An Approach to Regular Separability in Vector Addition Systems

Wojciech Czerwiński
University of Warsaw
wczerwin@mimuw.edu.pl

Georg Zetzsche
Max Planck Institute for Software Systems (MPI-SWS), Germany
georg@mpi-sws.org

Abstract

We study the problem of regular separability of languages of vector addition systems with states (VASS). It asks whether for two given VASS languages K and L, there exists a regular language R that includes K and is disjoint from L. While decidability of the problem in full generality remains an open question, there are several subclasses for which decidability has been shown.

We propose a general approach to deciding regular separability. We use it to obtain decidability for two subclasses. The first is regular separation of general VASS languages from languages of one-dimensional VASS. The second is regular separation of general VASS languages from integer VASS languages. Together, these two results generalize several of the previous decidability results for subclasses.

2012 ACM Subject Classification Theory of computation → Models of computation; Theory of computation → Formal languages and automata theory

Keywords and phrases separability problem, vector addition systems, decidability

Digital Object Identifier 10.4230/LIPIcs...

1 Introduction

Vector addition systems with states Vector addition systems with states (VASS) [14] are one of the most intensively studied model for concurrent systems. They can be seen as automata with finitely many counters, which can be increased or decreased whenever its values is nonnegative, but not tested for zero. Despite their fundamental nature and the extensive interest, core aspects remain obscure. A prominent example is the reachability problem, which was shown decidable in the early 1980s [26]. However, its complexity remains unsettled. The best known upper bounds are non-primitive-recursive [25], whereas the best known lower bound is tower hardness [6], and reachability seems far from being understood.

There is also a number of other natural problems concerning VASS where the complexity or even decidability remains unresolved. An example is the structural liveness problem, which asks whether there exists a configuration such that for every configuration c reachable from it and every transition t one can reach some configuration from c in which t is enabled. Its decidability status was settled only recently [17], but the complexity is still unknown. For closely related extensions of VASS, namely branching VASS and pushdown VASS even decidability status is unknown with the best lower bound being tower-hardness [21, 22]. This all suggests that there is still a lot to understand about VASS and related systems and one should seek ways to achieve that.

Separability problem One way to gain a fresh perspective and deeper understanding of the matter is to study a decision problems that generalize reachability. It seems to us that here, a natural choice is the problem of *regular separability*. It asks whether for two given

Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

languages K and L there exists a *regular separator*, i.e. a regular language R such that $K \subseteq R$ and $R \cap L = \emptyset$. Decidability of this problem for general VASS languages appears to be difficult. It has been shown decidable for several subclasses, namely for (i) *commutative VASS languages* [4] (equivalently, separability of sections of reachability sets by recognizable sets), for (ii) *one-counter nets* [7] i.e. VASS with one counter, (iii) integer VASS [3], i.e. VASS where we allow counters to become negative, and finally for (iv) *coverability languages*, which follows from the general decidability for well-structured transition systems [8]. However, in full generality, decidability remains a challenging open question. It should be mentioned that this line of research has already led to unforeseen insights: The closely related problem of separability by bounded regular languages prompted methods that turned out to yield decidability results that were deeply unexpected to the authors [5].

Contribution We present a general approach to deciding separability by regular languages and prove two new results, which generalize three out of four regular separability results shown until now. Namely we show decidability of regular separability of (i) VASS languages from languages of one counter nets and (ii) VASS languages from integer VASS languages.

The starting point of our approach is the observation that for many language classes $\mathcal{C}$, deciding regular separability of a language L from a given language K in $\mathcal{C}$ can be reduced to deciding regular separability of L from some fixed language G in $\mathcal{C}$. In the two cases (i) and (ii), this allows us to interpret the words in L as walks in the grid $\mathbb{Z}^n$. For (i), we then have to decide separability from those walks in $\mathbb{Z} = \mathbb{Z}^1$ that remain in $\mathbb{N}$ and arrive at zero. For (ii), we want to separate from all walks in $\mathbb{Z}^n$ that end in the origin. The corresponding fixed languages are denoted D_1 (for (i)) and Z_n (for (ii)), respectively. In order to decide separability from D_1 (Z_n, resp.), we classify those regular languages that are disjoint from D_1 (Z_n, resp.). In the case of Z_n, this classification leads to a geometric characterization of regular separability. Finally, the classifications are used to decide whether a given VASS language L is included in such a regular language.

We hope that this approach can be used to decide regular separability for VASS in full generality in the future. This would amount to deciding regular separability of a given VASS language from the set of all walks in $\mathbb{Z}^n$ that remain in $\mathbb{N}^n$ and arrive in the origin. The corresponding language is denoted D_n. We emphasize that an algorithm along these lines might directly yield new insights concerning reachability: Classifying those regular languages that are disjoint from D_n would yield an algorithm for reachability because the latter reduces to intersection of a given regular language with D_n. Such an algorithm would look for a certificate for *non-reachability* (like Leroux's algorithm [23]) instead of a run.

Related work Aside from *regular* separability, separability problems in a more general sense have also attracted significant attention in recent years. Here, the class of sought separators can differ from the regular languages. A series of recent works has concentrated on separability of regular languages by separators from subclasses [27, 28, 29, 30, 31, 32], and work in this direction has been started for trees as well [2, 11].

In the case of non-regular languages as input languages, it was shown early that regular separability is undecidable for context-free languages [35, 15]. Moreover, aside from the above mentioned results on regular separability, infinite-state systems have also been studied with respect to separability by bounded regular languages [5] and piecewise testable languages [9] and generalizations thereof [36].

2 Preliminaries

Let Σ be an alphabet and let ε denote the empty word. If $\Sigma = \{x_1, \ldots, x_n\}$, then the *Parikh image* of a word $w \in \Sigma^*$ is defined as $\Psi(w) = (|w|_{x_1}, \ldots, |w|_{x_n})$, where $|w|_x$ denotes the number of occurrences of x in w. Then, for $L \subseteq \Sigma^*$, we define its *commutative closure* as $\Pi(L) = \{u \in \Sigma^* \mid \exists v \in L\colon \Psi(v) = \Psi(u)\}$. For a finite automaton $\mathcal{A}$ with input alphabet Σ, let $\mathrm{Loop}(\mathcal{A}) \subseteq \Sigma^*$ be the set of words that can be read on a cycle in $\mathcal{A}$.

A *(n-dimensional) vector addition system with states (VASS)* [14] is a tuple $V = (Q, T, s, t, h)$, where Q is a finite set of *states*, $T \subseteq Q \times \mathbb{Z}^n \times Q$ is a finite set of transitions, $s \in Q$ is its *source state*, $t \in Q$ is its *target state*, and $h\colon T \to \Sigma_\varepsilon$ is its *labeling*, where $\Sigma_\varepsilon = \Sigma \cup \{\varepsilon\}$. A *configuration* of V is a pair $(q, \mathbf{u}) \in Q \times \mathbb{N}^d$. For each transition $(q, \mathbf{v}, q') \in T$ and configurations $(q, \mathbf{u})$, $(q', \mathbf{u}')$ with $\mathbf{u}' = \mathbf{u} + \mathbf{v}$, we write $(q, \mathbf{u}) \xrightarrow{h(t)} (q', \mathbf{u}')$. For a word $w \in \Sigma^*$, we write $(q, \mathbf{u}) \xrightarrow{w} (q', \mathbf{u}')$ if there are $x_1, \ldots, x_n \in \Sigma_\varepsilon$ and configurations $(q_i, \mathbf{v}_i)$ for $i \in [0, n]$ with $(q_{i-1}, \mathbf{v}_{i-1}) \xrightarrow{x_i} (q_i, \mathbf{v}_i)$ for $i \in [1, n]$, $(q_0, \mathbf{v}_0) = (q, \mathbf{u})$, and $(q_n, \mathbf{v}_n) = (q', \mathbf{u}')$. The language of V is then $\mathsf{L}(V) = \{w \in \Sigma^* \mid (s, 0) \xrightarrow{w} (t, 0)\}$. An *(n-dimensional) integer vector addition system with states ($\mathbb{Z}$-VASS)* [13] is syntactically a VASS, but for $\mathbb{Z}$-VASS, the configurations are pairs in $Q \times \mathbb{Z}^n$. This difference aside, the language is defined verbatim. Let $\mathcal{V}_n$ ($\mathcal{Z}_n$) denote the class of languages of n-dim. VASS ($\mathbb{Z}$-VASS).

Let $\Sigma_n = \{a_i, \bar{a}_i \mid i \in [1, n]\}$ and define the homomorphism $\varphi_n\colon \Sigma_n^* \to \mathbb{Z}^n$ by $\varphi_n(a_i) = \mathbf{e}_i$ and $\varphi_n(\bar{a}_i) = -\mathbf{e}_i$. Here, $\mathbf{e}_i \in \mathbb{Z}^n$ is the vector with 1 in coordinate i and 0 everywhere else. By way of φ_n, we can regards words from Σ_n^* as walks in the grid $\mathbb{Z}^n$ that start in the origin. Later, we will only write φ when the n is clear from the context. With this, let $Z_n = \{w \in \Sigma_n^* \mid \varphi(w) = 0\}$. Hence, Z_n is the set of walks that start and end in the origin.

Moreover, for $w \in \Sigma_1^*$, let $\mathrm{drop}(w) = \min\{\varphi(v) \mid v \text{ is a prefix of } w\}$. Thus, w if is interpreted as walking along $\mathbb{Z}$, then $\mathrm{drop}(w)$ is the lowest value attained on the way. Note that $\mathrm{drop}(w) \in [-|w|, 0]$ for every $w \in \Sigma_1^*$. We define $D_1 = \{w \in \Sigma_1^* \mid \mathrm{drop}(w) = 0,\ \varphi(w) = 0\}$. For each $i \in [1, n]$, let $\lambda_i\colon \Sigma_n^* \to \Sigma_1^*$ be the homomorphism with $\lambda_i(a_i) = a_1$, $\lambda_i(a_j) = \varepsilon$ for $j \neq i$, and $\lambda_i(\bar{a}_j) = \overline{\lambda_i(a_j)}$ for every $j \in [1, n]$. Then we define $D_n = \bigcap_{i=1}^n \lambda_i^{-1}(D_1)$. Thus, D_n is the set of walks in $\mathbb{Z}^n$ that start in the origin, remain in the positive quadrant $\mathbb{N}^n$, and end in the origin. For a word $w \in \Sigma_n^*$, $w = a_1 \cdots a_n$, $a_1, \ldots, a_n \in \Sigma$, let $\bar{w} = \bar{a}_1 \cdots \bar{a}_n$ and $w^{\mathsf{rev}} = a_n \cdots a_1$. Here, we set $\bar{\bar{a}} = a$ for $a \in \Sigma_n$.

For alphabets Σ, Γ, a subset $T \subseteq \Sigma^* \times \Gamma^*$ is a *rational transduction* if it is a homomorphic image of a regular language, i.e. if there is an alphabet Δ, a regular $K \subseteq \Delta^*$, and a homomorpism $h\colon \Delta^* \to \Sigma^* \times \Gamma^*$ such that $T = h(K)$. For a language $L \subseteq \Sigma^*$ and a subset $T \subseteq \Sigma^* \times \Gamma^*$, we define $TL = \{v \in \Gamma^* \mid \exists u \in L\colon (u, v) \in T\}$. It is well-known that if $S \subseteq \Sigma^* \times \Gamma^*$ and $T \subseteq \Delta^* \times \Sigma^*$ are rational transductions, then the relation $ST = \{(u, v) \in \Delta^* \times \Gamma^* \mid \exists w \in \Sigma^*\colon (u, w) \in T,\ (w, v) \in S\}$ and also $T^{-1} = \{(v, u) \in \Sigma^* \times \Delta^* \mid (u, v) \in T\}$ are rational transductions as well [1]. A language class $\mathcal{C}$ is called *full trio* if for every $L \subseteq \Sigma^*$ from $\mathcal{C}$, and every rational transduction $T \subseteq \Sigma^* \times \Gamma^*$, we also have TL in $\mathcal{C}$. The full trio *generated by* L is the class of all languages TL, where $T \subseteq \Sigma^* \times \Gamma^*$ is a rational transduction for some Γ. It is well-known that $\mathcal{V}_n$ ($\mathcal{Z}_n$) is the full trio generated by D_n (Z_n) [12].

By $\mathbb{Q}$, we denote the set of rational numbers. For $\mathbf{u}, \mathbf{v} \in \mathbb{Q}^n$, $\mathbf{u} = (u_1, \ldots, u_n)$, $\mathbf{v} = (v_1, \ldots, v_n)$, we define $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + \cdots + u_n v_n$ and $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$. Finally, for a subset $U \subseteq \mathbb{Q}^n$ and $\mathbf{v} \in \mathbb{Q}^n$, we denote $d(\mathbf{v}, U) = \inf\{\|\mathbf{v} - \mathbf{x}\| \mid \mathbf{x} \in U\}$.

Let us now recall previous results on regular separability for subclasses of VASS languages. The following was shown in [4].

▶ **Theorem 2.1.** *Given VASS languages $K, L \subseteq \Sigma^*$, it is decidable whether* $\Pi(K) \mid \Pi(L)$.

As observed in [5], Theorem 2.1 also implies the following.

▶ **Corollary 2.2.** *Given VASS languages K, L and words $w_1, \ldots, w_m \in \Sigma^*$ such that $K, L \subseteq w_1^* \cdots w_m^*$, it is decidable whether $K \mid L$.*

After Theorem 2.1, the next investigated subclass was that of 1-dim. VASS [7]:

▶ **Theorem 2.3.** *Given 1-dim. VASS V_0 and V_1, it is decidable whether $\mathsf{L}(V_0) \mid \mathsf{L}(V_1)$.*

Moreover, the next theorem has been established in [3].

▶ **Theorem 2.4.** *Given $\mathbb{Z}$-VASS languages $K, L \subseteq \Sigma^*$, it is decidable whether $K \mid L$.*

3 Main Results

In this section, we record the main results of this work. Our first main result is that regular separability is decidable if one input language is a VASS language and the other is the language of a one-dimensional VASS.

▶ **Theorem 3.1.** *Given a VASS V_0 and a 1-VASS V_1, it is decidable whether $\mathsf{L}(V_0) \mid \mathsf{L}(V_1)$.*

This generalizes Theorem 2.3, because here, one of the input languages can be an arbitrary VASS language. Our second main result is decidability of regular separability of a given VASS language from a given $\mathbb{Z}$-VASS language.

▶ **Theorem 3.2.** *Given a VASS V_0 and a $\mathbb{Z}$-VASS V_1, it is decidable whether $\mathsf{L}(V_0) \mid \mathsf{L}(V_1)$.*

As before, this significantly generalizes Theorem 2.4. In fact, Theorem 3.2 also generalizes Theorem 2.1, which is due to the following fact. Let $\Gamma_n = \{a_1, \ldots, a_n\} \subseteq \Sigma_n$.

▶ **Proposition 3.3.** *For $K, L \subseteq \Gamma^*$, we have $\Pi(K) \mid \Pi(L)$ if and only if $\Pi(K)\overline{\Pi(L)} \mid Z_n$.*

For the "only if" direction of Proposition 3.3, suppose there is a regular R with $\Pi(K) \subseteq R$ and $R \cap \Pi(L) = \emptyset$ and consider the language $S = \Pi(R \cap a_1^* \cdots a_n^*)$. Since $\Pi(K)$ and $\Pi(L)$ are commutative, S separates $\Pi(K)$ and $\Pi(L)$. By a classic result of Ginsburg and Spanier [10], as the commutative closure of a regular subset of $a_1^* \cdots a_n^*$, S is regular. Thus, the regular set $S \cdot \overline{\{a_1, \ldots, a_n\}^* \setminus S}$ includes $\Pi(K)\overline{\Pi(L)}$ and is disjoint from Z_n.

For the "if" direction of Proposition 3.3, we employ a general observation about separability. If M is a monoid, then we write $X \mid Y$ for subsets $X, Y \subseteq M$ if there is a recognizable subset $R \subseteq M$ with $X \subseteq R$ and $Y \cap R = \emptyset$. Moreover, we denote $\Delta = \{(m, m) \mid m \in M\}$.

▶ **Proposition 3.4.** *For $X, Y \subseteq M$, we have $X \mid Y$ if and only if $(X \times Y) \mid \Delta$.*

If $\Pi(K)\overline{\Pi(L)} \mid Z_n$, then in particular $\Pi(K)\overline{\Pi(L)} \mid E$, where $E = \{w\bar{w} \mid w \in \Gamma^*\}$, because $E \subseteq Z_n$. Consider the map $\tau\colon \Gamma^*\bar{\Gamma}^* \to \Gamma^* \times \Gamma^*$ with $\tau(u\bar{v}) = (u, v)$. Then τ is a bijection that preserves recognizability in both directions. Thus, $\Pi(K)\overline{\Pi(L)} \mid E$ implies $\tau(\Pi(K)\overline{\Pi(L)}) \mid \tau(E)$, which means $\Pi(K) \times \Pi(L) \mid \Delta$ and hence $\Pi(K) \mid \Pi(L)$ by Proposition 3.4.

4 VASS vs. One-Dimensional VASS

In this section, we introduce our approach to regular separability together with the first application: Regular separability of VASS languages and one-dimensional VASS languages.

Our approach is inspired by the decision procedure for regular separability for 1-VASS [7]. There, given languages K and L, the idea is to construct *approximants* K_k and L_k for $k \in \mathbb{N}$. Here, K_k and L_k are regular languages with $K \subseteq K_k$ and $L \subseteq L_k$ for which one can show that $K \mid L$ if and only if there is a $k \in \mathbb{N}$ with $K_k \cap L_k = \emptyset$. The latter condition is then checked algorithmically.

We simplify this idea in two ways. First, we show that for many language classes, one may assume that one of the two input languages is fixed (or fixed up to a parameter). Roughly speaking, if a language class $\mathcal{C}$ is defined by machines with a finite-state control, then $\mathcal{C}$ is typically a full trio since a transduction can be applied using a product construction in the finite-state control. Moreover, there is often a simple set $\mathcal{G}$ of languages so that $\mathcal{C}$ is the full trio generated by $\mathcal{G}$. For example, as mentioned above, $\mathcal{V}_n$ is generated by D_n for each $n \geq 1$. This makes the following simple lemma very useful.

▶ **Lemma 4.1.** *Let T be a rational transduction. Then $L \mid TK$ if and only if $T^{-1}L \mid K$.*

Proof. Suppose $L \subseteq R$ and $R \cap TK = \emptyset$ for some regular R. Then clearly $T^{-1}L \subseteq T^{-1}R$ and $T^{-1}R \cap K = \emptyset$. Therefore, the regular set $T^{-1}R$ witnesses $T^{-1}L \mid K$. Conversely, if $T^{-1}L \mid K$, then $K \mid T^{-1}L$ and hence, by the first direction, $(T^{-1})^{-1}K \mid L$. Since $(T^{-1})^{-1} = T$, this reads $TK \mid L$ and thus $L \mid TK$. ◀

Suppose we have full trios $\mathcal{C}_0$ and $\mathcal{C}_1$ generated by languages G_0 and G_1, respectively. Then, to decide if $T_0G_0 \mid T_1G_1$, we can check whether $T_1^{-1}T_0G_0 \mid G_1$. Since $T_1^{-1}T_0$ is also a rational transduction and hence $T_1^{-1}T_0G_0$ belongs to $\mathcal{C}_0$, this means we may assume that one of the input languages is G_1. This effectively turns separability into a decision problem with one input language L where we ask whether $L \mid G_1$.

Going further in this direction, instead of considering approximants of two languages, we just consider regular overapproximations of G_1 and decide whether L intersects all of them. However, we find it more convenient to switch to the complement and think in terms of *basic separators of* G_1 instead of overapproximations of G_1: By this, we mean a collection of regular languages where (i) each is disjoint from G_1 and (ii) every regular language R with $R \cap G_1 = \emptyset$ is included in a finite union of basic separators.

Basic separators for one-dimensional VASS Let us see this approach in an example and prove Theorem 3.1. Since $\mathcal{V}_1$ is generated as a full trio by D_1, Lemma 4.1 tells us that it suffices to decide whether a given VASS language L fulfills $L \mid D_1$. Now the first step is to develop a notion of basic separators for D_1.

Since $D_1 \subseteq \Sigma_1^*$, we assume now that $n = 1$, meaning $\varphi\colon \Sigma_1^* \to \mathbb{Z}$. One way a finite automaton can guarantee non-membership in D_1 is by modulo counting. For $k \in \mathbb{N}$, let

$$M_k = \{w \in \Sigma_1^* \mid \varphi(w) \not\equiv 0 \bmod k\},$$

which is regular. Another option is for an automaton to make sure an input word w avoids D_1 is to guarantee (i) for prefixes v of w, $\varphi(v)$ does not exceed some k if $\mathrm{drop}(v) = 0$ and (ii) $\varphi(w) \neq 0$. For $w \in \Sigma^*$, let $\mu(w) = \max\{\varphi(v) \mid v \text{ is a prefix of } w \text{ and } \mathrm{drop}(v) = 0\}$ and

$$P_k = \{w \in \Sigma^* \mid w \notin D_1 \text{ and } \mu(w) \leq k\}$$

These sets are regular: A word w with $\mu(w) \leq k$ avoids D_1 if and only if (i) φ drops below zero after a prefix where φ is confined to $[0, k]$ or (ii) φ stays above zero and thus assumes values in $[0, k]$ throughout. The third type of separator is a symmetric version of P_k, namely $\bar{P}_k^{\mathsf{rev}} = \{\bar{w}^{\mathsf{rev}} \mid w \in P_k\}$. These three types of languages form indeed basic sepators for D_1:

▶ **Lemma 4.2.** *Let $R \subseteq \Sigma^*$ be a regular language. Then $R \cap D_1 = \emptyset$ if and only if R is included in $M_k \cup P_\ell \cup \bar{P}_m^{\mathsf{rev}}$ for some $k, \ell, m \in \mathbb{N}$.*

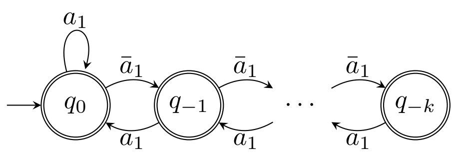


Figure 1 An automaton for $I_k \subseteq \Sigma_1^*$

The proof of Lemma 4.2 is very similar to the proof of Lemma 8 in [7], but phrased in a bit different setting. It is a relatively simple pumping argument: If R evades M_k, P_ℓ, and $\bar{P}_m^{\mathsf{rev}}$ for every $k, \ell, m \in \mathbb{N}$, then it has words w with $\varphi(w) \equiv 0 \bmod k$, $\mu(w) \geq \ell$, $\mu(\bar{w}^{\mathsf{rev}}) \geq m$ for arbitrarily high $k, \ell, m \in \mathbb{N}$. Then w has a prefix u with $\mathrm{drop}(u) = 0$ and $\varphi(u) \geq \ell$ and a suffix v with $\mathrm{drop}(\bar{v}^{\mathsf{rev}}) = 0$ $\varphi(\bar{w}^{\mathsf{rev}}) \geq m$. One can then pump a factor in the prefix and a factor in the suffix so as to (i) drive up the φ-values in the middle of the word and (ii) compensate the non-zero value of $\varphi(w)$. If k is a multiple of $n!$ and $\ell, m > n$, where n is the number of states of an automaton for R, then this yields a word in D_1. The full proof can be found in the appendix.

Deciding separability The next step in our approach is to decide whether a given VASS language L is contained in $M_k \cup P_\ell \cup \bar{P}_m^{\mathsf{rev}}$ for some $k, \ell, m \in \mathbb{N}$. Of course this is the case if and only if $L \subseteq M_k \cup P_k \cup \bar{P}_k^{\mathsf{rev}}$ for some $k \in \mathbb{N}$. Thus, Lemma 4.2 essentially tells us that whether $L \mid D_1$ holds only depends on three numbers associated to each word from L. Consider the function $\sigma \colon \Sigma^* \to \mathbb{N}^3$ with $\sigma(w) = (\mu(w), \varphi(w), \mu(\bar{w}^{\mathsf{rev}}))$. We call a subset $S \subseteq \mathbb{N}^3$ *separable* if there is a $k \in \mathbb{N}$ so that for every $(x_1, x_2, x_3) \in S$, we have $x_1 \leq k$ or $x_3 \leq k$ or $x_2 \not\equiv 0 \bmod k$. Then, Lemma 4.2 can be formulated as:

▶ **Lemma 4.3.** *Let $L \subseteq \Sigma^*$. If $L \cap D_1 = \emptyset$, then $L \mid D_1$ if and only if $\sigma(L)$ is separable.*

This enables us to modify L into a bounded language $\hat{L}$ that behaves the same in terms of separability from D_1. Let

$$\hat{L} = \{a^m \bar{a}^{m+1} a^r \bar{a}^s a^{n+1} \bar{a}^n \mid \exists w \in L \colon m \leq \mu(w),\ n \leq \mu(\bar{w}^{\mathsf{rev}}),\ r - s = \varphi(w)\}.$$

Note that if $v = a^m \bar{a}^{m+1} a^r \bar{a}^s a^{n+1} \bar{a}^n$, then $\mu(v) = m$, $\mu(\bar{v}^{\mathsf{rev}}) = n$, and $\varphi(v) = r - s$. Therefore, the set $\sigma(\hat{L})$ is separable if and only if $\sigma(L)$ is separable. Hence, we have:

▶ **Lemma 4.4.** *For every $L \subseteq \Sigma^*$ with $L \cap D_1 = \emptyset$, we have $L \mid D_1$ if and only if $\hat{L} \mid D_1$.*

Using standard VASS constructions, we can turn L into $\hat{L}$. Details are in the appendix.

▶ **Lemma 4.5.** *Given a VASS language $L \subseteq \Sigma^*$, one can construct a VASS for $\hat{L}$.*

This leaves us with the task of deciding whether $\hat{L} \mid D_1$. Since $\hat{L} \subseteq B$ with $B = a^* \bar{a}^* a^* \bar{a}^* a^* \bar{a}^*$, we have $\hat{L} \mid D_1$ if and only if $\hat{L} \mid (D_1 \cap B)$. As subsets of B, both $\hat{L}$ and $D_1 \cap B$ are bounded languages and we can decide whether $\hat{L} \mid (D_1 \cap B)$ using Corollary 2.2.

5 VASS vs. Integer VASS

In this section, we apply our approach to solving regular separability between a VASS language and a $\mathbb{Z}$-VASS language. Here, the collection of basic separators serves as a geometric characterization of separability and proving that it is a set of basic separators is more involved than in Section 4.

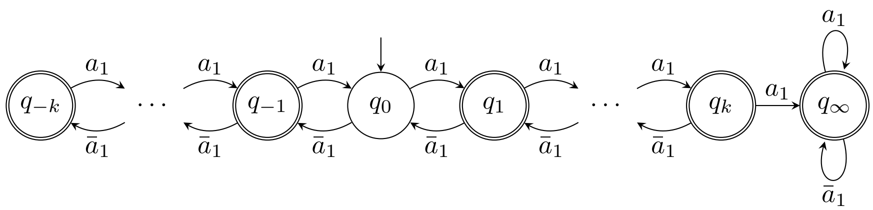


Figure 2 Automaton $\mathcal{A}_k$ with $\mathsf{L}(\mathcal{A}_k) \cap I_k = D_{1,k}$.

5.1 A geometric characterization

Lemma 4.1 tells us that regular separability between a VASS language and a $\mathbb{Z}$-VASS language amounts to checking whether a given VASS language $L \subseteq \Sigma_n^*$ is included in some regular language $R \subseteq \Sigma_n^*$ with $R \cap Z_n = \emptyset$. Therefore, in this section, we classify the regular languages $R \subseteq \Sigma_n^*$ with $R \cap Z_n = \emptyset$.

A very simple type of such languages is given by modulo counting. For $\mathbf{u}, \mathbf{v} \in \mathbb{Z}^n$, we write $\mathbf{u} \equiv \mathbf{v} \bmod k$ if $\mathbf{u}$ and $\mathbf{v}$ are component-wise congruent modulo k. The language

$$M_k = \{w \in \Sigma_n^* \mid \varphi(w) \not\equiv 0 \bmod k\}$$

is clearly regular and disjoint from Z_n.

Since Z_n is commutative (i.e. $\Pi(Z_n) = Z_n$), one might expect that it suffices to consider commutative separators. This is not the case: The language $L = (a_1\bar{a}_1)^* a_1^+$ is regularly separable from Z_1, but every commutative regular language including L intersects Z_1. Therefore, our second type of regular languages disjoint from Z_n is non-commutative and we start to describe it in the case $n = 1$. Consider the language

$$D_{1,k} = \{w \in \Sigma_1^* \mid \varphi(w) \neq 0 \text{ and for every infix } v \text{ of } w\text{: } \varphi(v) \geq -k\}.$$

The set $D_{1,k}$ is of course disjoint from Z_1. To see that $D_{1,k}$ is regular, observe first that the set $I_k = \{w \in \Sigma_1^* \mid \text{for every infix } v \text{ of } w\text{: } \varphi(v) \geq -k\}$ is regular, because it is accepted by the automaton in Figure 1: After reading a word w, the automaton's state reflects the difference $M - \varphi(w)$, where M is the maximal value $\varphi(v)$ for prefixes v of w. Second, the automaton $\mathcal{A}_k$ in Figure 2 satisfies $\mathsf{L}(\mathcal{A}_k) \cap I_k = D_{1,k}$: As long as the seen prefix w satisfies $\varphi(w) \in [-k, k]$, the state of $\mathcal{A}_k$ reflects $\varphi(w)$ exactly. However, as soon as $\mathcal{A}_k$ encounters a prefix w with $\varphi(w) > k$, it enters q_∞. From there, it accepts every following suffix, because an input from I_k can never reach 0 under φ with such a prefix w. Thus, $D_{1,k}$ is regular.

The language $D_{1,k}$ has analogs in higher dimension. Instead of making sure the value of φ never drops more than k along one particular axis, one can impose this condition in an arbitrary direction $\mathbf{u} \in \mathbb{Z}^n$. For every vector $\mathbf{u} \in \mathbb{Z}^n$ and $k \in \mathbb{N} \setminus \{0\}$, let

$$D_{\mathbf{u},k} = \{w \in \Sigma_n^* \mid \langle \varphi(w), \mathbf{u} \rangle \neq 0 \text{ and for every infix } v \text{ of } w\text{: } \langle \varphi(v), \mathbf{u} \rangle \geq -k\}.$$

To see that $D_{\mathbf{u},k}$ is regular, consider the morphism $h \colon \Sigma_n^* \to \{a_1, \bar{a}_1\}^*$ with $x \mapsto x^{\langle \varphi(x), \mathbf{u} \rangle}$ for $x \in \Sigma_n$. Here, we mean $a_i^\ell = \bar{a}_i^{|\ell|}$ and $\bar{a}_i^\ell = a_i^{|\ell|}$ in case $\ell \in \mathbb{Z}$, $\ell < 0$. Then $\langle \varphi(w), \mathbf{u} \rangle = \varphi(h(w))$ for any $w \in \Sigma_n^*$ and hence $D_{\mathbf{u},k} = h^{-1}(D_{1,k})$, meaning $D_{\mathbf{u},k}$ inherits regularity from $D_{1,k}$.

The main result of this section is that the sets M_k and $D_{\mathbf{u},k}$ suffice to explain disjointness of regular languages from Z_n in the following sense.

▶ **Theorem 5.1.** *Let $R \subseteq \Sigma_n^*$ be a regular language. Then $R \cap Z_n = \emptyset$ if and only if R is included in a finite union of languages of the form M_k and $D_{\mathbf{u},k}$ for $k \in \mathbb{N}$ and $\mathbf{u} \in \mathbb{Z}^n$.*

We therefore say that $L \subseteq \Sigma_n^*$ is *geometrically separable* if L is contained in a finite union of languages of the form M_k and $D_{\mathbf{u},k}$. Then, we can formulate Theorem 5.1 as a geometric characterization of separability from Z_n.

▶ **Corollary 5.2.** *For $L \subseteq \Sigma_n^*$, we have $L \mid Z_n$ if and only if L is geometrically separable.*

The remainder of Section 5.1 is devoted to proving Theorem 5.1.

Mapping to lower dimension Suppose we are given a vector space $U \subseteq \mathbb{Q}^n$ (represented by a basis) with $m = \dim U < n$ and a bound $\ell \geq 0$. Then we define the set

$$S_{U,\ell} = \{w \in \Sigma_n^* \mid \text{for every prefix } v \text{ of } w\text{: } d(\varphi(v), U) \leq \ell\}.$$

Hence, $S_{U,\ell}$ collects those words whose prefixes stay close to the subspace U. If a language L is contained in $S_{U,\ell}$, then in many respects, we will be able to treat L as a language in Σ_m^*, where $m = \dim U$. More concretely, we translate L to a subset of Σ_m^* using a transducer.

The transducer will consist of three parts, *coordinate transformation* (f), *intersection* ($R_{V,p}$), and *projection* (π_m). For the coordinate transformation, we choose an orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{Z}^n$ of $\mathbb{Q}^n$ such that $\mathbf{b}_1, \ldots, \mathbf{b}_m$ is a basis for U. This can be done, e.g. using Gram-Schmidt orthogonalisation [20]. If $B \in \mathbb{Z}^{n\times n}$ is the matrix whose columns are $\mathbf{b}_1, \ldots, \mathbf{b}_n$, then B is invertible and maps $V = \{(v_1, \ldots, v_n) \in \mathbb{Q}^n \mid v_{m+1} = \cdots = v_n = 0\}$ to U. Thus the inverse $B^{-1} \in \mathbb{Q}^n$ maps U to V. We can clearly choose an $\alpha \in \mathbb{Z}$ such that $\alpha B^{-1} \in \mathbb{Z}^{n\times n}$ and we set $A = \alpha B^{-1}$. For each $i \in [1, n]$, choose a word $w_i \in \Sigma_n^*$ with $\varphi(w_i) = A\varphi(a_i)$ and let $f\colon \Sigma_n^* \to \Sigma_n^*$ be the morphism with $f(a_i) = w_i$ and $f(\bar{a}_i) = \bar{w}_i$.

▶ **Lemma 5.3.** *If $f(L)$ is geometrically separable for $L \subseteq \Sigma_n^*$, then so is L. Moreover, we have $L \cap Z_n = \emptyset$ if and only if $f(L) \cap Z_n = \emptyset$.*

The first statement holds because $f^{-1}(M_k) \subseteq M_k$ and $f^{-1}(D_{\mathbf{u},k}) \subseteq D_{A^\top\mathbf{u},k}$, where $A^\top$ is the transpose of A. The second is due to A being injective and linear.

For the second step of our transformation (intersection), we observe that prefixes of words in $S_{V,p}$ can only assume finitely many values in the last $n - m$ coordinates under φ. To make this precise, let $\bar{\pi}_\ell\colon \mathbb{Q}^n \to \mathbb{Q}^\ell$ denote the projection on the last ℓ coordinates, $\bar{\pi}_\ell(v_1, \ldots, v_n) = (v_{n-\ell+1}, \ldots, v_n)$. Then we have $d(\mathbf{v}, V) = \|\bar{\pi}_{n-m}(\mathbf{v})\|$ for every $\mathbf{v} \in \mathbb{Q}^n$. If v is a prefix of $w \in S_{V,p}$, then $d(\varphi(v), V) \leq p$ implies $\|\bar{\pi}_{n-m}(\varphi(v))\| \leq p$ and hence there is a finite set $F \subseteq \mathbb{Z}^{n-m}$ such that $\bar{\pi}_\ell(\varphi(v)) \in F$ for every prefix v of some $w \in S_{V,p}$. Thus, the set $R_{V,p} = \{w \in S_{V,p} \mid \varphi(w) \in V\}$ is regular. The second step of our transformation is to intersect with $R_{V,p}$.

▶ **Lemma 5.4.** *Let $L \subseteq S_{V,p}$. If $L \cap R_{V,p}$ is geometrically separable, then so is L. Moreover, $L \cap Z_n = \emptyset$ if and only if $(L \cap R_{V,p}) \cap Z_n = \emptyset$.*

The first statement is due to $L \setminus R_{V,p}$ consisting of words w where for every prefix v, we have $\|\bar{\pi}_{n-m}(\varphi(w))\| \leq p$ and also $\bar{\pi}_{n-m}(\varphi(w)) \neq 0$. Thus, $L \setminus R_{V,p}$ is included in M_k for some $k \in \mathbb{N}$. The same reasoning gives the second statement.

In our third step, we project onto the first m coordinates: Let $\pi_m\colon \Sigma_n^* \to \Sigma_m^*$ be the morphism with $\pi_m(a_i) = a_i$, $\pi_m(\bar{a}_i) = \bar{a}_i$, $i \in [1, m]$, and $\pi_m(a_i) = \pi_m(\bar{a}_i) = \varepsilon$, $i \in [m+1, n]$.

▶ **Lemma 5.5.** *Let $L \subseteq R_{V,p}$. If $\pi_m(L)$ is geometrically separable, then so is L. Moreover, $L \cap Z_n = \emptyset$ if and only if $\pi_m(L) = \emptyset$.*

This is because the coordinates $1, \ldots, m$ are always zero in these walks, so it is easy to translate the basic separators from L to $\pi_m(L)$.

We are now prepared to define our transformation: Let $T_{U,\ell} \subseteq \Sigma_n^* \times \Sigma_m^*$ the transduction with $T_{U,\ell}L = \pi_m(f(L) \cap R_{V,p})$. Lemmas 5.3 to 5.5 clearly imply:

▶ **Proposition 5.6.** *Let $L \subseteq S_{U,\ell}$. If $T_{U,\ell}L \subseteq \Sigma_m^*$ is geometrically separable, then so is L. Also, $L \cap Z_n = \emptyset$ if and only if $(T_{U,\ell}L) \cap Z_m = \emptyset$. Thus, $L \mid Z_n$ if and only if $(T_{U,\ell}L) \mid Z_m$.*

Cones of automata For a set $S \subseteq \mathbb{Q}^n$, the *cone generated by S* consists of all vectors $x_1\mathbf{u}_1 + \cdots + x_\ell\mathbf{u}_\ell$ where $x_1, \ldots, x_\ell \in \mathbb{Q}_+$ and $\mathbf{u}_1, \ldots, \mathbf{u}_\ell \in S$. To each automaton $\mathcal{A}$ over Σ_n, we associate a cone as follows. If $w \in \Sigma_n^*$ labels a path in $\mathcal{A}$, then $\varphi(w)$ is the *effect* of that path. Let $\mathrm{cone}(\mathcal{A})$ be the cone generated by the effects of cycles of $\mathcal{A}$. Since every cycle effect is the sum of effects of simple cycles, we know that $\mathrm{cone}(\mathcal{A})$ is generated by the effects of simple cycles. In particular, $\mathrm{cone}(\mathcal{A})$ is finitely generated and the set of simple cycle effects can serve as a representation of $\mathrm{cone}(\mathcal{A})$. A key ingredient in our proof is a dichotomy of cones (Lemma 5.7), which is a direct consequence of the well-known Farkas' lemma [34].

▶ **Lemma 5.7.** *For every $\mathcal{A}$, either $\mathrm{cone}(\mathcal{A}) = \mathbb{Q}^n$ or $\mathrm{cone}(\mathcal{A})$ is included in some half-space.*

Linear automata For an automaton $\mathcal{A}$, consider the directed acyclic graph (dag) consisting of strongly connected components of $\mathcal{A}$. If this dag is a path, then $\mathcal{A}$ is called *linear*. Given an automaton $\mathcal{A}$, we can construct linear automata $\mathcal{A}_1, \ldots, \mathcal{A}_\ell$ with $\mathsf{L}(\mathcal{A}) = \mathsf{L}(\mathcal{A}_1) \cup \cdots \cup \mathsf{L}(\mathcal{A}_\ell)$.

▶ **Lemma 5.8.** *Let $\mathcal{A}$ be a linear automaton such that $\mathrm{cone}(\mathcal{A}) = \mathbb{Q}^n$. If $\mathsf{L}(\mathcal{A}) \cap Z_n = \emptyset$, then $\mathsf{L}(\mathcal{A})$ is included in M_k for some k.*

Proof. Since $\mathrm{cone}(\mathcal{A}) = \mathbb{Q}^n$, we have in particular $\mathbf{e}_1, -\mathbf{e}_1, \ldots, \mathbf{e}_n, -\mathbf{e}_n \in \mathrm{cone}(\mathcal{A})$, meaning that there are cycles labeled $w_1, \ldots, w_p$ such that $\mathbf{e}_i, -\mathbf{e}_i \in \textsc{Lin}_{\mathbb{Q}_+}(\varphi(w_1), \ldots, \varphi(w_p))$ for every $i \in [1, n]$. Therefore, there is a $k \in \mathbb{N}$ such that $k \cdot \mathbf{e}_i, -k \cdot \mathbf{e}_i \in \textsc{Lin}_{\mathbb{N}}(\varphi(w_1), \ldots, \varphi(w_p))$ for every $i \in [1, n]$. We claim that $\mathsf{L}(\mathcal{A}) \subseteq M_k$. Towards a contradiction, suppose $w \in \mathsf{L}(\mathcal{A})$ with $\varphi(w) \equiv 0 \bmod k$. Since $\mathcal{A}$ is linear, we can take the run for w and insert cycles so that the resulting run visits every state in $\mathcal{A}$. Instead of inserting every cycle once, we insert it k times, so that the resulting run (i) visits every state in $\mathcal{A}$ and (ii) reads a word $w' \in \Sigma_n^*$ with $\varphi(w') \equiv \varphi(w) \bmod k$. Now since $\varphi(w') \equiv \varphi(w) \equiv 0 \bmod k$, we can write $-\varphi(w') = x_1\varphi(w_1) + \cdots + x_p\varphi(w_p)$ with coefficients $x_1, \ldots, x_p \in \mathbb{N}$. Since in the run for w', every state of $\mathcal{A}$ is visited, we can insert cycles corresponding to the $w_1, \ldots, w_p$: For each $i \in [1, p]$, insert the cycle for w_i exactly x_i times. Let w'' be the word read by the resulting run and note that $w'' \in \mathsf{L}(\mathcal{A})$. Then we have $\varphi(w'') = \varphi(w') + x_1\varphi(w_1) + \cdots + x_p\varphi(w_p) = 0$ and thus $w'' \in Z_n$, contradicting $\mathsf{L}(\mathcal{A}) \cap Z_n = \emptyset$. ◀

▶ **Lemma 5.9.** *Let $\mathcal{A}$ be an automaton such that $\mathrm{cone}(\mathcal{A})$ is contained in some half-space. One can compute $k, \ell \in \mathbb{N}$, $\mathbf{u} \in \mathbb{Z}^n$, and a strict subspace $U \subseteq \mathbb{Q}^n$ such that $\mathsf{L}(\mathcal{A}) \subseteq D_{\mathbf{u},k} \cup S_{U,\ell}$.*

Proof. Suppose $\mathrm{cone}(A) \subseteq H$, where $H = \{\mathbf{v} \in \mathbb{Q}^n \mid \langle \mathbf{v}, \mathbf{u} \rangle \geq 0\}$ for some vector $\mathbf{u} \in \mathbb{Q}^n \setminus \{0\}$. Without loss of generality, we may assume $\mathbf{u} \in \mathbb{Z}^n \setminus \{0\}$. Let $U = \{\mathbf{v} \in \mathbb{Q}^n \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0\}$. Then clearly $\dim U = n - 1$. Observe that since $\mathrm{cone}(\mathcal{A}) \subseteq H$, we have $\langle \mathbf{v}, \mathbf{u} \rangle \geq 0$ for every cycle effect $\mathbf{v} \in \mathbb{Z}^n$ of $\mathcal{A}$. Let k be the number of states in $\mathcal{A}$. Now, whenever $w \in \mathsf{L}(\mathcal{A})$ and v is an infix of w, then $\langle \varphi(v), \mathbf{u} \rangle \geq -k$: If $\langle \varphi(v), \mathbf{u} \rangle < -k$, then the path reading v must contain a cycle reading $v' \in \Sigma_n^*$ with $\langle \varphi(v'), \mathbf{u} \rangle < 0$, which contradicts $\mathrm{cone}(\mathcal{A}) \subseteq H$.

We claim that $\mathsf{L}(\mathcal{A}) \subseteq D_{\mathbf{u},k} \cup S_{U,k}$. Let $w \in \mathsf{L}(\mathcal{A})$. We distinguish two cases. *Case 1:* Suppose w has a prefix v with $\langle \varphi(v), \mathbf{u} \rangle > k$. Write $w = vv'$. As argued above, we have

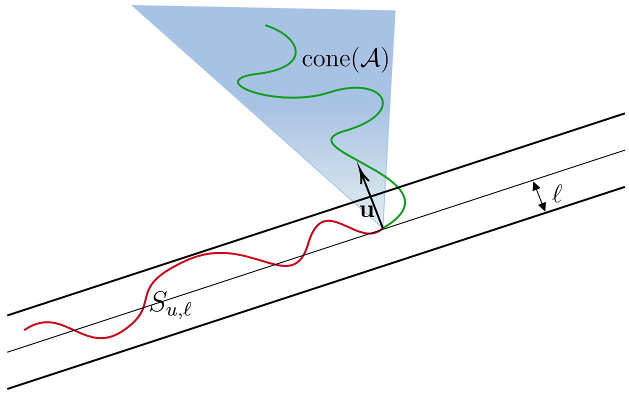


Figure 3 Two runs (red and green) inside $D_{\mathbf{u},k} \cup S_{U,\ell}$.

$\langle \varphi(v'), \mathbf{u} \rangle \geq -k$. Hence, $\langle \varphi(w), \mathbf{u} \rangle = \langle \varphi(v), \mathbf{u} \rangle + \langle \varphi(v'), \mathbf{u} \rangle > 0$. Thus, we have $w \in D_{\mathbf{u},k}$.
Case 2: Suppose for every prefix v of w, we have $\langle \varphi(v), \mathbf{u} \rangle \leq k$. Then, for every prefix v of w, we have $-k \leq \langle \varphi(v), \mathbf{u} \rangle \leq k$ and thus $d(\varphi(v), U) = |\langle \varphi(v), \mathbf{u} \rangle| / \|\mathbf{u}\| \leq k$ (see Lemma C.3). In particular, $w \in S_{U,k}$. ◀

Les us now prove Theorem 5.1. Suppose $R \subseteq \Sigma_n^*$ and $R \cap Z_n = \emptyset$. We show by induction on the dimension n that then, R is included in a finite union of sets of the form M_k and $D_{\mathbf{u},k}$. Let $R = \mathsf{L}(\mathcal{A})$ for an automaton $\mathcal{A}$. Since $\mathcal{A}$ can be decomposed into a finite union of linear automata, it suffices to prove the claim in the case that $\mathcal{A}$ is linear. If $\text{cone}(\mathcal{A}) = \mathbb{Q}^n$, then Lemma 5.8 tells us that $\mathsf{L}(\mathcal{A}) \subseteq M_k$ for some $k \in \mathbb{N}$. If $\text{cone}(\mathcal{A})$ is contained in some half-space, then according to Lemma 5.9, we have $R \subseteq D_{\mathbf{u},k} \cup S_{U,\ell}$ for some $\mathbf{u} \in \mathbb{Q}^n \setminus \{0\}$, $k, \ell \in \mathbb{N}$, and strict subspace $U \subseteq \mathbb{Q}^n$. This implies that the regular language $R \setminus D_{\mathbf{u},k}$ is included in $S_{U,\ell}$. We may therefore apply Proposition 5.6, which yields $T_{U,\ell}(R \setminus D_{\mathbf{u},k}) \cap Z_m = \emptyset$. Since $T_{U,\ell}(R \cap S_{U,\ell}) \subseteq \Sigma_m^*$ with $m = \dim U < n$, induction tells us that $T_{U,\ell}(R \setminus D_{\mathbf{u},k})$ is geometrically separable and hence, by Proposition 5.6, $R \setminus D_{\mathbf{u},k}$ is geometrically separable. Since $R \subseteq D_{\mathbf{u},k} \cup (R \setminus D_{\mathbf{u},k})$, R is geometrically separable.

5.2 Vector addition systems

In this section, we apply Corollary 5.2 to prove Theorem 3.2. Corollary 5.2 tells us that in order to decide whether $L \mid Z_n$ for $L \subseteq \Sigma_n^*$, it suffices to check whether there are $k \in \mathbb{N}$, $\ell_1, \ldots, \ell_m \in \mathbb{N}$, and vectors $\mathbf{u}_1, \ldots, \mathbf{u}_m \in \mathbb{Z}^n$ such that $L \subseteq M_k \cup D_{\mathbf{u}_1,\ell_1} \cup \cdots \cup D_{\mathbf{u}_m,\ell_m}$. It is tempting to conjecture that there is a finite collection of direction vectors $F \subseteq \mathbb{Z}^n$ (such as a basis together with negations) so that for a given language L, such an inclusion holds only if it holds with some $\mathbf{u}_1, \ldots, \mathbf{u}_m \in F$. In that case we would only need to consider scalar products of words in L with vectors in F and thus reformulate the problem over sections of reachability sets of VASS. However, this is not the case. For $\mathbf{u}, \mathbf{v} \in \mathbb{Q}^n$, we write $\mathbf{u} \sim \mathbf{v}$ if $\mathbb{Q}_+\mathbf{u} = \mathbb{Q}_+\mathbf{v}$. Since $\sim$ has infinitely many equivalence classes and every class intersects $\mathbb{Z}^n$, the following shows that there is no fixed set of directions.

▶ **Proposition 5.10.** *For each $\mathbf{u} \in \mathbb{Z}^n$, there is a $k_0 \in \mathbb{N}$ such that for $k \geq k_0$, the following holds. For every $\ell, \ell_1 \ldots, \ell_m \geq 1$, $\mathbf{u}_1, \ldots, \mathbf{u}_m \in \mathbb{Z}^n$ with $\mathbf{u}_i \not\sim \mathbf{u}$ for $i \in [1, n]$, we have $D_{\mathbf{u},k} \not\subseteq M_\ell \cup D_{\mathbf{u}_1,\ell_1} \cup \cdots D_{\mathbf{u}_m,\ell_m}$.*

We now turn to the proof of Theorem 3.2. According to Corollary 5.2, we have to decide whether a given VASS language $L \subseteq \Sigma_n^*$ satisfies $L \subseteq M_k \cup D_{\mathbf{u}_1,\ell_1} \cup \cdots \cup D_{\mathbf{u}_m,\ell_m}$ for some $k \in \mathbb{N}$,

$\mathbf{u}_1, \ldots, \mathbf{u}_m \in \mathbb{Z}^n$ and $\ell_1, \ldots, \ell_m \in \mathbb{N}$. Our algorithm employs the KLMST decomposition used by Sacerdote and Tenney [33], Mayr [26], Kosaraju [18], and Lambert [19] and recently cast in terms of ideal decompositions by Leroux and Schmitz [25]. The decomposition yields VASS languages $L_1, \ldots, L_p$ so that $L = L_1 \cup \cdots \cup L_p$ so that it suffices to check whether L_i is geometrically separable for each $i \in [1, n]$. This is easier than for L, because the decomposition also provides a structure for each L_i that will guide our algorithm.

Modular envelopes Before we dive into the KLMST decomposition, let us describe what information it yields. We say that an automaton $\mathcal{A}$ is a *modular envelope* for a language $L \subseteq \Sigma^*$ if (i) $L \subseteq L(\mathcal{A})$ and (ii) for every selection $u_1, \ldots, u_m \in \Sigma^*$ of words from $\text{Loop}(\mathcal{A})$ and every $w \in L$ and every $k \in \mathbb{N}$, there is a word $w' \in L$ so that each u_j is a factor of w' and $\Psi(w') \equiv \Psi(w) \bmod k$. In other words, $\mathcal{A}$ describes a regular overapproximation that is small enough that we can place every selection of loops in a word from L whose Parikh image is congruent modulo k to a given word. Using the KLMST decomposition, we prove:

▶ **Theorem 5.11.** *Given a VASS language L, one can construct VASS languages $L_1, \ldots, L_p$, together with a modular envelope $\mathcal{A}_i$ for each L_i such that $L = L_1 \cup \cdots \cup L_p$.*

We postpone the proof of Theorem 5.11 until later and first show how Theorem 5.11 can be used to decide whether L is geometrically separable.

Modular envelopes with cone $\mathbb{Q}^n$ By Lemma 5.7 we know that every cone either equals $\mathbb{Q}^n$ or is included in some half-space. The following lemma will be useful in the first case.

▶ **Lemma 5.12.** *Let $L \subseteq \Sigma_n^*$ be a language with a modular envelope $\mathcal{A}$. If* $\text{cone}(\mathcal{A}) = \mathbb{Q}^n$ *then the following are equivalent: (i) $L \mid Z_n$, (ii) $L \subseteq M_k$ for some $k \in \mathbb{N}$, (iii)* $\Pi(L) \mid Z_n$.

Proof. Note that (ii) implies (iii) immediately and that (iii) implies (i) because $L \subseteq \Pi(L)$. Thus, we only need to show that (i) implies (ii). By Corollary 5.2 if $L \mid Z_n$, then L is included in some $M_k \cup D_{\mathbf{u}_1,k} \cup \cdots \cup D_{\mathbf{u}_m,k}$. We show that in our case actually $L \subseteq M_k$.

Take any $w \in L$. We aim at constructing $w' \in L$ such that $w' \notin D_{\mathbf{u}_i,k}$ for every $i \in [1, m]$ and additionally $\varphi(w') \equiv \varphi(w) \bmod k$. Since $\text{cone}(\mathcal{A}) = \mathbb{Q}^n$, for every $\mathbf{u}_j$, there exist a loop v_j in $\mathcal{A}$ such that $\langle \varphi(v_j), \mathbf{u}_j \rangle < 0$. Since $\varphi(v_j), \mathbf{u}_j \in \mathbb{Z}^n$, we even have $\langle \varphi(v_j), \mathbf{u}_j \rangle \leq -1$. Since each v_j belongs to $\text{Loop}(\mathcal{A})$, the words v_j^{k+1} also belong to $\text{Loop}(\mathcal{A})$.

Since $\mathcal{A}$ is a modular envelope, there exists a word $w' \in L$ such that $\Psi(w) \equiv \Psi(w') \bmod k$ and all the words $v_1^{k+1}, \ldots, v_m^{k+1}$ are factors of w'. Recall that every factor u of every word in $D_{\mathbf{u}_i,k}$ has $\langle \varphi(u), \mathbf{u}_j \rangle \geq -k$. However, we have $\langle \varphi(v_j^{k+1}), \mathbf{u}_j \rangle \leq -(k+1)$. Thus, w' cannot belong to $D_{\mathbf{u}_i,k}$ for $i \in [1, m]$. Since $L \subseteq M_k \cup D_{\mathbf{u}_1,k} \cup \cdots \cup D_{\mathbf{u}_m,k}$, this only leaves $w' \in M_k$. Since $\Psi(w') \equiv \Psi(w) \bmod k$, we also have $\varphi(w') \equiv \varphi(w) \bmod k$ and thus $w \in M_k$. ◀

We are now prepared to explain the decision procedure for Theorem 3.2. The algorithm is illustrated in Algorithm 1. If $n = 0$, then $\Sigma_n = \emptyset$ and thus either $L = \emptyset$ or $L = \{\varepsilon\}$, meaning $L \mid Z_n$ if and only if $L \neq \emptyset$. Otherwise, we perform the KLMST decomposition, which, as explained in the next section, yields languages $L_1 \cup \cdots \cup L_p$ and modular envelopes $\mathcal{A}_1, \ldots, \mathcal{A}_p$ such that $L = L_1 \cup \cdots \cup L_n$. Since then $L \mid Z_n$ if and and only if $L_i \mid Z_n$ for each $i \in [1, p]$, we check for the latter. For each $i \in [1, p]$, the dichotomy of Lemma 5.7 guides a case distinction: If $\text{cone}(\mathcal{A}_i) = \mathbb{Q}^n$, then we know from Lemma 5.12 that $L_i \mid Z_n$ if and only if $\Pi(L_i) \mid Z_n$, which can be checked via Theorem 2.1.

If, however, $\text{cone}(\mathcal{A}_i)$ is contained in some half-space $H = \{\mathbf{x} \in \mathbb{Q}^n \mid \langle x, \mathbf{u} \rangle \geq 0\}$ for some $\mathbf{u} \in \mathbb{Z}^n$, then Lemma 5.9 tells us that $L_i \subseteq \mathsf{L}(\mathcal{A}_i) \subseteq D_{\mathbf{u},k} \cup S_{U,\ell}$ for some computable $k, \ell \in \mathbb{N}$

Algorithm 1: Deciding separability of a VASS language L from Z_n

```
Input: n ∈ ℕ and VASS language L = L(V) ⊆ Σ_n^*
if n = 0 and L = ∅ then return "yes"
if n = 0 and L ≠ ∅ then return "no"
Use KLMST decomposition to compute VASS languages L_1, ..., L_p, together with
  modular envelopes A_1, ..., A_p.
for i ∈ [1, p] do
    if cone(A_i) = ℚ^n then
        Check whether Π(L_i)|Z_n
        if not Π(L_i)|Z_n then return "no"
    end
    if cone(A_i) ⊆ H = {x ∈ ℚ^n | ⟨x, u⟩ ≥ 0} for some u ∈ ℤ^n \ {0} then
        Let U = {x ∈ ℚ^n | ⟨x, u⟩ = 0}.                      /* dim U = n − 1 */
        Compute k, ℓ ∈ ℕ with L_i ⊆ L(A_i) ⊆ D_{u,k} ∪ S_{U,ℓ}     /* L_i \ D_{u,k} ⊆ S_{U,ℓ} */
        Compute rational transduction T_{U,ℓ} ⊆ Σ_n^* × Σ_{n−1}^*
        Check recursively whether T_{U,ℓ}(L_i \ D_{u,k})|Z_{n−1} /* T_{U,ℓ}(L_i \ D_{u,k}) ⊆ Σ_{n−1}^* */
        if not T_{U,ℓ}(L_i \ D_{u,k})|Z_{n−1} then return "no"
    end
end
return "yes"
```

and $U = \{\mathbf{x} \in \mathbb{Q}^n \mid \langle \mathbf{x}, \mathbf{u} \rangle = 0\}$. In particular, we have $L_i \mid Z_n$ if and only if $L_i \setminus D_{\mathbf{u},k} \mid Z_n$. Note that $L_i \setminus D_{\mathbf{u},k} = L \cap (\Sigma_n^* \setminus D_{\mathbf{u},k})$ is a VASS language and is included in $S_{U,\ell}$. Thus, the walks corresponding to the words in L_i always stay close to the hyperplane U, which has dimension $n-1$. We can therefore use the transduction $T_{U,\ell}$ to transform L_i into a set of walks in $(n-1)$-dimensional space and decide separability recursively for the result: We have $T_{U,\ell}L_i \subseteq \Sigma_{n-1}^*$ and Proposition 5.6 tells us that $L_i \mid Z_n$ if and only if $T_{U,\ell}L_i \mid Z_{n-1}$.

Constructing modular envelopes As mentioned above, we use the KLMST decomposition (so named by Leroux and Schmitz [25] after its inventors Mayr [26], Kosaraju [18], Lambert [19] and Sacerdote and Tenney [33]) to construct the decomposition of L into $L_1, \ldots, L_p$ with modular envelopes $\mathcal{A}_1, \ldots, \mathcal{A}_p$. This decomposition is relatively unwieldy in terms of required concepts and terminology. Moreover, the construction of modular envelopes is not the main innovation of this paper but in large part a combination of existing methods. Therefore, we decided to include only a very high-level overview and keep the details in the appendix.

The decomposition yields perfect marked graph-transition sequences (MGTS) $\mathcal{N}_1, \ldots, \mathcal{N}_p$. Each MGTS $\mathcal{N}_i$ defines a VASS language L_i so that $L = L_1 \cup \cdots \cup L_p$. The MGTS are readily translated into finite automata $\mathcal{A}_1, \ldots, \mathcal{A}_p$ for which it is known that $\mathsf{L}(\mathcal{A}_i)$ overapproximates L_i [5]. For the second property of modular envelopes, we use a concept of *run amalgamations* as introduced by Leroux and Schmitz [25]. It permits the amalgamation of two runs into which a third run embeds via the order $\trianglelefteq$ on runs introduced by Jančar [16] and Leroux [24].

Given a word $w \in L_i$ and $u_1, \ldots, u_m \in \text{Loop}(\mathcal{A})$, we use Lambert's pumping lemma [19] to construct a run ρ_1 in $\mathcal{N}_i$ that contains $u_1, \ldots, u_m$ as factors, but that also embeds the run ρ for w. We argue that this embedding has a special property that guarantees that amalgamating ρ and ρ_1 yields a run in $\mathcal{N}_i$ that still embeds ρ with this property and still contains the u_i as factors. We can then repeat this process $(k-1)$-fold to obtain a run ρ_k reading a word w_k where $\Psi(w_k) = k \cdot (\Psi(w_1) - \Psi(w)) + \Psi(w)$ and thus $\Psi(w_k) \equiv \Psi(w) \bmod k$.

References

1 J. Berstel. *Transductions and context-free languages*. Teubner Studienbücher, Stuttgart, 1979.

2 Mikołaj Bojańczyk. It is undecidable if two regular tree languages can be separated by a deterministic tree-walking automaton. *Fundam. Inform.*, 154(1-4):37–46, 2017.

3 Lorenzo Clemente, Wojciech Czerwiński, Sławomir Lasota, and Charles Paperman. Regular separability of Parikh automata. In *Proceedings of ICALP '17*, pages 117:1–117:13, 2017.

4 Lorenzo Clemente, Wojciech Czerwiński, Sławomir Lasota, and Charles Paperman. Separability of reachability sets of vector addition systems. In *Proceedings of STACS '17*, pages 24:1–24:14, 2017.

5 Wojciech Czerwińsk, Piotr Hofman, and Georg Zetzsche. Unboundedness problems for languages of vector addition systems. In *Proceedings of ICALP '18*, pages 119:1–119:15, 2018.

6 Wojciech Czerwińsk, Slawomir Lasota, Ranko Lazic, Jérôme Leroux, and Filip Mazowiecki. The reachability problem for petri nets is not elementary (extended abstract). *CoRR*, abs/1809.07115, 2018.

7 Wojciech Czerwiński and Sławomir Lasota. Regular separability of one counter automata. In *Proceedings of LICS '17*, pages 1–12, 2017.

8 Wojciech Czerwiński, Slawomir Lasota, Roland Meyer, Sebastian Muskalla, K. Narayan Kumar, and Prakash Saivasan. Regular separability of well-structured transition systems. In *Proceedings of CONCUR '18*, pages 35:1–35:18, 2018.

9 Wojciech Czerwiński, Wim Martens, Lorijn van Rooijen, Marc Zeitoun, and Georg Zetzsche. A characterization for decidable separability by piecewise testable languages. *Discrete Mathematics & Theoretical Computer Science*, 19(4), 2017.

10 Seymour Ginsburg and Edwin H. Spanier. Bounded regular sets. *Proceedings of the American Mathematical Society*, 17(5):1043–1049, 1966.

11 Jean Goubault-Larrecq and Sylvain Schmitz. Deciding piecewise testable separability for regular tree languages. In *Proceedings of ICALP '16*, pages 97:1–97:15, 2016.

12 S. A. Greibach. Remarks on blind and partially blind one-way multicounter machines. *Theoretical Computer Science*, 7(3):311 – 324, 1978.

13 Christoph Haase and Simon Halfon. Integer vector addition systems with states. In *Proceedings of RP '14*, pages 112–124, 2014.

14 John E. Hopcroft and Jean-Jacques Pansiot. On the reachability problem for 5-dimensional vector addition systems. *Theor. Comput. Sci.*, 8:135–159, 1979.

15 Harry B. Hunt III. On the decidability of grammar problems. *Journal of the ACM*, 29(2):429–447, 1982.

16 Petr Jančar. Decidability of a temporal logic problem for Petri nets. *Theor. Comput. Sci.*, 74(1):71–93, 1990.

17 Petr Jančar. Deciding structural liveness of petri nets. In *Proceedings of SOFSEM '17*, pages 91–102, 2017.

18 S. Rao Kosaraju. Decidability of reachability in vector addition systems (preliminary version). In *Proceedings of STOC '82*, pages 267–281, 1982.

19 Jean-Luc Lambert. A structure to decide reachability in Petri nets. *Theor. Comput. Sci.*, 99(1):79–104, 1992.

20 Serge Lang. *Linear Algebra*. Addison-Wesley, 1966.

21 Ranko Lazić and Sylvain Schmitz. Nonelementary complexities for branching VASS, MELL, and extensions. *ACM Trans. Comput. Log.*, 16(3):20:1–20:30, 2015.

22 Ranko Lazić and Patrick Totzke. What makes Petri nets harder to verify: Stack or data? In *Concurrency, Security, and Puzzles — Essays Dedicated to Andrew William Roscoe on the Occasion of His 60th Birthday*, volume 10160 of *LNCS*, pages 144–161. Springer, 2017.

23 Jérôme Leroux. The general vector addition system reachability problem by Presburger inductive invariants. In *Proceedings of LICS '09*, pages 4–13, 2009.

24 Jérôme Leroux. Vector addition system reachability problem: a short self-contained proof. In *Proceedings of POPL '11*, pages 307–316, 2011.

25 Jérôme Leroux and Sylvain Schmitz. Demystifying reachability in vector addition systems. In *Proceedings of LICS'15*, pages 56–67, 2015.

26 Ernst W. Mayr. An algorithm for the general Petri net reachability problem. In *Proceedings of STOC'81*, pages 238–246, 1981.

27 Thomas Place. Separating regular languages with two quantifiers alternations. In *Proceedings of LICS '15*, pages 202–213, 2015.

28 Thomas Place, Lorijn van Rooijen, and Marc Zeitoun. Separating regular languages by piecewise testable and unambiguous languages. In *Proceedings of MFCS '13*, pages 729–740, 2013.

29 Thomas Place and Marc Zeitoun. Separation and the successor relation. In *Proceedings of STACS '15*, pages 662–675, 2015.

30 Thomas Place and Marc Zeitoun. Separating regular languages with first-order logic. *Logical Methods in Computer Science*, 12(1), 2016.

31 Thomas Place and Marc Zeitoun. Concatenation hierarchies: New bottle, old wine. In *Proceedings of CSR '17*, pages 25–37, 2017.

32 Thomas Place and Marc Zeitoun. Separation for dot-depth two. In *Proceedings of LICS '17*, pages 1–12, 2017.

33 George S. Sacerdote and Richard L. Tenney. The decidability of the reachability problem for vector addition systems (preliminary version). In *Proceedings of STOC '77*, pages 61–76. ACM, 1977.

34 Alexander Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, 1986.

35 Thomas G. Szymanski and John H. Williams. Noncanonical extensions of bottom-up parsing techniques. *SIAM Journal on Computing*, 5(2), 1976.

36 Georg Zetzsche. Separability by piecewise testable languages and downward closures beyond subwords. In *Proceedings of LICS '18*, pages 929–938, 2018.

A Missing proofs from Section 3

A.1 Proof of Proposition 3.4

In our proof of Proposition 3.4, we use a well-known fact about regular separability of unions.

▶ **Lemma A.1.** *Let $X = \bigcup_{i=1}^n X_i$ and $Y = \bigcup_{j=1}^m Y_j$ for subsets $X, Y \subseteq M$. Then $X \mid Y$ if and only if $X_i \mid Y_j$ for every $i \in [1, n]$ and $j \in [1, m]$.*

Proof. A separator witnessing $X \mid Y$ also witnesses $X_i \mid Y_j$ for every $i \in [1, n]$ and $j \in [1, m]$. This shows the "only if" direction.

For the "if" direction, suppose $R_{i,j} \subseteq M$ satisfies $X_i \subseteq R_{i,j}$ and $R_{i,j} \cap Y_j = \emptyset$. We claim that $R = \bigcup_{i=1}^n \bigcap_{j=1}^m R_{i,j}$ witnesses $X \mid Y$.

Since $X_i \subseteq R_{i,j}$ for every $i \in [1, n]$, we have $X_i \subseteq \bigcap_{j=1}^m R_{i,j}$ and hence $X = \bigcup_{i=1}^n X_i \subseteq R$. On the other hand, for every $i \in [1, n]$ and $k \in [1, m]$, we have $Y_k \cap R_{i,k} = \emptyset$ and thus $Y_k \cap \bigcap_{j=1}^m R_{i,j} = \emptyset$. This implies

$$Y \cap R = \bigcup_{k=1}^m Y_k \cap \bigcup_{i=1}^n \bigcap_{j=1}^m R_{i,j} = \bigcup_{k=1}^m \bigcup_{i=1}^n \underbrace{Y_k \cap \bigcap_{j=1}^m R_{i,j}}_{=\emptyset} = \emptyset.$$

◀

Proof of Proposition 3.4. One direction is immediate: If $X \mid Y$ with a recognizable separator $S \subseteq M$, then $X \times Y$ is separated from Δ with the recognizable set $S \times (M \setminus S)$.

Suppose $(X \times Y) \mid \Delta$ is witnessed by $S \subseteq M \times M$ with $X \times Y \subseteq S$ and $S \cap \Delta = \emptyset$. We can write $S = \bigcup_{i=1}^{n} R_i \times T_i$ for recognizable subsets $R_i, T_i \subseteq M$ for $i \in [1,n]$. Note that then $(R_i \times T_i) \cap \Delta = \emptyset$ and thus $T_i \subseteq M \setminus R_i$. Moreover, we have $X \subseteq \bigcup_{i=1}^{n} R_i$ and $Y \subseteq \bigcup_{i=1}^{n} T_i \subseteq \bigcup_{i=1}^{n} (M \setminus R_i)$.

For any $I \subseteq [1,n]$, let $R_I = \bigcap_{i\in I} R_i \cap \bigcap_{i\in[1,n]\setminus I} (M \setminus R_i)$, $X_I = X \cap R_I$, and $Y_I = Y \cap R_I$. We claim that for any $I, J \subseteq [1,n]$, we have $X_I | Y_J$. Since $X = \bigcup_{\emptyset\neq I\subseteq[1,n]} X_I$ and $Y = \bigcup_{I\subsetneq[1,n]} Y_I$, the proposition then follows from Lemma A.1.

Suppose $I = J$. We shall prove that then either $X_I = \emptyset$ or $Y_J = \emptyset$, which clearly implies $X_I \mid Y_J$. Toward a contradiction, assume that there are $x \in X_I$ and $y \in Y_J$. Since $X \times Y \subseteq S$, there is an $i \in [1,n]$ with $x \in R_i$ and $y \in T_i \subseteq (M \setminus R_i)$. The former implies $i \in I$, and the latter $i \notin J$, contradicting $I = J$.

Suppose $I \neq J$. If there is an $i \in I \setminus J$, then $X_I \subseteq R_I \subseteq R_i$ and $Y_J \subseteq R_J \subseteq M \setminus R_i$, meaning that R_i witnesses $X_I \mid Y_J$. On the other hand, if $i \in J \setminus I$, then $X_I \subseteq M \setminus R_i$ and $Y_I \subseteq R_i$, so that $M \setminus R_i$ witnesses $X_I \mid Y_J$. ◀

B Missing proofs from Section 4

B.1 Proof of Lemma 4.2

Proof. The "if" direction is obvious, so let us prove the "only if". Suppose $R \cap D_1 = \emptyset$ and $R = \mathsf{L}(\mathcal{A})$ for an automaton $\mathcal{A}$ with n states. We claim that then $R \subseteq M_{n!} \cup P_n \cup \bar{P}_n^{\mathsf{rev}}$.

Towards a contradiction, we assume that there is a word $w \in R$ with $w \notin M_{n!} \cup P_n \cup \bar{P}_n^{\mathsf{rev}}$. This means $\varphi(w) \equiv 0 \mod n!$ and w has a prefix u' with $\mathrm{drop}(u') = 0$ and $\varphi(u') = \mu(w) > n$ and a suffix v' with $\mathrm{drop}(\bar{v'}^{\mathsf{rev}}) = 0$ and $\varphi(\bar{v'}^{\mathsf{rev}}) = \mu(\bar{v'}^{\mathsf{rev}}) > n$, meaning $\varphi(v') < -n$. Let u be the shortest prefix of w with $\varphi(u) = \varphi(u')$ and let v be the shortest suffix with $\varphi(v') = \varphi(v)$. Then $|u| \leq |u'|$ and $|v| \leq |v'|$, which means in particular $\mathrm{drop}(u), \mathrm{drop}(v) = 0$.

Let us show that u and v do not overlap in w, i.e. $|w| \geq |u| + |v|$. If they do overlap, we can write $w = xyz$ so that $u = xy$ and $v = yz$ with $y \neq \varepsilon$. Then by minimality of u, we have $\varphi(x) < \varphi(xy)$ and thus $\varphi(y) > 0$. Symmetrically, minimality of v yields $\varphi(\bar{z}^{\mathsf{rev}}) < \varphi(\bar{yz}^{\mathsf{rev}})$ and thus $-\varphi(y) = \varphi(\bar{y}^{\mathsf{rev}}) > 0$, contradicting $\varphi(y) > 0$. Thus u and v do not overlap and we can write $w = uw'v$.

Since $\varphi(u) > n$, we can decompose $u = u_1u_2u_3$ so that $1 \leq \varphi(u_2) \leq n$ and in the run of $\mathcal{A}$ for w, u_2 is read on a cycle. Analogously, since $\varphi(v) < -n$, we can decompose $v = v_1v_2v_3$ so that $-n \leq \varphi(v_2) \leq -1$ and v_2 is read on a cycle.

Since $\varphi(w) \equiv 0 \bmod n!$ and $\varphi(u_2) \in [1,n]$ and $\varphi(v_2) \in [-n,-1]$, there are $p, q \in \mathbb{N}$ with $\varphi(w) + p\varphi(u_2) + q\varphi(v_2) = 0$. Moreover, we also have

$$\varphi(w) + (p + r|\varphi(v_2)|)\varphi(u_2) + (q + r|\varphi(u_2)|)\varphi(v_2) = 0 \tag{1}$$

for every $r \in \mathbb{N}$. Consider the word

$$w'' = u_1 u_2^{p+r|\varphi(v_2)|} u_3 w' v_1 v_2^{q+r|\varphi(u_2)|} v_3.$$

Since u_2 and v_2 are read on cycles, we have $w'' \in R$. Moreover, Equation (1) tells us that $\varphi(w'') = 0$. Finally, since $\mathrm{drop}(u) = 0$ and $\varphi(u_2) > 0$, for large enough r, we have $\mathrm{drop}(w'') = 0$ and hence $w'' \in D_1$. This is in contradiction to $R \cap D_1 = \emptyset$. ◀

B.2 Proof of Lemma 4.5

To prove Lemma 4.5, it is convenient to have a notion of subsets of $\Sigma^* \times \mathbb{N}^m$ described by vector addition systems. First, a *vector addition system* (VAS) is a VASS that has only one state. Since it has only one state, it is not mentioned in the configurations or the transitions. We say that $R \subseteq \Sigma^* \times \mathbb{N}^m$ is a *VAS relation* if there is a $d+m$-dimensional VAS V and vector $\mathbf{s}, \mathbf{t} \in \mathbb{N}^d$ such that $R = \{(w, \mathbf{u}) \in \Sigma^* \times \mathbb{N}^m \mid (\mathbf{s}, 0) \xrightarrow{w} (\mathbf{t}, \mathbf{u})\}$. Here, $\mathbf{s}$ and $\mathbf{t}$ are called *source* and *target vector*, respectively.

However, sometimes it is easier to describe a relation by a VASS than by a VAS. We say that $R \subseteq \Sigma^* \times \mathbb{N}^m$ is *decribed by the $d+m$-dimensional VASS $V = (Q, T, s, t, h)$* if $R = \{(w, \mathbf{u}) \in \Sigma^* \times \mathbb{N}^m \mid (s, 0, 0) \xrightarrow{w} (t, 0, \mathbf{u})\}$. Of course, a relation is a VAS relation if and only if it is described by some VASS and these descriptions are easily translated.

▶ **Lemma B.1.** *If $R \subseteq \Sigma^* \times \mathbb{N}^m$ and $S \subseteq \Sigma^* \times \mathbb{N}^n$ are VAS relations, then so is the relation $R \oplus S := \{(w, \mathbf{u}, \mathbf{v}) \mid (w, \mathbf{u}) \in R \wedge (w, \mathbf{v}) \in S\}$.*

Proof. We employ a simple product construction. Suppose V_0 describes R and V_1 describes S. Without loss of generality, let V_0 and V_1 be $d+m$-dimensional and $d+n$-dimensional, respectively. The new VAS V is $2d+m+n$-dimensional and has three types of transitions: First, for any letter $a \in \Sigma$, every transition $(\mathbf{u}_0, \mathbf{v}_0) \in \mathbb{Z}^{d+m}$ of V_0 with label a and $\mathbf{u}_0 \in \mathbb{Z}^d$ and $\mathbf{v}_0 \in \mathbb{Z}^m$, every transition $(\mathbf{u}_1, \mathbf{v}_1) \in \mathbb{Z}^{d+n}$ of V_1 with label a and $\mathbf{u}_1 \in \mathbb{Z}^d$ and $\mathbf{v}_1 \in \mathbb{Z}^n$, V has a transition $(\mathbf{u}_0, \mathbf{u}_1, \mathbf{v}_0, \mathbf{v}_1) \in \mathbb{Z}^{2d+m+n}$ with label a.

Second, for every transition $(\mathbf{u}, \mathbf{v}) \in \mathbb{Z}^{d+m}$ from V_0 labeled ε with $\mathbf{u} \in \mathbb{Z}^d$ and $\mathbf{v} \in \mathbb{Z}^m$, V has an ε-labeled transition $(\mathbf{u}, 0^d, \mathbf{v}, 0^n) \in \mathbb{Z}^{2d+m+n}$. Here, in slight abuse of notation, 0^k is meant to be a vector of zeros that occupies k coordinates. Third, for every transition $(\mathbf{u}, \mathbf{v}) \in \mathbb{Z}^{d+n}$ labeled ε from V_1 with $\mathbf{u} \in \mathbb{Z}^d$ and $\mathbf{v} \in \mathbb{Z}^m$, V has an ε-labeled transition $(0^d, \mathbf{u}, 0^m, \mathbf{v})$. If $\mathbf{s}_i$ and $\mathbf{t}_i$ are start and target vector of V_i for $i \in \{0, 1\}$, then $\mathbf{s} = (\mathbf{s}_0, \mathbf{s}_1)$ and $\mathbf{t} = (\mathbf{t}_0, \mathbf{t}_1)$ are used as start and target vectors for V. Then, it is routine to check that indeed $\{(w, \mathbf{u}, \mathbf{v}) \mid \mathbf{u} \in \mathbb{Z}^m, \mathbf{v} \in \mathbb{Z}^n, (\mathbf{s}, 0) \xrightarrow{w} (\mathbf{t}, \mathbf{u}, \mathbf{v})\} = R \oplus S$. ◀

▶ **Lemma B.2.** *Given a VAS language $L \subseteq \Sigma^*$ and a VAS relation $R \subseteq \Sigma^* \times \mathbb{N}^m$ one can construct a VAS for the language $\{a_1^{x_1} \cdots a_m^{x_m} \mid \exists w \in L \colon R(w, x_1, \ldots, x_m)\}$.*

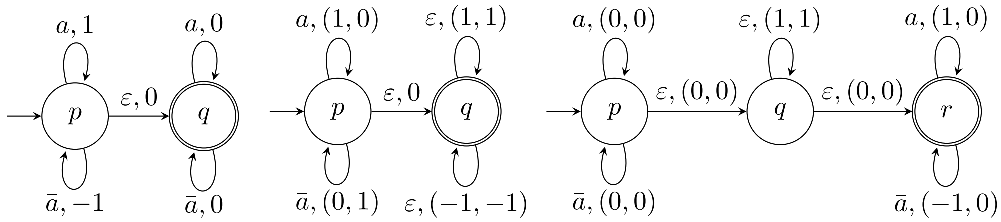
Proof. Suppose V is a d-dimensional VAS accepting L and V' is a $d+m$-dimensional VAS for R. We construct the $2d+m$-dimensional VAS V'', which has four types of transitions. First, for every transition $\mathbf{u} \in \mathbb{Z}^d$ labeled $a \in \Sigma$ and every a-labeled transition $\mathbf{v} \in \mathbb{Z}^{d+m}$ in V', we have an ε labeled transition $(\mathbf{u}, \mathbf{v})$ in V''. Second, for every ε-labeled transition $\mathbf{u} \in \mathbb{Z}^d$ in V, we have an ε-labeled transition $(\mathbf{u}, 0) \in \mathbb{Z}^{2d+m}$ in V''. Third, for every ε-labeled transition $\mathbf{u} \in \mathbb{Z}^{d+m}$ in V', we have an ε-labeled transition $(0, \mathbf{u}) \in \mathbb{Z}^{2d+m}$ transition in V''. Fourth, for every $i \in [1, m]$, we have an a_i-labeled transition $(0, -\mathbf{e}_i) \in \mathbb{Z}^{2d+m}$, where $\mathbf{e}_i \in \mathbb{Z}^m$ is the m-dimensional unit vector with 1 in coordinate i and 0 everywhere else. It is now easy construct a VASS V''' with $\mathsf{L}(V''') = \mathsf{L}(V'') \cap a_1^* \cdots a_n^*$. Then clearly, we have $\mathsf{L}(V''') = \{a_1^{x_1} \cdots a_m^{x_m} \mid \exists w \in L \colon R(w, x_1, \ldots, x_m)\}$. ◀

Proof of Lemma 4.5. First, let us show that the following relations are VAS relations:

$$R_1 = \{(w, n) \in \Sigma^* \times \mathbb{N} \mid n \leq \mu(w)\},$$
$$R_2 = \{(w, r, s) \in \Sigma^* \times \mathbb{N}^2 \mid r - s = \varphi(w)\},$$
$$R_3 = \{(w, n) \in \Sigma^* \times \mathbb{N} \mid n \leq \mu(\bar{w}^{\mathsf{rev}})\}.$$

In Figures 4a to 4c, we show vector addition systems with states for the relations R_1, R_2, and R_3 (it is easy to translate them to VAS for the relations). From the VASS for R_1 and R_3, one can readily build VAS for the relations

$$R_1' = \{(w, m, m+1) \in \Sigma^* \times \mathbb{N}^2 \mid m \leq \mu(w)\},$$
$$R_3' = \{(w, n+1, n) \in \Sigma^* \times \mathbb{N}^2 \mid m \leq \mu(\bar{w}^{\mathsf{rev}})\}.$$



(a) VASS for R_1 **(b)** VASS for R_2 **(c)** VASS for R_3

According to Lemma B.1, we can construct a VAS for $R = R'_1 \oplus R_2 \oplus R'_3 \subseteq \Sigma^* \times \mathbb{N}^6$. Applying Lemma B.2 to L and R yields a VAS for the language

$$\{a_1^m a_2^{m+1} a_3^r a_4^s a_5^{n+1} a_6^n \mid \exists w \in L\colon m \le \mu(w),\ r - s = \varphi(w),\ n \le \mu(\bar{w}^{\mathsf{rev}})\}.$$

Now appropriately renaming the symbols $a_1, \ldots, a_6$ to a or $\bar{a}$ yields a VAS for $\hat{L}$. ◀

C Missing proofs from Section 5

C.1 Proof of Lemma 5.3

Proof. For the first statement, we prove that $f^{-1}(M_k) \subseteq M_k$ and $f^{-1}(D_{\mathbf{u},k}) \subseteq D_{A^\top \mathbf{u},k}$, which clearly suffices. Note that if $w \in \Sigma_n^*$ with $\varphi(w) \equiv 0 \bmod k$, then also $\varphi(f(w)) = A\varphi(w) \equiv 0 \bmod k$. This implies $f^{-1}(M_k) \subseteq M_k$. For the second inclusion, suppose $w \in \Sigma_n^*$ with $f(w) \in D_{\mathbf{u},k}$ and let v be a prefix of of w. Then $f(v)$ is a prefix of $f(w)$ and thus

$$\langle \varphi(v), A^\top \mathbf{u}\rangle = \varphi(v)^\top A^\top \mathbf{u} = (A\varphi(v))^\top \mathbf{u} = \langle A\varphi(v), \mathbf{u}\rangle = \langle \varphi(f(v)), \mathbf{u}\rangle.$$

In particular, we have $\langle \varphi(v), A^\top \mathbf{u}\rangle = \langle \varphi(f(v)), \mathbf{u}\rangle \ge -k$ and $\langle \varphi(w), A^\top \mathbf{u}\rangle = \langle \varphi(f(w)), \mathbf{u}\rangle > 0$, which implies $w \in D_{A^\top \mathbf{u},k}$.

For the second statement, note that A is invertible, meaning $\varphi(f(w)) = A\varphi(w)$ vanishes if and only if $\varphi(w)$ vanishes. ◀

C.2 Proof of Lemma 5.4

Proof. Suppose $L \cap R_{V,p}$ is geometrically separable. Let $\bar{R}_{V,p} = \{w \in S_{V,p} \mid \varphi(w) \notin V\}$. Then $S_{V,p} = \bar{R}_{V,p} \cup R_{V,p}$. It suffices to show that $\bar{R}_{V,p} \subseteq M_k$ for some $k \in \mathbb{N}$, because then

$$L = (L \cap \bar{R}_{V,p}) \cup (L \cap R_{V,p}) \subseteq M_k \cup (L \cap R_{V,p}) \tag{2}$$

and $L \cap R_{V,p}$ being geometrically separable implies that L is geometrically separable as well.

To show that $\bar{R}_{V,p} \subseteq M_k$, let $F \subseteq \mathbb{Z}^{n-m}$ be a finite set such that $\bar{\pi}_{n-m}(\varphi(v)) \in F$ for every prefix v of a word $w \in S_{V,p}$. Moreover, choose $k \in \mathbb{N}$ so that $k > \|\mathbf{v}\|$ for every $\mathbf{v} \in F$. We claim that then $\bar{R}_{V,p} \subseteq M_k$. To this end, suppose $w \in \bar{R}_{V,p}$. Then $d(\varphi(w), V) \neq 0$ and hence $\bar{\pi}_{n-m}(\varphi(w)) \in F \setminus \{0\}$. In particular, we have $\varphi(w) \not\equiv 0 \bmod k$ and thus $w \in M_k$. This proves $\bar{R}_{V,p} \subseteq M_k$.

For the second statement, note that $L \cap Z_n = \emptyset$ clearly implies $(L \cap R_{V,p}) \cap Z_n = \emptyset$. Conversely, if $(L \cap R_{V,p}) \cap Z_n = \emptyset$, then Equation (2) entails $L \cap Z_n \subseteq (L \cap R_{V,p}) \cap Z_n = \emptyset$ because $M_k \cap Z_n = \emptyset$. ◀

C.3 Proof of Lemma 5.5

Proof. It suffices to show that for $w \in R_{V,p}$, two implications hold: (i) if $\pi_m(w) \in M_k$ for some $k \in \mathbb{N}$, then $w \in M_k$ and (ii) if $\pi_m(w) \in D_{\mathbf{u},k}$ for some $\mathbf{u} \in \mathbb{Z}^m$ and $k \in \mathbb{N}$, then $w \in D_{\mathbf{u}',k}$ for some $\mathbf{u}' \in \mathbb{Z}^n$.

Suppose $w \in R_{V,p}$ and $\pi_m(w) \in M_k$. Since $w \in R_{V,p}$, the last $n-m$ components of $\varphi(w)$ are zero. Thus, we have $\varphi(w) \equiv 0 \bmod k$ if and only if $\varphi(\pi_m(w)) \equiv 0 \bmod k$. This implies $w \in M_k$.

Now suppose $w \in R_{V,p}$ with $\pi_m(w) \in D_{\mathbf{u},k}$ for some $\mathbf{u} \in \mathbb{Z}^m$ and $k \in \mathbb{N}$. Let $\mathbf{u} = (u_1, \dots, u_m)$ and define $\mathbf{u}' = (u_1, \dots, u_m, 0, \dots, 0)$. Then clearly, $\langle \varphi(v), \mathbf{u}' \rangle = \langle \varphi(\pi_m(v)), \mathbf{u} \rangle$ for every word $v \in \Sigma_n^*$. In particular, we have $w \in D_{\mathbf{u}',k}$. ◀

C.4 Proof of Proposition 5.6

Proposition 5.6 almost follows from Lemmas 5.3 to 5.5. We only have to show that if $L \subseteq S_{U,\ell}$, then $T_{U,\ell}L \subseteq S_{V,p}$ for some computable $p \in \mathbb{N}$. For this, we have the following lemma. It is intuitively clear: Walks that stay close to U are mapped to walks that are close to $AU = V$.

▶ **Lemma C.1.** *We can compute some $p \in \mathbb{N}$ with $f(S_{U,\ell}) \subseteq S_{V,p}$.*

Proof. Choose $k \in \mathbb{N}$ so that $k \geq |f(a_i)|$ and $k \geq |f(\bar{a}_i)|$ for $i \in [1,n]$ and let $p = \|A\| \cdot \ell + k$. We claim that $f(S_{U,\ell}) \subseteq S_{V,p}$. Let $w \in S_{U,\ell}$.

Consider a prefix v of $f(w)$. Let us first consider the case that $v = f(u)$ for some prefix u of w. Since $w \in S_{U,\ell}$, we have $d(\varphi(u), U) \leq \ell$. Therefore,

$$\begin{aligned} d(\varphi(f(u)), V) &= d(A\varphi(u), AU) = \inf\{\|A\varphi(u) - A\mathbf{u}\| \mid \mathbf{u} \in U\} \\ &\leq \|A\| \cdot \inf\{\|\varphi(u) - \mathbf{u}\| \mid \mathbf{u} \in U\} = \|A\| \cdot d(\varphi(u), U) = \|A\| \cdot \ell. \end{aligned}$$

Now if v is any prefix of $f(w)$, then $v = f(u)v'$, where u is a prefix of w and $|v'| \leq k$. This implies that $d(\varphi(v), V) \leq d(\varphi(u), V) + k \leq \|A\| \cdot \ell + k = p$. ◀

Proof of Proposition 5.6. With Lemma C.1, the first two statements of Proposition 5.6 follow directly from Lemmas 5.3 to 5.5. Let us prove the conclusion in the second statement.

If $L \mid Z_n$ with a regular R with $L \subseteq R$ and $R \cap Z_n = \emptyset$, then by Proposition 5.6, we have $T_{U,\ell}R \cap Z_m = \emptyset$. Hence, $T_{U,\ell}R$ separates $T_{U,\ell}L$ and Z_m. Conversely, if $T_{U,\ell}L \mid Z_m$, then by Corollary 5.2, the language $T_{U,\ell}L$ is geometrically separable. According to Proposition 5.6, that implies that L is geometrically separable and in particular $L \mid Z_n$. ◀

C.5 Proof of Lemma 5.7

Let us recall the Farkas' lemma from linear programming [34, Corollary 7.1d].

▶ **Lemma C.2** (Farkas' Lemma)**.** *For every $A \in \mathbb{Q}^{n\times m}$ and $\mathbf{b} \in \mathbb{Q}^n$, exactly one of the following holds:*

1. *There exists an $\mathbf{x} \in \mathbb{Q}^n$, $\mathbf{x} \geq 0$, with $A\mathbf{x} = \mathbf{b}$.*
2. *There exists a $\mathbf{y} \in \mathbb{Q}^n$ with $\mathbf{y}^\top A \geq 0$ and $\langle \mathbf{y}, \mathbf{b} \rangle < 0$.*

Proof of Lemma 5.7. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{Z}^n$ be the effects of all simple cycles of $\mathcal{A}$ and let $C \in \mathbb{Z}^{n\times k}$ be the matrix with columns $\mathbf{u}_1, \dots, \mathbf{u}_k$. Then cone($\mathcal{A}$) consists of those vectors of the form $C\mathbf{x}$ with $\mathbf{x} \in \mathbb{Q}^n_+$. If cone($\mathcal{A}$) $\neq \mathbb{Q}^n$, then there is a vector $\mathbf{v} \in \mathbb{Q}^n$ with $\mathbf{v} \notin$ cone($\mathcal{A}$). This means the system of inequalities $C\mathbf{x} = \mathbf{v}$, $\mathbf{x} \geq 0$, does not have a solution. By Farkas' lemma, there exists a vector $\mathbf{y} \in \mathbb{Q}^n$ with $\mathbf{y}^\top C \geq 0$ and $\langle \mathbf{y}, \mathbf{v} \rangle < 0$. Hence, for every element $C\mathbf{x}$, $\mathbf{x} \in \mathbb{Q}^n_+$, of cone($\mathcal{A}$), we have $\langle \mathbf{y}, C\mathbf{x} \rangle = \mathbf{y}^\top C\mathbf{x} \geq 0$. Since $\mathbf{y} \in \mathbb{Q}^n$, there is a $k \in \mathbb{N}$ so that $\mathbf{u} = k\mathbf{y} \in \mathbb{Z}^n$. Then we have cone($\mathcal{A}$) $\subseteq \{\mathbf{x} \in \mathbb{Q}^n \mid \langle \mathbf{x}, \mathbf{u} \rangle \geq 0\}$. ◀

C.6 Distance from hyperplanes

▶ **Lemma C.3.** *Let* $\mathbf{u} \in \mathbb{Q}^n$ *and* $U = \{\mathbf{v} \in \mathbb{Q}^n \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0\}$. *Then* $d(\mathbf{v}, U) = \frac{|\langle \mathbf{v}, \mathbf{u} \rangle|}{\|\mathbf{u}\|}$ *for* $\mathbf{v} \in \mathbb{Q}^n$.

Proof. We extend $\mathbf{u}$ to an orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of $\mathbb{Q}^n$, meaning $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$ if $i \neq j$ and $\mathbf{b}_1 = \mathbf{u}$. Because of orthogonality, we may express $\|\mathbf{v}\|$ for any vector $\mathbf{v} \in \mathbb{Q}^n$ with $\mathbf{v} = v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n$ as

$$\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\langle v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n, v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n \rangle} = \sqrt{\sum_{i=1}^{n} v_i^2 \langle \mathbf{b}_i, \mathbf{b}_i \rangle} = \sqrt{\sum_{i=1}^{n} v_i^2 \|\mathbf{b}_i\|^2}.$$

Let $\mathbf{x} \in \mathbb{Q}^n$ be a vector with $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$. Since $\mathbf{b}_1 = \mathbf{u}$ and thus $\langle \mathbf{x}, \mathbf{u} \rangle = x_1$, the vector $\mathbf{x}$ belongs to U if and only if $x_1 = 0$. Therefore, for $\mathbf{v} \in \mathbb{Q}^n$ with $\mathbf{v} = v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n$ and $\mathbf{x} \in U$, we have

$$\|\mathbf{v} - \mathbf{x}\| = \sqrt{v_1^2\|\mathbf{u}\|^2 + \sum_{i=2}^{n} (v_i - x_i)^2 \|\mathbf{b}_i\|^2}.$$

This distance is minimal with $x_i = v_i$ for $i \in [2, n]$ and in that case, the distance is $d(\mathbf{v}, U) = \sqrt{v_1^2\|\mathbf{u}\|^2} = |v_1| \cdot \|\mathbf{u}\|$. Since $\langle \mathbf{v}, \mathbf{u} \rangle = v_1\|\mathbf{u}\|^2$, that implies $d(\mathbf{v}, U) = |\langle \mathbf{v}, \mathbf{u} \rangle| / \|\mathbf{u}\|$. ◀

C.7 Proof of Proposition 5.10

Proof. The idea of the proof is as follows. We construct a word w corresponding to a walk in $\mathbb{Z}^n$, which traverses long distances in many directions orthogonal to $\mathbf{u}$. That way, w cannot belong to any of the languages $D_{\mathbf{u}_i,\ell_i}$ for $i \in [1, m]$. Moreover, we carefully design the construction such that $w \notin M_\ell$. Furthermore, the walk never moves far in the direction of $-\mathbf{u}$ because that would imply $w \notin D_{\mathbf{u},k}$.

We begin by choosing $k_0 \in \mathbb{N}$. We extend the vector $\mathbf{u}$ to an orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{Z}^n$ of $\mathbb{Q}^n$, meaning that $\mathbf{b}_1 = \mathbf{u}$ and $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$ if $i \neq j$. Note that since $\mathbf{b}_i \neq 0$, we then have $\langle \mathbf{b}_i, \mathbf{b}_i \rangle = \|\mathbf{b}_i\|^2 \neq 0$. In particular, this means for every $\mathbf{v} \in P = \{\mathbf{b}_1, \mathbf{b}_2, -\mathbf{b}_2, \ldots, \mathbf{b}_n, -\mathbf{b}_n\}$, we have $\langle \mathbf{v}, \mathbf{u} \rangle \geq 0$. Note that except for $\mathbf{b}_1$, the set P contains every vector $\mathbf{b}_i$ positively and negatively.

For each $i \in [1, 2n-1]$, we pick a word $v_i \in \Sigma_n^*$ so that $\{\varphi(v_1), \ldots, \varphi(v_{2n-1})\} = P$. Note that then, we have $\langle \varphi(v_i), \mathbf{u} \rangle \geq 0$ for every $i \in [1, 2n-1]$. Choose $k_0 \in \mathbb{N}$ so that $k_0 \geq 2|v_i|$ for each $i \in [1, 2n-1]$.

To show $D_{\mathbf{u},k} \not\subseteq M_\ell \cup D_{\mathbf{u}_1,\ell_1} \cup \cdots \cup D_{\mathbf{u}_m,\ell_m}$, suppose $k \geq k_0$. We shall construct a word $w \in D_{\mathbf{u},k}$ so that $w \notin M_\ell$ and $w \notin D_{\mathbf{u}_i,\ell_i}$ for every $i \in [1, m]$. Pick $s \in \mathbb{N}$ with $s > \ell_i$ for $i \in [1, m]$. Write $\mathbf{u} = (x_1, \ldots, x_n)$ and let $u = a_1^{x_1} \cdots a_n^{x_n}$. Here, in slight abuse of notation, if $x_i < 0$, we mean $\bar{a}_i^{|x_i|}$ instead of $a_i^{x_i}$. Then clearly, we have $\varphi(u) = \mathbf{u}$ and every infix z of u satisfies $\langle \varphi(z), \mathbf{u} \rangle > 0$. Let

$$w = u^\ell v_1^{\ell \cdot s} v_2^{\ell \cdot s} \cdots v_{2n-1}^{\ell \cdot s}.$$

Let us first show that $w \in D_{\mathbf{u},k}$. Since $\langle \varphi(v_i), \mathbf{u} \rangle = \langle \mathbf{b}_i, \mathbf{u} \rangle = 0$, we have $\langle \varphi(w), \mathbf{u} \rangle = \ell \cdot \langle \varphi(u), \mathbf{u} \rangle = \ell \cdot \|u\|^2 > 0$. Let z be an infix of w. Since $\langle \varphi(y), \mathbf{u} \rangle > 0$ for every infix y of u and also $\langle \varphi(v_i), \mathbf{u} \rangle = 0$ for $i \in [1, 2n-1]$, we have $\langle \varphi(z), \mathbf{u} \rangle \geq -k_0 \geq -k$. Thus, we have $w \in D_{\mathbf{u},k}$.

Finally, we prove that $w \notin M_\ell$ and $w \notin D_{\mathbf{u}_i,\ell_i}$ for $i \in [1, m]$. First, note that $\varphi(w) \equiv 0 \bmod \ell$, so that $w \notin M_\ell$. Let us now show that for $i \in [1, m]$, we have $w \notin D_{\mathbf{u}_i,\ell_i}$. Since

$\mathbf{b}_1, \ldots, \mathbf{b}_n$ is a basis of $\mathbb{Q}^n$, we can write $\mathbf{u}_i = \alpha_1\mathbf{b}_1 + \cdots + \alpha_n\mathbf{b}_n$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{Q}$. Since the basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is an orthogonal basis, we have

$$\langle \mathbf{u}_i, \mathbf{b}_j\rangle = \langle \alpha_1\mathbf{b}_1+\cdots+\alpha_n\mathbf{b}_n, \mathbf{b}_j\rangle = \alpha_1\langle \mathbf{b}_1, \mathbf{b}_j\rangle+\cdots+\alpha_n\langle \mathbf{b}_n, \mathbf{b}_j\rangle = \alpha_j\langle \mathbf{b}_j, \mathbf{b}_j\rangle = \alpha_j\cdot\|\mathbf{b}_j\|^2$$

and thus $\langle \mathbf{u}_i, \mathbf{b}_j\rangle > 0$ if and only if $\alpha_j > 0$.

Observe that now either $\alpha_1 < 0$ or $\alpha_j \neq 0$ for some $j \in [2, n]$: Otherwise, we would have $\mathbf{u}_i = \alpha_1\mathbf{b}_1 = \alpha_1\mathbf{u}$ and thus $\mathbb{Q}_+\mathbf{u}_i = \mathbb{Q}_+\mathbf{u}$. Therefore, there is a vector $\mathbf{p} \in P$ with $\langle \mathbf{u}_i, \mathbf{p}\rangle < 0$. Let $\mathbf{p} = \varphi(v_p)$ with $p \in [1, 2n-1]$. Hence, the infix $v_p^{\ell\cdot s}$ of w satisfies $\langle \varphi(v_p^{\ell\cdot s}), \mathbf{u}_i\rangle < -\ell s < -\ell_i$ and hence $w \notin D_{\mathbf{u}_i,\ell_i}$. ◀

C.8 Proof of Theorem 5.11

A *Petri net* $N = (P, T, \textsc{Pre}, \textsc{Post})$ consists of a finite set P of *places*, a finite set T of *transitions* and two mappings $\textsc{Pre}, \textsc{Post}\colon T \to \mathbb{N}^P$. Configurations of Petri net are elements of $\mathbb{N}^P$, called *markings*. If for every place $p \in P$ we have $\textsc{Pre}(t)[p] \leq \mathbf{M}[p]$ for a transition $t \in T$ then t is *fireable* in $\mathbf{M}$ and the result of firing t in marking $\mathbf{M}$ is $\mathbf{M}' = \mathbf{M}+(\textsc{Post}(t)-\textsc{Pre}(t))$, we write $\mathbf{M} \xrightarrow{t} \mathbf{M}'$. We extend notions of fireability and firing naturally to sequences of transitions, we also write $\mathbf{M} \xrightarrow{w} \mathbf{M}'$ for $w \in T^*$. For a Petri net $N = (P, T, \textsc{Pre}, \textsc{Post})$ and markings $\mathbf{M}_0, \mathbf{M}_1$, we define the language $L(N, \mathbf{M}_0, \mathbf{M}_1) = \{w \in T^* \mid \mathbf{M}_0 \xrightarrow{w} \mathbf{M}_1\}$. Hence, $L(N, \mathbf{M}_0, \mathbf{M}_1)$ is the set of transition sequences leading from $\mathbf{M}_0$ to $\mathbf{M}_1$. A *labeled Petri net* is a Petri net $N = (P, T, \textsc{Pre}, \textsc{Post})$ together with an *initial marking* $\mathbf{M}_I$, a *final marking* $\mathbf{M}_F$, and a *labeling*, i.e. a homomorphism $h : T^* \to \Sigma^*$. The language *recognized by* the labeled Petri net is then defined as $L_h(N, \mathbf{M}_I, \mathbf{M}_F) = h(L(N, \mathbf{M}_I, \mathbf{M}_F))$.

It is folklore (and easy to see) that a language is a VAS language if and only if it is recognized by a labeled Petri net (and the translation is effective). Thus, it suffices to show Theorem 5.11 for languages of the form $L = h(L(N, \mathbf{M}_I, \mathbf{M}_F))$. Moreover, it is already enough to prove Theorem 5.11 for languages of the form $L(N, \mathbf{M}_I, \mathbf{M}_F)$: If $\mathcal{A}$ is a modular envelope for L, then applying h to the edges of $\mathcal{A}$ yields a modular envelope for $h(L)$. Thus from now on, we assume $L = L(N, \mathbf{M}_I, \mathbf{M}_F)$ for a fixed Petri net $N = (P, T, \textsc{Pre}, \textsc{Post})$.

Basic notions Let us introduce some notions used in Lambert's proof. We extend the set of configurations $\mathbb{N}^d$ into $\mathbb{N}_\omega^d$, where $\mathbb{N}_\omega = \mathbb{N} \cup \{\omega\}$. We extend the notion of transition firing into $\mathbb{N}_\omega^d$, by defining $\omega - k = \omega = \omega + k$ for every $k \in \mathbb{N}$. For $\mathbf{u}, \mathbf{v} \in \mathbb{N}_\omega^d$ we write $\mathbf{u} \leq_\omega \mathbf{v}$ if $\mathbf{u}[i] = \mathbf{v}[i]$ or $\mathbf{v}[i] = \omega$. Intuitively reaching a configuration with ω at some places means that it is possible to reach configurations with values ω substituted by arbitrarily high values.

A key notion in [19] is that of MGTS, which formulate restrictions on paths in Petri nets. A *marked graph-transition sequence (MGTS)* for our Petri net $N = (P, T, \textsc{Pre}, \textsc{Post})$ is a finite sequence $C_0, t_1, C_1 \ldots C_{n-1}, t_n, C_n$, where t_i are transitions from T and C_i are precovering graphs, which are defined next. A *precovering graph* is a quadruple $C = (G, \mathbf{m}, \mathbf{m}^{\text{init}}, \mathbf{m}^{\text{fin}})$, where $G = (V, E, h)$ is a finite, strongly connected, directed graph with $V \subseteq \mathbb{N}_\omega^P$ and labeling $h\colon E \to T$, and three vectors: a *distinguished* vector $\mathbf{m} \in V$, an *initial* vector $\mathbf{m}^{\text{init}} \in \mathbb{N}_\omega^P$, and a *final* vector $\mathbf{m}^{\text{fin}} \in \mathbb{N}_\omega^P$. A precovering graph has to meet two conditions: First, for every edge $e = (\mathbf{m}_1, \mathbf{m}_2) \in E$, there is an $\mathbf{m}_3 \in \mathbb{N}_\omega^P$ with $\mathbf{m}_1 \xrightarrow{h(e)} \mathbf{m}_3 \leq_\omega \mathbf{m}_2$. Second, we have $\mathbf{m}^{\text{init}}, \mathbf{m}^{\text{fin}} \leq_\omega \mathbf{m}$. Additionally we impose the restriction on MGTS that the initial vector of C_0 equals $\mathbf{M}_I$ and the final vector of C_n equals $\mathbf{M}_F$.

Languages of MGTS Each precovering graph can be treated as a finite automaton. For $\mathbf{m}_1, \mathbf{m}_2 \in V$, we denote by $L(C, \mathbf{m}_1, \mathbf{m}_2)$ the set of all $w \in T^*$ read on a path from $\mathbf{m}_1$ to

$\mathbf{m}_2$. Moreover, let $L(C) = L(C, \mathbf{m}, \mathbf{m})$. MGTS have associated languages as well. Let $\mathcal{N} = C_0, t_1, C_1 \ldots C_{n-1}, t_n, C_n$ be an MGTS of a Petri net N, where $C_i = (G_i, \mathbf{m}_i, \mathbf{m}_i^{\mathrm{init}}, \mathbf{m}_i^{\mathrm{fin}})$. Its language $L(\mathcal{N})$ is the set of all words of the form $w = w_0 t_1 w_1 \cdots w_{n-1} t_n w_n \in T^*$ where: $w_i \in L(C_i)$ for each $i \in [0, n]$ and (ii) there exist markings $\mathbf{u}_0, \mathbf{u}'_0, \mathbf{u}_1, \mathbf{u}'_1, \ldots, \mathbf{u}_n, \mathbf{u}'_n \in \mathbb{N}^P$ such that $\mathbf{u}_i \leq_\omega \mathbf{m}_i^{\mathrm{init}}$ and $\mathbf{u}'_i \leq_\omega \mathbf{m}_i^{\mathrm{fin}}$ and $\mathbf{u}_0 \xrightarrow{w_0} \mathbf{u}'_0 \xrightarrow{t_1} \mathbf{u}_1 \xrightarrow{w_1} \ldots \xrightarrow{w_{n-1}} \mathbf{u}'_{n-1} \xrightarrow{t_n} \mathbf{u}_n \xrightarrow{w_n} \mathbf{u}'_n$. Notice that by (ii) and the restriction that $\mathbf{m}_0^{\mathrm{init}} = \mathbf{M}_I$ and $\mathbf{m}_n^{\mathrm{fin}} = \mathbf{M}_F$, we have $L(\mathcal{N}) \subseteq L(N, \mathbf{M}_I, \mathbf{M}_F)$ for any MGTS $\mathcal{N}$.

Hence roughly speaking, $L(\mathcal{N})$ is the set of runs that contain the transitions $t_1, \ldots, t_n$ and additionally markings before and after firing these transitions are prescribed on some places: this is exactly what the restrictions $\mathbf{u}_i \leq_\omega \mathbf{m}_i^{\mathrm{init}}$, $\mathbf{u}'_i \leq_\omega \mathbf{m}_i^{\mathrm{fin}}$ impose.

As an immediate consequence of the definition, we observe that for every MGTS $\mathcal{N} = C_0, t_1, C_1 \ldots C_{n-1}, t_n, C_n$ we have

$$L(\mathcal{N}) \subseteq L(C_0) \cdot \{t_1\} \cdot L(C_1) \cdots L(C_{n-1}) \cdot \{t_n\} \cdot L(C_n). \tag{3}$$

Perfect MGTS Lambert calls MGTS with a particular property *perfect* [19]. Since the precise definition is involved and we do not need all the details, it is enough for us to mention a selection of properties of perfect MGTS. Intuitively, in perfect MGTSes, the value ω on place p in $\mathbf{M}_i$ means that inside of the component C_i, the token count in place p can be made arbitrarily high. In [19] it is shown (Theorem 4.2 (page 94) together with the preceding definition) that

▶ **Proposition C.4** ([19]). *For a Petri net N one can compute finitely many perfect MGTS $\mathcal{N}_1, \ldots, \mathcal{N}_m$ such that $L(N, \mathbf{M}_I, \mathbf{M}_F) = \bigcup_{i=1}^{m} L(\mathcal{N}_i)$.*

Building the automata According to Proposition C.4, it suffices to construct a modular envelope for each $L(\mathcal{N}_i)$. Hence, we consider a single perfect MGTS $\mathcal{N} = C_0, t_1, C_1, \ldots, t_n, C_n$ and shall construct a modular envelope $\mathcal{A}$ for $L(\mathcal{N})$. The automaton $\mathcal{A}$ is obtained by glueing together all the precovering graphs C_i along the transitions t_i. In other words, $\mathcal{A}$ is the disjoint union of all the graphs C_i and has an edge labeled t_i from $\mathbf{m}_{i-1}$ to $\mathbf{m}_i$ for each $i \in [1, n]$. The initial state of $\mathcal{A}$ is $\mathbf{m}_0$ and its final state is $\mathbf{m}_n$.

Run amalgamation In order to show that $\mathcal{A}$ is indeed a modular envelope, we will use an embedding between runs introduced by Jančar [16] and Leroux [24] and run amalgamation [24]. A triple $(\mathbf{u}, t, \mathbf{v}) \in \mathbb{N}^P \times T \times \mathbb{N}^P$ is a *transition triple* if $\mathbf{v} = \mathbf{u} + \mathrm{eff}(t)$. If there is no danger of confusion, we somtimes call $(\mathbf{u}, t, \mathbf{v})$ a transition. A triple $(\mathbf{u}, w, \mathbf{v})$ with $\mathbf{u}, \mathbf{v} \in \mathbb{N}^P$ and $w \in (\mathbb{N}^P \times T \times \mathbb{N}^P)^*$ is called a *prerun*. Let $\rho = (\mathbf{u}, w, \mathbf{v})$ and $\rho' = (\mathbf{u}', w', \mathbf{v}')$ be preruns with $w = (\mathbf{u}_0, t_1, \mathbf{v}_1)(\mathbf{u}_1, t_2, \mathbf{v}_2) \cdots (\mathbf{u}_{r-1}, t_r, \mathbf{v}_r)$ and $w' = (\mathbf{u}'_0, t'_1, \mathbf{v}'_1)(\mathbf{u}'_1, t'_2, \mathbf{v}'_2) \cdots (\mathbf{u}'_{s-1}, t'_s, \mathbf{v}'_s)$. An *embedding of ρ in ρ'* is a monotone map $\sigma\colon [1, r] \to [1, s]$ if $t'_{\sigma(i)} = t_i$, $\mathbf{u}_i \leq \mathbf{u}'_{\sigma(i)}$ and $\mathbf{v}_i \leq \mathbf{v}'_{\sigma(i)}$ for $i \in [1, r]$, and $\mathbf{u} \leq \mathbf{u}'$ and $\mathbf{v} \leq \mathbf{v}'$. In this case, the words $t'_1 \cdots t'_{\sigma(1)-1}$, $t'_{\sigma(i)+1} \cdots t'_{\sigma(i+1)-1}$ for $i \in [1, r-1]$, and $t'_{\sigma(r)+1} \cdots t'_s$ are called the *inserted words of σ*. By $F(\sigma) \subseteq T^*$, we denote the set of all factors of inserted words of σ. Furthermore, by $\Psi(\rho)$, we denote the Parikh image $\Psi(t_1 \cdots t_r) \in \mathbb{N}^T$.

Moreover, ρ is called a *run* if each $(\mathbf{u}_i, t_i, \mathbf{v}_i)$ is a transition and also $\mathbf{u} = \mathbf{u}_0$, $\mathbf{u}_i = \mathbf{v}_i$ for $i \in [1, r]$, and $\mathbf{v} = \mathbf{v}_r$. Note that this is equivalent to $\mathbf{u}_i = \mathbf{v}_i$ for $i \in [1, r]$ and $\mathbf{u} = \mathbf{u}_0 \xrightarrow{t_1} \mathbf{u}_1 \cdots \mathbf{u}_{r-1} \xrightarrow{t_r} \mathbf{u}_r$ and we sometimes use the latter notation to denote runs.

If the runs ρ and ρ' are runs in our MGTS, then we can associate to each marking $\mathbf{u}_i$ ($\mathbf{u}'_i$) in ρ (in ρ') a node v_i (v'_i) in some C_j. If σ is an embedding of ρ in ρ' and we have

$v'_{\sigma(i)} = v_i$ for $i \in [1, r]$, then σ is *node preserving*. In other words, σ maps each marking in ρ to a marking in ρ' that visits the same node in $\mathcal{N}$.

Suppose we have three runs ρ_0, ρ_1, ρ_2 and there are embeddings σ_1 of ρ_0 in ρ_1 and σ_2 of ρ_0 in ρ_2. As observed in [25, Prop. 5.1] one can define a new run ρ_3 in which both ρ_1 and ρ_2 embed. Let ρ_0 be the run $\mathbf{u}_0 \xrightarrow{t_1} \mathbf{u}_1 \xrightarrow{t_2} \cdots \xrightarrow{t_r} \mathbf{u}_r$. Then ρ_1 and ρ_2 can be written as

$$\mathbf{u}_0 + \mathbf{v}_0 \xrightarrow{w_0} \mathbf{u}_0 + \mathbf{v}_1 \xrightarrow{t_1} \mathbf{u}_1 + \mathbf{v}_1 \xrightarrow{w_1} \cdots \xrightarrow{t_r} \mathbf{u}_r + \mathbf{v}_r \xrightarrow{w_r} \mathbf{u}_r + \mathbf{v}_{r+1}$$

$$\mathbf{u}_0 + \mathbf{v}'_0 \xrightarrow{w'_0} \mathbf{u}_0 + \mathbf{v}'_1 \xrightarrow{t_1} \mathbf{u}_1 + \mathbf{v}'_1 \xrightarrow{w'_1} \cdots \xrightarrow{t_r} \mathbf{u}_r + \mathbf{v}'_r \xrightarrow{w'_r} \mathbf{u}_r + \mathbf{v}'_{r+1}.$$

for some $\mathbf{v}_i, \mathbf{v}'_i \in \mathbb{N}^P$, $i \in [0, r+1]$. Then the *amalgamated run* ρ_3 is defined as

$$\mathbf{u}_0 + \mathbf{v}_0 + \mathbf{v}'_0 \xrightarrow{w_0} \mathbf{u}_0 + \mathbf{v}_1 + \mathbf{v}'_0 \xrightarrow{w'_0} \mathbf{u}_0 + \mathbf{v}_1 + \mathbf{v}'_1 \xrightarrow{t_1} \mathbf{u}_1 + \mathbf{v}_1 + \mathbf{v}'_1 \xrightarrow{w_1} \cdots$$

$$\cdots \xrightarrow{w'_{r-1}} \mathbf{u}_{r-1} + \mathbf{v}_r + \mathbf{v}'_r \xrightarrow{t_r} \mathbf{u}_r + \mathbf{v}_r + \mathbf{v}'_r \xrightarrow{w_r} \mathbf{u}_r + \mathbf{v}_{r+1} + \mathbf{v}'_r \xrightarrow{w'_r} \mathbf{u}_r + \mathbf{v}_{r+1} + \mathbf{v}'_{r+1}.$$

and the embedding τ of ρ in ρ_3 is defined in the obvious way. Note that the embedding τ of ρ_0 in ρ_3 satisfies $F(\sigma_1) \cup F(\sigma_2) \subseteq F(\tau)$ and $\Psi(\rho_3) - \Psi(\rho_0) = (\Psi(\rho_1) - \Psi(\rho_0)) + (\Psi(\rho_2) - \Psi(\rho_0))$.

We are now prepared to prove that $\mathcal{A}$ is a modular envelope for $L = L(\mathcal{N})$. Suppose $w = w_0 t_1 w_1 \cdots t_r w_n \in L$ with a run $\rho\colon \mathbf{u}_0 \xrightarrow{w_0} \mathbf{u}'_0 \xrightarrow{t_1} \mathbf{u}_1 \xrightarrow{w_1} \cdots \xrightarrow{w_{n-1}} \mathbf{u}'_{n-1} \xrightarrow{t_n} \mathbf{u}_n \xrightarrow{w_n} \mathbf{u}'_n$. Moreover, let $u_1, \ldots, u_m \in \mathrm{Loop}(\mathcal{A})$. Since inside each strongly connected component of $\mathcal{A}$, one can combine several cycles into one, we may assume that $m = n$ and there is one word u_i for each component C_i of $\mathcal{A}$. We can deduce the following using Lambert's pumping lemma (Lemma 4.1 in [19]).

▶ **Lemma C.5.** *There is a run $\rho_1\colon \mathbf{v}_0 \xrightarrow{w'_0} \mathbf{v}'_0 \xrightarrow{t_1} \mathbf{u}_1 \xrightarrow{w'_1} \cdots \xrightarrow{w'_{n-1}} \mathbf{v}'_{n-1} \xrightarrow{t_n} \mathbf{u}_{n-1} \xrightarrow{w'_n} \mathbf{v}'_n$ in $\mathcal{N}$ so that there is a node preserving embedding σ_1 of ρ in ρ_1 with $u_i \in F(\sigma_1)$ for $i \in [1, n]$.*

Proof. We employ Lambert's iteration lemma, which involves covering sequences. Let C be a precovering graph for a Petri net $N = (P, T, \textsc{Pre}, \textsc{Post})$ with a distinguished vector $\mathbf{m} \in \mathbb{N}_\omega^P$ and initial vector $\mathbf{m}^{\text{init}} \in \mathbb{N}_\omega^P$. For a marking $\mathbf{M}_0$ let $L(N, \mathbf{M}_0) = \bigcup_{M \in \mathbb{N}^P} L(N, \mathbf{M}_0, \mathbf{M})$, i.e. the set of all the transition sequences fireable in $\mathbf{M}_0$. A sequence $u \in L(C) \cap L(N, \mathbf{m}^{\text{init}})$ is called a *covering sequence for* C if for every place $p \in P$ we have either 1) $\mathbf{m}^{\text{init}}[p] = \omega$, or 2) $\mathbf{m}[p] = \mathbf{m}^{\text{init}}[p]$ and $\mathrm{eff}(u)[p] = 0$, or 3) $\mathbf{m}[p] = \omega$ and $\mathrm{eff}(u)[p] > 0$.

Recall that we consider the MGTS $\mathcal{N} = C_0, t_1, C_1 \ldots C_{n-1}, t_n, C_n$. Let $C_i = (V_i, E_i, h_i)$ be a precovering graph, and let the distinguished vertex be $\mathbf{m}_i$ and initial vertex be $\mathbf{m}_i^{\text{init}}$. If $\mathcal{N}$ is a perfect MGTS, then according to the definition from [19] (page 92), for every $i \in [0, n]$ there exists a covering sequence $s_i \in L(C_i) \cap L(N, \mathbf{m}_i^{\text{init}})$.

We want to construct a run where ρ embeds via node preserving embedding and so that for each $i \in [1, n]$, the word u_i is factor of an insertion word. Moreover, let $w'_i \in L(C_i)$ so that w_i is a factor of w'_i. This is possible because C_i is strongly connected.

Observe that for sufficiently high $\ell \in \mathbb{N}$, the word $x_i = s_i^\ell u_i w'_i$ is also a covering sequence. Now Lambert's pumping lemma (Lemma 4.1 in [19] (page 92)) shows that for large enough k, there are runs with transition sequences

$$v_k = x_0^k \beta_0 y_0^k z_0^k \cdot t_1 \cdot x_1^k \beta_1 y_1^k z_1^k \cdot t_2 \cdots t_n \cdot x_n^k \beta_n y_n^k z_n^k$$

in $\mathcal{N}$. Now by construction of x_i, the proof of the pumping lemma yields that for large enough k, the run ρ embeds into the run for v_k so that nodes are preserved. ◀

Now ρ_1 is almost what we are looking for: Its transition word contains each u_i as a factor, but it is not clear whether $\Psi(\rho_1) \equiv \Psi(\rho) \bmod k$. However, we can use amalgamation to achieve this: Since ρ embeds in ρ_1 via σ_1, we can consider the run ρ_2 obtained by amalgamating ρ_1 with itself. We use the following simple observation:

▶ **Lemma C.6.** *Let τ, τ_1, τ_2 be runs in $\mathcal{N}$. If τ embeds in τ_1 and τ_2 via node preserving embeddings, then the amalgamation of τ_1 and τ_2 again belongs to $\mathcal{N}$.*

Proof. Since the result τ_3 of amalgamating τ_1 and τ_2 is a run of the Petri net, we only have to show that $\mathcal{N}$ contains the necessary vertices to accommodate τ_3. To this end, observe that in each component C_i, all vertices have ω in exactly the same components.

Since the embedding of τ in τ_i is node preserving, we know that whenever the marking in τ_i is larger in some coordinate, then that coordinate has to be an ω-coordinate in that node. Therefore, if we add these two difference vectors, as in the amalgamation, the resulting marking with still agree with the node content on the non-ω-coordinates. Therefore, the amalgamated run is a run of $\mathcal{N}$. ◀

Now for $j \geq 2$, let ρ_j be the run obtained by amalgamating ρ_{j-1} and ρ_1, where ρ is embedded in ρ_1 via σ_1 and in ρ_{j-1} via σ_{j-1}. Moreover, let σ_j be the resulting embedding. Then Lemma C.6 tells us that ρ_j belongs to $\mathcal{N}$ for every $j \geq 2$. Moreover, we have $u_i \in F(\sigma_{j-1}) \subseteq F(\sigma_j)$ for every $i \in [1, n]$. Finally, recall that $\Psi(\rho_j) - \Psi(\rho) = \Psi(\rho_{j-1}) - \Psi(\rho) + \Psi(\rho_1) - \Psi(\rho)$ and hence by induction $\Psi(\rho_j) = \Psi(\rho) + j \cdot (\Psi(\rho_1) - \Psi(\rho))$. In particular $\Psi(\rho_k) \equiv \Psi(\rho) \bmod k$. Therefore, the transition sequence of ρ_k is a word w' that has each u_i as a factor and satisfies $\Psi(w') \equiv \Psi(w) \bmod k$. This proves that $\mathcal{A}$ is a modular envelope.