The General Vector Addition System Reachability Problem by Presburger Inductive Separators

Jérôme Leroux

The General Vector Addition System Reachability Problem by Presburger Inductive Separators

Jérôme Leroux
Laboratoire Bordelais de Recherche en Informatique
CNRS UMR 5800, Talence, France
leroux@labri.fr

ABSTRACT

The reachability problem for Vector Addition Systems (VAS) or equivalently for Petri Nets is a central problem of net theory. The general problem is known decidable by algorithms exclusively based on the classical Kosaraju-Lambert-Mayr-Sacerdote-Tenney (KLMST) decomposition. This decomposition is difficult and it just has a non-primitive recursive upper-bound complexity. In this paper, we prove that if a final configuration is not reachable from an initial configuration, there exists a pair of Presburger formulas that denotes an inductive separator proving this property. We deduce an easy algorithm for deciding the reachability problem based on two semi-algorithms. A first one that tries to prove the reachability by fairly enumerating the possible paths and a second one that tries to prove the non-reachability by fairly enumerating pairs of Presburger formulas denoting inductive separators. This algorithm is the very first one that does not require the KLMST decomposition. In particular, this algorithm is the first candidate to obtain a precise (eventually elementary) upper-bound complexity for the VAS reachability problem.

1. INTRODUCTION

Vector Addition Systems (VAS) 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 the decidability of this problem. The proof was completed in 1981 by Mayr [9] and simplified by Kosaraju [7] from [10, 9]. Ten years later [8], Lambert provided a more simplified version based on [7]. This last proof still remains difficult and the upper-bound complexity of the corresponding algorithm is just known non-primitive recursive. Nowadays, the exact complexity of the reachability problem for VAS is still an open-problem. Even an elementary upper-bound complexity is open. In fact, the known general reachability algorithms are exclusively based on the Kosaraju-Lambert-Mayr-Sacerdote-Tenney (KLMST) decomposition.

In this paper, we prove that if a final configuration is not reachable from an initial configuration, there exists a we prove that if a configuration is not reachable from an initial configuration, there exists a pair of Presburger formulas that denotes an inductive separator proving this property. We deduce an easy algorithm for deciding the reachability problem based on two semi-algorithms. A first one that tries to prove the reachability by fairly enumerating the possible paths and a second one that tries to prove the non-reachability by fairly enumerating by fairly enumerating pairs of Presburger formulas denoting inductive separators. This algorithm is the very first one that does not require the KLMST decomposition. In particular, this algorithm should be a good candidate to obtain a precise (eventually elementary) upper-bound complexity for the VAS reachability problem. Note [5] that in general, reachability sets are not semi-linear. Semi-linear inductive invariants are obtained by observing that reachability sets can be precisely over-approximated by semi-linear sets.

Outline of the paper: In section 2, the non-reachability problem for VAS is reduced to the existence of inductive separators (a pair of inductive invariants). In section 3 we introduce the class of semi-pseudo-linear sets, a class of sets that can be precisely over-approximated by semi-linear sets. In section 4, reachability sets are proved semi-pseudo-linear. Finally in section 5 we show the existence of semi-linear inductive separators proving the non-reachability of a pair of configurations. *In order to simplify the presentation of this* *paper, the independent parts of sections 3 and 4 with the remaining of the paper are presented in some sub-sections. That means the reader may safely skip these sub-sections in order to read other sections of the paper.*

2. VECTOR ADDITION SYSTEMS

In this section, the non-reachability problem for Vector Addition Systems is reduced to the existence of *inductive separators.*

```
Reachability (s ∈ ℕ^n, 𝒱 = (Σ, n, T) a VAS, s' ∈ ℕ^n)
  repeat forever
     fairly  select  σ ∈ Σ*
       if s -σ->_𝒱 s'
          return "reachable"
     fairly  select  (ψ(x), ψ'(x)) formulas in FO (ℕ, +, ≤)
       if (ψ(x), ψ'(x)) inductive separator for ({s}, {s'})
          return "unreachable"
```

Some notations : As usual 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 *cardinal* of a finite set X is denoted by $|X|$. The *components of a vector* $\mathbf{x} \in \mathbb{Q}^n$ are denoted by $(\mathbf{x}[1], \dots, \mathbf{x}[n])$. Given a function $f : E \to F$ where E, F are sets, we denote by $f(X) = \{f(x) \mid x \in X\}$ for any subset $X \subseteq E$. This definition naturally defines sets $X_1 + X_2$ where $X_1, X_2 \subseteq \mathbb{Q}^n$. With slightly abusing notations, $\{\mathbf{x_1}\} + X_2$ and $X_1 + \{\mathbf{x_2}\}$ are simply denoted by $\mathbf{x_1} + X_2$ and $X_1 + \mathbf{x_2}$. The total order $\leq$ over $\mathbb{Q}$ is extended component-wise to the partial order $\leq$ satisfying $\mathbf{x} \leq \mathbf{x}'$ if and only if $\mathbf{x}[i] \leq \mathbf{x}'[i]$ for any $1 \leq i \leq n$. A *function* $f : \mathbb{Q}^n \to \mathbb{Q}^d$ is said *rational linear* if there exists a matrix $M \in \mathbb{Q}^{n \times d}$ and a vector $\mathbf{v} \in \mathbb{Q}^d$ such that $f(\mathbf{x}) = M\mathbf{x} + \mathbf{v}$ for any $\mathbf{x} \in \mathbb{Q}^n$. A function $f : \mathbb{Z}^n \to \mathbb{Z}^d$ is said *integral linear* if there exists a matrix $M \in \mathbb{Z}^{n \times d}$ and a vector $\mathbf{v} \in \mathbb{Z}^d$ such that $f(\mathbf{x}) = M\mathbf{x}+\mathbf{v}$ for any $\mathbf{x} \in \mathbb{Z}^n$. 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, recall that $\min(X)$ is finite and $X \subseteq \min(X)+\mathbb{N}^n$ for any $X \subseteq \mathbb{N}^n$. An *alphabet* is a non-empty finite set Σ. Set of words over Σ are denoted by Σ^*. The number of occurrences $a \in \Sigma$ in a word $\sigma \in \Sigma^*$ is denoted by $|\sigma|_a$. A *Parikh image* of a language $\mathcal{L} \subseteq \Sigma^*$ is a set $X = \{(|\sigma|_{a_1}, \dots, |\sigma|_{a_n}) \mid \sigma \in \mathcal{L}\}$ where $a_1, \dots, a_n$ is a finite sequence in Σ.

A *Vector Addition System (VAS)* is a tuple $\mathcal{V} = (\Sigma, n, \delta)$ where Σ is a non-empty finite alphabet, $n \in \mathbb{N}$ is the *dimension* and $\delta \in \Sigma \to \mathbb{Z}^n$ is the *displacement function.* A *configuration* is a vector in $\mathbb{N}^n$. The binary relation $\xrightarrow{a}_{\mathcal{V}}$ where $a \in \Sigma$ over the set of configurations is defined by $\mathbf{s} \xrightarrow{a}_{\mathcal{V}} \mathbf{s}'$ if and only if $\mathbf{s}' = \mathbf{s} + \delta(a)$. Given a word $\sigma = a_1 \dots a_k$ of $k \in \mathbb{N}$ elements $a_i \in \Sigma$, we denote by $\xrightarrow{\sigma}_{\mathcal{V}}$ the binary relation over the set of configurations that is equal to the concatenation $\xrightarrow{a_1}_{\mathcal{V}} \cdots \xrightarrow{a_k}_{\mathcal{V}}$ if $k \geq 1$ and that is equal to the identity binary relation if $k = 0$. We also denote by $\to_{\mathcal{V}}$ the *reachability binary relation* over the set of configurations defined by $\mathbf{s} \to_{\mathcal{V}} \mathbf{s}'$ if and only if there exists $\sigma \in \Sigma^*$ such that $\mathbf{s} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{s}'$. Given two sets S, S' of configurations, we denote by $\mathrm{post}^*_{\mathcal{V}}(S)$ and $\mathrm{pre}^*_{\mathcal{V}}(S')$ respectively the set of *reachable states from* S and the set of *co-reachable states from* S' formally defined by:

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

The *language accepted* by a tuple $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ where $(\mathbf{s}, \mathbf{s}')$ are two configurations of a VAS $\mathcal{V}$ is the set $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ of words $\sigma \in \Sigma^*$ such that $\mathbf{s} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{s}'$. The reachability problem for such a tuple $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ consists to decide if the language accepted is non-empty or equivalently $\mathbf{s} \to_{\mathcal{V}} \mathbf{s}'$. This problem can be reformulated by introducing the definition of separators. A pair (S, S') of configuration sets called a *separator* if $\mathrm{post}^*_{\mathcal{V}}(S) \cap \mathrm{pre}^*_{\mathcal{V}}(S') = \emptyset$. Naturally, a pair $(\mathbf{s}, \mathbf{s}')$ is in the complement of the reachability relation $\to_{\mathcal{V}}$ if and only if the pair $(\{\mathbf{s}\}, \{\mathbf{s}'\})$ is a separator. A separator (I, I') is said *inductive* if I is a *forward invariant* $\mathrm{post}^*_{\mathcal{V}}(I) = I$ and I' is a *backward invariant* $\mathrm{pre}^*_{\mathcal{V}}(I') = I'$. As $(\mathrm{post}^*_{\mathcal{V}}(S), \mathrm{pre}^*_{\mathcal{V}}(S'))$ is an inductive separator for any separator (S, S'), we deduce that a pair (S, S') is a separator if and only if there exists an inductive separator (I, I') for (S, S').

We are interested in inductive separators definable in the decidable logic $\mathrm{FO}(\mathbb{N}, +, \leq)$. Note that a pair $(\psi(\mathbf{x}), \psi'(\mathbf{x}))$ of formulas in this logic denotes an inductive separator (I, I') if and only if $\psi(\mathbf{x}) \wedge \psi'(\mathbf{x})$ and the following formulas are unsatisfiable for any $a \in \Sigma$. In particular we can effectively decide if $(\psi(\mathbf{x}), \psi'(\mathbf{x}))$ denotes an inductive separator.

$$\psi(\mathbf{x}) \quad \wedge \quad \mathbf{x}' = \mathbf{x} + \delta(a) \quad \wedge \quad \neg\psi(\mathbf{x}')$$
$$\psi'(\mathbf{x}') \quad \wedge \quad \mathbf{x}' = \mathbf{x} + \delta(a) \quad \wedge \quad \neg\psi'(\mathbf{x})$$

In this paper we prove that for any separator (S, S') definable in $\mathrm{FO}(\mathbb{N}, +, \leq)$, there exists an inductive separator (I, I') definable in $\mathrm{FO}(\mathbb{N}, +, \leq)$ containing (S, S'). We deduce that algorithm Reachability($\mathbf{s}$,$\mathcal{V}$,$\mathbf{s}'$) decides the reachability problem. The termination is guaranteed by the previous result. Note [5] that in general, the inductive separator $(\mathrm{post}^*_{\mathcal{V}}(S), \mathrm{pre}^*_{\mathcal{V}}(S'))$ is not definable in $\mathrm{FO}(\mathbb{N}, +, \leq)$ even if S and S' are reduced to single vectors $S = \{\mathbf{s}\}$ and $S' = \{\mathbf{s}'\}$. That means, this inductive separator must be over-approximated by another inductive separator (I, I') definable in $\mathrm{FO}(\mathbb{N}, +, \leq)$. Intuitively, the approximation is obtained by observing that $\mathrm{post}^*_{\mathcal{V}}(S) \cap S'$ and $S \cap \mathrm{pre}^*_{\mathcal{V}}(S')$ are *semi-pseudo-linear* for any pair (S, S') of sets definable in $\mathrm{FO}(\mathbb{N}, +, \leq)$, a class of sets that can be precisely over-approximated by sets definable in $\mathrm{FO}(\mathbb{N}, +, \leq)$.

3. SEMI-PSEUDO-LINEAR SETS

In this section we introduce the class of *pseudo-linear sets* and *semi-pseudo-linear sets.* We show that a pseudo-linear set X can be precisely over-approximated by a *linear set* L called a *linearization* of X. We also introduce a monotonic function $\dim : (P(\mathbb{Z}^n), \subseteq) \to (\{-\infty, 0, \dots, n\}, \leq)$ that associates to any set $X \subseteq \mathbb{Z}^n$ a *dimension* $\dim(X)$. We show that $\dim(X_1 \cup X_2) = \max\{\dim(X_1), \dim(X_2)\}$ for any $X_1, X_2 \subseteq \mathbb{Z}^n$. Essentially, in this section, we prove that any linearizations L_1, L_2 of pseudo-linear sets X_1, X_2 with an empty intersection $X_1 \cap X_2 = \emptyset$ satisfy $\dim(L_1 \cap L_2) < \max\{\dim(X_1), \dim(X_2)\}$.

We first associate a dimension to sets $X \subseteq \mathbb{Z}^n$. The *dimension* $\dim(X)$ of a non-empty set $X \subseteq \mathbb{Z}^n$ is the minimal integer $d \in \{0, \dots, n\}$ such that:

$$\sup_{k \geq 0} \frac{|X \cap \{-k, \dots, k\}^n|}{(1+2k)^d} < +\infty$$

The dimension of the empty-set set is denoted by $\dim(\emptyset) = -\infty$. Let us observe some immediate properties satisfied by the dimension function. First of all, we have $\dim(X) \leq 0$ if and only if X is finite. The dimension function is monotonic $\dim(X_1) \leq \dim(X_2)$ for any $X_1 \subseteq X_2$. Moreover it satisfies $\dim(X_1 \cup X_2) = \max\{\dim(X_1), \dim(X_2)\}$ and $\dim(X_1 + X_2) \leq \dim(X_1) + \dim(X_2)$. In particular $\dim(\mathbf{v} + X) = \dim(X)$ for any $\mathbf{v} \in \mathbb{Z}^n$ and for any $X \subseteq \mathbb{Z}^n$.

Now, let us recall the definition of *semi-linear sets*. A *monoïd* M of $\mathbb{Z}^n$ is a subset $M \subseteq \mathbb{Z}^n$ that contains the zero vector $\mathbf{0} \in M$ and that is stable by addition $M + M \subseteq M$. Given any subset $X \subseteq \mathbb{Z}^n$, observe that $X^* = \{\sum_{i=1}^{k} \mathbf{x_i} \mid k \in \mathbb{N}\ \mathbf{x_i} \in X\}$ is the unique minimal for the inclusion monoïd that contains X. It is called the monoïd *generated* by X. A finite set $P \subseteq \mathbb{Z}^n$ is called a *set of periods.* A set $L \subseteq \mathbb{Z}^n$ is said *linear* [3] if there exists a vector $\mathbf{b} \in \mathbb{Z}^n$ and a set of periods $P \subseteq \mathbb{Z}^n$ such that $L = \mathbf{b} + P^*$. 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 $\mathrm{FO}(\mathbb{Z}, +, \leq)$ are exactly the semi-linear sets and sets definable in $\mathrm{FO}(\mathbb{N}, +, \leq)$ are exactly the non-negative semi-linear sets.

The definition of semi-pseudo-linear sets requires the definition of *attractors* of a monoïd. Given a vector $\mathbf{x}$ in a monoïd M, we observe that $\mathbf{x} + M$ is a subset of M. We are interested in vectors $\mathbf{a} \in M$ such that for any set $\mathbf{x} + M$ where $\mathbf{x} \in M$ there exists an integer $N \in \mathbb{N}$ such that $N\mathbf{a} \in \mathbf{x} + M$. More formally, an *attractor* of a monoïd M is a vector $\mathbf{a} \in M$ such that $(\mathbb{N}\mathbf{a}) \cap (\mathbf{x} + M) \neq \emptyset$ for any $\mathbf{x} \in M$. We denote by $\mathcal{A}(M)$ the *set of attractors* of M. The following Lemma 1 characterizes the set $\mathcal{A}(P^*)$ where P is a set of periods. In particular, this lemma shows that $\mathcal{A}(P^*)$ is non empty.

LEMMA 1. *We have* $\mathcal{A}(P^*) = \{\mathbf{0}\}$ *if* $k = 0$ *and* $\mathcal{A}(P^*) = P^* \cap ((\mathbb{Q}_+ \backslash \{0\})\mathbf{p_1} + \dots + (\mathbb{Q}_+ \backslash \{0\})\mathbf{p_k})$ *if* $k \geq 1$ *for any set of periods* $P = \{\mathbf{p_1}, \dots, \mathbf{p_k}\}$.

PROOF. Since the case $k = 0$ is immediate, we assume that $k \geq 1$. Let us first consider an attractor $\mathbf{a} \in \mathcal{A}(P^*)$. As $\sum_{j=1}^{k} \mathbf{p_j} \in P^*$ and $\mathbf{a} \in \mathcal{A}(P^*)$, there exists $N \in \mathbb{N}$ 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 $\mathbf{p} + \mathbf{a} \in P^*$, there exists a sequence $(N_j)_{1 \leq j \leq k}$ of elements in $\mathbb{N}$ such that $\mathbf{p} + \mathbf{a} = \sum_{j=1}^{k} N_j \mathbf{p_j}$. Combining this equality with the previous one provides $\mathbf{a} = \sum_{j=1}^{k} \frac{1+N_j}{1+N} \mathbf{p_j}$. Thus $\mathbf{a} \in (\mathbb{Q}_+ \backslash \{0\})\mathbf{p_1} + \dots + (\mathbb{Q}_+ \backslash \{0\})\mathbf{p_k}$. Conversely, let us consider $\mathbf{a} \in P^* \cap ((\mathbb{Q}_+ \backslash \{0\})\mathbf{p_1} + \dots + (\mathbb{Q}_+ \backslash \{0\})\mathbf{p_k})$. Observe that there exists an integer $d \geq 1$ large enough such that $d\mathbf{a} \in (\mathbb{N} \backslash \{0\})\mathbf{p_1} + \dots + (\mathbb{N} \backslash \{0\})\mathbf{p_k}$. In particular for any $\mathbf{x} \in P^*$ there exists $N \in \mathbb{N}$ such that $Nd\mathbf{a} \in \mathbf{x} + P^*$. $\square$

EXAMPLE 1. *Let* $P = \{(1,1), (1,0)\}$. *The monoïd generated by* P *is equal to* $P^* = \{\mathbf{x} \in \mathbb{N}^2 \mid \mathbf{x}[2] \leq \mathbf{x}[1]\}$, *and the set of attractors of* P^* *is equal to* $\mathcal{A}(P^*) = \{\mathbf{x} \in \mathbb{N}^2 \mid 0 < \mathbf{x}[2] < \mathbf{x}[1]\}$.

A set $X \subseteq \mathbb{Z}^n$ is said *pseudo-linear* if there exists $\mathbf{b} \in \mathbb{Z}^n$ and a set of periods $P \subseteq \mathbb{Z}^n$ such that $X \subseteq \mathbf{b} + P^*$ and such that for any finite set $R \subseteq \mathcal{A}(P^*)$ there exists $\mathbf{x} \in X$ such that $\mathbf{x} + R^* \subseteq X$. In this case, P is called a *linearizator* of X and the linear set $L = \mathbf{b} + P^*$ is called a *linearization* of X. A *semi-pseudo-linear set* is a finite union of *pseudo-linear sets.*

EXAMPLE 2. *The set* $P = \{(1,1), (1,0)\}$ *is a linearizator of the pseudo-linear set* $X = \{\mathbf{x} \in \mathbb{N}^2 \mid \mathbf{x}[2] \leq \mathbf{x}[1] \leq 2^{\mathbf{x}[2]}\}$. *Moreover* P^* *is a linearization of* X.

All other results and notations introduced in this section are not used in the sequel. The reader may safely skip the remaining of this section to read the other ones. In sub-section 3.1 we characterize the dimension of linear sets and pseudo-linear sets. This characterization is used in the next sub-section 3.2 to 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 $\dim(L_1 \cap L_2) < \max\{\dim(X_1), \dim(X_2)\}$.

In these two sub-sections, *vector spaces* are used. A *vector space* V of $\mathbb{Q}^n$ is a subset $V \subseteq \mathbb{Q}^n$ that contains the zero vector $\mathbf{0} \in V$, that is stable by addition $V + V \subseteq V$ and that is stable by rational product $\lambda\mathbf{v} \in V$ for any $\lambda \in \mathbb{Q}$ and for any $\mathbf{v} \in V$. Observe that for any set $X \subseteq \mathbb{Q}^n$ the set $V = \{\sum_{i=1}^{k} \lambda_i \mathbf{x_i} \mid k \in \mathbb{N}\ \lambda_i \in \mathbb{Q}\ \mathbf{x_i} \in X\}$ is the unique minimal for the inclusion vector space that contains X. This vector space is called the vector space generated by X. Recall that for any vector space V of $\mathbb{Q}^n$ there exists a finite set $X \subseteq V$ that generates V. The minimal for $\leq$ integer $d \in \mathbb{N}$ such that there exists a finite set X that generated V is called the *rank* of V and it is denoted by $\mathrm{rank}(V)$.

3.1 Dimension of (pseudo-)linear sets

In this section, we prove that the dimension of a pseudo-linear set X is equal to the rank of the vector space V generated by any linearizator P of X.

We first prove the following Lemmas 2.

LEMMA 2. *We have* $\dim(M) = rank(V)$ *where* V *is the vector space generated by a monoïd* M.

PROOF. Since $M \subseteq \mathbb{Z}^n \cap V$ it is sufficient to prove that $\dim(M) \geq \mathrm{rank}(V)$ and $\dim(\mathbb{Z}^n \cap V) \leq \mathrm{rank}(V)$. Let us denote by $||\mathbf{x}||_\infty = \max\{|\mathbf{x}[1]|, \dots, |\mathbf{x}[k]|\}$ the usual ∞-*norm* of a vector $\mathbf{x} \in \mathbb{Q}^n$. As M generates the vector space V recall that there exists a sequence $\mathbf{m_1}, \dots, \mathbf{m_d} \in M$ with $d = \mathrm{rank}(V)$ that generates V. Since the case $d = 0$ is immediate we assume that $d \geq 1$. We denote by $f : \mathbb{Q}^d \to V$ the rational linear function $f(\mathbf{x}) = \sum_{i=1}^{d} \mathbf{x}[i]\mathbf{m_i}$.

Let us first prove that $\dim(M) \geq d$. By minimality of $d = \text{rank}(V)$ note that f is injective. In particular the cardinal of $f(\{0, \dots, k\}^d)$ is equal to $(1+k)^d$. Note that a vector $\mathbf{m}$ in this set satisfies $||\mathbf{m}||_\infty \leq k \sum_{i=1}^{d} ||\mathbf{m_i}||_\infty$ and $\mathbf{m} \in M$. We deduce that $\dim(M) \geq d$.

Now, let us prove that $\dim(\mathbb{Z}^n \cap V) \leq d$. Since for any matrix, the rank of the column vectors is equal to the rank of the line vectors, there exists a sequence $1 \leq j_1 < \dots < j_d \leq n$ such that the rational linear function $g : \mathbb{Q}^n \to \mathbb{Q}^d$ defined by $g(\mathbf{x}) = (\mathbf{x}[j_1], \dots, \mathbf{x}[j_d])$ satisfies $h = g \circ f$ is a bijective rational linear function. In particular we deduce that for any $\mathbf{v} \in \mathbb{Z}^n \cap V \cap \{-k, \dots, k\}^n$ there exists a vector $\mathbf{x} = g(\mathbf{v}) \in \{-k, \dots, k\}^d$ such that $\mathbf{v} = f \circ h^{-1}(\mathbf{x})$. Therefore $|\mathbb{Z}^n \cap V \cap \{-k, \dots, k\}^n| \leq (1+2k)^d$ for any $k \in \mathbb{N}$. We deduce that $\dim(\mathbb{Z}^n \cap V) \leq d$. □

LEMMA 3. *For any pseudo-linear set $X \subseteq \mathbb{Z}^n$, we have $\dim(X) = \text{rank}(V)$ where V is the vector space generated by any linearizator P of X.*

PROOF. Let P be a linearizator of a pseudo-linear set X and let V be the vector space generated by P. Note that there exists a vector $\mathbf{b} \in \mathbb{Z}^n$ such that $X \subseteq \mathbf{b} + P^*$. From Lemma 2 we have $\dim(\mathbf{b} + P^*) = \text{rank}(V)$. In particular $\dim(X) \leq \text{rank}(V)$. Conversely, let us consider an attractor $\mathbf{a} \in \mathcal{A}(P^*)$ and observe that $R = \{\mathbf{a}\} \cup (\mathbf{a} + P) \subseteq \mathcal{A}(P^*)$. 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 V. Thus, from Lemma 2 we deduce that $\dim(\mathbf{x} + R^*) = \text{rank}(V)$. In particular $\dim(X) \geq \text{rank}(V)$. We have proved the equality $\dim(X) = \text{rank}(V)$. □

3.2 Pseudo-linear sets intersection

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 $\dim(L_1 \cap L_2) < \max\{\dim(X_1), \dim(X_2)\}$.

We first characterize the intersection of two linear sets.

LEMMA 4. *For any set of periods P_1, P_2 there exists a set of periods P such that $P_1^* \cap P_2^* = P^*$. Moreover, for any $\mathbf{b_1}, \mathbf{b_2} \in \mathbb{Z}^n$, 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 $\mathbf{p_{i,1}}, \dots, \mathbf{p_{i,k_i}}$ 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 $P_1^* \cap P_2^* = P^*$. □

In order to prove the following proposition, we introduce the definition of *groups*. A *group* G of $\mathbb{Z}^n$ is a monoïd of $\mathbb{Z}^n$ such that any element admits an inverse $-G \subseteq G$. Observe that for any set $X \subseteq \mathbb{Z}^n$, the set $G = X^* - X^*$ is the unique minimal for the inclusion group that contains X. This group is called the group generated by X. Now, let us consider the group $G = M - M$ generated by a monoïd M and observe that a vector $\mathbf{a}$ is an attractor of M if and only if for any $\mathbf{g} \in G$ there exists $N \in \mathbb{N}$ such that $\mathbf{g} + N\mathbf{a} \in M$.

LEMMA 5. *For any 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. A vector $\mathbf{v} \in V$ can be decomposed into a sum $\mathbf{v} = \sum_{i=1}^{k} \lambda_i \mathbf{g_i}$ with $k \in \mathbb{N}$, $\lambda_i \in \mathbb{Q}$ and $\mathbf{g_i} \in G$. Let us consider an integer $d \geq 1$ such that $d\lambda_i \in \mathbb{Z}$ and observe that $d\mathbf{v} \in G$. □

Now, we prove the main result of this section.

PROPOSITION 1. *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:*

$$\dim(L_1 \cap L_2) < \max\{\dim(X_1), \dim(X_2)\}$$

PROOF. Let L_1, L_2 be linearizations of two pseudo-linear sets $X_1, X_2 \subseteq \mathbb{Z}^n$. For the moment, we do not assume that $X_1 \cap X_2$ is empty. There exists some linearizators P_1, P_2 of the pseudo-linear sets X_1, X_2 and vectors $\mathbf{b_1}, \mathbf{b_2} \in \mathbb{Z}^n$ such that $L_1 = \mathbf{b_1} + P_1^*$ and $L_2 = \mathbf{b_2} + P_2^*$ are linearizations of X_1, X_2. Let us denote by V_1, V_2 the vector spaces generated by P_1, P_2. Lemma 3 shows that $\dim(X_1) = \text{rank}(V_1)$ and $\dim(X_2) = \text{rank}(V_2)$. From Lemma 4 there exists a set of periods P and a finite set $B \subseteq \mathbb{Z}^n$ such that $P_1^* \cap P_2^* = P^*$ and $L_1 \cap L_2 = B + P^*$. Observe that if $B = \emptyset$ the proposition is immediate. Thus, we can assume that there exists $\mathbf{b} \in B$. Let V be the vector space generated by P. Lemma 2 shows that $\dim(B + P^*) = \text{rank}(V)$. Observe that $V \subseteq V_1 \cap V_2$. Thus, if there exists $j \in \{1, 2\}$ such that V is strictly included in V_j then $\text{rank}(V) < \text{rank}(V_j)$ and in this case $\dim(L_1 \cap L_2) < \max\{\dim(X_1), \dim(X_2)\}$.

So we can assume that $V_1 = V = V_2$. We prove in the sequel that $X_1 \cap X_2 \neq \emptyset$ providing the proposition. 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 attractor of P^* and let us prove that $\mathbf{a} \in \mathcal{A}(P_j^*)$. Note that $\mathbf{a} \in P^* \subseteq P_j^*$. Let $\mathbf{p} \in \mathcal{A}(P_j^*)$. Since $-\mathbf{p} \in V$

and V is the vector space generated by G, Lemma 5 shows that there exists an integer $d \geq 1$ such that $-d\mathbf{p} \in G$. From $\mathbf{a} - d\mathbf{p} \in G$ and $\mathbf{a} \in \mathcal{A}(P^*)$ we deduce that there exists $N \in \mathbb{N}$ such that $\mathbf{a} - d\mathbf{p} + N\mathbf{a} \in P^*$. From $P^* \subseteq P_j^*$ we deduce that $\mathbf{a} \in \frac{1}{1+N}(d\mathbf{p} + P_j^*)$. From $\mathbf{p} \in \mathcal{A}(P_j^*)$ and Lemma 1 we get $\mathbf{a} \in \mathcal{A}(P_j^*)$.

Let $R_j = \{\mathbf{a}\} \cup (\mathbf{a} + P_j)$. From $\mathbf{a} \in \mathcal{A}(P_j^*)$, Lemma 1 shows that $R_j \subseteq \mathcal{A}(P_j^*)$. As X_j is pseudo-linear, there exists $\mathbf{x_j} \in X_j$ such that $\mathbf{x_j} + R_j^* \subseteq X_j$. 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 5 shows that there exists an integer $d_1 \geq 1$ such that $d_1\mathbf{r_2'} \in G_1$. As $\mathbf{a} \in \mathcal{A}(P_1^*)$, there exists an integer $N_1 \geq 0$ 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 by $\mathbf{r_1''}$ this vector. Symmetrically, there exist some integers $d_2 \geq 1$ and $N_2, N_2' \geq 0$ 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:

$$\begin{aligned}
&\mathbf{x_1} + \mathbf{r_1} + (d_2 - 1)\mathbf{r_1'} + \mathbf{r_1''} + (N_2 + N_2')\mathbf{a} \\
&= \mathbf{b} + d_2\mathbf{r_1'} + d_1\mathbf{r_2'} + (N_1 + N_1' + N_2 + N_2')\mathbf{a} \\
&\mathbf{x_2} + \mathbf{r_2} + (d_1 - 1)\mathbf{r_2'} + \mathbf{r_2''} + (N_1 + N_1')\mathbf{a} \\
&= \mathbf{b} + d_1\mathbf{r_2'} + d_2\mathbf{r_1'} + (N_2 + N_2' + N_1 + N_1')\mathbf{a}
\end{aligned}$$

We have proved that this last vector is in $(\mathbf{x_1} + R_1^*) \cap (\mathbf{x_2} + R_2^*)$. In particular $X_1 \cap X_2 \neq \emptyset$. □

4. REACHABILITY SETS

In this section we prove that $\mathrm{post}_{\mathcal{V}}^*(S) \cap S'$ and $S \cap \mathrm{pre}_{\mathcal{V}}^*(S')$ are semi-pseudo-linear for any semi-linear sets $S, S' \subseteq \mathbb{N}^n$. All other results and notations introduced in this section are not used in the sequel. The reader may safely skip the remaining of this section in order to read the other ones. In sub-section 4.1 we recall the classical Kosaraju-Lambert-Mayr-Sacerdote-Tenney (KLMST) decomposition. This decomposition is used in the next sub-section 4.2 to establish the semi-pseudo-linearity of Parikh images of $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$. Finally, the main result of this section is proved in sub-section 4.3.

The following lemma 6 is used in different sub-sections.

LEMMA 6. *Images $X' = f(X)$ of pseudo-linear sets X by integral linear functions f are pseudo-linear. Moreover the linear set $L' = f(L)$ is a linearization of X' for any linearization L of X.*

PROOF. Let us consider an integral linear function $f : \mathbb{Z}^n \to \mathbb{Z}^{n'}$ defined by an integral matrix $M \in \mathbb{Z}^{n \times n'}$ and an integral 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 linearization L of X. We are going to prove that $L' = f(L)$ is a linearization of $X' = f(X)$. There exists a vector $\mathbf{b} \in \mathbb{Z}^n$ and a set of periods $P \subseteq \mathbb{Z}^n$ such that $L = \mathbf{b} + P^*$. Let us consider $\mathbf{b}' = f(\mathbf{b})$ and $P' = \{M\mathbf{p} \mid \mathbf{p} \in P\}$ 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 $R' = \{\mathbf{r_1'}, \dots, \mathbf{r_d}'\}$ of attractors of $(P')^*$. As $\mathbf{r_i'} \in (P')^*$ there exists $\mathbf{p_i} \in P^*$ such that $\mathbf{r_i'} = M\mathbf{p_i}$. Lemma 1 shows that $\mathbf{r_i'}$ is a sum of vectors of the form $\lambda_{i,\mathbf{p}}M\mathbf{p}$ over all $\mathbf{p} \in P$ where $\lambda_{i,\mathbf{p}} > 0$ is a rational value that naturally depends on $\mathbf{p}$. There exists an integer $n_i \geq 1$ large enough such that $n_i\lambda_{i,\mathbf{p}} \in \mathbb{N}\backslash\{0\}$ for any i and $\mathbf{p}$. 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, form Lemma 1 we deduce that $\mathbf{r_i}$ is an attractor of P^*. Observe that $n_i\mathbf{r_i'} = M\mathbf{r_i}$. 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 $\mathbf{r_i}$ is an attractor of P^* and $\mathbf{p_i} \in P^*$ we deduce that $\mathbf{r_i} + k_i\mathbf{p_i}$ is also an attractor of P^*. We have proved that $R \subseteq \mathcal{A}(P^*)$. As L is a linearization of X, there exists $\mathbf{x} \in X$ such that $\mathbf{x} + R^* \subseteq X$. We deduce that $f(\mathbf{x}) + MR^* \subseteq X'$. Let us consider $\mathbf{x}' = f(\mathbf{x}) + M(\sum_{i=1}^{d} \mathbf{r_i})$ 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 $n_i\mathbf{r_i'} = M\mathbf{r_i}$ we deduce that $\mathbf{x}' + \mathbf{r}' = f(\mathbf{x}) + M(\sum_{i=1}^{d}(\mathbf{r_i} + k_i\mathbf{p_i}) + \sum_{i=1}^{d} \mu_i\mathbf{r_i})$.Observe that $\mathbf{r_i} + k_i\mathbf{p_i}$ and $\mathbf{r_i}$ are both in R. We have proved that $\mathbf{x}' + \mathbf{r}' \in f(\mathbf{x}) + MR^*$. Thus $\mathbf{x}' + (R')^* \subseteq X'$. We have proved that L' is a linearization of X'. □

4.1 The KLMST decomposition

In this section we recall the KLMST decomposition by following notations introduced by Lambert [8].

We first extend the set of non-negative integers $\mathbb{N}$ with an additional element $\top$. In the sequel, this element is either interpreted as a "very large integer" or a "don't care integer". More formally, we denote by $\mathbb{N}_\top$ the set $\mathbb{N} \cup \{\top\}$. The total order $\leq$ over $\mathbb{N}$ is extended over $\mathbb{N}_\top$ by $x_1 \leq x_2$ if and only if $x_2 = \top \vee (x_1, x_2 \in \mathbb{N} \wedge x_1 \leq x_2)$. The equality $=$ over $\mathbb{N}$ is also extended to a partial order $\trianglelefteq$ over $\mathbb{N}_\top$ by $x_1 \trianglelefteq x_2$ if and only if $x_2 = \top \vee (x_1, x_2 \in \mathbb{N} \wedge x_1 = x_2)$. Intuitively element $\top$ denotes a "very large integer" for the total order $\leq$ whereas it denotes a "don't care integer" for the partial order $\trianglelefteq$. Given a sequence $(x_i)_{i \geq 0}$ in $\mathbb{N}_\top$, we denote by $x = \lim_{i \to +\infty} x_i$ the element $x = \top$ if for any $r \in \mathbb{N}$ there exists $i_0 \geq 0$ such that $x_i \geq r$ for any $i \geq i_0$ and the element $x \in \mathbb{N}$ if there exists $i_0 \geq 0$ such that $x_i = x$ for any $i \geq i_0$. When $x = \lim_{i \to +\infty} x_i$ exists we say that $(x_i)_{i \geq 0}$ *converges toward* x.

We also extends the semantics of VAS. A vector in $\mathbb{N}_\top^n$ is called an *extended configuration* of $\mathcal{V}$. The addition function $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ is extended to the totally-defined function in $(\mathbb{Z} \cup \{\top\}) \times (\mathbb{Z} \cup \{\top\} \to (\mathbb{Z} \cup \{\top\})$ satisfying $x_1 + x_2 = \top$ if $x_1 = \top$ or $x_2 = \top$. With slightly abusing notations, the binary relation $\xrightarrow{a}_{\mathcal{V}}$ where $a \in \Sigma$ over the set of extended configurations is defined by $\mathbf{x} \xrightarrow{a}_{\mathcal{V}} \mathbf{x}'$ if and only if $\mathbf{x}' = \mathbf{x} + \delta(a)$. Given a word $\sigma = a_1 \dots a_k$ of $k \in \mathbb{N}$ elements $a_i \in \Sigma$, we denote by $\xrightarrow{\sigma}_{\mathcal{V}}$ the binary relation over the set of extended configurations that is equal to the concatenation $\xrightarrow{a_1}_{\mathcal{V}} \cdots \xrightarrow{a_k}_{\mathcal{V}}$ if $k \geq 1$ and that is equal to the identity binary relation if $k = 0$. Given an extended configuration $\mathbf{x}$ we denote by $\mathbf{x} \xrightarrow{\sigma}_{\mathcal{V}}$ if there exists an extended configuration $\mathbf{x}'$ such that $\mathbf{x} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{x}'$ and symmetrically for any extended configuration $\mathbf{x}'$ we denote by $\xrightarrow{\sigma}_{\mathcal{V}} \mathbf{x}'$ if there exists an extended configuration $\mathbf{x}$ such that $\mathbf{x} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{x}'$.

Next we recall some elements of graph theory. A *graph* G is a tuple $G = (Q, \Sigma, T)$ where Q is a non-empty finite set of *states*, Σ is an alphabet, and $T \subseteq Q \times \Sigma \times Q$ is a finite set of *transitions.* A *path* π is a word $\pi = t_1 \dots t_k$ of $k \in \mathbb{N}$ transitions $t_i \in T$ such that there exists $q_0, \dots, q_k \in Q$ and $a_1, \dots, a_k \in \Sigma$ such that $t_i = (q_{j-1}, a_j, q_j)$ for any $1 \leq j \leq k$. In this case we say that π is a path from q_0 to q_k labelled by $\sigma = a_1 \dots a_k$ and we denote π by $q_0 \xrightarrow{\sigma}_G q_k$ or simply $q_0 \rightarrow_G q_k$. Given a transition $t \in T$, we denote by $|\pi|_t$ the number of occurrences of t in π. When $q_0 = q_k$, the path π is called a *cycle*. Let us recall the following lemma.

Lemma 7 (Euler Cycles). *Let $G = (Q, \Sigma, T)$ be a strongly connected graph. For any sequence $(\mu_t)_{t \in T}$ of integers $\mu_t > 0$ satisfying the following equality for any state $q_0 \in Q$, there exists a cycle π such that $|\pi|_t = \mu_t$ for any transition $t \in T$:*

$$\sum_{t=(q,a,q_0) \in T} \mu_t = \sum_{t'=(q_0,a,q') \in T} \mu_{t'}$$

A *graph vector* $G = (Q, \Sigma, T)$ for $\mathcal{V}$ is a graph such that $Q \subseteq \mathbb{N}^n_\top$ is a non-empty finite set of extended configurations, and $T \subseteq Q \times \Sigma \times Q$ is a finite set of transitions $(\mathbf{x}, a, \mathbf{x}')$ such that $\mathbf{x} \xrightarrow{a}_{\mathcal{V}} \mathbf{x}'$. Even if the proof of the following lemma is immediate by induction over the length of σ, it is central in the KLMST decomposition. In fact a path $\mathbf{x} \xrightarrow{\sigma}_G \mathbf{x}'$ implies the relation $\mathbf{x} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{x}'$.

Lemma 8 (Graph vector paths). *For any $\mathbf{x} \xrightarrow{\sigma}_{\mathcal{V}} \mathbf{x}'$, for any sequences $(\mathbf{x_c})_{c \in \mathbb{N}}$ and $(\mathbf{x'_c})_{c \in \mathbb{N}}$ of extended configurations that converge toward $\mathbf{x} = \lim_{c \rightarrow +\infty} \mathbf{x_c}$ and $\mathbf{x}' = \lim_{c \rightarrow +\infty} \mathbf{x'_c}$, there exists $c_0 \in \mathbb{N}$ such that $\mathbf{x_c} \xrightarrow{\sigma}_{\mathcal{V}}$ and $\xrightarrow{\sigma}_{\mathcal{V}} \mathbf{x'_c}$ for any $c \geq c_0$.*

A *marked graph vector* for $\mathcal{V}$ is a tuple $(\mathbf{m}, \mathbf{x}, G, \mathbf{x}', \mathbf{m}')$ where G is a graph vector, $\mathbf{x}, \mathbf{x}'$ are two states of this graph vector, and $\mathbf{m} \trianglelefteq \mathbf{x}$ and $\mathbf{m}' \trianglelefteq \mathbf{x}'$ are two extended configurations.

A *marked graph vector sequences (MGVS)* for $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ is an alternating sequence of marked graph vectors for $\mathcal{V}$ and actions of the following form where $\mathbf{m_0} = \mathbf{s}$ and $\mathbf{m'_k} = \mathbf{s}'$:

$$\mathcal{U} = (\mathbf{m_0}, \mathbf{x_0}, G_0, \mathbf{x'_0}, \mathbf{m'_0}), a_1, \dots, a_k, (\mathbf{m_k}, \mathbf{x_k}, G_k, \mathbf{x'_k}, \mathbf{m'_k})$$

The *language accepted* by a MGVS $\mathcal{U}$ is the set $\mathcal{L}(\mathcal{U})$ of words of the form $\sigma_0 a_1 \sigma_1 \dots a_k \sigma_k$ such that for any $0 \leq j \leq k$ there exists a path $\mathbf{x_j} \xrightarrow{\sigma_j}_{G_j} \mathbf{x'_j}$ and there exists two configurations $\mathbf{s_j} \trianglelefteq \mathbf{m_j}$ and $\mathbf{s'_j} \trianglelefteq \mathbf{m'_j}$ such that:

$$\mathbf{s_0} \xrightarrow{\sigma_0}_{\mathcal{V}} \mathbf{s'_0} \xrightarrow{a_1}_{\mathcal{V}} \mathbf{s_1} \xrightarrow{\sigma_1}_{\mathcal{V}} \mathbf{s'_1} \dots \mathbf{s'_{k-1}} \xrightarrow{a_k}_{\mathcal{V}} \mathbf{s_k} \xrightarrow{\sigma_k}_{\mathcal{V}} \mathbf{s'_k}$$

We observe that $\mathcal{L}(\mathcal{U}) \subseteq \mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}')$ since $(\mathbf{s_0}, \mathbf{s'_k}) = (\mathbf{s}, \mathbf{s}')$.

We now associate a characteristic linear system to a MGVS $\mathcal{U}$. Denoting by $\mu_{j,t}$ the number of occurrences of a transition $t \in T_j$ in the path $\mathbf{x_j} \xrightarrow{\sigma_j}_{G_j} \mathbf{x'_j}$ we get a non-negative sequence $(\mu_{j,t})_t$ indexed by $t \in T_j$. We also obtain a sequence ξ of the form $\xi = (\mathbf{s_j}, (\mu_{j,t})_t, \mathbf{s'_j})_j$ indexed by $0 \leq j \leq k$ said *associated* to σ. We observe that ξ is a non-negative integral solution of the following linear system called the *characteristic system* of the MGVS $\mathcal{U}$ where $\chi_{\mathbf{x}}(q) = 1$ if $q = \mathbf{x}$ and where $\chi_{\mathbf{x}}(q) = 0$ otherwise:

$$\begin{cases} \underline{\text{for all } 1 \leq j \leq k} \\ \mathbf{s'_{j-1}} + \delta(a_j) = \mathbf{s_j} \\ \\ \underline{\text{for all } 0 \leq j \leq k} \\ \mathbf{s_j} + \displaystyle\sum_{t=(q,a,q') \in T_j} \mu_{j,t} \delta(a) = \mathbf{s'_j} \\ \\ \underline{\text{for all } 0 \leq j \leq k \text{ and for all } 1 \leq i \leq n} \\ \mathbf{s_j}[i] = \mathbf{m_j}[i] \text{ if } \mathbf{m_j}[i] \in \mathbb{N} \\ \mathbf{s'_j}[i] = \mathbf{m'_j}[i] \text{ if } \mathbf{m'_j}[i] \in \mathbb{N} \\ \\ \underline{\text{for all } 0 \leq j \leq k \text{ and for all } q_j \in Q_j} \\ \chi_{\mathbf{x_j}}(q_j) + \displaystyle\sum_{t=(q,a,q_j) \in T} \mu_{j,t} = \chi_{\mathbf{x'_j}}(q_j) + \displaystyle\sum_{t'=(q_j,a,q') \in T} \mu_{j,t'} \end{cases}$$

Naturally there exists non-negative integral solutions ξ of the characteristic system that are not associated to accepted words. In particular even if there exists non-negative integral solutions of the characteristic linear system we cannot conclude that $\mathcal{L}(\mathcal{U}) \neq \emptyset$. However, under the following *perfect* condition, we can prove that $\mathcal{L}(\mathcal{U}) \neq \emptyset$.

The homogeneous form of the characteristic system, obtained by replacing constants by zero is called the *homogeneous characteristic system of* $\mathcal{U}$. In the sequel, a solution of the homogeneous characteristic system is denoted by $\xi_{\mathbf{0}} = (\mathbf{s_{0,j}}, (\mu_{0,j,t})_t, \mathbf{s_{0,j}}')_j$.

A *perfect* MGVS $\mathcal{U}$ is an MGVS such that the graph G_j is strongly connected and $\mathbf{x_j} = \mathbf{x'_j}$ for any $0 \leq j \leq k$, the characteristic system has an integral solution, there exists a non-negative rational solution $\xi_{\mathbf{0}} = (\mathbf{s_{0,j}}, (\mu_{0,j,t})_t, \mathbf{s_{0,j}}')_j$ of the homogeneous characteristic system satisfying the following additional inequalities where $0 \leq j \leq k$ and $1 \leq i \leq n$:

- $\mathbf{s_{0,j}}[i] > 0$ if $\mathbf{m_j}[i] = \top$, and
- $\mathbf{s'_{0,j}}[i] > 0$ if $\mathbf{m'_j}[i] = \top$, and
- $\mu_{0,j,t} > 0$ for any $t \in T_j$.

and such that for any $0 \leq j \leq k$ and $1 \leq i \leq n$:

- there exists a cycle $\theta_j = (\mathbf{x_j} \xrightarrow{w_j}_{G_j} \mathbf{x_j})$ such that $\mathbf{m_j} \xrightarrow{w_j}_{\mathcal{V}}$ and such that $\mathbf{m_j} + \delta(w_j) \geq \mathbf{m_j}$ and $\delta(w_j)[i] > 0$ if $\mathbf{m_j}[i] \in \mathbb{N}$ and $\mathbf{x_j}[i] = \top$, and
- there exists a cycle $\theta'_j = (\mathbf{x'_j} \xrightarrow{w'_j}_{G_j} \mathbf{x'_j})$ such that $\xrightarrow{w'_j}_{\mathcal{V}} \mathbf{m'_j}$ and such that $\mathbf{m'_j} - \delta(w'_j) \geq \mathbf{m'_j}$ and $-\delta(w'_j)[i] > 0$ if $\mathbf{m'_j}[i] \in \mathbb{N}$ and $\mathbf{x'_j}[i] = \top$.

In the sequel, even if $\mathbf{x_j} = \mathbf{x'_j}$ for any $0 \leq j \leq k$, we still use both notations $\mathbf{x_j}$ and $\mathbf{x'_j}$ in order to keep results symmetrical. Let us recall without proof the fundamental decomposition theorem.

THEOREM 1 (FUNDAMENTAL DECOMPOSITION[8]). *For any tuple $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$, we can effectively compute a finite sequence of perfect MGVS $\mathcal{U}_1, \ldots, \mathcal{U}_l$ for this tuple such that:*

$$\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s}') = \mathcal{L}(\mathcal{U}_1) \cup \ldots \cup \mathcal{L}(\mathcal{U}_l)$$

In the remaining of this section, we associate to a perfect MGVS $\mathcal{U}$, a non-negative integral solution ξ of its characteristic system and a non-negative integral solution $\xi_\mathbf{0}$ of its homogeneous characteristic system that explains why $\mathcal{L}(\mathcal{U}) \neq \emptyset$. This two solutions ξ and $\xi_\mathbf{0}$ are respectively defined in Lemma 9 and Lemma 10. These two lemmas are independent of each other and can be read in any order.

LEMMA 9. *There exists a non-negative integral solution $\xi = (\mathbf{s_j}, (\mu_{j,t})_t, \mathbf{s'_j})_j$ of the characteristic system of a perfect MGVS such that $\mu_{j,t} > 0$ for any $0 \leq j \leq k$ and $t \in T_j$ and such that $\mathbf{s_j} \xrightarrow{w_j}_{\mathcal{V}}$ and $\xrightarrow{w'_j}_{\mathcal{V}} \mathbf{s'_j}$ for any $0 \leq j \leq k$.*

PROOF. The definition of perfect MGVS requires that there exists an integral solution $\xi = (\mathbf{s_j}, (\mu_{j,t})_t, \mathbf{s'_j})_j$ of its characteristic system. This solution is non-necessary non-negative. However, there exists a non-negative rational solution $\xi_\mathbf{0} = (\mathbf{s_{0,j}}, (\mu_{0,j,t})_t, \mathbf{s'_{0,j}})_j$ of the homogeneous characteristic system satisfying the perfect MGVS condition. Naturally, by replacing $\xi_\mathbf{0}$ by a sequence in $(\mathbb{N}\backslash\{0\})\xi_\mathbf{0}$ we can assume that $\xi_\mathbf{0}$ is a non-negative integral solution also satisfying the perfect MGVS condition. Now, just observe that there exists an integer $c_0 \geq 0$ large enough such that $\xi + c_0\xi_\mathbf{0}$ is a non-negative integral solution of the characteristic system satisfying $\mu_{j,t} + c_0\mu_{0,j,t} > 0$ for any $t \in T_j$ and for any $0 \leq j \leq k$. Moreover, as $\lim_{c\rightarrow+\infty}(\mathbf{s_j} + c\mathbf{s_{0,j}}) = \mathbf{m_j}$ and $\lim_{c\rightarrow+\infty}(\mathbf{s'_j} + c\mathbf{s'_{0,j}}) = \mathbf{m'_j}$, the relations $\mathbf{m_j} \xrightarrow{\sigma_j}_{\mathcal{V}}$ and $\xrightarrow{\sigma_j}_{\mathcal{V}} \mathbf{m'_j}$ and Lemma 8 shows that there exists an integer c large enough such that $(\mathbf{s_j}+c\mathbf{s_{0,j}}) \xrightarrow{w_j}_{\mathcal{V}}$ and $\xrightarrow{w'_j}_{\mathcal{V}} (\mathbf{s'_j}+c\mathbf{s'_{0,j}})$. Therefore $\xi + c\xi_\mathbf{0}$ is a non-negative integral solution of the characteristic system satisfying the lemma. □

LEMMA 10. *There exists a non-negative integral solution $\xi_\mathbf{0} = (\mathbf{s_{0,j}}, (\mu_{0,j,t})_t, \mathbf{s'_{0,j}})_j$ of the homogeneous characteristic system of a perfect MGVS $\mathcal{U}$ such that $\mu_{0,j,t} > |\theta_j|_t + |\theta'_j|_t$ for any $0 \leq j \leq k$ and $t \in T_j$, and such that for any $0 \leq j \leq k$ and $1 \leq i \leq n$:*

- $\mathbf{s_{0,j}} \geq \mathbf{0}$ *and* $\mathbf{s_{0,j}}[i] > 0$ *if and only if* $\mathbf{m_j}[i] = \top$.
- $\mathbf{s'_{0,j}} \geq \mathbf{0}$ *and* $\mathbf{s'_{0,j}}[i] > 0$ *if and only if* $\mathbf{m'_j}[i] = \top$.
- $\mathbf{s_{0,j}} + \delta(w_j) \geq \mathbf{0}$ *and* $(\mathbf{s_{0,j}} + \delta(w_j))[i] > 0$ *if and only if* $\mathbf{x_j}[i] = \top$.
- $\mathbf{s'_{0,j}} - \delta(w'_j) \geq \mathbf{0}$ *and* $(\mathbf{s_{0,j}}' - \delta(w'_j))[i] > 0$ *if and only if* $\mathbf{x'_j}[i] = \top$.

PROOF. Let $\xi_\mathbf{0} = (\mathbf{s_{0,j}}, (\mu_{0,j,t})_t, \mathbf{s'_{0,j}})_j$ be a non-negative rational solution of the homogeneous characteristic system satisfying the perfect MGVS condition. By replacing $\xi_\mathbf{0}$ by $(\mathbb{N}\backslash\{0\})\xi_\mathbf{0}$ we can assume that $\xi_\mathbf{0}$ is a non-negative integral solution satisfying the perfect condition. We are going to prove that there exists an integer $c \in \mathbb{N}$ large enough such that $c\xi_\mathbf{0}$ satisfies the lemma.

First of all, observe that for any $c \geq 1$ and for any $1 \leq i \leq n$, we have:

- $c\mathbf{s_{0,j}} \geq \mathbf{0}$ and $c\mathbf{s_{0,j}}[i] > 0$ if and only if $\mathbf{m_j}[i] = \top$.
- $c\mathbf{s'_{0,j}} \geq \mathbf{0}$ and $c\mathbf{s'_{0,j}}[i] > 0$ if and only if $\mathbf{m'_j}[i] = \top$.

Let us consider $1 \leq i \leq n$ and let us prove that there exists an integer $c_i \geq 0$ such that for any $c \geq c_i$ we have $(c\mathbf{s_{0,j}} + \delta(w_j))[i] \geq 0$ and $(c\mathbf{s_{0,j}} + \delta(w_j))[i] > 0$ if and only if $\mathbf{x_j}[i] = \top$. 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 $\xi_\mathbf{0}$ is solution of the homogeneous characteristic system, we get $\mathbf{s_{0,j}}[i] = 0$. Moreover the cycle $\theta_j = (\mathbf{x_j} \xrightarrow{w_j}_{G_j} \mathbf{x_j})$ 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 $c_i = 0$. Let us consider the second case $(\mathbf{m_j}[i], \mathbf{x_j}[i]) \in \mathbb{N} \times \{\top\}$. As in the previous case, since $\mathbf{m_j}[i] \in \mathbb{N}$ we deduce that $\mathbf{s_{0,j}}[i] = 0$. Note that the perfect condition shows that $\delta(w_j)[i] > 0$ in this case. In particular for any $c \geq 0$ 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 $c_i = 0$. 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 $c_i \geq 0$ large enough such that $(c\mathbf{s_{0,j}} + \delta(w_j))[i] > 0$ for any $c \geq c_i$. We have proved the three cases.

Symmetrically, for any $1 \leq i \leq n$, there exists an integer $c'_i \geq 0$ such that for any $c \geq c'_i$ we have $(c\mathbf{s'_{0,j}} - \delta(w'_j))[i] \geq 0$ and $(c\mathbf{s'_{0,j}} - \delta(w_j))[i] > 0$ if and only if $\mathbf{x'_j}[i] = \top$.

Finally, as $\mu_{0,j,t} > 0$ for any $t \in T_j$ and for any $0 \leq j \leq k$, we deduce that there exists an integer $c \geq 0$ large enough such that $c\mu_{0,j,t} > |\theta_j|_t + |\theta_{j'}|_t$ for any $t \in T_j$ for any $0 \leq j \leq k$. Naturally, we can also assume that $c \geq 1$, $c \geq c_i$ and $c \geq c'_i$ for any $1 \leq i \leq n$. We deduce that $c\xi_\mathbf{0}$ satisfies the lemma. □

Now, let us consider a perfect MGVS $\mathcal{U}$ and let us fix two tuples $\xi = (\mathbf{s_j}, (\mu_{j,t})_t, \mathbf{s'_j})_j$ and $\xi_\mathbf{0} = (\mathbf{s_{0,j}}, (\mu_{0,j,t})_t, \mathbf{s'_{0,j}})_j$ satisfying respectively Lemma 9 and Lemma 10. As $\mu_{j,t} > 0$ for any $t \in T_j$ and G_j is strongly connected, Lemma 7 shows that there exists a cycle $\pi_j = (\mathbf{x_j} \xrightarrow{\sigma_j}_{G_j} \mathbf{x'_j})$ such that $\mu_{j,t} = |\pi_j|_t$ for any $t \in T_j$. Note that we have $\mathbf{s_j}+\delta(\sigma_j) = \mathbf{s'_j}$. Moreover, as $\mu_{j,t} - |\theta_j|_t + |\theta_{j'}|_t > 0$ for any $t \in T_j$ we also deduce that there exists a cycle $\pi_{0,j} = (\mathbf{x_j} \xrightarrow{\sigma_{0,j}}_{G_j} \mathbf{x'_j})$ such that $|\pi_{0,j}|_t = \mu_{0,j,t} - (|\theta_j|_t + |\theta_{j'}|_t)$ for any $t \in T_j$. Note that we have $\mathbf{s_{0,j}} + \delta(w_j) + \delta(\sigma_{0,j}) + \delta(w'_j) = \mathbf{s'_{0,j}}$.

In the sequel we provide technical lemmas that prove together that $\mathcal{L}(\mathcal{U}) \neq \emptyset$. These lemmas are also used in the next sub-section 4.2.

LEMMA 11. *For any $c \geq 0$ we have:*

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

PROOF. Since the two relations are symmetrical, we just prove the first one. The choice of ξ satisfying Lemma 9 shows that $\mathbf{s_j} \xrightarrow{w_j}_{\mathcal{V}}$. The conditions $\mathbf{s_{0,j}} \geq \mathbf{0}$, $\mathbf{s_{0,j}} + \delta(w_j) \geq \mathbf{0}$ and $\mathbf{s_j} \xrightarrow{w_j}_{\mathcal{V}}$ with an immediate induction on the integer $c \geq 0$ provides the required relation. □

LEMMA 12. *There exists $c_0 \geq 0$ such that for any $c \geq c_0$:*

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

PROOF. Let us recall that $\mathbf{x_i} = \mathbf{x'_i}$. We denote by $\mathbf{u}$ the vector in $\{0,1\}^n$ satisfying $\mathbf{u}[i] = 1$ if $\mathbf{x_i}[i] = \top = \mathbf{x'_i}[i]$ and satisfying $\mathbf{u}[i] = 0$ otherwise. From the choice of $\xi_\mathbf{0}$ satisfying Lemma 10, we observe that $\mathbf{s_{0,j}} + \delta(w_j) \geq \mathbf{u}$ and $\mathbf{s'_{0,j}} - \delta(w'_j) \geq \mathbf{u}$. Note that $\lim_{c \to +\infty}(\mathbf{s_j} + c\mathbf{u}) = \mathbf{x_j}$. As $\mathbf{x_j} \xrightarrow{\sigma_{0,j}}_{G_j} \mathbf{x'_j}$, Lemma 8 proves that there exists an integer $c_0 \geq 0$ such that $\mathbf{s_j} + c_0\mathbf{u} \xrightarrow{\sigma_{0,j}}_{\mathcal{V}}$. Now, let us consider integers $c \geq 1$ and $c' \geq 0$ such that $c + c' \geq c_0$ and let us prove the relation:

$$\mathbf{s_j} + c(\mathbf{s_{0,j}} + \delta(w_j)) + c'(\mathbf{s'_{0,j}} - \delta(w'_j))$$
$$\xrightarrow{\sigma_{0,j}}_{\mathcal{V}}$$
$$\mathbf{s_j} + (c-1)(\mathbf{s_{0,j}} + \delta(w_j)) + (c'+1)(\mathbf{s'_{0,j}} - \delta(w'_j))$$

From $\mathbf{s_{0,j}} + \delta(w_j) \geq \mathbf{u}$ and $\mathbf{s'_{0,j}} - \delta(w'_j) \geq \mathbf{u}$ we deduce that $c(\mathbf{s_{0,j}} + \delta(w_j)) + c'(\mathbf{s'_{0,j}} - \delta(w'_j)) \geq (c+c')\mathbf{u} \geq c_0\mathbf{u}$. Thus, the previous relation directly comes from $\mathbf{s_j} + c_0\mathbf{u} \xrightarrow{\sigma_{0,j}}_{\mathcal{V}}$ and $\mathbf{s_{0,j}} + \delta(w_j) + \delta(\sigma_{0,j}) + \delta(w'_j) = \mathbf{s'_{0,j}}$. Now, an immediate induction provides $\mathbf{s_j} + c(\mathbf{s_{0,j}} + \delta(w_j)) \xrightarrow{\sigma_{0,j}^c}_{\mathcal{V}} \mathbf{s_j} + c(\mathbf{s'_{0,j}} - \delta(w'_j))$ for any $c \geq c_0$. □

LEMMA 13. *There exists $c' \geq 0$ such that for any $c \geq c'$:*

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

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 8 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 any $c \geq c'$. Since $\mathbf{s_j} + \delta(\sigma_j) = \mathbf{s'_j}$ we are done. □

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 12 and Lemma 13. Note that we have proved the following relation:

$$\mathbf{s_j} + c\mathbf{s_{0,j}} \xrightarrow{w_j^c \sigma_{0,j}^c \sigma_j (w'_j)^c}_{\mathcal{V}} \mathbf{s'_j} + c\mathbf{s'_{0,j}}$$

Therefore there exists a word in $\mathcal{L}(\mathcal{U})$ associated to $\xi + c\xi_\mathbf{0}$. In particular we have proved that $\mathcal{L}(\mathcal{U}) \neq \emptyset$.

4.2 Parikh images

In this section, we prove that Parikh images of $\mathcal{L}(\mathcal{U})$ are pseudo-linear for any perfect MGVS $\mathcal{U}$ for $(\mathbf{s}, \mathcal{V}, \mathbf{s'})$. From Theorem 1 we deduce that Parikh images of $\mathcal{L}(\mathbf{s}, \mathcal{V}, \mathbf{s'})$ are semi-pseudo-linear.

Let us consider a perfect MGVS $\mathcal{U}$ for $(\mathbf{s}, \mathcal{V}, \mathbf{s'})$:

$$\mathcal{U} = (\mathbf{m_0}, \mathbf{x_0}, G_0, \mathbf{x'_0}, \mathbf{m'_0}), a_1, \ldots, a_k, (\mathbf{m_k}, \mathbf{x_k}, G_k, \mathbf{x'_k}, \mathbf{m'_k})$$

We denote by H the non-negative integral solutions of the characteristic system and we denote by H' the subset of H corresponding to the sequence ξ associated to a word in $\mathcal{L}(\mathcal{U})$. Since the Parikh image of $\mathcal{L}(\mathcal{U})$ is equal to the image of H' by an integral linear function, from Lemma 6 it is sufficient to prove that H' is pseudo-linear. Intuitively, a linearizator for H' is obtained by considering the set H_0 of non-negative integral solutions of the homogeneous characteristic system. More formally, we are going to prove that $P_0 = \min(H_0 \backslash \{\mathbf{0}\})$ is a linearizator for H'. Note that $H_0 = P_0^*$.

Let us consider $\xi \in H$ and $\xi_\mathbf{0} \in H_0$ satisfying the following Lemma 14. We deduce that H' is included in the linear set $\xi - \xi_\mathbf{0} + P_0^*$.

LEMMA 14. *There exists $\xi \in H$ and $\xi_\mathbf{0} \in H_0$ such that $\xi_\mathbf{0} + H \subseteq \xi + H_0$.*

PROOF. As the MGVS $\mathcal{U}$ is perfect the set H is non empty. Let us consider the set I of components i such that $\xi[i]$ is independent of $\xi \in H$. As the MGVS is perfect we deduce that for any integer $c \geq 0$ there exists $\xi \in H$ such that $\xi[i] \geq c$ for any $i \notin I$. As $\min(H)$ is finite, we deduce that there exists $\xi \in H$ such that $\xi \geq \xi'$ for any $\xi' \in \min(H)$. In particular $\xi_\mathbf{0} = \sum_{\xi' \in \min(H)}(\xi - \xi')$ is in H_0. Let us prove that $\xi_\mathbf{0} + H \subseteq \xi + H_0$. Consider $\xi'' \in H$. By definition of $\min(H)$, there exists $\xi''' \in \min(H)$ such that $\xi''' \leq \xi''$. The definition of $\xi_\mathbf{0}$ shows that $\xi_\mathbf{0} - (\xi - \xi''')$ is equal to a sum of terms $(\xi - \xi')$ indexed by $\xi' \in \min(H) \backslash \{\xi'''\}$. Therefore $\xi_\mathbf{0} - (\xi - \xi''') \in H_0$. As $\xi'' - \xi''' \in H_0$ we have proved that the sum of $\xi_\mathbf{0} - (\xi - \xi''')$ and $\xi'' - \xi'''$ is also in H_0. Note that this sum is equal to $\xi_\mathbf{0} - \xi + \xi''$. We have proved that $\xi_\mathbf{0} + \xi'' \in \xi + H_0$. Therefore $\xi_\mathbf{0} + H \subseteq \xi + H_0$. □

Now, let us consider a set $R_0 = \{\xi_\mathbf{1}, \ldots, \xi_\mathbf{d}\}$ of attractors 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 15. *For any $\xi_\mathbf{i} = ((\mathbf{s_{i,j}})_j, (\mu_{i,j,t})_{j,t}, (\mathbf{s'_{i,j}})_j)$ attractor of H_0 there exists a cycle $\pi_{i,j} = (\mathbf{x_j} \xrightarrow{\sigma_{i,j}}_{G_j} \mathbf{x'_j})$ such that $\mu_{i,j,t} = |\pi_{i,j}|_t$ for any $t \in T_j$ and any $0 \leq j \leq k$.*

PROOF. Since $\mathcal{U}$ is perfect, for any $t \in T_j$, there exists a solution $\xi_\mathbf{0} = (\mathbf{s_{0,j}}, (\mu_{0,j,t})_t, \mathbf{s'_{0,j}})_j$ in H_0 such that $\mu_{0,j,t} > 0$. As $H_0 = P_0^*$, for any $t \in T_j$ there exists $\xi_\mathbf{0} \in P_0$ satisfying the same property. Lemma 1 shows that $\xi_\mathbf{i} = (\mathbf{s_{i,j}}, (\mu_{i,j,t})_t, \mathbf{s'_{i,j}})_j$ is a sum over all solutions $\xi_\mathbf{0} \in P_0$ of terms of the form $\lambda \xi_\mathbf{0}$ where $\lambda > 0$ is a rational value that naturally depends on $\xi_\mathbf{0}$. In particular we deduce that $\mu_{i,j,t} > 0$ for any $t \in T_j$ and for any $0 \leq j \leq k$. Lemma 7 shows that there exists a cycle $\pi_{i,j} = (\mathbf{x_j} \xrightarrow{\sigma_{i,j}}_{G_j} \mathbf{x'_j})$ such that $\mu_{i,j,t} = |\pi_{i,j}|_t$ for any $t \in T_j$ and any $1 \leq j \leq k$. □

Now, let us consider a solution ξ of the characteristic system and a solution $\xi_\mathbf{0}$ of the homogeneous characteristic system satisfying respectively Lemma 9 and Lemma 10.

LEMMA 16. *There exists $c_i \geq 0$ such that for any $1 \leq j \leq k$ and $c \geq c_i$:*

$$\mathbf{s_j} + c(\mathbf{s_{0,j}} + \delta(w_j)) + \mathbf{s_{i,j}} \xrightarrow{\sigma_{i,j}}_{\mathcal{V}} \mathbf{s_j} + c(\mathbf{s_{0,j}} + \delta(w_j)) + \mathbf{s'_{i,j}}$$

PROOF. As $\lim_{c \to +\infty} (\mathbf{s_j} + c(\mathbf{s_{0,j}} + \delta(w_j))) = \mathbf{x_j}$ and $\mathbf{x_j} \xrightarrow{\sigma_{i,j}}_{G_j} \mathbf{x_j}$, Lemma 8 proves that there exists an integer $c_i \geq 0$ such that $(\mathbf{s_j} + c(\mathbf{s_{0,j}} + \delta(w_j))) \xrightarrow{\sigma_{i,j}}_{\mathcal{V}}$ for any $c \geq c_i$ and for any $0 \leq j \leq k$. As $\mathbf{s_{i,j}} \geq \mathbf{0}$ and $\mathbf{s_{i,j}} + \delta(\sigma_{i,j}) = \mathbf{s'_{i,j}} \geq \mathbf{0}$ we deduce the lemma. □

Now, let us consider an integer $c \geq 0$ such that $c \geq c_0$, $c \geq c'$ and $c \geq c_i$ for any $1 \leq i \leq d$ where c_0, c', c_i are respectively defined in Lemma 12, Lemma 13 and Lemma 16. From these lemmas and Lemma 11 we deduce that for any sequence $n_1, \ldots, n_d \in \mathbb{N}$ we have the following relation:

$$\mathbf{s_j} + c\mathbf{s_{0,j}} + \sum_{i=1}^{d} n_i \mathbf{s_{i,j}} \xrightarrow{w_j^c \sigma_{1,j}^{n_1} \ldots \sigma_{d,j}^{n_d} \sigma_j \sigma_{0,j}^c (w'_j)^c}_{\mathcal{V}} \mathbf{s'_j} + c\mathbf{s'_{0,j}} + \sum_{i=1}^{d} n_i \mathbf{s'_{i,j}}$$

We have proved that there exists a word $\sigma \in \mathcal{L}(\mathcal{U})$ associated with $\xi + c\xi_\mathbf{0} + \sum_{i=1}^{d} n_i \xi_\mathbf{i}$. Let $\xi' = \xi + c\xi_\mathbf{0}$. We have proved that $\xi' + R_0^* \subseteq H'$. We deduce the following proposition:

PROPOSITION 2. *Parikh images of $\mathcal{L}(\mathcal{U})$ are pseudo-linear for any perfect MGVS $\mathcal{U}$ for $(\mathbf{s}, \mathcal{V}, \mathbf{s}')$.*

4.3 Reachability sets

In this section we prove that $\mathrm{post}_{\mathcal{V}}^*(S) \cap S'$ and $S \cap \mathrm{pre}_{\mathcal{V}}^*(S')$ are semi-pseudo-linear for any semi-linear sets $S, S' \subseteq \mathbb{N}^n$.

Since a semi-linear sets is a finite union of linear sets we prove this result for two linear sets $S = \mathbf{s} + P^*$ and $S' = \mathbf{s}' + (P')^*$ where $\mathbf{s}, \mathbf{s}' \in \mathbb{N}^n$ and where P, P' are two set of periods of $\mathbb{N}^n$. 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 $\bar{\mathcal{V}} = (\bar{\Sigma}, n, \bar{\delta})$. From the previous subsection, we deduce that Parikh images of $\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$ are semi-pseudo-linear. Let us consider the displacement functions $\bar{\delta}_P$ and $\bar{\delta}_{P'}$ defined over $\bar{\Sigma}$ by:

$$\bar{\delta}_P(a) = \begin{cases} \bar{\delta}(a) & \text{if } a \in \Sigma_P \\ \mathbf{0} & \text{otherwise} \end{cases} \qquad \bar{\delta}_{P'}(a) = \begin{cases} -\bar{\delta}(a) & \text{if } a \in \Sigma_{P'} \\ \mathbf{0} & \text{otherwise} \end{cases}$$

Just observe that $\mathrm{post}_{\mathcal{V}}^*(S) \cap S' = \mathbf{s}' + \bar{\delta}_{P'}(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$ and $S \cap \mathrm{pre}_{\mathcal{V}}^*(S') = \mathbf{s} + \bar{\delta}_P(\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}'))$. In particular these two sets are semi-pseudo-linear as Parikh images of $\mathcal{L}(\mathbf{s}, \bar{\mathcal{V}}, \mathbf{s}')$. Therefore, we have proved the following theorem:

THEOREM 2. *For any pair $S, S' \subseteq \mathbb{N}^n$ of semi-linear sets, the sets $\mathrm{post}_{\mathcal{V}}^*(S) \cap S'$ and $S \cap \mathrm{pre}_{\mathcal{V}}^*(S')$ are semi-pseudo-linear.*

5. SEMI-LINEAR SEPARATORS

In this section we prove that there exists a semi-linear inductive separator for any semi-linear separator.

Given a separator (S, S'), the set $D = \mathbb{N}^n \backslash (S \cup S')$ is called the *co-domain* of (S, S'). Note that a separator with an empty co-domain is inductive even if there exists inductive separators with non-empty co-domains.

The semi-linear indutive separator for a semi-linear separator (S_0, S'_0) is obtain inductively. We build a non-decreasing sequence $(S_j, S'_j)_{j \geq 0}$ of semi-linear separators starting from the initial semi-linear separator (S_0, S'_0) such that the dimension of the co-domain $D_j = \mathbb{N}^n \backslash (S_j \cup S'_j)$ is strictly decreasing. In order to obtain this sequence, observe that it is sufficient to show that for any semi-linear separator (S_0, S'_0) with a non-empty co-domain D_0, there exists a semi-linear separator $(S, S') \supseteq (S_0, S'_0)$ with a co-domain D such that $\dim(D) < \dim(D_0)$.

We first define a set S' that over-approximates S'_0 and such that (S_0, S') remains a separator. As S_0 and D_0 are semi-linear, the main result of section 4 shows 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 semi-linear 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 $\mathrm{post}_{\mathcal{V}}^*(S_0) \cap D_0 \subseteq \bigcup_{j=1}^{k} L_j$. Thus $\mathrm{post}_{\mathcal{V}}^*(S_0) \cap \mathrm{pre}_{\mathcal{V}}^*(S') = \emptyset$ and 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') remains a separator. As D_0 and S' are semi-linear, the main result proved in section 4 shows that $D_0 \cap \text{pre}^*_{\mathcal{V}}(S')$ 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 semi-linear set S.

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

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

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

$$D = D_0 \cap \left(\bigcup_{\substack{1 \leq j_1 \leq k \\ 1 \leq j_2 \leq k'}} (L_{j_1} \cap L'_{j_2}) \right)$$

From $X_{j_1}, X'_{j_2} \subseteq D_0$ we get $\max\{\dim(X_{j_1}), \dim(X'_{j_2})\} \leq \dim(D_0)$. As $X_{j_1} \subseteq \text{post}^*_{\mathcal{V}}(S_0) \subseteq \text{post}^*_{\mathcal{V}}(S)$ and $X'_{j_2} \subseteq \text{pre}^*_{\mathcal{V}}(S')$ and (S, S') is a separator, we deduce that X_{j_1} and X'_{j_2} are two pseudo-linear sets with an empty intersection. From the main result proved in section 3, we get $\dim(L_{j_1} \cap L'_{j_2}) < \max\{\dim(X_{j_1}), \dim(X'_{j_2})\}$. We deduce $\dim(D) < \dim(D_0)$. We have proved the following theorem.

THEOREM 3. *There exists a semi-linear inductive separator for any semi-linear separator.*

6. CONCLUSION

We have proved the termination of the algorithm Reachability. Even tough the proof is based on the classical KLMST decomposition, its complexity does not depend on this decomposition. In fact, the complexity of this algorithm depends on the size of the minimal pair of Presburger formulas denoting an inductive separator. This algorithm is the *very first one* that does not require the KLMST decomposition. In particular, this algorithm is the first candidate to obtain a precise (eventually elementary) upper-bound complexity for the VAS reachability problem.

We left as an open question the problem of computing a lower bound and a upper bound of the size of a pair of Presburger formulas denoting an inductive separator. Note that the VAS exhibiting a large (Ackermann size) but finite reachability set given in [4] does not directly provide a lower-bound for this size since inductive separators can over-approximate reachability sets.

We also left as an open question the problem of adapting such an algorithm to obtain a complete *Counter Example Guided Abstract Refinement* approach [1] for the VAS reachability problem based on *interpolants* [6] for FO $(\mathbb{N}, +, \leq)$. In practice, such an algorithm should be more efficient than the enumeration-based algorithm provided in this paper.

Acknowledgment: We thank *Jean Luc Lambert* for a fruitful discussion during a Post-doc in 2005 at IRISA (INRIA Rennes, France) and for his work on semi-linear VAS.

7. 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 Math.*, 16(2):285–296, 1966.

[4] M. Hack. The recursive equivalence of the reachability problem and the liveness problem for petri nets and vector addition systems. In *15th Annual Symposium on Switching and Automata Theory, 14-16 October 1974, The University of New Orleans, USA*, pages 156–164. IEEE, 1974.

[5] J. E. Hopcroft and J.-J. Pansiot. On the reachability problem for 5-dimensional vector addition systems. *Theor. Comput. Sci.*, 8:135–159, 1979.

[6] R. Jhala and K. L. McMillan. A Practical and Complete Approach to Predicate Refinement. In *Proc. of 12th Conf. on Tools and Algorithms for the Construction and Analysis of Systems (TACAS'06)*, volume 3920 of *LNCS*, pages 459–473. Springer, 2006.

[7] 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.

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

[9] 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.

[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.