X with respect to the inclusion. This vector space is called the *vector space generated* by X. Recall that for every vector space V there exists a finite set $B \subseteq V$ that generates V. The minimal integer $d \in \mathbb{N}$ such that there exists a finite set $B \subseteq V$ with d elements that generates V is called the *rank* of V and denoted $\text{rank}(V)$. Note that $\text{rank}(V) \in \{0, \ldots, n\}$ and for every set $X \subseteq \mathbb{Q}^n$ there exists a finite set $B \subseteq X$ such that the vector space V generated by B is equal to the vector space generated by X and such that $|B| = \text{rank}(V)$.

The *dimension of a non empty set* $X \subseteq \mathbb{Q}^n$ is the minimal integer $d \in \{0, \ldots, n\}$ such that there exist $k \in \mathbb{N}$, a sequence $(V_1, \ldots, V_k)$ of vector spaces $V_j \subseteq \mathbb{Q}^n$, and a sequence $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ of vectors $\mathbf{a}_j \in \mathbb{Q}^n$ such that $X \subseteq \bigcup_{j=1}^k (\mathbf{a}_j + V_j)$ and $\text{rank}(V_j) \leq d$. We denote by $\dim(X)$ the dimension of X. By definition $\dim(\emptyset) = -\infty$.

Example 7.1. *Let* $X_0 = \{(0,0)\}$, $X_1 = \{\mathbf{x} \in \mathbb{N}^2 \mid \mathbf{x}[1] = \mathbf{x}[2]\}$ *and* $X_2 = \{\mathbf{x} \in \mathbb{N}^2 \mid \mathbf{x}[2] \leq \mathbf{x}[1]\}$ *be the sets depicted in Figure 6. We have* $\dim(X_0) = 0$, $\dim(X_1) = 1$ *and* $\dim(X_2) = 2$.

Let us show some immediate properties satisfied by the dimension function. Observe that $\dim(X) = -\infty$ if and only if X is empty. The dimension function is monotonic $\dim(X_1) \leq \dim(X_2)$ for every $X_1 \subseteq X_2$. Moreover it satisfies $\dim(X_1 \cup X_2) =$
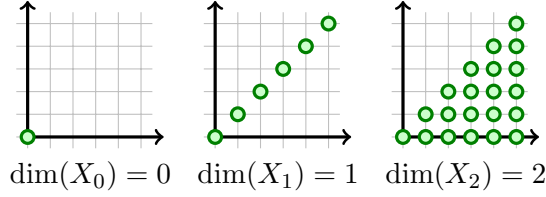
$$\dim(X_0) = 0 \quad \dim(X_1) = 1 \quad \dim(X_2) = 2$$

Figure 6: Dimension of some sets.

$\max\{\dim(X_1), \dim(X_2)\}$ and $\dim(X_1 + X_2) \le \dim(X_1) + \dim(X_2)$. Note also that $\dim(\mathbf{a} + X) = \dim(X)$ for every $\mathbf{a} \in \mathbb{Q}^n$. In the sequel, we prove that (1) $\mathrm{rank}(V) = \dim(V)$ for every vector space V, (2) $\dim(M) = \dim(V)$ for every vector space V generated by a monoid M, and (3) $\dim(L) = \dim(X)$ for every linearization L of a pseudo linear set X. We first prove the following lemma.

Lemma 7.2. *Let $M \subseteq \mathbb{Q}^n$ be a monoid and let $(V_1, \dots, V_k)$ be a sequence of vector spaces $V_j \subseteq \mathbb{Q}^n$ and let $(\mathbf{a}_1, \dots, \mathbf{a}_k)$ be a sequence of vectors $\mathbf{a}_j \in \mathbb{Q}^n$. If $M \subseteq \bigcup_{j=1}^{k}(\mathbf{a}_j + V_j)$ then there exists j such that $\mathbf{a}_j \in V_j$ and $M \subseteq \mathbf{a}_j + V_j$.*

Proof. Let us observe that $\mathbf{a}_j \in V_j$ implies $\mathbf{a}_j + V_j = V_j$. We first prove that $M \subseteq \bigcup_{j=1}^{k}(\mathbf{a}_j + V_j)$ implies $M \subseteq \bigcup_{j \in J} V_j$ where J is the set of $j \in \{1, \dots, k\}$ such that $\mathbf{a}_j \in V_j$. Let us consider $\mathbf{m} \in M$. Since M is a monoid we deduce that $\lambda \mathbf{x} \in M$ for every $\lambda \in \mathbb{N}$. In particular there exists $j \in \{1, \dots, k\}$ such that $\lambda \mathbf{x} \in \mathbf{a}_j + V_j$ for infinitely many λ. In particular there exist $\lambda < \lambda'$ in $\mathbb{N}$ and $\mathbf{v}, \mathbf{v}' \in V_j$ such that $\lambda \mathbf{x} = \mathbf{a}_j + \mathbf{v}$ and $\lambda' \mathbf{x} = \mathbf{a}_j + \mathbf{v}'$. Now, just observe that $\mathbf{a}_j = \frac{\lambda \mathbf{v}' - \lambda' \mathbf{v}}{\lambda' - \lambda}$ implies $\mathbf{a}_j \in V_j$ and $\mathbf{x} = \frac{\mathbf{a}_j + \mathbf{v}'_j}{\lambda'}$ implies that $\mathbf{x} \in V_j$. Therefore $M \subseteq \bigcup_{j \in J} V_j$.

Now let us prove by induction over $k \in \mathbb{N}_{>0}$ that for every sequence $(V_1, \dots, V_k)$ of vector spaces $V_j \subseteq \mathbb{Q}^n$, if $M \subseteq \bigcup_{j=1}^{k} V_j$ then there exists j such that $M \subseteq V_j$. The case $k = 1$ is immediate. Assume that the lemma is already proven for an integer $k \in \mathbb{N}_{>0}$. Let us consider a monoid $M \subseteq \mathbb{Q}^n$, a sequence $(V_1, \dots, V_{k+1})$ of vector spaces $V_j \subseteq \mathbb{Q}^n$ such that $M \subseteq \bigcup_{j=1}^{k+1} V_j$. Let us prove that there exists $j \in \{1, \dots, k+1\}$ such that $M \subseteq V_j$. Naturally if $M \subseteq V_{k+1}$ we are done. Thus, we can assume that M is not included in V_{k+1} and we can pick a vector $\mathbf{m} \in M \backslash V_{k+1}$. Let $\mathbf{x} \in M$ and let us prove that $\mathbf{x} \in \bigcup_{j=1}^{k} V_j$. Note that if $\mathbf{x} \notin V_{k+1}$ we are done. Thus, we can assume without loss of generality that $\mathbf{x} \in V_{k+1}$. Let us introduce $\mathbf{y}_\lambda = \mathbf{x} + \lambda \mathbf{m}$ where $\lambda \in \mathbb{N}$. Since M is a monoid that contains $\mathbf{x}$ and $\mathbf{m}$ we deduce that $\mathbf{y}_\lambda \in M$. Assume by contradiction that $\mathbf{y}_\lambda \in V_{k+1}$ for $\lambda \ne 0$. Since $\mathbf{x}$ and $\mathbf{y}_\lambda$ are both in V_{k+1}, and V_{k+1} is a vector space, we deduce from $\mathbf{m} = \frac{1}{\lambda}(\mathbf{y}_\lambda - \mathbf{x})$ that $\mathbf{m} \in V_{k+1}$. We get a contradiction with $\mathbf{m} \notin V_{k+1}$. Thus $\mathbf{y}_\lambda \in \bigcup_{j=1}^{k} V_j$ for every $\lambda \in \mathbb{N}_{>0}$. Hence there exists $j \in \{1, \dots, k\}$ such that $\mathbf{y}_\lambda \in V_j$ for infinitely many $\lambda \in \mathbb{N}_{>0}$. In particular there exists $\lambda < \lambda'$ in $\mathbb{N}_{>0}$ such that $\mathbf{y}_\lambda, \mathbf{y}'_\lambda \in V_j$. As V_j is a vector space, from $\mathbf{x} = \frac{\lambda \mathbf{y}_{\lambda'} - \lambda' \mathbf{y}_\lambda}{\lambda' - \lambda}$ we deduce that $\mathbf{x} \in V_j$. We have proved that $M \subseteq \bigcup_{j=1}^{k} V_j$. From the induction hypothesis, we deduce that there exists $j \in \{1, \dots, k\}$ such that $M \subseteq V_j$. We have proved the property by induction. $\square$

Now, we can prove the following results.

Lemma 7.3. *We have $\dim(V) = \mathrm{rank}(V)$ for every vector space V.*

Proof. Since $V \subseteq \mathbf{0} + V$ we get $\dim(V) \leq \operatorname{rank}(V)$. Conversely, there exists a sequence $(V_1, \ldots, V_k)$ of vector spaces $V_j \subseteq \mathbb{Q}^n$ and a sequence $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ of vectors $\mathbf{a}_j \in \mathbb{Q}^n$ such that $V \subseteq \bigcup_{j=1}^{k}(\mathbf{a}_j + V_j)$ and $\operatorname{rank}(V_j) \leq \dim(V)$. As V is a vector space and in particular a monoid, Lemma 7.2 shows that there exists j such that $V \subseteq \mathbf{a}_j + V_j$ and $\mathbf{a}_j \in V_j$. From $\mathbf{a}_j + V_j = V_j$ we deduce that $V \subseteq V_j$. In particular $\operatorname{rank}(V) \leq \operatorname{rank}(V_j)$ and we have proved the other relation $\operatorname{rank}(V) \leq \dim(V)$. $\square$

Proposition 7.4. *We have* $\dim(M) = rank(V)$ *where V is the vector space generated by a monoid M.*

Proof. Since $M \subseteq V$ we get $\dim(M) \leq \operatorname{rank}(V)$. Conversely, there exists a sequence a sequence $(V_1, \ldots, V_k)$ of vector spaces $V_j \subseteq \mathbb{Q}^n$ and a sequence $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ of vectors $\mathbf{a}_j \in \mathbb{Q}^n$ such that $M \subseteq \bigcup_{j=1}^{k}(\mathbf{a}_j + V_j)$ and $\operatorname{rank}(V_j) \leq \dim(M)$. From Lemma 7.2 there exists j such that $\mathbf{a}_j \in V_j$ and $M \subseteq \mathbf{a}_j + V_j$. As $\mathbf{a}_j \in V_j$ we get $\mathbf{a}_j + V_j = V_j$. We deduce that $M \subseteq V_j$. By minimality of the vector space generated by M, we deduce that $V \subseteq V_j$. In particular, $\operatorname{rank}(V) \leq \operatorname{rank}(V_j)$. Since $\operatorname{rank}(V_j) \leq \dim(M)$ we deduce the other relation $\operatorname{rank}(V) \leq \dim(M)$. $\square$

As expected, the dimension of a pseudo-linear set is equal to the dimension of every linearization.

Lemma 7.5. *We have* $\dim(X) = \dim(L)$ *for every linearization L of a pseudo-linear set $X \subseteq \mathbb{Z}^n$.*

Proof. There exists $\mathbf{b} \in \mathbb{Z}^n$ and a linearizator M for X such that $L = \mathbf{b} + M$. From $X \subseteq L$ we deduce that $\dim(X) \leq \dim(L)$. Let us prove the converse. Let us consider an interior vector $\mathbf{a} \in \mathcal{I}(M)$. Since M is finitely generated, there exists a finite set P such that $M = P^*$. Observe that $R = \{\mathbf{a}\} \cup (\mathbf{a} + P)$ is a finite subset of $\mathcal{I}(M)$. As X is pseudo-linear, there exists $\mathbf{x} \in X$ such that $\mathbf{x} + R^* \subseteq X$. Note that the vector space generated by R is equal to the vector space generated by P. Thus, from Proposition 7.4 we deduce that $\dim(R^*) = \dim(P^*)$. As $\dim(\mathbf{x} + R^*) = \dim(R^*)$ and $\dim(\mathbf{b} + P^*) = \dim(P^*)$ we deduce that $\dim(\mathbf{x} + R^*) = \dim(L)$. Since $\mathbf{x} + R^* \subseteq X$ we deduce that $\dim(L) \leq \dim(X)$. $\square$

7.2. **Pseudo-linear sets with empty intersections.** In this section we prove that linearizations L_1, L_2 of two pseudo-linear sets X_1, X_2 with an empty intersection $X_1 \cap X_2 = \emptyset$ satisfy the strict inequality $\dim(L_1 \cap L_2) < \dim(X_1 \cup X_2)$. Note that even if $X_1 \cap X_2 = \emptyset$, the intersection $L_1 \cap L_2$ may be non empty since L_1, L_2 are over-approximations of X_1, X_2.

Example 7.6. *Let us consider the pseudo-linear set X described in Example 4.3 and a linearization $L = \{\mathbf{x} \in \mathbb{Z}^2 \mid 0 \leq \mathbf{x}[2] \leq \mathbf{x}[1]\}$ of X. We also consider the linear set $X' = (8, 2) + \{(1, 0), (3, -1)\}^*$. Sets X and X' are depicted together in Figure 7. Note that $L' = X'$ is a linearization of the linear set X'. Notice that $X \cap X' = \emptyset$. The set $L \cap L'$ is depicted in gray in Figure 7. Observe that $L \cap L' = \{(8, 2), (11, 1), (14, 0)\} + \{(1, 0)\}^*$. Therefore $1 = \dim(L \cap L') < \dim(X \cup X') = 2$.*

We first introduce the class of groups. A *group* of $\mathbb{Q}^n$ is a set $Z \subseteq \mathbb{Q}^n$ such that $\mathbf{0} \in Z$, $Z + Z \subseteq Z$ and $-Z \subseteq Z$. Observe that for every $X \subseteq \mathbb{Q}^n$, the set $G = X^* - X^*$ is the minimal group that contains X with respect to the inclusion. This group is said to be generated by X. Let us consider the group $G = M - M$ generated by a monoid M and
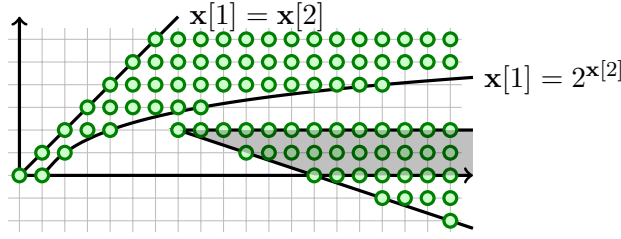
Figure 7: Two pseudo-linear sets with an empty intersection.

$\mathbf{a} \in \mathbb{Z}^n$. Observe that $\mathbf{a} \in \mathcal{I}(M)$ if and only if for every $\mathbf{g} \in G$ there exists an integer $N \geq 1$ such that $\mathbf{g} + N\mathbf{a} \in M$.

Lemma 7.7. *For every vector $\mathbf{v} \in V$ where V is the vector space generated by a group G, there exists an integer $d \geq 1$ such that $d\mathbf{v} \in G$.*

Proof. As $\mathbf{v} \in V$, either $\mathbf{v} = \mathbf{0}$ or $\mathbf{v}$ can be decomposed into a finite sum $\mathbf{v} = \sum_{i=1}^{k} \lambda_i \mathbf{g}_i$ with $k \geq 1$, $\lambda_i \in \mathbb{Q}$ and $\mathbf{g}_i \in G$. The case $\mathbf{v} = \mathbf{0}$ is immediate with $d = 1$ and the second case is obtained by considering an integer $d \geq 1$ such that $d\lambda_i \in \mathbb{Z}$ for every i. ∎

Lemma 7.8 ([3]). *For every finite sets $P_1, P_2 \subseteq \mathbb{Z}^n$ there exists a finite set $P \subseteq \mathbb{Z}^n$ such that $P_1^* \cap P_2^* = P^*$. Moreover, for every $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{Z}^n$, there exists a finite set $B \subseteq \mathbb{Z}^n$ such that $(\mathbf{b}_1 + P_1^*) \cap (\mathbf{b}_2 + P_2^*) = B + (P_1^* \cap P_2^*)$.*

Proof. Let us consider an enumeration $\mathbf{p}_{i,1}, \ldots, \mathbf{p}_{i,k_i}$ of the $k_i \geq 0$ vectors in P_i where $i \in \{1, 2\}$. If $k_1 = 0$ or if $k_2 = 0$ then $P_1^* = \{\mathbf{0}\}$ or $P_2^* = \{\mathbf{0}\}$ and the lemma is immediate. Thus, we can assume that $k_1, k_2 \geq 1$.

Let us consider the set X of vectors $(\lambda_1, \lambda_2) \in \mathbb{N}^{k_1} \times \mathbb{N}^{k_2}$ such that $\mathbf{b}_1 + \sum_{j=1}^{k_1} \lambda_1[j]\mathbf{p}_{1,j} = \mathbf{b}_2 + \sum_{j=1}^{k_2} \lambda_2[j]\mathbf{p}_{2,j}$. Let us also consider the set X_0 of vectors $(\lambda_1, \lambda_2) \in \mathbb{N}^{k_1} \times \mathbb{N}^{k_2}$ such that $\sum_{j=1}^{k_1} \lambda_1[j]\mathbf{p}_{1,j} = \sum_{j=1}^{k_2} \lambda_2[j]\mathbf{p}_{2,j}$. Observe that $X = Z + X_0$ where Z is the finite set $Z = \min(X)$ and $X_0 = Z_0^*$ where Z_0 is the finite set $Z_0 = \min(X_0 \backslash \{\mathbf{0}\})$.

Let us denote by B the finite set of vectors $\mathbf{b} \in \mathbb{Z}^n$ such that there exists $(\lambda_1, \lambda_2) \in Z$ satisfying $\mathbf{b}_1 + \sum_{j=1}^{k_1} \lambda_1[j]\mathbf{p}_{1,j} = \mathbf{b} = \mathbf{b}_2 + \sum_{j=1}^{k_2} \lambda_2[j]\mathbf{p}_{2,j}$. Let us also denote by P the finite set of vectors $\mathbf{p} \in \mathbb{Z}^n$ such that there exists $(\lambda_1, \lambda_2) \in Z_0$ satisfying $\sum_{j=1}^{k_1} \lambda_1[j]\mathbf{p}_{1,j} = \mathbf{p} = \sum_{j=1}^{k_2} \lambda_2[j]\mathbf{p}_{2,j}$. Remark that $(\mathbf{b}_1 + P_1^*) \cap (\mathbf{b}_2 + P_2^*) = B + P^*$ and $P_1^* \cap P_2^* = P^*$. ∎

We say that two linear sets L_1, L_2 have a *non-degenerate intersection* if $\dim(L_1) = \dim(L_1 \cap L_2) = \dim(L_2)$.

Lemma 7.9. *Let $L_1 = \mathbf{b}_1 + M_1$ and $L_2 = \mathbf{b}_2 + M_2$ be two linear sets with a non-degenerate intersection. There exist finite sets $R_1 \subseteq \mathcal{I}(M_1)$ and $R_2 \subseteq \mathcal{I}(M_2)$ such that $(\mathbf{x}_1 + R_1^*) \cap (\mathbf{x}_2 + R_2^*) \neq \emptyset$ for every $(\mathbf{x}_1, \mathbf{x}_2) \in (L_1, L_2)$.*

Proof. As M_1, M_2 are finitely generated, there exists some finite sets $P_1, P_2 \subseteq \mathbb{Z}^n$ such that $M_1 = P_1^*$ and $M_2 = P_2^*$. From Lemma 7.8 there exists a finite set $P \subseteq \mathbb{Z}^n$ and a finite set $B \subseteq \mathbb{Z}^n$ such that $P_1^* \cap P_2^* = P^*$ and $L_1 \cap L_2 = B + P^*$. Note that $B = \emptyset$ is not possible since in this case $\dim(L_1 \cap L_2) = -\infty$. Thus there exists a vector $\mathbf{b} \in B$.

Let us denote by V_1, V, V_2 the vector spaces generated respectively by P_1, P, P_2 and let us prove that $V_1 = V = V_2$. Proposition 7.4 shows that $\dim(L_1) = \text{rank}(V_1)$, $\dim(L_1 \cap$

$L_2) = \mathrm{rank}(V)$ and $\dim(L_2) = \mathrm{rank}(V_2)$. From $\dim(L_1 \cap L_2) = \dim(L_1)$ we deduce that $\mathrm{rank}(V) = \mathrm{rank}(V_1)$. Moreover as $P^* \subseteq P_1^*$ we deduce that $V \subseteq V_1$. The inclusion $V \subseteq V_1$ and the relation $\mathrm{rank}(V) = \mathrm{rank}(V_1)$ prove together that $V = V_1$. Symmetrically we deduce that $V = V_2$.

We denote by G_1, G, G_2 the groups generated respectively by P_1, P, P_2. Note that the vector spaces generated by G_1, G, G_2 are equal to V_1, V, V_2.

Let $\mathbf{a}$ be an interior vector of P^* and let us prove that $\mathbf{a} \in \mathcal{I}(P_1^*) \cap \mathcal{I}(P_2^*)$. Let $j \in \{1, 2\}$. Note that $\mathbf{a} \in P^* \subseteq P_j^*$. Let $\mathbf{p} \in \mathcal{I}(P_j^*)$. Since $-\mathbf{p} \in V$ and V is the vector space generated by G, Lemma 7.7 shows that there exists an integer $d \geq 1$ such that $-d\mathbf{p} \in G$. From $\mathbf{a} \in \mathcal{I}(P^*)$ we deduce that there exists $N \geq 1$ such that $-d\mathbf{p} + N\mathbf{a} \in P^*$. From $P^* \subseteq P_j^*$ we deduce that $\mathbf{a} \in \frac{1}{N}(d\mathbf{p} + P_j^*)$. From $\mathbf{p} \in \mathcal{I}(P_j^*)$ and Lemma 4.2 we get $\mathbf{a} \in \mathcal{I}(P_j^*)$.

We define R_1 and R_2 by $R_j = \{\mathbf{a}\} \cup (\mathbf{a} + P_j)$ for $j \in \{1, 2\}$. Since $\mathbf{a} \in \mathcal{I}(P_j^*)$, Lemma 4.2 shows that $R_j \subseteq \mathcal{I}(P_j^*)$. Let us consider $\mathbf{x}_1 \in L_1$ and $\mathbf{x}_2 \in L_2$ and let us prove that $(\mathbf{x}_1 + R_1^*) \cap (\mathbf{x}_2 + R_2^*) \neq \emptyset$.

From $\mathbf{b}, \mathbf{x}_j \in \mathbf{b}_j + P_j^*$ we deduce that $\mathbf{x}_j - \mathbf{b} \in G_j$. As the group generated by R_j is equal to G_j, there exists $\mathbf{r}_j, \mathbf{r}_j' \in R_j^*$ such that $\mathbf{x}_j + \mathbf{r}_j = \mathbf{b} + \mathbf{r}_j'$.

As V is the vector space generated by G_1 and $\mathbf{r}_2' \in R_2^* \subseteq V_2 = V$, Lemma 7.7 shows that there exists an integer $d_1 \geq 1$ such that $d_1\mathbf{r}_2' \in G_1$. As $\mathbf{a} \in \mathcal{I}(P_1^*)$, there exists an integer $N_1 \geq 1$ such that $d_1\mathbf{r}_2' + N_1\mathbf{a} \in P_1^*$. As $P_1^* \subseteq R_1^* - \mathbb{N}\mathbf{a}$, we deduce that there exists an integer $N_1' \geq 0$ such that $d_1\mathbf{r}_2' + (N_1 + N_1')\mathbf{a} \in R_1^*$. We denote this vector by $\mathbf{r}_1''$. Symmetrically, there exist some integers $d_2 \geq 1$, $N_2 \geq 1$ and $N_2' \geq 0$ such that the vector $d_2\mathbf{r}_1' + (N_2 + N_2')\mathbf{a}$ denoted by $\mathbf{r}_2''$ is in R_2^*. We get:

$$\mathbf{x}_1 + \mathbf{r}_1 + (d_2 - 1)\mathbf{r}_1' + \mathbf{r}_1'' + (N_2 + N_2')\mathbf{a} = \mathbf{b} + d_2\mathbf{r}_1' + d_1\mathbf{r}_2' + (N_1 + N_1' + N_2 + N_2')\mathbf{a}$$

$$\mathbf{x}_2 + \mathbf{r}_2 + (d_1 - 1)\mathbf{r}_2' + \mathbf{r}_2'' + (N_1 + N_1')\mathbf{a} = \mathbf{b} + d_1\mathbf{r}_2' + d_2\mathbf{r}_1' + (N_2 + N_2' + N_1 + N_1')\mathbf{a}$$

We have proved that these vectors are equal. Therefore $(\mathbf{x}_1 + R_1^*) \cap (\mathbf{x}_2 + R_2^*) \neq \emptyset$. $\qquad\square$

Proposition 7.10. *Let L_1, L_2 be linearizations of pseudo-linear sets $X_1, X_2 \subseteq \mathbb{Z}^n$ with an empty intersection $X_1 \cap X_2 = \emptyset$. We have:*

$$\dim(L_1 \cap L_2) < \dim(X_1 \cup X_2)$$

Proof. Let us consider linearizations L_1, L_2 of two pseudo-linear sets X_1, X_2 such that $\dim(L_1 \cap L_2) \geq \dim(X_1 \cup X_2)$ and let us prove that $X_1 \cap X_2 \neq \emptyset$. Lemma 7.5 shows that $\dim(X_1) = \dim(L_1)$ and $\dim(X_2) = \dim(L_2)$. By monotonicity of the dimension function, we deduce that $\dim(L_1) = \dim(L_1 \cap L_2) = \dim(L_2)$. Thus L_1 and L_2 have a non-degenerate intersection. As L_1, L_2 are two linear sets, there exists $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{Z}^n$ and two finitely generated monoids M_1, M_2 such that $L_1 = \mathbf{b}_1 + M_1$ and $L_2 = \mathbf{b}_2 + M_2$. Lemma 7.9 shows that there exist finite sets $R_1 \subseteq \mathcal{I}(M_1)$ and $R_2 \subseteq \mathcal{I}(M_2)$ such that $(\mathbf{x}_1 + R_1^*) \cap (\mathbf{x}_2 + R_2^*) \neq \emptyset$ for every $(\mathbf{x}_1, \mathbf{x}_2) \in (L_1, L_2)$. As L_1, L_2 are linearizations of the pseudo-linear sets X_1, X_2 there exists $(\mathbf{x}_1, \mathbf{x}_2) \in (X_1, X_2)$ such that $\mathbf{x}_1 + R_1^* \subseteq X_1$ and $\mathbf{x}_2 + R_2^* \subseteq X_2$. As $(\mathbf{x}_1, \mathbf{x}_2) \in (L_1, L_2)$ we deduce that $(\mathbf{x}_1 + R_1^*) \cap (\mathbf{x}_2 + R_2^*) \neq \emptyset$. We have proved that $X_1 \cap X_2 \neq \emptyset$. $\qquad\square$

8. Presburger Complete Separators

The VAS reachability problem can be reformulated by introducing the definition of separators. A pair (S, S') of configuration sets is called a *separator* for a VAS $\mathcal{V}$ if $S \times S'$ has an empty intersection with the reachability binary relation $\xrightarrow{*}_{\mathcal{V}}$. The set $D = \mathbb{N}^n \backslash (S \cup S')$ is called the *(free) domain* of (S, S'). A separator with an empty domain is said to be *complete*. We extend the inclusion relation over separators by $(S_0, S_0') \subseteq (S, S')$ if $S_0 \subseteq S$ and $S_0' \subseteq S'$.

Complete separators can be characterized by introducing the forward and backward invariants. Let us consider the following sets for every pair (S, S') of configurations sets and for every $a \in \Sigma$:

$$\mathrm{post}_{\mathcal{V}}^a(S) = \{\mathbf{s}' \in \mathbb{N}^n \mid \exists \mathbf{s} \in S \quad \mathbf{s} \xrightarrow{a}_{\mathcal{V}} \mathbf{s}'\}$$
$$\mathrm{pre}_{\mathcal{V}}^a(S') = \{\mathbf{s} \in \mathbb{N}^n \mid \exists \mathbf{s}' \in S' \quad \mathbf{s} \xrightarrow{a}_{\mathcal{V}} \mathbf{s}'\}$$

A set $S \subseteq \mathbb{N}^n$ is called a *forward invariant* if $\mathrm{post}_{\mathcal{V}}^a(S) \subseteq S$ for every $a \in \Sigma$. A set $S' \subseteq \mathbb{N}^n$ is called a *backward invariant* if $\mathrm{pre}_{\mathcal{V}}^a(S') \subseteq S'$ for every $a \in \Sigma$. Note that a pair (S, S') of configuration sets is a complete separator if and only if (S, S') is a partition of $\mathbb{N}^n$, S is a forward invariant and S' is a backward invariant.

In this section we prove that Presburger separators are included in Presburger complete separators. In general $(\mathrm{post}_{\mathcal{V}}^*(S), \mathrm{pre}_{\mathcal{V}}^*(S'))$ is a separator that is neither complete nor Presburger (see Example 2.4). That means, this separator must be over-approximated by another one.

Remark 8.1. *In the sequel, we often use the fact that a pair (S, S') of subsets of $\mathbb{N}^n$ is a separator if and only if $\mathrm{post}_{\mathcal{V}}^*(S) \cap \mathrm{pre}_{\mathcal{V}}^*(S') = \emptyset$ if and only if $\mathrm{post}_{\mathcal{V}}^*(S) \cap S' = \emptyset$ if and only if $S \cap \mathrm{pre}_{\mathcal{V}}^*(S') = \emptyset$.*

Lemma 8.2. *Let (S_0, S_0') be a Presburger separator with a non-empty domain D_0. There exists a Presburger separator (S, S') with a domain D such that $S_0 \subseteq S$, $S_0' \subseteq S'$, and such that:*

$$\dim(D) < \dim(D_0)$$

Proof. We first define a set S' that over-approximates S_0' and such that (S_0, S') is a separator. As S_0 is semi-linear, Theorem 6.4 shows that $\mathrm{post}_{\mathcal{V}}^*(S_0)$ is a Petri set. As D_0 is semi-linear, we deduce that $\mathrm{post}_{\mathcal{V}}^*(S_0) \cap D_0$ is equal to a finite union of pseudo-linear sets $X_1, \ldots, X_k$. Let us consider some linearizations $L_1, \ldots, L_k$ of these pseudo-linear sets and let us define the following Presburger set S'.

$$S' = S_0' \cup (D_0 \backslash (\bigcup_{j=1}^k L_j))$$

We observe that $\mathrm{post}_{\mathcal{V}}^*(S_0) \cap S' = \emptyset$ since $\mathrm{post}_{\mathcal{V}}^*(S_0) \cap S_0' = \emptyset$ and $\mathrm{post}_{\mathcal{V}}^*(S_0) \cap D_0 \subseteq \bigcup_{j=1}^k L_j$. We have proved that S' contains S_0' and (S_0, S') is a separator.

Now we define symmetrically a set S that over-approximates S_0 and such that (S, S') is a separator. As S' is semi-linear, Theorem 6.4 shows that $\mathrm{pre}_{\mathcal{V}}^*(S')$ is a Petri set. As D_0 is semi-linear we deduce that $D_0 \cap \mathrm{pre}_{\mathcal{V}}^*(S')$ is equal to a finite union of pseudo-linear sets

$X'_1, \ldots, X'_{k'}$. Let us consider some linearizations $L'_1, \ldots, L'_{k'}$ of these pseudo-linear sets and let us define the following Presburger set S.

$$S = S_0 \cup (D_0 \backslash (\bigcup_{j'=1}^{k'} L'_{j'}))$$

Once again, note that $S \cap \mathrm{pre}^*_{\mathcal{V}}(S') = \emptyset$. Thus S contains S_0 and (S, S') is a separator.

Let D be the domain of the separator (S, S'). From $D_0 = \mathbb{N}^n \backslash (S_0 \cup S'_0)$, we get the following equality:

$$D = D_0 \cap \left(\bigcup_{\substack{1 \leq j \leq k \\ 1 \leq j' \leq k'}} (L_j \cap L'_{j'}) \right)$$

From $X_j, X'_{j'} \subseteq D_0$ we get $\dim(X_j \cup X'_{j'}) \leq \dim(D_0)$. As $X_j \subseteq \mathrm{post}^*_{\mathcal{V}}(S_0) \subseteq \mathrm{post}^*_{\mathcal{V}}(S)$ and $X'_{j'} \subseteq \mathrm{pre}^*_{\mathcal{V}}(S')$ and (S, S') is a separator, we deduce that X_j and $X'_{j'}$ are two pseudo-linear sets with an empty intersection. Proposition 7.10 provides $\dim(L_j \cap L'_{j'}) < \dim(X_j \cup X'_{j'})$. We deduce $\dim(D) < \dim(D_0)$. $\qquad\square$

An induction over the dimension of the domain D of a Presburger separator provides the following Theorem 8.3 thanks to Lemma 8.2.

Theorem 8.3. *Presburger separators are included in Presburger complete separators.*

As $(\{\mathbf{s}\}, \{\mathbf{s}'\})$ is a Presburger separator if $(\mathbf{s}, \mathbf{s}') \notin \xrightarrow{*}_{\mathcal{V}}$, the previous theorem shows that there exists a Presburger complete separator (S, S') that contains $(\{\mathbf{s}\}, \{\mathbf{s}'\}$. By considering $I = S$, the following Corollary 8.4 is proved.

Corollary 8.4. *Let (s, s') be a pair of configurations of a VAS $\mathcal{V}$. We have $(\mathbf{s}, \mathbf{s}') \notin \xrightarrow{*}_{\mathcal{V}}$ if and only if there exists a Presburger formula denoting a forward invariant I such that $\mathbf{s} \in I$ and $\mathbf{s}' \notin I$.*

9. Conclusion

Thanks to the classical KLMST decomposition we have proved that the Parikh Images of languages recognized by VASs are semi-pseudo-linear. As an application, we have proved that for every pair $(\mathbf{s}, \mathbf{s}')$ of configurations in the complement of the reachability relation there exists a Presburger formula $\psi(\mathbf{x})$ denoting a forward invariant I such that $\mathbf{s} \in I$ and $\mathbf{s}' \notin I$. We deduce that the following algorithm decides the reachability problem.

```
1  Reachability( s , V , s' )
2      k ← 0
3      repeat forever
4          for each word σ ∈ Σᵏ
5              if s ─σ→_V s'
6                  return "reachable"
7          for each Presburger formula ψ(x) of length k
8              if ψ(s) and ¬ψ(s') are true and
9                  ψ(x) ∧ y = x + δ(a) ∧ ¬ψ(y) unsat ∀a ∈ Σ
```

```
10              return ''unreachable''
11          k ← k + 1
```

The correctness is immediate and the termination is guaranteed by Corollary 8.4. This algorithm is the *very first one* that does not require the classical KLMST decomposition for its implementation. Even though the termination proof is based on the KLMST decomposition, the complexity of the algorithm does not depend on this decomposition. In fact, the complexity depends on the minimal size of a word $\sigma \in \Sigma^*$ such that $\mathbf{s} \xrightarrow{\sigma}_\mathcal{V} \mathbf{s}'$ if $\mathbf{s} \xrightarrow{*}_\mathcal{V} \mathbf{s}'$, and the minimal size of a Presburger formula $\psi(\mathbf{x})$ denoting a forward invariant I such that $\mathbf{s} \in I$ and $\mathbf{s}' \notin I$ otherwise. We left as an open question the problem of computing lower and upper bounds for these sizes. Note that the VAS exhibiting a large (Ackermann size) but finite reachability set given in [8] does not directly provide an Ackerman lower-bound for these sizes since inductive separators can over-approximate reachability sets.

We also left as an open question the problem of adapting the *Counter Example Guided Abstraction Refinement* approach [1] to obtain an algorithm for the VAS reachability problem with termination guarantee. In practice, such an algorithm should be more efficient than the previously given enumeration-based algorithm.

Acknowledgment

References

[1] E. M. Clarke, O. Grumberg, S. Jha, Y. Lu, and H. Veith. Counterexample-guided abstraction refinement. In E. A. Emerson and A. P. Sistla, editors, *Computer Aided Verification, 12th International Conference, CAV 2000, Chicago, IL, USA, July 15-19, 2000, Proceedings*, volume 1855 of *Lecture Notes in Computer Science*, pages 154–169. Springer, 2000.

[2] J. Esparza and M. Nielsen. Decidability issues for Petri nets - a survey. *Bulletin of the European Association for Theoretical Computer Science*, 52:245–262, 1994.

[3] S. Ginsburg and E. H. Spanier. Semigroups, Presburger formulas and languages. *Pacific Journal of Mathematics*, 16(2):285–296, 1966.

[4] J. E. Hopcroft and J.-J. Pansiot. On the reachability problem for 5-dimensional vector addition systems. *Theoritical Computer Science*, 8:135–159, 1979.

[5] S. R. Kosaraju. Decidability of reachability in vector addition systems (preliminary version). In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing, (STOC 1982), 5-7 May 1982, San Francisco, California, USA*, pages 267–281. ACM, 1982.

[6] J. L. Lambert. A structure to decide reachability in Petri nets. *Theoretical Computer Science*, 99(1):79–104, 1992.

[7] E. W. Mayr. An algorithm for the general Petri net reachability problem. In *Conference Proceedings of the Thirteenth Annual ACM Symposium on Theory of Computation, (STOC 1981), 11-13 May 1981, Milwaukee, Wisconsin, USA*, pages 238–246. ACM, 1981.

[8] E. W. Mayr and A. R. Meyer. The complexity of the finite containment problem for Petri nets. *J. ACM*, 28(3):561–576, 1981.

[9] C. Rackoff. The covering and boundedness problems for vector addition systems. *Theoretical Computer Science*, 6(2), 1978.

[10] G. S. Sacerdote and R. L. Tenney. The decidability of the reachability problem for vector addition systems (preliminary version). In *Conference Record of the Ninth Annual ACM Symposium on Theory of Computing, 2-4 May 1977, Boulder, Colorado, USA*, pages 61–76. ACM, 1977.

APPENDIX A. PROOFS OF PROPOSITION 3.6

An MRGS is said to be *original-perfect* if it satisfies the large solution condition and its marked reachability graphs satisfy the input and output loop conditions.

Even if the proof of the following lemma is immediate by induction over the length of w, w', it is central in the KLMST decomposition.

Lemma A.1 (Continuity)**.**

- *For every* $\mathbf{x} \xrightarrow{w}_{\mathcal{V}}$ *there exists an integer* $c \geq 0$ *such that* $\mathbf{y} \xrightarrow{w}_{\mathcal{V}}$ *for every extended configuration* $\mathbf{y}$ *satisfying* $\mathbf{y}[i] \geq c$ *if* $\mathbf{x}[i] = \top$ *and* $\mathbf{y}[i] = \mathbf{x}[i]$ *otherwise for every* i.
- *For every* $\xrightarrow{w'}_{\mathcal{V}} \mathbf{x}'$ *there exists an integer* $c' \geq 0$ *such that* $\xrightarrow{w'}_{\mathcal{V}} \mathbf{y}'$ *for every extended configuration* $\mathbf{y}'$ *satisfying* $\mathbf{y}'[i] \geq c'$ *if* $\mathbf{x}'[i] = \top$ *and* $\mathbf{y}'[i] = \mathbf{x}'[i]$ *otherwise for every* i.

Lemma A.2. *Perfect MRGSs are original-perfect.*

Proof. Let us consider a perfect MRGS $\mathcal{U}$. Notice that $\mathcal{U}$ satisfies the large solution condition since from every accepted sequence $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_j$ we deduce a solution $(\mathbf{s}_j, ||\pi_j||, \mathbf{s}'_j)_c$. Since the input loop condition and the output loop condition are symmetrical, we just prove that the marked reachability graph $\mathcal{M}_j$ satisfies the input loop condition. We consider an integer $c \in \mathbb{N}$ satisfying $c > \mathbf{m}_j[i]$ for every i such that $\mathbf{m}_j[i] < \mathbf{x}_j[i]$. Since $\mathcal{U}$ is perfect, there exists an accepted sequence $(\mathbf{s}_j, \pi_j, \mathbf{s}'_j)_{0 \leq j \leq k}$, a prefix $\mathbf{x}_j \xrightarrow{w_j}_{\mathcal{V}} \mathbf{x}_j$ of π_j, an extended configuration $\mathbf{r}_j$ such that $\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}} \mathbf{r}_j$ and such that $\mathbf{r}_j[i] \geq c$ for every i such that $\mathbf{x}_j[i] = \top$. Since $\mathbf{s}_j \trianglelefteq \mathbf{m}_j$ we deduce that $\mathbf{s}_j \leq \mathbf{m}_j$. As $\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}} \mathbf{r}_j$ and $\mathbf{s}_j \leq \mathbf{m}_j$ there exists an extended configuration $\mathbf{y}_j$ such that $\mathbf{m}_j \xrightarrow{w_j}_{\mathcal{V}} \mathbf{y}_j$. Let us prove that $\mathbf{y}_j[i] > \mathbf{m}_j[i]$ for every i such that $\mathbf{m}_j[i] < \mathbf{x}_j[i]$. Let i be such an integer. Since $\mathbf{m}_j \trianglelefteq \mathbf{x}_j$ and $\mathbf{m}_j[i] < \mathbf{x}_j[i]$ we deduce that $\mathbf{m}_j[i] \in \mathbb{N}$ and $\mathbf{x}_j[i] = \top$. From $\mathbf{x}_j[i] = \top$ we deduce that $\mathbf{r}_j[i] \geq c$. From $\mathbf{m}_j[i] \in \mathbb{N}$ we deduce that $\mathbf{s}_j[i] = \mathbf{m}_j[i]$. Thus $\mathbf{y}_j[i] = \mathbf{r}_j[i] \geq c > \mathbf{s}_j[i] = \mathbf{m}_j[i]$. Lemma 3.3 shows that $\mathcal{M}_j$ satisfies the input loop condition. $\square$

Now, let us consider an original-perfect MRGS $\mathcal{U}$ and let us prove that $\mathcal{U}$ is perfect. Since $\mathcal{M}_j$ satisfies the input and output loop conditions, Lemma 3.3 and Lemma 3.4 show that:

- there exist a cycle $\theta_j = (\mathbf{x}_j \xrightarrow{w_j}_{G_j} \mathbf{x}_j)$ and an extended configuration $\mathbf{y}_j$ satisfying both $\mathbf{m}_j \xrightarrow{w_j}_{\mathcal{V}} \mathbf{y}_j$ and $\mathbf{y}_j[i] > \mathbf{m}_j[i]$ for every i such that $\mathbf{m}_j[i] < \mathbf{x}_j[i]$,
- there exist a cycle $\theta'_j = (\mathbf{x}'_j \xrightarrow{w'_j}_{G_j} \mathbf{x}'_j)$ and an extended configuration $\mathbf{y}'_j$ satisfying both $\mathbf{y}'_j \xrightarrow{w'_j}_{\mathcal{V}} \mathbf{m}'_j$ and $\mathbf{y}'_j[i] > \mathbf{m}'_j[i]$ for every i such that $\mathbf{m}'_j[i] < \mathbf{x}'_j[i]$.

The proof that $\mathcal{U}$ is perfect is obtained by first exhibiting a solution ξ with components in $\mathbb{N}$ of the characteristic system and a solution ξ_0 with components in $\mathbb{N}$ of the homogeneous characteristic system satisfying some particular properties. These two solutions ξ and ξ_0 are respectively defined in Lemma A.3 and Lemma A.4.

Lemma A.3. *There exists a solution* $\xi = (\mathbf{s}_j, \mu_j, \mathbf{s}'_j)_j$ *of the characteristic system such that for every* j:

- $\mathbf{s}_j$ *is a configuration satisfying* $\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}}$,

- μ_j *is the Parikh image of a path* $\pi_j = (\mathbf{x}_j \xrightarrow{\sigma_j}_{G_j} \mathbf{x}'_j)$,
- $\mathbf{s}'_j$ *is a configuration satisfying* $\xrightarrow{w'_j}_{\mathcal{V}} \mathbf{s}'_j$.

Proof. As $\mathbf{m}_j \xrightarrow{w_j}_{\mathcal{V}}$, Lemma A.1 shows that there exists an integer $c \geq 0$ such that $\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}}$ for every configuration $\mathbf{s}_j$ satisfying $\mathbf{s}_j[i] \geq c$ if $\mathbf{m}_j[i] = \top$ and $\mathbf{s}_j[i] = \mathbf{m}_j[i]$ otherwise for every i. Symmetrically, as $\xrightarrow{w'_j}_{\mathcal{V}} \mathbf{m}'_j$, Lemma A.1 shows that there exists an integer $c' \geq 0$ such that $\xrightarrow{w'_j}_{\mathcal{V}} \mathbf{s}'_j$ for every configuration $\mathbf{s}'_j$ satisfying $\mathbf{s}'_j[i] \geq c'$ if $\mathbf{m}'_j[i] = \top$ and $\mathbf{s}'_j[i] = \mathbf{m}'_j[i]$ otherwise for every i. Since $\mathcal{U}$ satisfies the large solution condition there exists a solution $\xi = (\mathbf{s}_j, \mu_j, \mathbf{s}'_j)_j$ with components in $\mathbb{N}$ of the characteristic system such that $\mathbf{s}_j$ and $\mathbf{s}'_j$ satisfies the previous conditions and such that $\mu_j(t) \geq 1$ for every $t \in T_j$. As G_j is strongly connected, Euler's Lemma shows that μ_j is the Parikh image of a path $\pi_j = (\mathbf{x}_j \xrightarrow{\sigma_j}_{G_j} \mathbf{x}'_j)$. $\square$

Lemma A.4. *There exists a solution* $\xi_0 = (\mathbf{s}_{0,j}, \mu_{0,j}, \mathbf{s}'_{0,j})$ *of the homogeneous characteristic system such that for every* j:
- *the value* $\mathbf{s}_{0,j}[i]$ *is strictly positive if* $\mathbf{m}_j[i] = \top$ *and it is equal to* 0 *otherwise for every* i,
- *the value* $(\mathbf{s}_{0,j} + \delta(w_j))[i]$ *is strictly positive if* $\mathbf{x}_j[i] = \top$ *and it is equal to* 0 *otherwise for every* i,
- $\mu_{0,j} - (\|\theta_j\| + \|\theta'_j\|)$ *is the Parikh image of a cycle* $\pi_{0,j} = (\mathbf{x}_j \xrightarrow{\sigma_{0,j}}_{G_j} \mathbf{x}_j)$ *and* $|\pi_{0,j}|_t > 0$ *for every* $t \in T_j$,
- *the value* $(\mathbf{s}'_{0,j} - \delta(w'_j))[i]$ *is strictly positive if* $\mathbf{x}'_j[i] = \top$ *and it is equal to* 0 *otherwise for every* i, *and*
- *the value* $\mathbf{s}'_{0,j}[i]$ *is strictly positive if* $\mathbf{m}'_j[i] = \top$ *and it is equal to* 0 *otherwise for every* i.

Proof. As $\mathcal{U}$ satisfies the large solution condition, Lemma 3.5 shows that there exists a solution $\xi_0 = (\mathbf{s}_{0,j}, \mu_{0,j}, \mathbf{s}'_{0,j})_j$ with components in $\mathbb{Q}$ of the homogeneous characteristic system satisfying the additional constraints $\mathbf{s}_{0,j}[i] > 0$ if $\mathbf{m}_j[i] = \top$, $\mathbf{s}'_{0,j}[i] > 0$ if $\mathbf{m}'_j[i] = \top$, and $\mu_{0,j}(t) > 0$ for every $t \in T_j$. By multiplying ξ_0 by a positive integer, we can assume that ξ_0 is a solution with components in $\mathbb{Z}$ satisfying the additional constraints. We are going to prove that there exists a positive integer $c \geq 1$ such that $c\xi_0$ satisfies the lemma.

First of all, observe that for every $c \geq 1$ and for every j:
- the value $c\mathbf{s}_{0,j}[i]$ is strictly positive if $\mathbf{m}_j[i] = \top$ and it is equal to 0 otherwise for every i,
- the value $c\mathbf{s}'_{0,j}[i]$ is strictly positive if $\mathbf{m}'_j[i] = \top$ and it is equal to 0 otherwise for every i.

Let us consider $1 \leq i \leq n$. Let us prove that there exists a positive integer $c_i \geq 1$ such that for every $c \geq c_i$ the value $(c\mathbf{s}_{0,j} + \delta(w_j))[i]$ is strictly positive if $\mathbf{x}_j[i] = \top$ and it is equal to 0 otherwise. Note that $\mathbf{m}_j[i] \trianglelefteq \mathbf{x}_j[i]$ thus either $\mathbf{m}_j[i] = \mathbf{x}_j[i] \in \mathbb{N}$, or $(\mathbf{m}_j[i], \mathbf{x}_j[i]) \in \mathbb{N} \times \{\top\}$, or $\mathbf{m}_j[i] = \mathbf{x}_j[i] = \top$. We separate the proof following these three cases. Let us first consider the case $\mathbf{m}_j[i] = \mathbf{x}_j[i] \in \mathbb{N}$. As $\mathbf{m}_j[i] \in \mathbb{N}$ and ξ_0 is a solution of the homogeneous characteristic system, we get $\mathbf{s}_{0,j}[i] = 0$. The cycle θ_j shows that $\mathbf{x}_j + \delta(w_j) = \mathbf{x}_j$. From $\mathbf{x}_j[i] \in \mathbb{N}$ we deduce that $\delta(w_j)[i] = 0$. In particular $(c\mathbf{s}_{0,j} + \delta(w_j))[i] = 0$ and we have proved the case $\mathbf{m}_j[i] = \mathbf{x}_j[i] \in \mathbb{N}$ by considering $c_i = 1$. Let us consider the second case $(\mathbf{m}_j[i], \mathbf{x}_j[i]) \in \mathbb{N} \times \{\top\}$. As $\mathbf{m}_j[i] \in \mathbb{N}$ we deduce that $\mathbf{s}_{0,j}[i] = 0$. Since $\mathbf{m}_j[i] < \mathbf{x}_j[i]$ the condition satisfied by the loop θ_j shows that $\mathbf{y}_j[i] > \mathbf{m}_j[i]$. As $\mathbf{y}_j[i] = \mathbf{m}_j[i] + \delta(w_j)[i]$, we deduce that $\delta(w_j)[i] > 0$. In particular for every $c \geq 1$ we have $(c\mathbf{s}_{0,j} + \delta(w_j))[i] > 0$ and we have proved the case $(\mathbf{m}_j[i], \mathbf{x}_j[i]) \in \mathbb{N} \times \{\top\}$ by considering $c_i = 1$. Finally, let us consider

the case $\mathbf{m}_j[i] = \mathbf{x}_j[i] = \top$. As $\mathbf{m}_j[i] = \top$ we deduce that $\mathbf{s}_{0,j}[i] > 0$ in particular there exists an integer $c_i \geq 1$ large enough such that $(c\mathbf{s}_{0,j} + \delta(w_j))[i] > 0$ for every $c \geq c_i$. We have proved the three cases.

Symmetrically, for every $1 \leq i \leq n$, there exists an integer $c_i' \geq 0$ such that for every $c \geq c_i'$ the value $(c\mathbf{s}_{0,j}' - \delta(w_j'))[i]$ is strictly positive if $\mathbf{x}_j'[i] = \top$ and it is equal to 0 otherwise.

Finally, as $\mu_{0,j}(t) > 0$ for every $t \in T_j$ and for every $0 \leq j \leq k$, we deduce that there exists an integer $c \geq 1$ large enough such that $c\mu_{0,j}(t) > |\theta_j|_t + |\theta_{j'}|_t$ for every $t \in T_j$ and for every $0 \leq j \leq k$. Naturally, we can also assume that $c \geq c_i$ and $c \geq c_i'$ for every $1 \leq i \leq n$. Let us replace ξ_0 by $c\xi_0$. As $\mu_{0,j}(t) - |\theta_j|_t + |\theta_{j'}|_t > 0$ for every $t \in T_j$, Euler's Lemma shows that $\mu_{0,j} - (\|\theta_j\| + \|\theta_j'\|)$ is the Parikh image of a cycle $\pi_{0,j} = (\mathbf{x}_j \xrightarrow{\sigma_{0,j}}_{G_j} \mathbf{x}_j)$. $\square$

Let us fix notations satisfying both Lemma A.3 and Lemma A.4. We now provide technical lemmas that prove together that $\mathcal{U}$ is perfect.

Lemma A.5. *For every $c \geq 0$ we have:*

$$\mathbf{s}_j + c\mathbf{s}_{0,j} \quad \xrightarrow{w_j^c}_{\mathcal{V}} \quad \mathbf{s}_j + c(\mathbf{s}_{0,j} + \delta(w_j))$$

$$\mathbf{s}_j' + c(\mathbf{s}_{0,j}' - \delta(w_j')) \quad \xrightarrow{(w_j')^c}_{\mathcal{V}} \quad \mathbf{s}_j' + c\mathbf{s}_{0,j}'$$

Proof. Since the two relations are symmetrical, we just prove the first one. The choice of ξ satisfying Lemma A.3 shows that $\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}}$. Let us consider $c \in \mathbb{N}$ and let us prove by induction over c' that for every $0 \leq c' \leq c$ we have:

$$\mathbf{s}_j + c\mathbf{s}_{0,j} \xrightarrow{w_j^{c'}}_{\mathcal{V}} \mathbf{s}_j + (c - c')\mathbf{s}_{0,j} + c'(\mathbf{s}_{0,j} + \delta(w_j))$$

Naturally, the case $c' = 0$ is immediate. The induction is obtained just by observing that $\mathbf{s}_{0,j} \geq \mathbf{0}$, $\mathbf{s}_{0,j} + \delta(w_j) \geq \mathbf{0}$ and $\mathbf{s}_j \xrightarrow{w_j}_{\mathcal{V}}$. $\square$

Lemma A.6. *There exists $c_0 \geq 0$ such that for every $c \geq c_0$:*

$$\mathbf{s}_j + c(\mathbf{s}_{0,j} + \delta(w_j)) \quad \xrightarrow{\sigma_{0,j}^c}_{\mathcal{V}} \quad \mathbf{s}_j + c(\mathbf{s}_{0,j}' - \delta(w_j'))$$

Proof. Since there exists a path in G_j from $\mathbf{x}_j$ to $\mathbf{x}_j'$ we deduce that $\mathbf{x}_j[i] = \top$ if and only if $\mathbf{x}_j'[i] = \top$. We denote by $\mathbf{u}_j$ the vector in $\{0,1\}^n$ satisfying $\mathbf{u}_j[i] = 1$ if $\mathbf{x}_j[i] = \top = \mathbf{x}_j'[i]$ and satisfying $\mathbf{u}_j[i] = 0$ otherwise. From the choice of ξ_0 satisfying Lemma A.4, we observe that $\mathbf{s}_{0,j} + \delta(w_j) \geq \mathbf{u}_j$ and $\mathbf{s}_{0,j}' - \delta(w_j') \geq \mathbf{u}_j$. Note that $\lim_{c \to +\infty}(\mathbf{s}_j + c\mathbf{u}_j) = \mathbf{x}_j$. As $\mathbf{x}_j \xrightarrow{\sigma_{0,j}}_{G_j} \mathbf{x}_j$, Lemma A.1 proves that there exists an integer $c_0 \geq 0$ such that $\mathbf{s}_j + c_0\mathbf{u}_j \xrightarrow{\sigma_{0,j}}_{\mathcal{V}}$. Now, let us consider an integer $c \geq c_0$. Let us prove by induction over c' that for every $0 \leq c' \leq c$, we have:

$$\mathbf{s}_j + c(\mathbf{s}_{0,j} + \delta(w_j))$$

$$\xrightarrow{\sigma_{0,j}^{c'}}_{\mathcal{V}}$$

$$\mathbf{s}_j + (c - c')(\mathbf{s}_{0,j} + \delta(w_j)) + c'(\mathbf{s}_{0,j}' - \delta(w_j'))$$

Naturally, the case $c' = 0$ is immediate. Assume the previous relation holds for an integer c' such that $0 \leq c' < c$ and let us consider $c'' = c' + 1$. From $\mathbf{s}_{0,j} + \delta(w_j) \geq \mathbf{u}_j$ and

$\mathbf{s}'_{0,j} - \delta(w'_j) \geq \mathbf{u}_j$ we deduce that $(c-c')(\mathbf{s}_{0,j}+\delta(w_j)) + c'(\mathbf{s}'_{0,j} - \delta(w'_j)) \geq c\mathbf{u}_j \geq c_0\mathbf{u}_j$. Thus, the induction directly comes from $\mathbf{s}_j + c_0\mathbf{u}_j \xrightarrow{\sigma_{0,j}}_{\mathcal{V}}$ and $\mathbf{s}_{0,j} + \delta(w_j) + \delta(\sigma_{0,j}) + \delta(w'_j) = \mathbf{s}'_{0,j}$. $\square$

Lemma A.7. *There exists $c' \geq 0$ such that for every $c \geq c'$:*

$$\mathbf{s}_j + c(\mathbf{s}'_{0,j} - \delta(w'_j)) \quad \xrightarrow{\sigma_j}_{\mathcal{V}} \quad \mathbf{s}'_j + c(\mathbf{s}'_{0,j} - \delta(w'_j))$$

Proof. As $\lim_{c \to +\infty}(\mathbf{s}'_j + c(\mathbf{s}'_{0,j} - \delta(w'_j))) = \mathbf{x}'_j$ and $\mathbf{x}_j \xrightarrow{\sigma_j}_{G_j} \mathbf{x}'_j$, Lemma A.1 proves that there exists $c' \geq 0$ such that $\xrightarrow{\sigma_j}_{\mathcal{V}} (\mathbf{s}'_j + c(\mathbf{s}'_{0,j} - \delta(w'_j)))$ for every $c \geq c'$. Since $\mathbf{s}_j + \delta(\sigma_j) = \mathbf{s}'_j$ we are done. $\square$

Now, let us consider an integer $c \geq 0$ satisfying $c \geq c_0$ and $c \geq c'$ where c_0 and c' are respectively defined by Lemma A.6 and Lemma A.7. For each $0 \leq j \leq k$, we consider the following path:

$$\pi_{j,c} = (\mathbf{x}_j \xrightarrow{w_j^c}_{G_j} \mathbf{x}_j \xrightarrow{\sigma_{0,j}^c \sigma_j}_{G_j} \mathbf{x}'_j \xrightarrow{(w'_j)^c}_{G_j} \mathbf{x}'_j)$$

We have proved that $(\mathbf{s}_j + c\mathbf{s}_{0,j}, \pi_{j,c}, \mathbf{s}'_j + c\mathbf{s}'_{0,j})_j$ is an accepted sequence for $\mathcal{U}$. Thus $\mathcal{U}$ is perfect.