



The Complexity of Reachability Problems for Flat Counter Machines with Periodic Loops

Marius Bozga, Radu Iosif, Filip Konečný

► To cite this version:

Marius Bozga, Radu Iosif, Filip Konečný. The Complexity of Reachability Problems for Flat Counter Machines with Periodic Loops. 15th International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI) , Jan 2014, San Diego, CA, United States. hal-01418891

HAL Id: hal-01418891

<https://hal.science/hal-01418891>

Submitted on 30 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain

The Complexity of Reachability Problems for Flat Counter Machines with Periodic Loops

Marius Bozga¹ Radu Iosif¹
 Filip Konečný²

¹ Verimag/CNRS/Université de Grenoble Alpes (France)

² NetSuite Brno (Czech Republic)

February 16, 2016

Abstract

This paper proves the NP-completeness of the reachability problem for the class of flat counter machines with difference bounds and, more generally, octagonal relations, labeling the transitions on the loops. The proof is based on the fact that the sequence of powers $\{R^n\}_{n=0}^{\infty}$ of such relations can be encoded as a periodic sequence of matrices, and that both the prefix and the period of this sequence are simply exponential in the size of the binary representation of a relation R . This result allows to characterize the complexity of the reachability problem for one of the most studied class of counter machines [6, 10], and has a potential impact on other problems in program verification.

1 Introduction

Counter machines are powerful abstractions of programs, commonly used in software verification. Due to their expressive power, counter machines can simulate Turing machines [29], thus their decision problems (reachability, termination) are undecidable. This early negative result motivated researchers to define classes of systems with decidable reachability problems, such as: vector addition systems [27, 22, 23], reversal-bounded counter machines [20], Datalog programs with gap-order constraints [32], and flat counter machines [4, 10, 6]. Despite the fact that the reachability problem is, in principle, decidable for these classes, few of these results are actually supported by tools, and used for real-life verification purposes. The main reason is that the complexity of the reachability problems for these systems is, in general, prohibitive. As a practical consequence, most software verifiers rely on incomplete algorithms, which, due to the loss of precision, may raise large numbers of false alarms. Improving the precision of these tools requires mixed techniques such as combinations of model checking, static analysis and acceleration, that rely on identifying subproblems for which the set

of reachable states, or the transitive closure of the transition relation, can be computed precisely [18, 16], by cost-effective algorithms.

In this paper, we study the complexity of the reachability problems for a class of *flat counter machines*, whose control structure forbids nested loops and the transitions occurring inside loops are labeled with *octagons*, i.e. conjunctions of inequalities of the form $\pm x \pm y \leq c$ where x, y denote the current or next values of the counters and c is an integer constant. Our main result states that the reachability problem for this class of counter machines is NP-complete. This result is a direct generalization of the NP-completeness of the reachability problem for the subclass of *difference bounds constraints*, which are finite conjunctions of inequalities of the form $x - y \leq c$, with c being an integer constant.

Due to the particular syntax of the octagonal constraints, in which the variables are always multiplied by coefficients from the set $\{-1, 0, 1\}$, such relations can be represented by square matrices of a fixed dimension, called *difference bounds matrices* (DBM). The main idea of the NP-completeness proof is that sequence of DBMs corresponding to the sequence of relations $\{R^n\}_{n=0}^\infty$ is *periodic*, in the sense that the matrices situated at equal distance in the sequence, beyond a certain prefix, differ by equal quantities. If the prefix and the period of this sequence are known, one can build a quantifier-free formula of Presburger arithmetic that characterizes this sequence, and reduce the reachability problem to an instance of the satisfiability problem in the quantifier-free fragment of Presburger arithmetic, known to be NP-complete [35].

The main technical difficulty is, given an octagonal constraint that defines a relation R , building such a Presburger formula in polynomial time. To this end, we show that the prefix b and the period c of the sequence of DBMs representing $\{R^n\}_{n=0}^\infty$ are *simply exponential* in the size of the octagon defining R . Using this argument, one can (i) guess the prefix b and the period c of the relation, (ii) compute the powers R^b and R^c , using exponentiation by squaring, (iii) verify the validity of the guess for b and c , and (iv) build the needed Presburger constraint in polynomial time.

Proving the simply exponential bounds for the prefix and the period of an octagonal relation uses insights from the theory of weighted graphs and tropical algebra [15, 33]. We use the classical representation of DBMs by weighted constraint graphs, such that the n -th power of a difference bounds relation R is defined by the minimal weight paths of a constraint graph of width n , called an *unfolding graph* [10]. Then we define a weighted automaton, called *zigzag automaton*, that recognizes the set of constraint paths in the unfolding graph. The minimal weight paths, needed to define the n -times composition of R with itself, are given by the n -th power of the incidence matrix of the transition table of the zigzag automaton, where matrix multiplication is defined in the tropical semiring $\langle \mathbb{Z}_{\pm\infty}, \min, +, \infty, 0 \rangle$. Since the sequence of tropical powers of any given matrix is periodic, we obtain that the sequence of DBMs representing the relations $\{R^n\}_{n=0}^\infty$ is periodic as well.

We first prove the existence of simply exponential bounds on the prefix and the period of the sequence of DBMs that represent the sequence of relations $\{R^n\}_{n=0}^\infty$, where R is a relation defined by a difference bounds constraint. These bounds are then generalized to octagonal relations. The most technical part is proving the bound on the period of such sequences, which requires an insight on the particular structure of loops in the zigzag automaton describing the powers (w.r.t. composition) of a difference

bounds relations. The crucial point is restricting the zigzag automaton to recognize a subset of constraint paths, with a bounded number of direction changes, whose set of weights is sufficient to define the relation R^n . This idea originates in the work of Comon and Jurski [10] on defining transitive closures of difference bounds relations by formulae of Presburger arithmetic, in order to prove decidability of the reachability problem for flat counter machines with difference bounds relations.

1.1 Related Work

The study of the computational complexity of the reachability problem (and other related problems, such as coverability and boundedness) for various classes of counter machines has recently received much attention.

An important class of counter machines with decidable reachability problems are Vector Addition Systems with States (VASS). This problem has been shown to be EXSPACE-hard by Lipton [26] and currently no upper bound has been found. On the other hand, the problems of coverability and boundedness for VASS are shown to be EXSPACE-complete [31]. Because the transition relation of VASS can be defined as a finite disjunction of difference bounds constraints, these counter machines are, in principle, not flat. However, when restricting the number of counters to two, Hopcroft and Pansiot [19] have shown that the set of reachable configurations of a VASS is semi-linear, thus definable in Presburger arithmetic. Along this line, Leroux and Sutre [24] showed that it is possible to build a flat counter machine, with the same transitions as the original 2-counter VASS and same reachable set of configurations. A close analysis of their construction revealed that reachability of 2-counter VASS (mostly known as 2-dimensional VASS) is a PSPACE-complete problem [3].

In their work, Ibarra and Gurari [17] study the reachability problem for counter machines with increment, decrement and zero test, in the *reversal-bounded* case, where the counters are allowed to switch between non-decreasing and non-increasing modes a number of times, bounded by a constant. It is found that, when the number of counters and reversals are fixed constants (i.e. not part of the representation of the counter machines) the emptiness problem is decidable in logarithmic space, and hence, in polynomial time. Moreover, if the machines under consideration are all deterministic, the emptiness problem is NLOGSPACE-complete. On the other hand, if the number of counters and reversals are part of the input, the emptiness problem is in PSPACE. Our model of computation is incomparable, since flat counter machines with non-deterministic counter updates are not reversal-bounded, in general. For instance, if the future value of a counter x is chosen to be $x - 1 \leq x' \leq x + 1$ within a loop, the counter can switch any number of times between increasing and decreasing modes.

The class of *gap-order* constraints, initially introduced by Revesz [32], consists of finite conjunctions of difference bounds constraints $x - y \leq c$, where c is a positive constant. Counter machines with gap-order constraints have been studied by Bozzelli and Pinchinat [9] who coined their reachability problem to PSPACE-complete. Our result is incomparable to [9], as we show NP-completeness for flat counter machines with strictly more general¹ octagonal relations on loops.

¹The generalization of gap-order to difference bound constraints suffices to show undecidability of non-

The results which are probably closest to ours are the ones in [12, 11], where flat counter machines with linear affine guards and vector addition updates are considered. In [12] it is shown that model-checking for Linear Temporal Logic is NP-complete for these systems, matching thus our complexity for reachability with difference bounds constraints, while model-checking first-order logic and linear μ -calculus is PSPACE-complete [11], matching the complexity of CTL* model checking for gap-order constraints [9]. These results are again incomparable with ours, since (i) the linear affine guards are more general, while (ii) the vector addition updates are more restrictive (e.g. the direct transfer of values $x'_i = x_j$ for $i \neq j$ is not allowed).

The result of this paper is a refinement of earlier decidability proofs for the reachability problem concerning flat counter machines with loops labeled by difference bounds [10, 7] constraints. The first such result, due to Comon and Jurski [10] defines the sequence of n -times compositions $\{R^n\}_{n=0}^\infty$ of a difference bounds relation R by a formula in Presburger arithmetic. The essence of their proof is the definition, by a formula of Presburger arithmetic, of a subset of paths, in the constraint graph representing R^n , that encompasses the set of paths of minimal weight, relevant for the definition of the relation R^n . They show that only certain paths, that roughly go back and forth from one extremity of the graph to the other, without changing direction in between, are important in the definition of the closed form. The idea of considering only such *simple* paths is instrumental in our work, for establishing a simply exponential upper bound on the period of these relations.

By exploring further the structure of these simple paths, Konečný [21] showed that the sequence of the closed form of the power sequence of a difference bounds (respectively, octagonal) relation can be defined by a quantifier-free Presburger formula which, moreover, can be built in polynomial time by a deterministic algorithm. As a result, the reachability problem for flat counter machines can be proved to be in NPTIME directly, by polynomial reduction to the satisfiability of quantifier-free Presburger arithmetic. Unlike the proof given in this paper, Konečný's proof [21] does not use periodic sequences, relying on an enumeration of polynomially many minimal weight paths. Besides providing an alternative proof of NP-completeness to the reachability problem, the results in this paper define closed forms using only finite disjunctions of difference bounds constraints (respectively, octagons) whose coefficients are parameterized by n . This characterization of the closed forms is of particular interest for other problems, such as, e.g. the complexity of the termination problem for periodic classes of relations [8], or extensions of the model of flat counter machines with recursive calls [14].

2 Preliminaries

We denote by \mathbb{Z} , \mathbb{N} and \mathbb{N}_+ the sets of integers, positive (including zero) and strictly positive integers, respectively. We define $\mathbb{Z}_\infty = \mathbb{Z} \cup \{\infty\}$ and $\mathbb{Z}_{\pm\infty} = \mathbb{Z}_\infty \cup \{-\infty\}$. We write $[n]$ for the interval $\{0, \dots, n-1\}$, $\text{abs}(n)$ for the absolute value of the integer $n \in \mathbb{Z}$, and $\text{gcd}(n_1, \dots, n_k)$, $\text{lcm}(n_1, \dots, n_k)$ for the greatest common divisor and least common multiple of the natural numbers $n_1, \dots, n_k \in \mathbb{N}$, respectively. The cardinality of a finite set S is denoted by $\|S\|$.

flat counter machines, hence the restriction to flat control structures is crucial.

A *weighted graph* is a tuple $G = \langle V, E, w \rangle$, where V is a set of vertices, $E \subseteq V \times V$ is a set of edges, and $w : E \rightarrow \mathbb{Z}$ is a weight function. If G is clear from the context, we write $u \xrightarrow{\alpha} v$ for $(u, v) \in E$ and $w(u, v) = \alpha$. A *path* in G is a sequence of the form $\pi : v_0 \xrightarrow{w_1} v_1 \cdots v_{k-1} \xrightarrow{w_k} v_k$. We denote by $\text{src}(\pi) = v_0$, $\text{dst}(\pi) = v_k$ its source and destination vertices, by $|\pi| = k$ and by $w(\pi) = \sum_{i=1}^k w_i$ its weight. The path π is said to be (i) *elementary* if $v_i = v_j$ only if $i = 0$ and $j = k$, (ii) *a cycle* if $\text{src}(\pi) = \text{dst}(\pi)$, and (iii) *minimal* if, for any path π' such that $\text{src}(\pi') = \text{src}(\pi)$, $\text{dst}(\pi') = \text{dst}(\pi)$, we have $w(\pi) \leq w(\pi')$. We denote by $\mu(G) = \max(\{\text{abs}(\alpha) \mid u \xrightarrow{\alpha} v\} \cup \{1\})$ the maximum between the absolute values of the weights of G and 1.

The set of $n \times n$ square matrices with coefficients in \mathbb{Z}_∞ ($\mathbb{Z}_{\pm\infty}$) is denoted as $\mathbb{Z}_\infty^{n \times n}$ ($\mathbb{Z}_{\pm\infty}^{n \times n}$). Each matrix $M \in \mathbb{Z}_\infty^{n \times n}$ is the *incidence matrix* of a weighted graph $G_M = \langle V_M, E_M, w_M \rangle$, where $V_M = \{1, \dots, n\}$, $E_M = \{(i, j) \mid M_{ij} < \infty\}$ and $w(i, j) = M_{ij}$, for all $i, j \in \{1, \dots, n\}$. In this case, we also define $\mu(M) = \mu(G_M)$.

A *term* t over a set of variables $\mathbf{x} = \{x_1, \dots, x_N\}$ is a linear combination $a_0 + a_1x_1 + \dots + a_Nx_N$, for some integer constants $a_0, a_1, \dots, a_N \in \mathbb{Z}$. An *atomic proposition* is a predicate of the form $t \leq 0$ or $t \equiv_c 0$, where t is a term, $c \in \mathbb{N}_+$ is a constant, and \equiv_c denotes equality modulo c . The boolean constants *false* and *true* are denoted by \perp and \top , respectively. *Quantifier-free Presburger Arithmetic* (QFPA) is the set of boolean combinations of atomic propositions of the above form. For a QFPA formula ϕ , let $\text{Atom}(\phi)$ denote the set of atomic propositions in ϕ , and $\phi[t/x]$ denote the formula obtained by substituting the variable x with the term t in ϕ . We assume that all integers are encoded in binary and denote by $|\phi|$ the size of the binary encoding of a formula ϕ .

Let \mathbf{x} denote a nonempty set of integer variables. A *valuation* of \mathbf{x} is a function $\nu : \mathbf{x} \rightarrow \mathbb{Z}$. The set of valuations is denoted by $\mathbb{Z}^\mathbf{x}$. If $\nu \in \mathbb{Z}^\mathbf{x}$ is a valuation, we denote by $\nu \models \varphi$ the fact that the formula obtained from φ by replacing each occurrence of $x \in \mathbf{x}$ with the integer $\nu(x)$ is valid under the standard interpretation of the first-order arithmetic. A formula φ is said to be *consistent* if and only if there exists a valuation ν , such that $\nu \models \varphi$. The consistency problem (also known as the satisfiability problem) for QFPA is NP-complete [35, Lemma 5].

Let \mathbf{x}' denote the set $\{x' \mid x \in \mathbf{x}\}$ of *primed* variables. A formula $\phi(\mathbf{x}, \mathbf{x}')$ is evaluated with respect to two valuations $\nu, \nu' \in \mathbb{Z}^\mathbf{x}$, by replacing each occurrence of $x \in \mathbf{x}$ with $\nu(x)$ and each occurrence of $x' \in \mathbf{x}'$ with $\nu'(x')$ in ϕ . We write $(\nu, \nu') \models \phi$ when the formula obtained from these replacements is valid. A formula $\phi(\mathbf{x}, \mathbf{x}')$ is said to define a relation $R \subseteq \mathbb{Z}^\mathbf{x} \times \mathbb{Z}^\mathbf{x}$ whenever for all $\nu, \nu' \in \mathbb{Z}^\mathbf{x}$, we have $(\nu, \nu') \in R$ iff $(\nu, \nu') \models \phi$. The empty relation is denoted by \emptyset . The composition of two relations $R_1, R_2 \subseteq \mathbb{Z}^\mathbf{x} \times \mathbb{Z}^\mathbf{x}$ defined by formulae $\varphi_1(\mathbf{x}, \mathbf{x}')$ and $\varphi_2(\mathbf{x}, \mathbf{x}')$, respectively, is the relation $R_1 \circ R_2 \subseteq \mathbb{Z}^\mathbf{x} \times \mathbb{Z}^\mathbf{x}$, defined by the formula $\exists \mathbf{y} . \varphi_1(\mathbf{x}, \mathbf{y}) \wedge \varphi_2(\mathbf{y}, \mathbf{x}')$.

The *identity* on \mathbf{x} is the relation $I_\mathbf{x} \subseteq \mathbb{Z}^\mathbf{x} \times \mathbb{Z}^\mathbf{x}$ defined by the formula $\bigwedge_{x \in \mathbf{x}} x' = x$. For any relation $R \subseteq \mathbb{Z}^\mathbf{x} \times \mathbb{Z}^\mathbf{x}$, we define $R^0 = I_\mathbf{x}$ and $R^{n+1} = R^n \circ R = R \circ R^n$, for all $n \in \mathbb{N}$. R^n is called the *n-th power* of R in the sequel. The infinite sequence of relations $\{R^n\}_{n=0}^\infty$ is called the *power sequence* of R . With these notations, $R^+ = \bigcup_{n=1}^\infty R^n$ denotes the *transitive closure* of R , and $R^* = R^+ \cup I_\mathbf{x}$ denotes the *reflexive and transitive closure* of R . A relation R is said to be **-consistent* if and only if $R^n \neq \emptyset$, for all $n \in \mathbb{N}_+$. If R is not *-consistent, there exists $b > 0$ such that $R^n = \emptyset$, for all $n \geq b$.

Definition 1 A class of relations \mathcal{R} is the union of all monoids $\langle \mathcal{R}_\mathbf{x}, \circ, I_\mathbf{x} \rangle$, where

$\mathcal{R}_{\mathbf{x}} \subseteq 2^{\mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}}$ is a set of relations over \mathbf{x} closed under conjunction and composition, containing the relations $I_{\mathbf{x}}$ and \emptyset .

In this paper we will define classes of relations by a fragments of QFPA. In fact, any fragment of QFPA that contains equality and is closed under conjunction and quantifier elimination defines a class of relations.

3 The Reachability Problem for Flat Counter Machines

In this section we define *counter machines*, which are essentially a generalization of Rabin-Scott finite nondeterministic automata, extended with a set of integer counters, and transitions described by quantifier-free Presburger formulae. Formally, a counter machine (CM) is a tuple $M = \langle \mathbf{x}, \mathcal{L}, \ell_{\text{init}}, \ell_{\text{fin}}, \Delta \rangle$, where:

- \mathbf{x} is a set of *variables* (counters) ranging over \mathbb{Z} ,
- \mathcal{L} is a set of *control locations*,
- $\ell_{\text{init}}, \ell_{\text{fin}} \in \mathcal{L}$ are the *initial* and *final* control locations, respectively,
- Δ is a set of *transition rules* of the form $\ell \xrightarrow{\phi(\mathbf{x}, \mathbf{x}')} \ell'$, where $\ell, \ell' \in \mathcal{L}$ are control locations and $\phi(\mathbf{x}, \mathbf{x}')$ is a QFPA formula defining both (i) the conditions on the current values \mathbf{x} that enable the transition, and (ii) the updates of the current values \mathbf{x} to the next values \mathbf{x}' .

The size of the binary representation of a counter machine is defined as $|M| = \sum_{\ell \xrightarrow{\phi} \ell'} |\phi|$, i.e. the sum of the sizes of all formulae labeling the transition rules of M .

A *configuration* of M is a pair (ℓ, ν) , where $\ell \in \mathcal{L}$ is a control location, and $\nu \in \mathbb{Z}^{\mathbf{x}}$ is a valuation of the variables. A *run* of M is a sequence of configurations $(\ell_0, \nu_0), \dots, (\ell_n, \nu_n)$, where $\ell_0 = \ell_{\text{init}}$, $\ell_n = \ell_{\text{fin}}$ and for each $i = 0, \dots, n-1$, there exists a transition rule $\ell_i \xrightarrow{\phi_i} \ell_{i+1}$ such that $(\nu_i, \nu_{i+1}) \models \phi_i$. The *reachability problem* asks, given a counter machine M , does M have a run?

Let us now define the flatness restriction on counter machines. The *control flow graph* of M is the labeled graph whose vertices are the control locations \mathcal{L} and whose edges are the transition rules in Δ . A cycle in this graph is *elementary* if it does not contain another cycle. A counter machine M is *flat* if and only if every control location belongs to at most one elementary cycle in its control flow graph.

For a set of relations \mathcal{R} , we denote by $\text{REACHFLAT}(\mathcal{R})$ the class of reachability problems for all flat counter machines M where, for each transition rule $\ell \xrightarrow{\phi(\mathbf{x}, \mathbf{x}')} \ell'$ belonging to a cycle in the control flow graph of M , the formula $\phi(\mathbf{x}, \mathbf{x}')$ defines a relation from \mathcal{R} . The main result is that $\text{REACHFLAT}(\text{OCT})$ is NP-complete, where OCT is the set of relations defined below.

Definition 2 A formula $\phi(\mathbf{x})$ is an *octagonal constraint* if it is a finite conjunction of atomic propositions of the form $\pm x_i \leq \alpha_i$ or $\pm x_i \pm x_j \leq \beta_{ij}$, where $\alpha_i, \beta_{ij} \in \mathbb{Z}$, for all $1 \leq i, j \leq N$. We denote by OCT the set of relations $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$ defined by octagonal constraints $\phi(\mathbf{x}, \mathbf{x}')$.

Example 1 Fig. 1 shows a flat counter machine $M = \langle \{i, j, b\}, \{\ell_0, \ell_1, \ell_2, \ell_3\}, \ell_0, \ell_3, \Delta \rangle$. The machine increments both counters i and j by executing the self-loop on state ℓ_1 a

number of times equal to the value of b , that was guessed on the transition $\ell_0 \rightarrow \ell_1$, then it will move to ℓ_2 and will increment i , while decrementing j , until $j = 0$. Finally, it moves to its final state if $i = 2b$. Observe that all transition rules, except for $\ell_2 \xrightarrow{i=2b} \ell_3$, are labeled with octagonal relations. ■

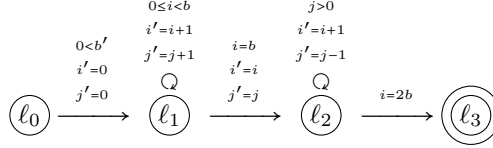


Figure 1: A flat counter machine

In the rest of this section we set the ground for the proof of the NP-completeness result by presenting necessary background notions concerning octagonal constraints, starting with a simpler class of formulae, called difference bounds constraints. In particular, we show that the set of octagonal relations are closed under compositions, thus OCT is a class of relations in the sense of Definition 1.

3.1 Difference Bounds Constraints

Let $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$ be a set of variables, for some $N \in \mathbb{N}_+$. Without losing generality, we consider only formulae in which each atomic proposition involves exactly two variables. Atomic propositions $x_i \leq \alpha_i$, $x_i \geq \alpha_i$, for $\alpha_i \in \mathbb{Z}$, are replaced by $x_i - \zeta \leq \alpha_i$, $\zeta - x_i \leq -\alpha_i$, respectively, for an extra variable ζ , with the implicit assumption $\zeta = 0$.

Definition 3 A difference bounds constraint $\phi(\mathbf{x})$ is a finite conjunction of atomic propositions of the form $x_i - x_j \leq \alpha_{ij}$, $1 \leq i, j \leq N$, where $\alpha_{ij} \in \mathbb{Z}$. A relation $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$ is a difference bounds relation if it is defined by a difference bounds constraint $\phi_R(\mathbf{x}, \mathbf{x}')$.

If $\phi(\mathbf{x})$ is a difference bounds constraint, the *difference bounds matrix* (DBM) representing ϕ is the matrix $M_\phi \in \mathbb{Z}_\infty^{N \times N}$, where $(M_\phi)_{ij} = \alpha_{ij}$ if $x_i - x_j \leq \alpha_{ij} \in \text{Atom}(\phi)$, and $(M_\phi)_{ij} = \infty$, otherwise (see Fig. 2 (a) for an example). In particular, any inconsistent difference bounds constraint is represented by the $N \times N$ matrix $[-\infty]_N$, whose coefficients are all $-\infty$. Dually, a matrix $M \in \mathbb{Z}_\infty^{N \times N}$ corresponds to the difference bounds constraint $\Phi(M) \equiv \bigwedge_{M_{ij} < \infty} x_i - x_j < M_{ij}$. M is *consistent* if $\Phi(M)$ is a consistent formula. The *constraint graph* \mathcal{G}_ϕ of a difference bounds constraint ϕ is the weighted graph whose incidence matrix is M_ϕ (see Fig. 2 (b) for an example).

Definition 4 A consistent DBM $M \in \mathbb{Z}_\infty^{N \times N}$ is said to be closed if and only if:

1. $M_{ii} = 0$, for all $1 \leq i \leq N$, and
2. all triangle inequalities $M_{ik} \leq M_{ij} + M_{jk}$ hold, for all $1 \leq i, j, k \leq N$.

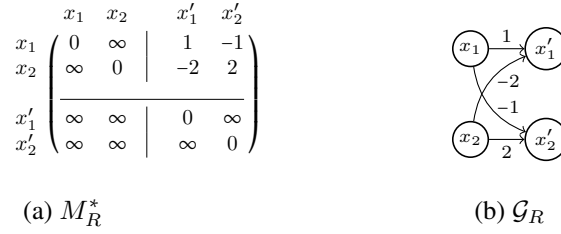


Figure 2: Let $\phi(x_1, x_2, x'_1, x'_2) \equiv x_1 - x'_1 \leq 1 \wedge x_1 - x'_2 \leq -1 \wedge x_2 - x'_1 \leq -2 \wedge x_2 - x'_2 \leq 2$ be a difference bounds constraint. **(a)** shows the closed DBM of M_ϕ^* and **(b)** shows the constraint graph \mathcal{G}_ϕ .

Given a consistent DBM M , the unique closed DBM which is logically equivalent to M is denoted by M^* . If M is an inconsistent DBM, we denote $M^* = [-\infty]_N$, by convention. Observe that the closed DBM is a canonical (unique) representation of a difference bounds constraint. Moreover, this canonical representation of a DBM can be computed in cubic time, using the classical Floyd-Warshall shortest path algorithm.

It is known that quantifier elimination for difference bounds constraints takes cubic time in the size of the binary representation of the constraint². Then the set of relations $R \subseteq \mathbb{Z}^x \times \mathbb{Z}^x$ defined by difference bounds constraints $\phi_R(\mathbf{x}, \mathbf{x}')$ is closed under relational composition. Since the identity relation I_x is definable by a difference constraint and the empty relation \emptyset is definable by any inconsistent constraint the set of difference bounds relations forms a class (Definition 1), denoted as DB in the following.

Because any difference bounds relation $R \subseteq \mathbb{Z}^x \times \mathbb{Z}^x$, defined by a formula $\phi(\mathbf{x}, \mathbf{x}')$, is uniquely represented by the difference bounds constraint $\Phi(M_\phi^*)$, we define the size of its binary representation as $|R| = |\Phi(M_\phi^*)|$, independently of the choice of ϕ . In principle, any algorithm that takes as input a difference bounds relation R can be considered w.l.o.g. to work directly on its canonical representation, because the time needed to compute the canonical representation of R is $\mathcal{O}(|R|^3)$, thus any super-polynomial bound derived with this assumption carry over to the general case.

3.2 Octagonal Constraints

Given a set of variables $\mathbf{x} = \{x_1, \dots, x_N\}$, an octagonal constraint $\phi(\mathbf{x})$ (Definition 2) is usually represented by a difference bounds constraints $\bar{\phi}(\mathbf{y})$, where $\mathbf{y} = \{y_1, \dots, y_{2N}\}$, y_{2i-1} stands for $+x_i$ and y_{2i} stands for $-x_i$, with the implicit requirement that $y_{2i-1} = -y_{2i}$, for each $1 \leq i \leq N$. Observe that the latter condition cannot be directly represented as a difference bounds constraint. Formally, we have:

$$\begin{aligned}
 (x_i - x_j \leq c) \in \text{Atom}(\phi) &\Leftrightarrow (y_{2i-1} - y_{2j-1} \leq c), (y_{2j} - y_{2i} \leq c) \in \text{Atom}(\bar{\phi}) \\
 (-x_i + x_j \leq c) \in \text{Atom}(\phi) &\Leftrightarrow (y_{2j-1} - y_{2i-1} \leq c), (y_{2i} - y_{2j} \leq c) \in \text{Atom}(\bar{\phi}) \\
 (-x_i - x_j \leq c) \in \text{Atom}(\phi) &\Leftrightarrow (y_{2i} - y_{2j-1} \leq c), (y_{2j} - y_{2i-1} \leq c) \in \text{Atom}(\bar{\phi}) \\
 (x_i + x_j \leq c) \in \text{Atom}(\phi) &\Leftrightarrow (y_{2i-1} - y_{2j} \leq c), (y_{2j-1} - y_{2i} \leq c) \in \text{Atom}(\bar{\phi})
 \end{aligned}$$

²To eliminate $\exists x.\phi(x)$, one computes the closed DBM M_ϕ^* in cubic time in the binary size of ϕ and eliminates the row and column corresponding to x from it.

In order to handle the \mathbf{y} variables in the following, we define $\bar{i} = i - 1$, if i is even, and $\bar{i} = i + 1$ if i is odd. Obviously, we have $\bar{\bar{i}} = i$, for all $i \in \mathbb{Z}$, $i \geq 1$.

An octagonal constraint $\phi(\mathbf{x})$ is represented by the matrix $M_{\phi}^{-} \in \mathbb{Z}_{\infty}^{2N \times 2N}$, corresponding to $\bar{\phi}(\mathbf{y})$. A matrix $M \in \mathbb{Z}_{\infty}^{2N \times 2N}$ is *coherent* if $M_{ij} = M_{\bar{j}\bar{i}}$ for all $1 \leq i, j \leq 2N$. This property is needed because an atomic proposition $x_i - x_j \leq c$ can be represented as both $y_{2i-1} - y_{2j-1} \leq c$ and $y_{2j} - y_{2i} \leq c$. Dually, a coherent matrix $M \in \mathbb{Z}_{\infty}^{2N \times 2N}$ corresponds to the following octagonal constraint:

$$\Omega(M) \equiv \bigwedge_{1 \leq i, j \leq N} x_i - x_j \leq M_{2i-1, 2j-1} \wedge \bigwedge_{1 \leq i, j \leq N} x_i + x_j \leq M_{2i-1, 2j} \wedge \bigwedge_{1 \leq i, j \leq N} -x_i - x_j \leq M_{2i, 2j-1}$$

A coherent DBM M is said to be *octagonal-consistent* if $\Omega(M)$ is consistent.

Definition 5 An octagonal-consistent coherent DBM $M \in \mathbb{Z}_{\infty}^{2N \times 2N}$ is said to be *tightly closed* if and only if it is closed and $M_{ij} \leq \lfloor \frac{M_{i\bar{i}}}{2} \rfloor + \lfloor \frac{M_{\bar{j}j}}{2} \rfloor$, for all $1 \leq i, j \leq N$.

The last condition from Definition 5 ensures that the knowledge induced by the implicit conditions $y_i + y_{\bar{i}} = 0$, which cannot be represented as difference constraints, has been propagated through the DBM. Since $2y_i = y_i - y_{\bar{i}} \leq M_{i\bar{i}}$ and $-2y_j = y_{\bar{j}} - y_j \leq M_{\bar{j}j}$, we have $y_i \leq \lfloor \frac{M_{i\bar{i}}}{2} \rfloor$ and $-y_j \leq \lfloor \frac{M_{\bar{j}j}}{2} \rfloor$, which implies $y_i - y_j \leq \lfloor \frac{M_{i\bar{i}}}{2} \rfloor + \lfloor \frac{M_{\bar{j}j}}{2} \rfloor$, thus $M_{ij} \leq \lfloor \frac{M_{i\bar{i}}}{2} \rfloor + \lfloor \frac{M_{\bar{j}j}}{2} \rfloor$ must hold, if M is supposed to be the most precise DBM representation of an octagonal constraint. If $j = \bar{i}$ in the previous, we obtain $M_{i\bar{i}} \leq 2\lfloor \frac{M_{i\bar{i}}}{2} \rfloor$, implying that $M_{i\bar{i}}$ is necessarily even, if M is tightly closed.

The following theorem [2] provides an effective way of testing octagonal-consistency and computing the tight closure of a coherent DBM. Moreover, it shows that the tight closure of a given DBM is unique and can also be computed within the same cubic time upper bound, as the DBM closure:

Theorem 1 Let $M \in \mathbb{Z}_{\infty}^{2N \times 2N}$ be a coherent DBM. Then M is octagonal-consistent iff M is consistent and $\lfloor \frac{M_{i\bar{i}}^*}{2} \rfloor + \lfloor \frac{M_{\bar{j}j}^*}{2} \rfloor \geq 0$, for all $1 \leq i \leq 2N$. Moreover, if M is octagonal-consistent, the tight closure of M is the DBM $M^t \in \mathbb{Z}_{\infty}^{2N \times 2N}$ defined as:

$$M_{ij}^t = \min \left\{ M_{ij}^*, \left\lfloor \frac{M_{i\bar{i}}^*}{2} \right\rfloor + \left\lfloor \frac{M_{\bar{j}j}^*}{2} \right\rfloor \right\}$$

for all $1 \leq i, j \leq 2N$, where $M^* \in \mathbb{Z}_{\infty}^{2N \times 2N}$ is the closure of M .

Proof: [2, Theorems 2 and 3]. □

The tight closure of DBMs is needed for checking entailment between octagonal constraints and for quantifier elimination, as shown by the following proposition.

Proposition 1 Let $\phi(\mathbf{x})$ and $\psi(\mathbf{x})$ be two consistent octagonal constraints. Then, the following hold:

1. $\phi \Rightarrow \psi$ if and only if $\left(M_{\phi}^t \right)_{ij} \leq \left(M_{\psi}^t \right)_{ij}$, for all $1 \leq i, j \leq 2N$.
2. $\exists x_k. \phi(\mathbf{x}) \Leftrightarrow \Omega(M')$, where M' is the DBM obtained by eliminating the lines and columns $2k$ and $2k + 1$ from M_{ϕ}^t .

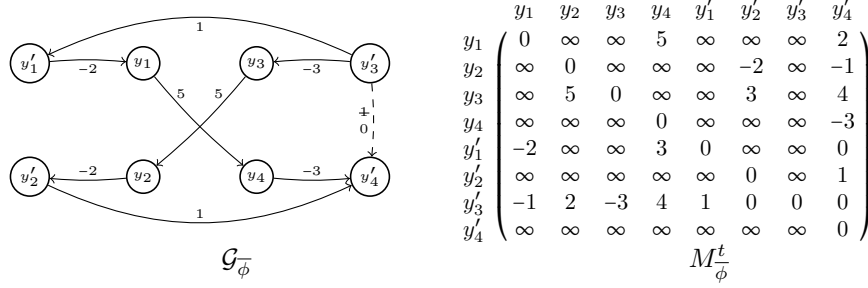


Figure 3: Graph and matrix representation of the difference bounds representation $\bar{\phi}(\mathbf{y}, \mathbf{y}')$ of an octagonal relation defined by $\phi(\mathbf{x}, \mathbf{x}') \equiv x_1 + x_2 \leq 5 \wedge x'_1 - x_1 \leq -2 \wedge x'_2 - x_2 \leq -3 \wedge x'_2 - x'_1 \leq 1$.

Proof: Point (1) is by [28, Theorem 4.4.1] and point (2) is by [5, Theorem 2]. \square

Since octagonal constraints have quantifier elimination, by Proposition 1 (2), the set OCT of octagonal relations forms a class, in the sense of Definition 1. Moreover, since a tightly closed DBM is a canonical representation of an octagonal relation, we can define w.l.o.g. the size of the binary representation of a relation $R \in \text{OCT}$ as $|R| = |\Omega(M_\phi^t)|$, where ϕ is any octagonal constraint that defines R . Again, this definition has no impact on the computational complexity of the decision problems involving octagonal relations, because the canonical representation of any $R \in \text{OCT}$ can be computed in time $\mathcal{O}(|R|^3)$.

Example 2 Consider the octagonal relation defined by $\phi(x_1, x_2, x'_1, x'_2) \equiv x_1 + x_2 \leq 5 \wedge x'_1 - x_1 \leq -2 \wedge x'_2 - x_2 \leq -3 \wedge x'_2 - x'_1 \leq 1$. Its difference bounds representation is $\bar{\phi}(\mathbf{y}, \mathbf{y}') \Leftrightarrow y_1 - y_4 \leq 5 \wedge y_3 - y_2 \leq 5 \wedge y'_1 - y_1 \leq -2 \wedge y_2 - y'_2 \leq -2 \wedge y'_3 - y_3 \leq -3 \wedge y_4 - y'_4 \leq -3 \wedge y'_3 - y'_1 \leq 1 \wedge y'_2 - y'_4 \leq 1$, where $\mathbf{y} = \{y_1, \dots, y_4\}$. Figure 3(a) shows the graph representation $\mathcal{G}_{\bar{\phi}}$. Note that the implicit constraint $y'_3 - y'_4 \leq 1$, represented by a dashed edge in Figure 3(a), is not tight. The tightening step replaces the bound 1, crossed in Figure 3(a), with 0. Figure 3(b) shows the tightly closed DBM representation of R , denoted $M_{\bar{\phi}}^t$. \blacksquare

3.3 Periodic Relations

As shown in the previous, both difference bounds and octagonal relations are closed under compositions and have canonical matrix representations. When studying the complexity of the reachability problem $\text{REACHFLAT}(\text{OCT})$, a crucial point concerns the behavior of the power sequence $\{R^k\}_{k=0}^\infty$, for a relation $R \in \text{OCT}$. A first observation is that, for any $k \geq 0$, we have that $R^k \in \text{OCT}$. If we denote by $\sigma(R)$ the canonical DBM representation of R , we show that the sequence of DBMs $\{\sigma(R^k)\}_{k=0}^\infty$ is *periodic*, in a sense that we define next. This fact is important for the definition of a nondeterministic algorithm that solves the above reachability problem.

We consider infinite sequences $\{s_k\}_{k=0}^\infty$ in $\mathbb{Z}_{\pm\infty}$, with the following extension of addition: (i) for all $x \in \mathbb{Z}_\infty$, $x + \infty = \infty + x = \infty$, and (ii) for all $x \in \mathbb{Z}_{\pm\infty}$, $x + (-\infty) = -\infty + x = -\infty$. A sequence $\{s_k\}_{k=0}^\infty$ is an *arithmetic progression* if there exists a constant $\lambda \in \mathbb{Z}_{\pm\infty}$, called *rate*, such that $s_{k+1} = s_k + \lambda$, for all $k \geq 0$. A generalization of this notion are *periodic* sequences, defined below.

Definition 6 An infinite sequence $\{s_k\}_{k=0}^\infty$, where $s_k \in \mathbb{Z}_{\pm\infty}$, for all $k \geq 0$, is said to be periodic if and only if there exist integer constants $b \geq 0$, $c > 0$ and $\lambda_0, \dots, \lambda_{c-1} \in \mathbb{Z}_{\pm\infty}$ such that $s_{b+(k+1)c+i} = s_{b+kc+i} + \lambda_i$, for all $k \geq 0$ and all $i \in [c]$. The smallest b , c and λ_i are called the prefix, period and rates of the sequence.

Note that an arithmetic progression is a periodic sequence with prefix 0 and period 1.

In the following, we consider sequences of square matrices and say that an infinite sequence $\{M_k\}_{k=0}^\infty$ of matrices $M_k \in \mathbb{Z}_{\pm\infty}^{n \times n}$ is *periodic* if every sequence $\{(M_k)_{ij}\}_{k=0}^\infty$ is periodic, for all $i, j \in [n]$. The next lemma provides a characterization of periodicity for a sequence of matrices, with an estimation of its prefix and period.

Lemma 1 A sequence of $\mathbb{Z}_{\pm\infty}^{n \times n}$ matrices $\{M_k\}_{k=0}^\infty$ is periodic iff there exist integers $b \geq 0$, $c > 0$ and matrices $\Lambda_0, \dots, \Lambda_{c-1} \in \mathbb{Z}_{\pm\infty}^{n \times n}$ such that:

$$\forall k \geq 0 \forall i \in [c] . M_{b+(k+1)c+i} = M_{b+kc+i} + \Lambda_i .$$

If, moreover, b_{ij} and c_{ij} are the prefix and period of the sequence $\{(M_k)_{ij}\}_{k=0}^\infty$, then $b = \max_{1 \leq i, j \leq n} (b_{ij})$, $c = \text{lcm}_{1 \leq i, j \leq n} (c_{ij})$ are the smallest such integers.

Proof. “ \Rightarrow ” Suppose that the sequence $\{M_k\}_{k=0}^\infty$ is periodic. Then, for each $i, j \in [n]$, the sequence $\{(M_k)_{ij}\}_{k=0}^\infty$ is periodic, and let $\lambda_0^{ij}, \dots, \lambda_{c_{ij}-1}^{ij}$ be the rates of this sequence. For all $\ell \in [c]$ and all $i, j \in [n]$, define $(\Lambda_\ell)_{ij} = \frac{c}{c_{ij}} \cdot \left(\lambda_{(b-b_{ij}+\ell) \bmod c_{ij}}^{ij} \right)$. The check $M_{b+(k+1)c+\ell} = M_{b+kc+\ell} + \Lambda_\ell$, for all $k \geq 0$ and $\ell \in [c]$ is straightforward. “ \Leftarrow ” For each $i, j \in [n]$, the sequence $\{(M_k)_{ij}\}_{k=0}^\infty$ is periodic: $(M_{b+(k+1)c+\ell})_{ij} = (M_{b+kc+\ell})_{ij} + (\Lambda_\ell)_{ij}$, for all $k \geq 0$ and $\ell \in [c]$.

For the last point, suppose first, by contradiction, that there exists $b' < \max_{1 \leq i, j \leq n} (b_{ij})$ such that $M_{b'+(k+1)c+\ell} = M_{b'+kc+\ell} + \Lambda_\ell$, for all $k \geq 0$ and $\ell \in [c]$. Let $1 \leq s, t \leq n$ be such that $b_{st} = \max_{1 \leq i, j \leq n} (b_{ij})$. Then the sequence $\{(M_k)_{st}\}_{k=0}^\infty$ is periodic with prefix $b' < b_{st}$, contradiction. Second, suppose by contradiction, that there exists $c' < \text{lcm}_{1 \leq i, j \leq n} (c_{ij})$ such that $M_{b+(k+1)c'+\ell} = M_{b+kc'+\ell} + \Lambda_\ell$, for all $k \geq 0$ and $\ell \in [c']$. Then there exists $1 \leq s, t \leq n$ such that c_{st} does not divide c' , which contradicts the fact that the sequence $\{(M_k)_{st}\}_{k=0}^\infty$ is periodic with period c_{st} . \square

Let us focus now on sequences of matrices that represent the power sequences $\{R^k\}_{k=0}^\infty$, where $R \in \text{OCT}$. Given a set of variables $\mathbf{x} = \{x_1, \dots, x_N\}$, we denote by $\text{OCT}_{\mathbf{x}}$ the class of octagonal relations $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$. Formally, let $\sigma : \text{OCT}_{\mathbf{x}} \rightarrow \mathbb{Z}_{\pm\infty}^{4N \times 4N} \cup \{[-\infty]_{4N}\}$ be the bijection that maps each consistent relation R into its canonical DBM $\sigma(R) \in \mathbb{Z}_{\pm\infty}^{4N \times 4N}$ and the inconsistent relation into $\sigma(\emptyset) = [-\infty]_{4N}$. Then R is said to be *periodic* if the matrix sequence $\{\sigma(R^k)\}_{k=0}^\infty$ is periodic. If every relation in a certain class is periodic, we call that class periodic as well.

Example 3 Consider the octagonal relation $R \subseteq \mathbb{Z}^{\{x,y\}} \times \mathbb{Z}^{\{x,y\}}$ defined by the formula $x' = y + 1 \wedge y' = x$, where for all $\ell \in \mathbb{N}$:

$$\sigma(R^{2\ell+1}) = \begin{array}{c|cccc} & x & y & x' & y' \\ \hline x & 0 & \infty & \infty & \ell \\ y & \infty & 0 & -\ell-1 & \infty \\ x' & \infty & \ell+1 & 0 & \infty \\ y' & -\ell & \infty & \infty & 0 \end{array} \quad \sigma(R^{2\ell+2}) = \begin{array}{c|cccc} & x & y & x' & y' \\ \hline x & 0 & \infty & -\ell-1 & \infty \\ y & \infty & 0 & \infty & -\ell-1 \\ x' & \ell+1 & \infty & 0 & \infty \\ y' & \infty & \ell+1 & \infty & 0 \end{array}$$

The sequence $\{\sigma(R^k)\}_{k=0}^\infty$ is periodic with prefix $b = 1$ and period $c = 2$, where:

$$\Lambda_0 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{vmatrix} \quad \Lambda_1 = \begin{vmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix}$$

■

One of the results in this paper is that the class OCT is periodic. The proof of this fact relies essentially on the fact that the class DB is periodic, and uses (a variant of) Theorem 1 to generalize this result from difference bounds to octagonal relations. In the next section we give a generic nondeterministic decision procedure for the problem REACHFLAT(\mathcal{R}), where \mathcal{R} is a periodic class of relations. Moreover, we identify certain conditions under which each branch of the procedure terminates in polynomial time, which provides an NP upper bound for the REACHFLAT(\mathcal{R}) problem.

4 An Algorithm for the Reachability Problem

In general, the decision procedures for the reachability problem for flat counter machines rely on *acceleration* [4, 13], which is defining the transitive closure of the relations that occur on the cycles of these machines by formulae from the quantifier-free fragment of Presburger arithmetic. To show that these reachability problems belong to the class NP, it is essential to build these QFPA formulae in polynomial time.

For the sake of simplicity, we explain the idea of a nondeterministic algorithm (Algorithm 1) for the reachability problem on the flat counter machine below:

$$\ell_{\text{init}} \xrightarrow{I(\mathbf{x})} \overset{\phi(\mathbf{x}, \mathbf{x}')}{\underset{\text{Q}}{\ell}} \xrightarrow{F(\mathbf{x})} \ell_{\text{fin}} \quad (1)$$

where $I(\mathbf{x})$ and $F(\mathbf{x})$ are QFPA formulae and $\phi(\mathbf{x}, \mathbf{x}')$ is an octagonal constraint defining a relation $R \in \text{OCT}_{\mathbf{x}}$, for a given set of variables $\mathbf{x} = \{x_1, \dots, x_N\}$.

Let us assume for now that this relation is periodic. The algorithm guesses candidate values for the prefix $b \geq 0$ and period $c > 0$ of R (line 2), computes a candidate rate Λ (line 3), and checks if b, c and Λ satisfy the following condition (line 4):

$$\text{IND}(B, C, \Lambda) : \forall n \geq 0. \sigma(\sigma^{-1}(B + n \cdot \Lambda) \circ \sigma^{-1}(C)) = B + (n+1) \cdot \Lambda \quad (2)$$

Algorithm 1 nondeterministic algorithm for the reachability problem (1)

input: $M = \langle \mathbf{x}, \{\ell_{\text{init}}, \ell, \ell_{\text{fin}}\}, \ell_{\text{init}}, \ell_{\text{fin}}, \Delta \rangle$ of the form (1), where $\mathbf{x} = \{x_1, \dots, x_N\}$
output: YES if and only if M has a run from ℓ_{init} to ℓ_{fin}

- 1: **let** R be the relation defined by $\phi(\mathbf{x}, \mathbf{x}')$
- 2: **chose** $b \geq 0$ and $c > 0$
- 3: **let** $\Lambda \in \mathbb{Z}_{\infty}^{4N \times 4N}$ be a matrix such that $\sigma(R^b) + \Lambda = \sigma(R^{b+c})$
- 4: **if** $\text{IND}(\sigma(R^b), \sigma(R^c), \Lambda)$ **then**
- 5: **chose** $i \in [b]$
- 6: $\phi^{<b} \leftarrow I(\mathbf{x}) \wedge \Omega(\sigma(R^i)) \wedge F(\mathbf{x}')$
- 7: **chose** $j \in [c]$
- 8: $\phi^{\geq b} \leftarrow k \geq 0 \wedge I(\mathbf{x}) \wedge \varsigma(\sigma(R^{b+j}) + k \cdot \Lambda) \wedge F(\mathbf{x}')$
- 9: **if** $\phi^{<b} \vee \phi^{\geq b}$ is satisfiable **then**
- 10: **return** YES
- 11: **fail**

where B, C and Λ are square matrices of equal dimension, in our case $B = \sigma(R^b)$, $C = \sigma(R^c)$ and Λ is such that $\sigma(R^b) + \Lambda = \sigma(R^{b+c})$. Intuitively, this means that b , c and Λ are valid choices for the prefix, period and rate of the sequence of matrices $\{\sigma(R^k)\}_{k=0}^{\infty}$, in the sense of Lemma 1.

In case the reachability problem for M has a positive answer, i.e. there exists a run from ℓ_{init} to ℓ_{fin} , two cases are possible. Either the number of iterations of the loop is (i) strictly smaller than b , or (ii) between $b + nc$ and $b + (n + 1)c$, for some $n \geq 0$. The first case is captured by the QFPA formula $\phi^{<b}$ (line 6), where $\Omega(\sigma(R^i))$ is the canonical octagonal constraint representing the relation R^i .

The second case is encoded by the QFPA formula $\phi^{\geq b}$ (line 8). Here $k \notin \mathbf{x}$ is a parameter variable and by $\mathbb{Z}[k]_{\infty}$ we denote the set of univariate linear terms of the form $a \cdot k + b$, with $a, b \in \mathbb{Z}_{\infty}$. Also $\mathbb{Z}[k]_{\infty}^{m \times m}$ denotes the set of $m \times m$ square matrices of such terms. With these notations, ς is a mapping of matrices $M[k] \in \mathbb{Z}[k]_{\infty}^{4N \times 4N}$ into parametric octagonal constraints consisting of atomic propositions of the form $\pm x \pm y \leq a \cdot k + b$, defined in the same way as the octagonal constraint $\Omega(M)$ is defined for a matrix $M \in \mathbb{Z}_{\infty}^{4N \times 4N}$. Moreover, ς satisfies the following condition:

$$\forall M \in \mathbb{Z}[k]_{\infty}^{4N \times 4N} \forall n \in \mathbb{N} . \varsigma(M)(n) = \sigma^{-1}(M[n/k]) \quad (3)$$

The final step is checking the satisfiability of the disjunction $\phi^{<b} \vee \phi^{\geq b}$ (line 9). If the formula produced by a nondeterministic branch of the algorithm is satisfiable, the reachability question has a positive answer. Otherwise, if no branch produces a satisfiable formula, the reachability question has a negative answer.

To prove that the class of reachability problems $\text{REACHFLAT}(\text{OCT})$ is in NP , it is enough to show that, for any machine M of the form (1), each branch of Algorithm 1 terminates in $\text{PTIME}(|M|)$. For this, the matrices $\sigma(R^c)$, $\sigma(R^i)$ and $\sigma(R^{b+j})$ must be computable in $\text{PTIME}(|R|)$, for all $i = 0, \dots, b$ and $j = 0, \dots, c$ and, moreover, the condition $\text{IND}(\sigma(R^b), \sigma(R^c), \Lambda)$ (2) must be decidable in $\text{NPTIME}(|R|)$. Under these conditions, the QFPA formulae $\phi^{<b}$ and $\phi^{\geq b}$ are of polynomial size in $|M|$, and the satisfiability of their disjunction is decidable in $\text{NPTIME}(|M|)$.

The following theorem generalizes this argument to arbitrary flat counter machines by giving sufficient conditions under which the class $\text{REACHFLAT}(\text{OCT})$ is NP-complete.

Theorem 2 $\text{REACHFLAT}(\text{OCT})$ is NP-complete if there exists a constant d , such that the following hold, for each relation $R \in \text{OCT}$:

1. $|R^n| = \mathcal{O}((|R| \cdot \log n)^d)$, for all $n > 0$,
2. R is periodic with prefix and period of the order of $2^{\mathcal{O}(|R|^d)}$.

Before giving the proof of Theorem 2, we show that the condition $\text{IND}(B, C, \Lambda)$ is decidable in nondeterministic polynomial time, by reduction to the satisfiability of a QFPA formula. The proof relies on a symbolic tight closure algorithm (Algorithm 2), which builds such a formula using a cubic number of steps.

Algorithm 2 Symbolic Tight Closure algorithm

input: a matrix $M \in \mathbb{Z}_\infty[k]^{m \times m}$ of univariate linear terms
output: a matrix $T \in \mathbb{Z}_\infty[k]^{m \times m}$ of univariate terms over $\min, +$, and $\lfloor \frac{\cdot}{2} \rfloor$

- 1: **function** SYMBFW(M)
- 2: **for** $i = 1, \dots, m$ **do**
- 3: **for** $j = 1, \dots, m$ **do**
- 4: $P_{ij} \leftarrow M_{ij}$
- 5: **for** $k = 1, \dots, m$ **do**
- 6: **for** $i = 1, \dots, m$ **do**
- 7: **for** $j = 1, \dots, m$ **do**
- 8: $P_{ij} \leftarrow \min(P_{ij}, P_{ik} + P_{kj})$
- 9: **return** P
- 1: $T \leftarrow \text{SYMBFW}(M)$
- 2: **for** $i = 1, \dots, m$ **do**
- 3: **for** $j = 1, \dots, m$ **do**
- 4: $T_{ij} \leftarrow \min(T_{ij}, \lfloor \frac{T_{ii}}{2} \rfloor + \lfloor \frac{T_{jj}}{2} \rfloor)$
- 5: **return** T

Lemma 2 Given $N > 0$ and matrices $B, C, \Lambda \in \mathbb{Z}_\infty^{4N \times 4N}$, the condition $\text{IND}(B, C, \Lambda)$ is decidable in nondeterministic polynomial time.

Proof. The condition $\text{IND}(B, C, \Lambda)$ checks the validity of the equivalence:

$$\forall k \geq 0. \sigma(\sigma^{-1}(B + k \cdot \Lambda) \circ \sigma^{-1}(C)) = B + (k + 1) \cdot \Lambda.$$

For a matrix $M \in \mathbb{Z}_\infty[k]^{4N \times 4N}$ of univariate linear terms in k , we define the labeled graph $\mathcal{H}_M = \langle \mathbf{y} \cup \mathbf{y}', \rightarrow \rangle$, where $\mathbf{y} = \{y_1, \dots, y_{2N}\}$ are the variables used in the difference bounds encoding of an octagonal relation with variables $\{x_1, \dots, x_N\}$, and whose labeled edges are $y_i \xrightarrow{M_{i,j}} y_j$, $y_i \xrightarrow{M_{i,j+2N}} y'_j$, $y'_i \xrightarrow{M_{i+2N,j}} y_j$, $y'_i \xrightarrow{M_{i+2N,j+2N}} y'_j$, for all $1 \leq i, j \leq 2N$. The left-hand side of the equivalence $\text{IND}(B, C, \Lambda)$ corresponds to the graph \mathcal{H}_{lhs} with vertices $\mathbf{y}^{(0)} \cup \mathbf{y}^{(1)} \cup \mathbf{y}^{(2)}$, such that:

- $\mathcal{H}_{\text{lhs}} \downarrow_{\mathbf{y}^{(0)} \cup \mathbf{y}^{(1)}}$ is the graph $\mathcal{H}_{B+k \cdot \Lambda}$, in which the vertices $\mathbf{y}^{(0)}$ and $\mathbf{y}^{(1)}$ replace \mathbf{y} and \mathbf{y}' , and
- $\mathcal{H}_{\text{lhs}} \downarrow_{\mathbf{y}^{(1)} \cup \mathbf{y}^{(2)}}$ is the constraint graph \mathcal{G}_C representing the difference bounds constraint $\Phi(C)$, in which $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ replace the vertices \mathbf{y} and \mathbf{y}' of \mathcal{G}_C , respectively.

The right-hand side of the equivalence $\text{IND}(B, C, \Lambda)$ is represented, in a similar way, by the graph \mathcal{H}_{rhs} , which equals the graph $\mathcal{H}_{B+(k+1) \cdot \Lambda}$, with vertices $\mathbf{y}^{(0)}$ and $\mathbf{y}^{(2)}$ replacing \mathbf{y} and \mathbf{y}' , respectively.

Since both graphs denote (parametric) octagonal constraints, by Proposition 1 (1) we need to prove that, for all $k \geq 0$ the path labels corresponding to the minimal paths within the *tight closures* of the incidence matrices of \mathcal{H}_{lhs} and \mathcal{H}_{rhs} are equal, for all $k \geq 0$. These tight closures can be expressed by univariate terms with variable k , built in time $\mathcal{O}(N^3)$, from constants $c \in \mathbb{Z}$ and the functions $\min, +$ and $\lfloor \frac{\cdot}{2} \rfloor$, by Algorithm 2, which implements the result of Theorem 1. Notice that the size of each such term is bounded by the time needed to build it, i.e. $\mathcal{O}(N^3)$. Finally, each term can be encoded in QFPA, because all constituent functions, i.e. $\min, +$ and $\lfloor \frac{\cdot}{2} \rfloor$ are QFPA-definable. As a direct consequence, the condition $\text{IND}(B, C, \Lambda)$ is decidable in NPTIME. \square

Proof of Theorem 2 NP-hardness is by reduction from the satisfiability problem for QFPA, and the fact that any transition rule of a flat CM, that is not part of a cycle, can be labeled by an arbitrary QFPA formula. Given an instance $\phi(\mathbf{x})$ of the QFPA satisfiability problem, we consider the CM $\ell_{\text{init}} \xrightarrow{\phi} \ell_{\text{fin}}$. The reachability problem has a positive answer iff ϕ has a satisfying assignment.

To prove that $\text{REACHFLAT}(\text{OCT})$ is contained in NP, let $M = \langle \mathbf{x}, \mathcal{L}, \ell_1, \ell_n, \Delta \rangle$ be a flat CM, where $\mathcal{L} = \{\ell_1, \dots, \ell_n\}$. First, we reduce the control flow graph $\langle \mathcal{L}, \Delta \rangle$ of M to a dag and several self-loops, by replacing each non-trivial cycle:

$$\ell_{i_0} \xrightarrow{\phi_0} \ell_{i_1} \xrightarrow{\phi_1} \ell_{i_2} \dots \ell_{i_{k-2}} \xrightarrow{\phi_{k-2}} \ell_{i_{k-1}} \xrightarrow{\phi_{k-1}} \ell_{i_0}$$

where $k > 1$, with the following sequence:

$$\begin{array}{ccccccc} \lambda_0(\mathbf{x}, \mathbf{x}') & \lambda_1(\mathbf{x}, \mathbf{x}') & & \lambda_{k-1}(\mathbf{x}, \mathbf{x}') & & & \\ \textcircled{} & \textcircled{} & & \textcircled{} & & & \\ \ell_{i_0} & \xrightarrow{\phi_0} \ell_{i_1} & \dots & \ell_{i_{k-1}} & \xrightarrow{\phi_{k-1}} \ell'_{i_0} & \xrightarrow{\lambda_0(\mathbf{x}, \mathbf{x}')} \dots & \xrightarrow{\phi_{k-1}} \ell'_{i_{k-1}} \end{array} \quad (4)$$

where $\lambda_j(\mathbf{x}, \mathbf{x}') = \exists \mathbf{x}_1, \dots, \mathbf{x}_{j+1} \cdot \phi_j(\mathbf{x}, \mathbf{x}_{j+1}) \wedge \dots \wedge \phi_{k-1}(\mathbf{x}_{k-1}, \mathbf{x}_k) \wedge \phi_0(\mathbf{x}_k, \mathbf{x}_1) \wedge \dots \wedge \phi_{j-1}(\mathbf{x}_{j-1}, \mathbf{x}')$, $\ell'_{i_1}, \dots, \ell'_{i_{k-1}}$ are fresh control locations not in \mathcal{L} , and for each rule $\ell_{i_j} \xrightarrow{\phi} \ell_m$ of M , where $m \neq i_{(j+1) \bmod k}$, we add a rule $\ell'_{i_j} \xrightarrow{\phi} \ell_m$, for each $j = 0, \dots, k-1$. This step doubles at most the number of control locations in \mathcal{L} . W.l.o.g., we can consider henceforth that each control location ℓ_i belongs to at most one self loop labeled by a formula ϕ_i , for $i = 1, \dots, 2n$.

For each cycle $\ell_i \xrightarrow{\phi_i} \ell_i$, where ϕ_i defines a relation $R_i \in \text{OCT}_{\mathbf{x}}$, for each $i = 1, \dots, 2n$, the nondeterministic algorithm performs the steps of Algorithm 1, namely:

1. Guess values $b_i \geq 0$ and $c_i > 0$, of the order of $2^{\mathcal{O}(|R_i|^d)}$, compute the powers $R_i^{b_i}$, $R_i^{c_i}$ and $R_i^{b_i+c_i}$ and find Λ_i such that $\sigma(R_i^{b_i+c_i}) = \sigma(R_i^{b_i}) + \Lambda_i$. By the hypotheses

- (1) and (2) this computation is possible in $\text{PTIME}(|R_i|)$, using exponentiation by squaring.
2. Check the validity of the condition $\text{IND}(\sigma(R_i^{b_i}), \sigma(R_i^{c_i}), \Lambda_i)$ in $\text{NPTIME}(|R_i|)$, which is possible by Lemma 2.
3. Build a QFPA formula $\phi_i(k, \mathbf{x}, \mathbf{x}') = \phi_i^{<b_i}(\mathbf{x}, \mathbf{x}') \vee \phi_i^{>b_i}(k, \mathbf{x}, \mathbf{x}')$ in $\text{PTIME}(|R_i|)$.

The second step uses a breadth-first dag traversal to label each control location in ℓ_i , for $i = 1, \dots, 2n$, with a QFPA formula $\theta_i(\mathbf{x}, \mathbf{x}')$ that captures the summary (effect) of the set of executions of M from the initial state ℓ_1 to ℓ_i . We assume w.l.o.g. that (i) for every location ℓ_i , $i = 2, \dots, 2n$, there exists a path in M from ℓ_1 to ℓ_i , and (ii) there is no self-loop involving ℓ_1 in M . We consider the sets of variables $\mathbf{k} = \{k_1, \dots, k_{2n}\}$ and $\mathbf{x}_i^t = \{x_{i,\ell}^t \mid \ell = 1, \dots, N\}$, where $i = 1, \dots, 2n$ and $t \in \{\text{in}, \text{out}\}$. We define $\theta_1 = \top$, and for all $j = 2, \dots, 2n$:

$$\theta_j(\mathbf{x}_{1,\dots,2n}^{\text{in}}, \mathbf{x}_{1,\dots,2n}^{\text{out}}) = \bigvee_{\substack{R_{ij} \\ \ell_i \Rightarrow \ell_j}} \theta_i(\mathbf{x}_{1,\dots,2n}^{\text{in}}, \mathbf{x}_{1,\dots,2n}^{\text{out}}) \wedge \phi_{ij}(\mathbf{x}_i^{\text{out}}, \mathbf{x}_j^{\text{in}}) \wedge \phi_j(k_j, \mathbf{x}_j^{\text{in}}, \mathbf{x}_j^{\text{out}})$$

where ϕ_{ij} is the formula defining R_{ij} . It is not difficult to prove that, for all $i = 1, \dots, 2n$ and $\nu, \nu' \in \mathbb{Z}^x$: $(\nu, \nu') \models \theta_i \Leftrightarrow M$ has a run $(\ell_1, \nu), \dots, (\ell_i, \nu')$.

Since for every location ℓ_i , $i = 2, \dots, 2n$, there exists a control path from ℓ_1 to it, the breadth-first traversal guarantees that each predecessor ℓ_i of a location ℓ_j is labeled with the summary θ_i before ℓ_j is visited by the algorithm, ensuring that the definition above is correct. Moreover, the dag structure (excepting the self-loops) of the CM guarantees that it is sufficient to visit each locations only once in order to label it with a summary. Thus, the labeling takes polynomial time and, consequently, $|\theta_i|$ is polynomially bounded by $|M|$, for each $i = 1, \dots, 2n$. Since the size of the summary labeling the final location is polynomial in $|M|$, and the satisfiability problem is in $\text{NPTIME}(|M|)$, it follows that $\text{REACHFLAT}(\text{OCT})$ is contained in NP. \square

In the next section we prove the point (1) from Theorem 2. The rest of the paper is dedicated to proving point (2), which requires a more complex technical argument.

4.1 The First Ingredients of the Proof

In order to apply Theorem 2 we start by proving that the first assumption from its statement holds for each octagonal relation $R \in \text{OCT}$, namely that the size of the binary representation of the n -th power R^n is bounded by a polynomial function with arguments $|R|$ and $\log n$. In this case, the binary size of an exponentially large power R^n , where $n = 2^{\mathcal{O}(|R|^d)}$ and d is a constant, is bounded by a polynomial in $|R|$. Moreover, such powers can be computed in a polynomial number of steps, using exponentiation by squaring. This is essential in proving that each branch of the nondeterministic Algorithm 1 terminates in polynomial time, and also in the generalization of this reasoning to arbitrary flat CM with cycles labeled by octagonal constraints (Theorem 2).

As in most proofs in the rest of this paper, it is useful to prove the statement first for the simpler class of difference bounds relations, and use Theorem 1 (or a variant thereof) to relate difference bounds with octagonal relations. Since difference bounds

relations can be represented by weighted graphs (see Fig. 2 (b) for an example), we use a weighted graph to represent the n -th power of a relation $R \in \text{DB}$.

We write \mathcal{G}_R for the weighted graph $\mathcal{G}_{\sigma(R)}$, in which each vertex $1 \leq i \leq N$ is replaced by the variable x_i , and each vertex $N < i \leq 2N$ is replaced by x'_i . For a matrix $M \in \mathbb{Z}_{\infty}^{2N \times 2N}$, we denote its top-left, top-right, bottom-left and bottom-right $N \times N$ corners as $\blacksquare M, M^{\blacksquare}, \blacktriangleleft M$, and M_{\blacksquare} , respectively (see Fig. 2 (a) for an example).

Definition 7 Let $R \in \text{DB}_{\mathbf{x}}$ be a relation and $n \in \mathbb{N}_+$ a constant. Let $\mathcal{G}_R^n = \langle \bigcup_{k=0}^n \mathbf{x}^{(k)}, \rightarrow, w \rangle$ be a weighted graph, where $\mathbf{x}^{(k)} = \{x_i^{(k)} \mid 1 \leq i \leq N\}$ and for all $k \in [n]$:

- $x_i^{(k)} \xrightarrow{c} x_j^{(k)}$ if and only if $x_i \xrightarrow{c} x_j$ is an edge of \mathcal{G}_R ,
- $x_i^{(k)} \xrightarrow{c} x_j^{(k+1)}$ if and only if $x_i \xrightarrow{c} x'_j$ is an edge of \mathcal{G}_R ,
- $x_i^{(k+1)} \xrightarrow{c} x_j^{(k)}$ if and only if $x'_i \xrightarrow{c} x_j$ is an edge of \mathcal{G}_R ,
- $x_i^{(k+1)} \xrightarrow{c} x_j^{(k+1)}$ if and only if $x'_i \xrightarrow{c} x'_j$ is an edge of \mathcal{G}_R .

The constraint graph \mathcal{G}_R^n is said to be an *unfolding* of the constraint graph \mathcal{G}_R . The key observation relating the power R^n of R and the unfolding graph \mathcal{G}_R^n is the following: each difference constraint defining R^n is given by a minimal path between the extremal vertices (from the set $\mathbf{x}^{(0)} \cup \mathbf{x}^{(n)}$) in \mathcal{G}_R^n . Formally, for all $i, j \in \{1, \dots, N\}$, the power R^n is defined by the conjunction of the following constraints:

$$\begin{aligned} x_i - x_j &\leq (\blacksquare \sigma(R^n))_{ij} = \min w_{\mathcal{G}_R^n}(x_i^{(0)}, x_j^{(0)}) \\ x_i - x'_j &\leq (\sigma(R^n)^{\blacksquare})_{ij} = \min w_{\mathcal{G}_R^n}(x_i^{(0)}, x_j^{(n)}) \\ x'_i - x_j &\leq (\blacktriangleleft \sigma(R^n))_{ij} = \min w_{\mathcal{G}_R^n}(x_i^{(n)}, x_j^{(0)}) \\ x'_i - x'_j &\leq (\sigma(R^n)_{\blacksquare})_{ij} = \min w_{\mathcal{G}_R^n}(x_i^{(n)}, x_j^{(n)}) \end{aligned} \quad (5)$$

where $\min w_{\mathcal{G}_R^n}(x_i^{(p)}, x_j^{(q)}) = \min_{\ell \in \mathbb{N}} \{ \min w_{\mathcal{G}_R^n}(x_i^{(p)}, x_j^{(q)}, \ell) \}$, and $\min w_{\mathcal{G}_R^n}(x_i^{(p)}, x_j^{(q)}, \ell)$ is the minimal weight among all paths of length ℓ between $x_i^{(p)}$ and $x_j^{(q)}$ in \mathcal{G}_R^n , or ∞ , if no such path exists. When the length is not important, we denote a path between $x_i^{(p)}$ and $x_j^{(q)}$ as $x_i^{(p)} \xrightarrow{*} x_j^{(q)}$.

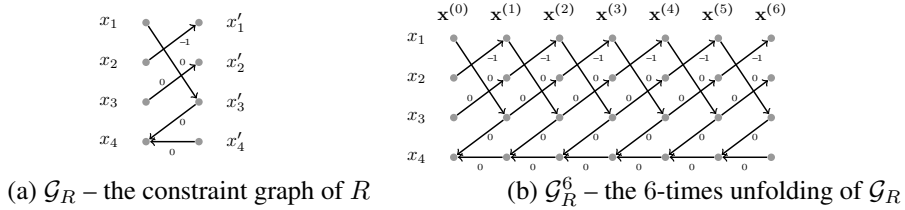


Figure 4: Constraint graphs for the DB relation $R \equiv x_2 - x'_1 \leq -1 \wedge x_3 - x'_2 \leq 0 \wedge x_1 - x'_3 \leq 0 \wedge x'_4 - x_4 \leq 0 \wedge x'_3 - x_4 \leq 0$.

Example 4 Consider the difference bounds relation R defined by the formula $\phi \equiv x_2 - x'_1 \leq -1 \wedge x_3 - x'_2 \leq 0 \wedge x_1 - x'_3 \leq 0 \wedge x'_4 - x_4 \leq 0 \wedge x'_3 - x_4 \leq 0$. Fig. 4 (a) shows the constraint graph \mathcal{G}_R and Fig. 4 (b) depicts the unfolding \mathcal{G}_R^6 of \mathcal{G}_R . ■

We are now ready to prove that the first condition of Theorem 2 holds for any difference bounds relation.

Lemma 3 *There exists a constant $d > 0$ such that, for every relation $R \in \text{DB}$, we have $|R^n| = \mathcal{O}((|R| \cdot \log n)^d)$, for all $n > 0$.*

Proof: We assume w.l.o.g. that R^n is consistent, otherwise $R^m = \emptyset$ for all $m \geq n$ and $|R^m| = |R^n|$. Then the unfolding \mathcal{G}_R^n does not contain cycles of negative weight, thus any minimal path in \mathcal{G}_R^n does not contain a cycle. Since \mathcal{G}_R^n has $(n+1) \cdot N$ nodes, any minimal path has weight at most $\mu \cdot (n+1) \cdot N$, where μ is the maximal value among the labels of \mathcal{G}_R . Since there are at most $4N^2$ such paths in the definition of R^n (5), we compute:

$$\begin{aligned} |R^n| &\leq 4N^2 \log(\mu \cdot (n+1) \cdot N) &\leq 4|R|^2 \log(|R| \cdot (n+1) \cdot |R|) \\ &\leq 4|R|^2 (2 \log |R| + \log n + 1) &\leq 16|R|^3 \cdot \log n \leq 16(|R| \cdot \log n)^3 \quad . \quad \square \end{aligned}$$

In order to establish a similar result for the more general class OCT, we must first relate the powers of an octagonal relation $R \in \text{OCT}$, defined by a constraint $\phi(\mathbf{x}, \mathbf{x}')$ with the powers of the difference bounds relation $\bar{R} \in \text{DB}$, defined by the constraint $\bar{\phi}(\mathbf{y}, \mathbf{y}')$, obtained from ϕ by doubling the number of variables. The following lemma establishes the needed correspondence between R^n and \bar{R}^n . We recall that a relation is said to be \star -consistent if each of its powers is consistent.

Lemma 4 *For any \star -consistent relation $R \in \text{OCT}$, the following holds, for any $n \geq 0$:*

$$\sigma(R^n)_{ij} = \min \left(\sigma(\bar{R}^n)_{ij}, \left\lfloor \frac{\sigma(\bar{R}^n)_{i\bar{i}}}{2} \right\rfloor + \left\lfloor \frac{\sigma(\bar{R}^n)_{\bar{j}j}}{2} \right\rfloor \right) .$$

Proof: [8, Lemma 4.30]. \square

The following lemma proves the validity of the first condition of Theorem 2, for every octagonal relation.

Lemma 5 *There exists a constant $d > 0$ such that, for every relation $R \in \text{OCT}$, we have $|R^n| = \mathcal{O}((|R| \cdot \log n)^d)$, for all $n > 0$.*

Proof: We assume w.l.o.g. that R is \star -consistent, otherwise there exists $n > 0$ such that $R^m = \emptyset$, thus $|R^m| = |R^n|$, for all $m \geq n$. By Lemma 4, we infer, for any $n > 0$:

$$|R^n| = \sum_{i,j=1}^{4N} \log(\sigma(R^n)_{ij}) \leq \sum_{i,j=1}^{4N} \log(\sigma(\bar{R}^n)_{ij}) = |\bar{R}^n| .$$

By Lemma 3, we have $|\bar{R}^n| = \mathcal{O}((|\bar{R}| \log n)^d)$ for a constant $d > 0$, not depending on R . Observe that $|R| \leq 2|\bar{R}|$, since every constraint of R is encoded by two constraints of \bar{R} . We obtain thus $|R^n| = \mathcal{O}((|R| \log n)^d)$, for any $n > 0$. \square

5 The Periodicity of Octagonal Relations

In this section we prove that the class of octagonal relations is periodic, that is, for each relation $R \in \text{OCT}$, the sequence of matrices $\{\sigma(R^k)\}_{k=0}^\infty$ is periodic. Moreover, we show that the prefix and the period of this sequence are of the order of $2^{\mathcal{O}(|R|^d)}$, for a constant $d > 0$ that does not depend on the choice of R . By Theorem 2, a consequence is that the class of problems $\text{REACHFLAT}(\text{OCT})$ is NP-complete.

The core of the proof is showing periodicity of difference bounds relations and establishing the upper bounds for the prefix and period of sequence $\{\sigma(R^k)\}_{k=0}^\infty$, where $R \in \text{DB}$. The main idea is that the coefficients of any matrix $\sigma(R^k)$ can be derived from the k -th power of a larger matrix \mathcal{M}_R , where the matrix product is defined using min as addition and $+$ as multiplication. Intuitively, a sequence $\{\mathcal{M}_R^k\}_{k=0}^\infty$ gives the minimal weights of the paths of length $k = 0, 1, \dots$ in a weighted graph \mathcal{A}_R , whose incidence matrix is \mathcal{M}_R . To obtain the simply exponential bounds on the period and prefix of a sequence $\{\sigma(R^k)\}_{k=0}^\infty$, we develop this periodicity result further, by exploiting the structure of the strongly connected components of \mathcal{A}_R .

In a nutshell, the set of paths from an unfolding \mathcal{G}_R^k of the constraint graph \mathcal{G}_R that represents the relation $R \in \text{DB}$ (Definition 7) is the language, consisting of words of length k , recognized by a weighted automaton \mathcal{A}_R (called *zigzag automaton* in the following). We use an idea of Comon and Jurski [10] that show that the set of minimal weights of these paths can be captured by a subset of paths, in which only a bounded number of direction changes may occur. Based on this fact, we define \mathcal{A}_R to recognize only these simple paths from \mathcal{G}_R^k , by considering a *saturated* relation R_{sat} , with the same periodic behavior as R . The simply exponential upper bound on the period of the sequence $\{\sigma(R^k)\}_{k=0}^\infty$ follows by a proof of the fact that, in each strongly connected component of \mathcal{A}_R , there is an elementary cycle of minimal weight/length ratio, whose length is of the order of $2^{\mathcal{O}(N)}$, where N is the number of variables from R .

5.1 Saturation of Difference Bounds Relations

We start by proving that the periodic behavior of a sequence of powers $\{R^k\}_{k=0}^\infty$ of a relation $R \in \text{DB}$ can be analyzed by considering a *saturated* version of R , denoted as R_{sat} . First we show that such a relation exists for each $R \in \text{DB}$ and that all powers of R , beyond a certain threshold, can be computed by a function taking as arguments powers of R_{sat} instead. As a consequence, R is periodic if R_{sat} is periodic and the period of R is bounded by the period of R_{sat} . The salient property of a saturated difference bounds relation is that, every power R_{sat}^k is defined by the weights of the minimal paths in the unfolding \mathcal{G}_R^k of the constraint graph defining R , with a bounded number of direction changes. This detail is instrumental in providing an accurate upper bound on the period of difference bounds relations.

Let $\mathbf{x} = \{x_1, \dots, x_N\}$ be a set of variables. We recall first the notion of *folded graph* introduced by Comon and Jurski [10]. Given a relation $R \in \text{DB}_{\mathbf{x}}$, we consider the weighted graph $\mathcal{G}_R^f = \langle \mathbf{x}, \rightarrow_f, w_f \rangle$ which has an edge $x_i \xrightarrow{\alpha}_f x_j$ for each edge $x_i \xrightarrow{\alpha} x_j$, $x_i \xrightarrow{\alpha} x'_j$, $x'_i \xrightarrow{\alpha} x_j$, or $x'_i \xrightarrow{\alpha} x'_j$ in \mathcal{G}_R . In other words, the folded graph \mathcal{G}_R^f is obtained from the weighted graph \mathcal{G}_R by merging all vertices x_i and x'_i , respectively.

Example 5 For example, Fig. 5 (b) shows the folded graph for the relation defined by the formula $x_2 - x'_1 \leq -1 \wedge x_3 - x'_2 \leq 0 \wedge x_1 - x'_3 \leq 0 \wedge x'_4 - x_4 \leq 0 \wedge x'_3 - x_4 \leq 0$, whose constraint graph is given in Fig. 5 (a). ■

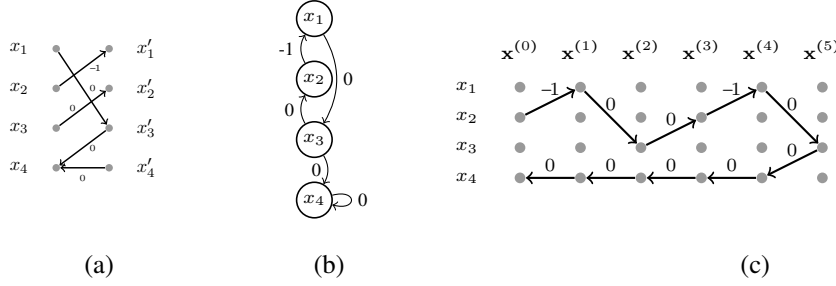


Figure 5: (a) Constraint graph \mathcal{G}_R , (b) folded graph \mathcal{G}_R^f and (c) zigzag paths in the \mathcal{G}_R^5 unfolding of the difference bounds relation R , defined by $x_2 - x'_1 \leq -1 \wedge x_3 - x'_2 \leq 0 \wedge x_1 - x'_3 \leq 0 \wedge x'_4 - x_4 \leq 0 \wedge x'_3 - x_4 \leq 0$.

The folded graph induces the following equivalence relation on indices of variables: $i \sim_R j$ iff x_i and x_j belong to the same strongly connected component of \mathcal{G}_R^f . For example, the equivalence classes of the \sim_R relation, induced by the folded graph in Fig. 5 (b), are $\{1, 2, 3\}$ and $\{4\}$.

The following *corner inequalities* are generalized triangle inequalities that occur in an unfolding of size two of the constraint graph \mathcal{G}_R of a relation $R \in \text{DB}$ (see Fig. 6).

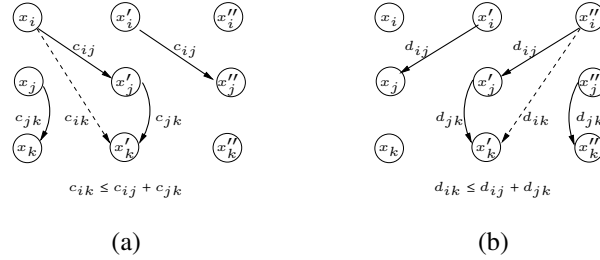


Figure 6: Corner inequalities

Definition 8 A relation $R \in \text{DB}_\mathbf{x}$ is saturated if, for all $1 \leq i, j, k \leq N$, such that $i \sim_R j \sim_R k$, the following hold:

$$\begin{aligned} (\sigma(R)^\blacksquare)_{ik} &\leq (\sigma(R)^\blacksquare)_{ij} + (\blacksquare\sigma(R))_{jk} \\ (\blacksquare\sigma(R))_{ik} &\leq (\blacksquare\sigma(R))_{ij} + (\sigma(R)_\blacksquare)_{jk} \end{aligned}$$

The first inequality is depicted in Fig. 6 (a) and the second case in Fig. 6 (b). The importance of the above definition lies in the fact that the powers R^n of a saturated relation $R \in \text{DB}$ can be defined using only a subset of the minimal paths between the

(extremal) vertices in the unfolding \mathcal{G}_R^n of the constraint graph of R (see the constraints (5) for a definition of R^n using the minimal paths of \mathcal{G}_R^n). Essentially, these are the minimal paths that, moreover, do not change direction while traversing variables from the same equivalence class of the \sim_R relation.

Definition 9 Let \mathcal{G}_R^n be the n -th unfolding of the constraint graph \mathcal{G}_R of a relation $R \in \text{DB}$. An edge $x_i^{(p)} \rightarrow x_j^{(q)}$ of \mathcal{G}_R^n is forward (backward) if $q = p + 1$ ($q = p - 1$). A path is forward (backward) if it consists only of forward (backward) edges. A path π is saturated if for each forward (backward) subpath $\rho : x_i^{(p)} \xrightarrow{*} x_j^{(q)}$ of π , of length $|\rho| > 1$, we have $i \sim_R j$.

We show that the n -th power of a saturated relation can be defined by considering only the saturated minimal paths in the n -th unfolding of its constraint graph. For instance, the path in Fig. 5 (c) is saturated.

Lemma 6 Given a saturated relation $R \in \text{DB}$, for each minimal path $\pi : x_i^{(p)} \xrightarrow{*} x_j^{(q)}$ in the unfolding \mathcal{G}_R^n of its constraint graph \mathcal{G}_R , there exists a saturated path $\pi_{\text{sat}} : x_i^{(p)} \xrightarrow{*} x_j^{(q)}$ such that $w(\pi) = w(\pi_{\text{sat}})$.

Proof: We define a sequence π_0, π_1, \dots of paths such that $\pi_0 = \pi$ and for each $t \geq 0$, π_{t+1} is obtained from π_t as follows. Because R is saturated:

- for every two adjacent edges $x_i^{(p)} \xrightarrow{\alpha} x_j^{(p+1)} \xrightarrow{\beta} x_\ell^{(p+1)}$ on π_t , where $i \sim_R \ell$, there exists an edge $x_i^{(p)} \xrightarrow{\gamma} x_\ell^{(p+1)}$, where $\gamma \leq \alpha + \beta$, and
- for every two adjacent edges $x_i^{(p+1)} \xrightarrow{\alpha} x_j^{(p)} \xrightarrow{\beta} x_\ell^{(p)}$, where $i \sim_R \ell$, there exists an edge $x_i^{(p+1)} \xrightarrow{\gamma} x_\ell^{(p)}$, where $\gamma \leq \alpha + \beta$.

π_{t+1} is obtained by replacing all pairs $x_i^{(p)} \xrightarrow{\alpha} x_j^{(p+1)} \xrightarrow{\beta} x_\ell^{(p+1)}$ with $x_i^{(p)} \xrightarrow{\gamma} x_\ell^{(p+1)}$ and all pairs $x_i^{(p+1)} \xrightarrow{\alpha} x_j^{(p)} \xrightarrow{\beta} x_\ell^{(p)}$, where $i \sim_R \ell$, with the edges $x_i^{(p)} \xrightarrow{\gamma} x_\ell^{(p+1)}$ and $x_i^{(p)} \xrightarrow{\gamma} x_\ell^{(p)}$, respectively. Clearly, π_{t+1} is a path in \mathcal{G}_R^n and, moreover, we have that $|\pi_{t+1}| < |\pi_t|$ and $w(\pi_{t+1}) \leq w(\pi_t)$. The sequence is finite, because $|\pi_t|$ decreases at each step, and the last path in the sequence is saturated. If, moreover, π is minimal, all paths in the sequence are minimal as well. \square

In the rest of this section, we prove that for each difference bounds relation R it is possible to find a saturated relation R_{sat} with the same periodic behavior as R . The problem of periodicity and the evaluation of the upper bounds for the prefix and period of R is carried out considering R_{sat} instead. This step is instrumental in proving a simply exponential upper bound on the period of R .

At this point, it is useful to distinguish between relations that are $*$ -consistent ($R^k \neq \emptyset$ for all $k \geq 0$) and the ones that are not. In the latter case, the period of the power sequence $\{R^k\}_{k=0}^\infty$ is 1 and the prefix is bounded by the cut-off result below.

Lemma 7 For any relation $R \in \text{DB}_x$, the following hold:

1. R is $*$ -consistent only if for every $n \in \mathbb{N}_+$ and $1 \leq i < j \leq N$, such that $i \sim_R j$, for any paths $\pi_i : x_i^{(0)} \xrightarrow{*} x_i^{(k)}$ and $\pi_j : x_j^{(k)} \xrightarrow{*} x_j^{(0)}$ in \mathcal{G}_R^n with $0 < k \leq n$, we have $w(\pi_i) + w(\pi_j) \geq 0$.

2. R is not $*$ -consistent only if $R^n = \emptyset$, for all $n \geq 6N^7 \cdot \mu(\mathcal{G}_R)$.

Proof: In the following, the *travel* of a path $\pi : x_{i_1}^{(n_1)} \rightarrow \dots \rightarrow x_{i_k}^{(n_k)}$ in the unfolding \mathcal{G}_R^n (Definition 7) of the constraint graph \mathcal{G}_R is defined by $\tau(\pi) = \max_{j=1}^k \{n_j\} - \min_{j=1}^{k-1} \{n_j\}$.

(1) By contradiction, suppose that R is $*$ -consistent and let $\pi_i : x_i^{(0)} \xrightarrow{*} x_i^{(k)}$ and $\pi_j : x_j^{(k)} \xrightarrow{*} x_j^{(0)}$ in \mathcal{G}_R^n , such that $i \sim j$ and $w(\pi_i) + w(\pi_j) < 0$. Because $i \sim j$, there exist paths $\zeta : x_j^{(d)} \rightarrow \dots \rightarrow x_i^{(d+p)}$ and $\xi : x_i^{(d)} \rightarrow \dots \rightarrow x_j^{(d+q)}$ in \mathcal{G}_R^n , for some $-N \leq p, q \leq N$ and $d \geq \max(\text{abs}(p), \text{abs}(q))$. For any $m \in \mathbb{N}$ we build the following path:

$$\zeta \cdot \pi_i^m \cdot \xi = x_j^{(d)} \rightarrow \dots \rightarrow x_i^{(d+p)} \rightarrow \dots \rightarrow x_i^{(d+p+mk)} \rightarrow \dots \rightarrow x_j^{(d+p+q+mk)}.$$

For $m \geq \lceil \frac{-(p+q)}{k} \rceil$, we have $p+q+mk > 0$, i.e. the above path has a positive travel. Let us now repeat this path k times, and concatenate it with the path

$$\pi_j^{p+q+km} : x_j^{(k(d+p+q+mk))} \rightarrow \dots \rightarrow x_j^{(d)}.$$

We obtain the cycle $(\zeta \cdot \pi_i^m \cdot \xi)^k \cdot \pi_j^{p+q+km}$, starting and ending in $x_j^{(d)}$. Observe that the unfolding \mathcal{G}_R^n has to be sufficiently large to accomodate this cycle. Since n can be taken arbitrarily large, this is not a restriction. The weight of this cycle is:

$$k \cdot (w(\zeta) + w(\xi) + m \cdot w(\pi_i)) + (p+q+km) \cdot w(\pi_j) = km \cdot (w(\pi_i) + w(\pi_j)) + k \cdot (w(\zeta) + w(\xi)) + (p+q) \cdot w(\pi_j).$$

For a sufficiently large $m \geq \lceil \frac{-(p+q)}{k} \rceil$, the weight of this cycle is negative, which contradicts with the assumption that R is $*$ -consistent. It follows that $w(\pi_i) + w(\pi_j) \geq 0$.

(2) If R is not $*$ -consistent, there exists $n \in \mathbb{N}_+$ such that $R^n = \emptyset$, thus there exists a cycle γ of negative weight in \mathcal{G}_R^n . Let $\gamma : x_{i_1}^{(n_1)} \rightarrow \dots \rightarrow x_{i_{k-1}}^{(n_{k-1})} \rightarrow x_{i_1}^{(n_1)}$ be a negative cycle of minimal travel in \mathcal{G}_R^n . If $\tau(\gamma) \leq N^2$, the cycle is present also in $\mathcal{G}_R^{N^2}$. Hence $R^n = \emptyset$, for every $n \geq N^2$. We consider thus that $\tau(\gamma) > N^2$. By the pigeonhole principle, there exist a pair of variables x_i, x_j such that the pairs of vertices $x_i^{(k)}, x_j^{(k)}$ and $x_i^{(\ell)}, x_j^{(\ell)}$ occur on the cycle, for some $0 \leq k < \ell \leq n$. Thus there exist paths $\pi_i : x_i^{(k)} \rightarrow \dots \rightarrow x_i^{(\ell)}$ and $\pi_j : x_j^{(\ell)} \rightarrow \dots \rightarrow x_j^{(k)}$. Observe that we can always chose the pairs $x_i^{(k)}, x_j^{(k)}$ and $x_i^{(\ell)}, x_j^{(\ell)}$ such that $\ell - k \leq N^2$. Clearly, we have $i \sim j$, since x_i and x_j occur on the cycle γ . Suppose now that $w(\pi_i) + w(\pi_j) \geq 0$. Since $\ell - k > 0$, we obtain a cycle of smaller travel, by eliminating π_i and π_j from γ , and concatenating the path $x_j^{(k)} \rightarrow \dots \rightarrow x_i^{(k)}$ with $x_i^{(\ell-k)} \rightarrow \dots \rightarrow x_j^{(\ell-k)}$. But then we obtain a cycle γ' of weight $w(\gamma') \leq w(\gamma) < 0$ and travel $\tau(\gamma') < \tau(\gamma)$. This would contradict the assumption that $\tau(\gamma)$ is the minimal travel of all negative weight cycles in \mathcal{G}_R^n . Hence it must be the case that $w(\pi_i) + w(\pi_j) < 0$.

Given π_i and π_j such that $i \sim j$ and $w(\pi_i) + w(\pi_j) < 0$, we apply the construction of point (1) to obtain a cycle $(\zeta \cdot \pi_i^m \cdot \xi)^{\ell-k} \cdot \pi_j^{p+q+(\ell-k)m}$ of weight:

$$m \cdot (\ell - k) \cdot (w(\pi_i) + w(\pi_j)) + (\ell - k) \cdot (w(\zeta) + w(\xi)) + (p+q) \cdot w(\pi_j)$$

where ζ , ξ , p and q are the ones from point (1). To obtain a negative cycle, it is thus sufficient to chose $m = (\ell - k) \cdot (w(\zeta) + w(\xi)) + (p + q) \cdot w(\pi_j)$. Since ζ and π are elementary paths, we have $w(\zeta) + w(\xi) \leq 2N \cdot \mu(\mathcal{G}_R)$ and $p + q \leq 2N$. Moreover, since $\ell - k \leq N^2$, we have $w(\pi_j) \leq N^2 \cdot \mu(\mathcal{G}_R)$. Then it is sufficient to take $m = 4N^3 \cdot \mu(\mathcal{G}_R)$. The travel of the cycle thus constructed is at most:

$$\begin{aligned} (\ell - k) \cdot (p + q + (\ell - k) \cdot m) &\leq N^2 \cdot (2N + N^2 \cdot 4N^3 \cdot \mu(\mathcal{G}_R)) \\ &\leq 4N^7 \cdot \mu(\mathcal{G}_R) + 2N^3 \leq 6N^7 \cdot \mu(\mathcal{G}_R) . \end{aligned}$$

The last inequality is obtained from the following observation: since $w(\pi_i) + w(\pi_j) < 0$, there must exist at least an edge of non-zero weight in \mathcal{G}_R , hence $\mu(\mathcal{G}_R) > 0$. \square

In the light of the previous lemma, we consider, from now on, that $R \in \text{DB}_x$ is a $*$ -consistent relation. We are now ready to define a saturated relation R_{sat} , which is periodic if and only if R is periodic, in which case the prefix and the period of R_{sat} bound the prefix and the period of R , respectively.

Let $\phi(\mathbf{x}, \mathbf{x}')$ be any difference bounds constraint (Definition 3) that defines R and $\tilde{\phi}$ be the conjunction of all atomic propositions from $\text{Atom}(\phi)$ involving \sim_R -equivalent variables. We define the following sequence of formulae:

$$\begin{aligned} \psi_0(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z}) &\equiv \tilde{\phi}(\mathbf{y}, \mathbf{x}) \wedge \tilde{\phi}(\mathbf{x}, \mathbf{x}') \wedge \tilde{\phi}(\mathbf{x}', \mathbf{z}) \\ \psi_{n+1}(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z}) &\equiv \exists \mathbf{y}' \exists \mathbf{z}' . \tilde{\phi}(\mathbf{y}, \mathbf{y}') \wedge \psi_n(\mathbf{y}', \mathbf{x}, \mathbf{x}', \mathbf{z}') \wedge \tilde{\phi}(\mathbf{z}', \mathbf{z}), \quad \text{for all } n \in \mathbb{N} \end{aligned}$$

where $\mathbf{y} = \{y_1, \dots, y_N\}$, $\mathbf{z} = \{z_1, \dots, z_N\}$. The following lemma shows that the sequence $\{\exists \mathbf{y} \exists \mathbf{z} . \psi_n\}_{n=0}^{\infty}$ converges in at most N^2 steps and its limit defines a saturated difference bounds relation.

Lemma 8 *For every $n \geq N^2$, we have $\exists \mathbf{y} \exists \mathbf{z} . \psi_{n+1}(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z}) \Leftrightarrow \exists \mathbf{y} \exists \mathbf{z} . \psi_n(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z})$ and the formula $\exists \mathbf{y} \exists \mathbf{z} . \psi_{N^2}(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z}) \wedge \phi$ defines a saturated relation R_{sat} .*

Proof: “ \Rightarrow ” We have, for all $n \geq 0$:

$$\begin{aligned} \psi_{n+1}(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z}) &\Leftrightarrow \exists \mathbf{y}' \exists \mathbf{z}' . \tilde{\phi}(\mathbf{y}, \mathbf{y}') \wedge \psi_n(\mathbf{y}', \mathbf{x}, \mathbf{x}', \mathbf{z}') \wedge \tilde{\phi}(\mathbf{z}', \mathbf{z}) \\ &\Rightarrow \exists \mathbf{y} \exists \mathbf{z} . \psi_n(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z}) . \end{aligned}$$

“ \Leftarrow ” For each $n \geq 0$, $\exists \mathbf{y} \exists \mathbf{z} . \psi_n$ is equivalent to a difference bounds constraint, obtained by eliminating the existential quantifiers and let M_n be the DBM of canonical the quantifier-free difference bounds constraint that is equivalent to $\exists \mathbf{y} \exists \mathbf{z} . \psi_n$. Suppose, by contradiction, that there exists $n \geq N^2$ such that $\exists \mathbf{y} \exists \mathbf{z} . \psi_n \not\equiv \exists \mathbf{y} \exists \mathbf{z} . \psi_{n+1}$. Then there exist $k, \ell \in \{1, \dots, 2N\}$ such that $(M_n^*)_{k\ell} > (M_{n+1}^*)_{k\ell}$. There are four cases, namely $k \leq N, k > N$ and $\ell \leq N, \ell > N$. We prove the case $k, \ell \leq N$, the other three cases being symmetric.

Because $(M_{n+1}^*)_{k\ell} < \infty$, there exists a path $\pi : x_k^{(n)} \rightarrow \dots \rightarrow x_\ell^{(n)}$ of weight $w(\pi) = (M_{n+1}^*)_{k\ell}$ in the unfolding graph $\mathcal{G}_{\tilde{\phi}}^{2(n+1)+1}$. But the only paths in $\mathcal{G}_{\tilde{\phi}}^{2(n+1)+1}$ are among \sim_R -equivalent variables, by the definition of $\tilde{\phi}$, thus it must be the case that $k \sim_R \ell$. Moreover, π is not a path in $\mathcal{G}_{\tilde{\phi}}^{2n+1}$, or else we would have $w(\pi) \geq (M_n^*)_{k\ell}$, hence $(M_{n+1}^*)_{k\ell} \geq (M_n^*)_{k\ell}$, which contradicts our assumption. Since $n + 1 > N^2$ and

there are at most N^2 pairs of variables in \mathbf{x} , by the pigeonhole principle there exists a pair i, j and two positions $0 \leq n_1 < n_2 \leq 2(n+1)+1$ such that π can be factorized as:

$$\begin{array}{ccccccc} x_k^{(n)} & \rightarrow & \dots & \rightarrow & x_i^{(n_1)} & \xrightarrow{\xi} & x_i^{(n_2)} \\ & & & & & & \vdots \\ & & & & & & \downarrow \\ x_\ell^{(n)} & \leftarrow & \dots & \leftarrow & x_j^{(n_1)} & \xleftarrow{\zeta} & x_j^{(n_2)} \end{array}$$

Let π' denote the path obtained from π by replacing the segment $\xi : x_i^{(n_1)} \rightarrow \dots \rightarrow x_i^{(n_2)}$ with the segment $\zeta : x_i^{(n_2)} \rightarrow \dots \rightarrow x_j^{(n_2)}$, in which each position is shifted by $n_1 - n_2 > 0$. Hence π' is a path between $x_k^{(n)}$ and $x_\ell^{(n)}$ in \mathcal{G}_R^{2n+1} , thus $w(\pi') \geq (M_n^*)_{k\ell} > w(\pi)$. Since $w(\pi) = w(\pi') + w(\xi) + w(\zeta)$, we obtain that $w(\xi) + w(\zeta) < 0$, and because $k \sim_R \ell$, we have that $i \sim_R j$ as well. By Lemma 7 (1), this contradicts the assumption that R is \star -consistent.

For the second point, $\exists \mathbf{y} \exists \mathbf{z} . \psi_R^{N^2}(\mathbf{y}, \mathbf{x}, \mathbf{x}', \mathbf{z}) \wedge \phi_R \Rightarrow \phi_R$, hence $R_{\text{sat}} \subseteq R$. Now suppose, by contradiction, that R^s is not saturated. Then there exist three indices $i \sim_R j \sim_R k$ that violate one of the corner inequalities from Def. 8. Assume w.l.o.g that $(\sigma(R)^\blacksquare)_{ik} > (\sigma(R)^\blacksquare)_{ij} + (\sigma(R)^\blacksquare)_{jk}$ the other case being symmetric. Then we obtain that $\exists \mathbf{y} \exists \mathbf{z} . \psi_R^{N^2} \not\Rightarrow \exists \mathbf{y} \exists \mathbf{z} . \psi_R^{N^2+1}$, contradicting the first point of the Lemma. \square

Below we relate the powers of R with those of the relation R_{sat} , defined in the statement of Lemma 8, in the sense that a proof of periodicity for R_{sat} constitutes a proof for the periodicity of R . Moreover, the prefix and period of R_{sat} are used to bound the prefix and period of R , respectively. We recall that, for a difference bounds relation $R \in \text{DB}$, $\sigma(R)$ is the closed DBM that defines R and $\mu(\sigma(R))$ is the maximum between the absolute values of the coefficients of $\sigma(R)$ and 1.

Lemma 9 *Given a \star -consistent relation $R \in \text{DB}_{\mathbf{x}}$, where $\mathbf{x} = \{x_1, \dots, x_N\}$, for every $n \geq 2N^2$, we have $R^n = R^{N^2} \circ R_{\text{sat}}^{n-2N^2} \circ R^{N^2}$. Moreover, R is periodic with prefix b and period c if R_{sat} is periodic with prefix b_{sat} and period c_{sat} , where:*

- $b = b_{\text{sat}} + \mathcal{O}(2^{N \log N}) \cdot \max(\mu(\sigma(R^{N^2})), \max_{0 \leq i < c_{\text{sat}}} \mu(\sigma(R_{\text{sat}}^{b_{\text{sat}}+i})))$ and
- c divides c_{sat} .

Proof: Since $R_{\text{sat}} \subseteq R$, by Lemma 8, we obtain that $R^n \supseteq R^{N^2} \circ R_{\text{sat}}^{n-2N^2} \circ R^{N^2}$, for each $n \geq 2N^2$. The dual inclusion follows by noticing that, for any difference bounds constraint ϕ that defines R , we have $\phi(\mathbf{x}, \mathbf{x}') \Rightarrow \tilde{\phi}(\mathbf{x}, \mathbf{x}')$, because $\tilde{\phi}$ is obtained by dropping several atomic propositions from ϕ . Also, since R is \star -consistent, it must be the case that R_{sat} is \star -consistent as well.

Assume that R_{sat} is periodic with prefix b_{sat} and period c_{sat} , thus the sequence of matrices $\{\sigma(R_{\text{sat}}^k)\}_{k=0}^\infty$ is periodic. By Lemma 1, there exist matrices $\Lambda_i \in \mathbb{Z}_{\infty}^{2N \times 2N}$, such that $\sigma(R_{\text{sat}}^{b_{\text{sat}}+k c_{\text{sat}}+i}) = \sigma(R_{\text{sat}}^{b_{\text{sat}}+i}) + k \cdot \Lambda_i$, for all $k \geq 0$ and $i \in [c_{\text{sat}}]$. Since $R^{b_{\text{sat}}+2N^2+\ell} = R^{N^2} \circ R^{b_{\text{sat}}+\ell} \circ R^{N^2}$ for all $\ell \geq 0$, it is sufficient to show that the sequence $\{\sigma(R^{b_{\text{sat}}+k c_{\text{sat}}+i+2N^2})\}_{k=0}^\infty$ is periodic, for each $i \in [c_{\text{sat}}]$. To this end, let us

observe that, for each $i \in [c_{\text{sat}}]$, the matrix $\sigma(R^{b_{\text{sat}}+kc_{\text{sat}}+i+2N^2}) \in \mathbb{Z}_{\infty}[k]^{2N \times 2N}$ is the incidence matrix of the labeled graph $\mathcal{G}^i = (\cup_{j=0}^3 \mathbf{x}^{(j)}, \rightarrow^i, w^i)$, defined as follows:

- $\mathcal{G}^i \downarrow_{\mathbf{x}^{(0)} \cup \mathbf{x}^{(1)}}$ is $\mathcal{G}_{R^{N^2}}$ with $\mathbf{x}^{(0)}$ and $\mathbf{x}^{(1)}$ replacing \mathbf{x} and \mathbf{x}' , respectively.
- $\mathcal{G}^i \downarrow_{\mathbf{x}^{(2)} \cup \mathbf{x}^{(3)}}$ is $\mathcal{G}_{R^{N^2}}$ with $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(3)}$ replacing \mathbf{x} and \mathbf{x}' , respectively.
- $\mathcal{G}^i \downarrow_{\mathbf{x}^{(1)} \cup \mathbf{x}^{(2)}}$ is defined by the following edges labeled by univariate linear terms with variable k , for each $s, t = 1, \dots, N$:

$$\begin{array}{ccc} x_s^{(1)} & \xrightarrow{\sigma(R^{b_{\text{sat}}+i})_{s,t} + k \cdot (\Lambda_i)_{s,t}} & x_t^{(1)} \\ x_s^{(1)} & \xrightarrow{\sigma(R^{b_{\text{sat}}+i})_{s,t+N} + k \cdot (\Lambda_i)_{s,t+N}} & x_t^{(2)} \end{array} \quad \begin{array}{ccc} x_s^{(2)} & \xrightarrow{\sigma(R^{b_{\text{sat}}+i})_{s+N,t} + k \cdot (\Lambda_i)_{s+N,t}} & x_t^{(1)} \\ x_s^{(2)} & \xrightarrow{\sigma(R^{b_{\text{sat}}+i})_{s+N,t+N} + k \cdot (\Lambda_i)_{s+N,t+N}} & x_t^{(2)} \end{array}$$

In other words, the middle graph $\mathcal{G}^i \downarrow_{\mathbf{x}^{(1)} \cup \mathbf{x}^{(2)}}$ corresponds to the matrix $\sigma(R^{b_{\text{sat}}+i}) + k \cdot \Lambda_i \in \mathbb{Z}_{\infty}[k]^{2N \times 2N}$, whereas the extremities $\mathcal{G}^i \downarrow_{\mathbf{x}^{(0)} \cup \mathbf{x}^{(1)}}$ and $\mathcal{G}^i \downarrow_{\mathbf{x}^{(1)} \cup \mathbf{x}^{(2)}}$ are constraint graphs defining the relation R^{N^2} . Clearly, for each $i \in [c_{\text{sat}}]$ and $k \geq 0$, the coefficients of the matrix $\sigma(R^{b_{\text{sat}}+kc_{\text{sat}}+i+2N^2})$ are given by the weights of the minimal paths from and to the vertices in the set $\mathbf{x}^{(0)} \cup \mathbf{x}^{(3)}$ in \mathcal{G}^i . Since R is $*$ -consistent, no cycle of negative weight can be found in \mathcal{G}^i , for any $k \geq 0$ and $i \in [c_{\text{sat}}]$. Thus the minimal paths in \mathcal{G}^i are necessarily elementary, thus of length at most $4N$, which is the number of vertices in \mathcal{G}^i . Consequently, there exist at most $(4N)^N = \mathcal{O}(2^{N \log N})$ such paths, and $(\sigma(R^{b_{\text{sat}}+kc_{\text{sat}}+i+2N^2}))_{st} = \min(w_{st}^{i,1}(k), \dots, w_{st}^{i,L_{st}}(k))$, for all $s, t \in \{1, \dots, 2N\}$, where $w_{st}^{i,1}(k), \dots, w_{st}^{i,L_{st}}(k)$ are the univariate linear terms labeling the elementary paths from \mathcal{G}^i of the form:

- $x_s^{(0)} \rightarrow^* x_t^{(0)}$ if $1 \leq s, t \leq N$,
- $x_s^{(0)} \rightarrow^* x_{t-N}^{(3)}$ if $1 \leq s \leq N$ and $N < t \leq 2N$,
- $x_{s-N}^{(3)} \rightarrow^* x_t^{(0)}$ if $N < s \leq 2N$ and $1 \leq t \leq N$,
- $x_{s-N}^{(3)} \rightarrow^* x_{t-N}^{(3)}$ if $N < s, t \leq 2N$,

and $L_{st} = \mathcal{O}(2^{N \log N})$ is the number of such paths. Moreover, because each term $w_{st}^{i,j}(k)$ denotes a path of length at most $4N$ in \mathcal{G}^i , we have that:

$$\text{abs}(w_{st}^{i,j}(0)) \leq 4N \cdot \max(\mu(\sigma(R^{N^2})), \max_{0 \leq i < c_{\text{sat}}} \mu(\sigma(R_{\text{sat}}^{b_{\text{sat}}+i}))) \quad (6)$$

The upper bound on the period of $\{\sigma(R^k)\}_{k=0}^{\infty}$ can be found by the following argument. Observe that $\{(\sigma(R_{\text{sat}}^{b_{\text{sat}}+kc_{\text{sat}}+i}))_{st}\}_{k=0}^{\infty}$ is an arithmetic progression, for all $1 \leq s, t \leq 2N$, thus each sequence $\{w_{st}^{\ell}(k)\}_{k=0}^{\infty}$ is an arithmetic progression. By Lemma 19, the sequence $\{\min(w_{st}^1(k), \dots, w_{st}^{L_{st}}(k))\}_{k=0}^{\infty}$ has period 1 and by Lemma 1, the sequence of matrices $\{\sigma(R^{b_{\text{sat}}+kc_{\text{sat}}+i+2N^2})\}_{k=0}^{\infty}$ has period 1, for each $i \in [c_{\text{sat}}]$. The period of the sequence $\{\sigma(R^k)\}_{k=0}^{\infty}$ is thus a divisor of c_{sat} . Considering that $L_{st} = \mathcal{O}(2^{N \log N})$, the upper bound on the prefix of R is obtained as follows:

$$\begin{aligned} b &\leq b_{\text{sat}} + (2N)^2 \cdot \max_{\substack{0 \leq i < c_{\text{sat}} \\ 1 \leq s, t \leq 2N}} \sum_{j=1}^{L_{st}} \text{abs}(w_{st}^{i,j}(0)) && \text{by Lemma 19} \\ &= b_{\text{sat}} + \mathcal{O}(2^{N \log N}) \cdot \max(\mu(\sigma(R^{N^2})), \max_{0 \leq i < c_{\text{sat}}} \mu(\sigma(R_{\text{sat}}^{b_{\text{sat}}+i}))) && \text{by (6)} \end{aligned}$$

□

We have reduced the problem of proving the periodicity of an arbitrary difference bounds relation R to proving the periodicity of a saturated difference bounds relation R_{sat} . In the next two sections, we show that any saturated relation R is periodic (5.2) and, moreover, that the prefix and the period of the sequence $\{\sigma(R^k)\}_{k=0}^{\infty}$ are of the order of $2^{O(|R|)}$ (5.4). The previous lemma generalizes these bounds to arbitrary difference bounds relations. Finally, section 5.5 extends these bounds to octagonal relations, concluding the proof of NP-completeness for the class of problems REACHFLATOCT.

5.2 Zigzag Automata

In this section we define *zigzag automata*, which are an important tool for reasoning about the powers of difference bounds relations. Consider an unfolding \mathcal{G}_R^n of the constraint graph \mathcal{G}_R , for some $n > 0$. We recall the constraints (5) which define the power R^n , using the minimal paths in \mathcal{G}_R^n between the vertices in the set $\mathbf{x}^{(0)} \cup \mathbf{x}^{(n)}$. Each such path can be seen as a word over the finite alphabet of subgraphs of \mathcal{G}_R , and the set of paths between two distinguished vertices is the language of a finite (weighted) automaton, called *zigzag automaton* [7]. Intuitively, a zigzag automaton reads at step i , all edges between $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(i+1)}$ simultaneously. The weight of a transition fired by the zigzag automaton is the sum of the weights of these edges. Each run of length n in a zigzag automaton recognizes a word consisting of a single path between two extremal vertices in \mathcal{G}_R^n , from the set $\mathbf{x}^{(0)} \cup \mathbf{x}^{(n)}$. Since we are interested in the minimal paths that occur in the constraints (5), we aim at computing the minimal weight among all runs of length n , as a function of n .

Formally, a *weighted automaton*³ [34] is a tuple $A = \langle \Sigma, \omega, Q, I, F, \Delta \rangle$, where Σ is a finite alphabet, $\omega : \Sigma \rightarrow \mathbb{Z}$ is a function associating integer weights to alphabet symbols, Q, I, F are the set of states, initial and final states, respectively, and $\Delta \subseteq Q \times \Sigma \times Q$ is a transition relation. The weight of a non-empty word $w = \sigma_1 \dots \sigma_n \in \Sigma^+$ is defined as $\omega(w) = \sum_{i=1}^n \omega(\sigma_i)$ and the weight of the empty word is $\omega(\varepsilon) = 0$. When A is clear from the context, we denote by $q \xrightarrow{\sigma} q'$ the fact that $(q, \sigma, q') \in \Delta$. A *run* of A is a sequence $q_0 \xrightarrow{\sigma_0} q_1 \xrightarrow{\sigma_1} \dots \xrightarrow{\sigma_{n-1}} q_n$, denoted $q_0 \xrightarrow{\sigma_0 \dots \sigma_{n-1}} q_n$. A state $q \in Q$ is *reachable* if there exists a run from an initial state to it, and *co-reachable* if there exists a run from it to a final state. A word $w \in \Sigma^*$ is accepted by A if there exists a run $q_0 \xrightarrow{w} q_n$ such that $q_0 \in I$ and $q_n \in F$. We denote by $\mathcal{L}(A)$ the set of words accepted by A , i.e. the *language* of A . Moreover, we define the function $\text{minw}_A(n) = \min \{\omega(w) \mid w \in \mathcal{L}(A), |w| = n\}$ yielding, for each $n \in \mathbb{N}$, the minimal weight among all words of length n recognized by A , or ∞ if no such word exists.

Given a difference bounds relation $R \in \text{DB}_{\mathbf{x}}$, where $\mathbf{x} = \{x_1, \dots, x_N\}$ is a set of variables, let $\mathcal{G}_R = \langle \mathbf{x} \cup \mathbf{x}', \rightarrow, w \rangle$ be the constraint graph that defines R . The alphabet Σ_R is the set of all subgraphs of \mathcal{G}_R such that (i) the in-degree and out-degree of each node are at most 1, and (ii) the difference between the number of edges from \mathbf{x} to \mathbf{x}' and the number of edges from \mathbf{x}' to \mathbf{x} is either $-1, 0$ or 1 . The weight of a graph symbol $\mathcal{G} \in \Sigma_R$ is the sum of the weights that occur on its edges $\omega(\mathcal{G}) = \sum_{x \xrightarrow{c} y} c$.

³We adopt a simplified version of the classical definition [34] that is sufficient for our purposes.

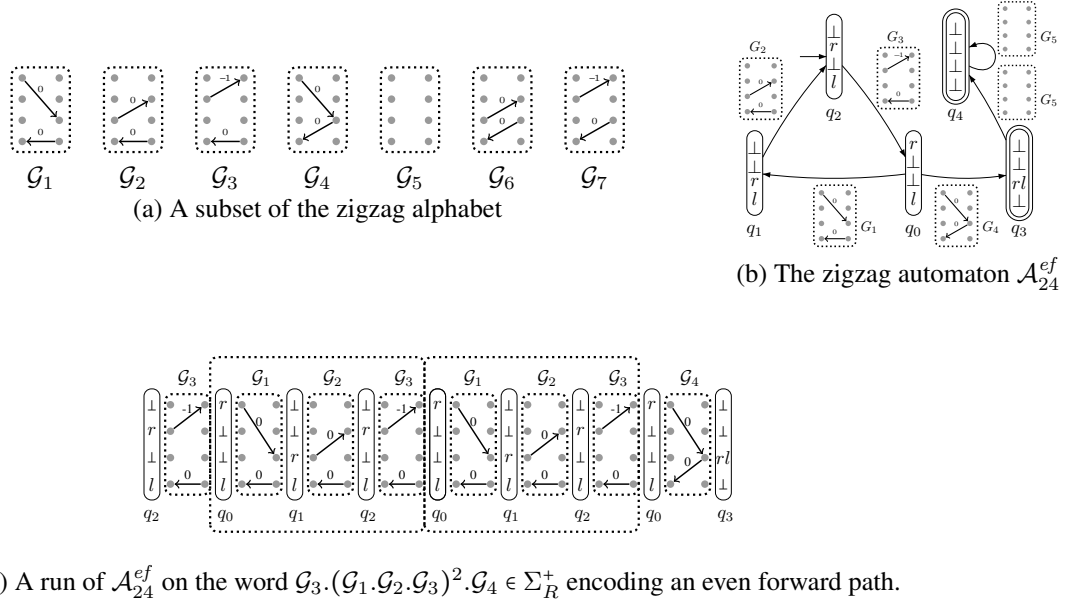


Figure 7: Zigzag automaton for the relation R defined by $x_2 - x'_1 \leq -1 \wedge x_3 - x'_2 \leq 0 \wedge x_1 - x'_3 \leq 0 \wedge x'_4 - x_4 \leq 0 \wedge x'_3 - x_4 \leq 0$.

Example 6 Fig. 7 (a) shows a subset of the zigzag alphabet Σ_R for the difference bounds relation $R \Leftrightarrow x_2 - x'_1 \leq -1 \wedge x_3 - x'_2 \leq 0 \wedge x_1 - x'_3 \leq 0 \wedge x'_4 - x_4 \leq 0 \wedge x'_3 - x_4 \leq 0$ from Ex. 4. The weights of the symbols in the word are $\omega(\mathcal{G}_1) = \omega(\mathcal{G}_2) = \omega(\mathcal{G}_4) = 0$, $\omega(\mathcal{G}_3) = -1$. Fig. 4 (c) shows a path $x_2^{(0)} \rightarrow \dots \rightarrow x_4^{(0)}$ from the unfolding graph \mathcal{G}_R^8 , encoded by the word $\mathcal{G}_3.(\mathcal{G}_1.\mathcal{G}_2.\mathcal{G}_3)^2.\mathcal{G}_4$. ■

The set of states of the zigzag automaton is $Q = \{\ell, r, \ell r, r\ell, \perp\}^N$, i.e. the set of N -tuples of symbols $\ell, r, \ell r, r\ell$ and \perp . Intuitively, these symbols capture the direction of the incoming and outgoing edges of the alphabet symbols: ℓ for a path traversing from right to left, r for a path traversing from left to right, ℓr for a right incoming and right outgoing path, $r\ell$ for a left incoming and left outgoing path, and \perp when there are no incoming nor outgoing edges from that node. As a remark, the number of states of a zigzag automaton is bounded by 5^N . For example, Figure 7 (c) shows the use of states in a zigzag automaton.

The transition relation $\Delta \subseteq Q \times \Sigma_R \times Q$ is defined as follows. For all $\mathbf{q}, \mathbf{q}' \in Q$ and $\mathcal{G} \in \Sigma_R$, we have $\mathbf{q} \xrightarrow{\mathcal{G}} \mathbf{q}'$, if and only if, for all $1 \leq i \leq N$:

- $\mathbf{q}_i = \ell$ iff \mathcal{G} has one edge to x_i and no other edge involving x_i ,
- $\mathbf{q}'_i = \ell$ iff \mathcal{G} has one edge from x'_i and no other edge involving x'_i ,
- $\mathbf{q}_i = r$ iff \mathcal{G} has one edge from x_i and no other edge involving x_i ,
- $\mathbf{q}'_i = r$ iff \mathcal{G} has one edge to x'_i and no other edge involving x'_i ,
- $\mathbf{q}_i = \ell r$ iff \mathcal{G} has exactly two edges involving $x_i, x_j^{(\ell)} \rightarrow x_i \rightarrow x_k^{(r)}$ and $j \not\vdash_R k$,
- $\mathbf{q}'_i = r\ell$ iff \mathcal{G} has exactly two edges involving $x'_i, x_j^{(r)} \rightarrow x'_i \rightarrow x_k^{(\ell)}$ and $j \not\vdash_R k$,

- $\mathbf{q}'_i \in \{\ell r, \perp\}$ iff G has no edge involving x'_i ,
- $\mathbf{q}_i \in \{r\ell, \perp\}$ iff G has no edge involving x_i .

Observe that the variables that occur on any path which traverses a vertex labeled ℓr or $r\ell$ may not belong to the same SCC of the folded graph \mathcal{G}_R^f . As a consequence, every path recognized by a zigzag automaton is saturated. For example, the path recognized by the run in Figure 4 (c) goes forward while crossing the variables x_1, x_2, x_3 and, after changing direction, goes backward while crossing only x_4 .

We distinguish four types of paths in \mathcal{G}_R^n . A path $x_i^{(k)} \xrightarrow{*} x_j^{(\ell)}$ is said to be *odd forward* if $k = 0$ and $\ell = n$, *even forward* if $k = \ell = 0$, *odd backward* if $k = n$ and $\ell = 0$, and *even backward* if $k = \ell = n$. The symbols needed to represent an odd path have an odd number of edges, while those encoding even paths have even numbers of edges.

The zigzag automaton for R is a union of four types of automata. Formally, for each $i, j \in \{1, \dots, N\}$ and $t \in \{of, ob, ef, eb\}$ (we use the abbreviations *of*=odd forward, *ob*=odd backward, *ef*=even forward and *eb*=even backward), the weighted automaton $\mathcal{A}_{ij}^t = \langle Q, \omega, I_{ij}^t, F_{ij}^t, \Delta \rangle$ recognizes the saturated paths $x_i^{(p)} \xrightarrow{*} x_j^{(q)}$ of type t , with $p, q \in \{0, n\}$. More precisely, we define the initial and final states as follows:

$$\begin{aligned}
I_{ij}^{of} &= \{\mathbf{q} \mid \mathbf{q}_i = r \text{ and } \mathbf{q}_h \in \{\ell r, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{i\}\} \\
F_{ij}^{of} &= \{\mathbf{q} \mid \mathbf{q}_j = r \text{ and } \mathbf{q}_h \in \{r\ell, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{j\}\} \\
I_{ij}^{ob} &= \{\mathbf{q} \mid \mathbf{q}_i = \ell \text{ and } \mathbf{q}_h \in \{\ell r, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{i\}\} \\
F_{ij}^{ob} &= \{\mathbf{q} \mid \mathbf{q}_j = \ell \text{ and } \mathbf{q}_h \in \{r\ell, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{j\}\} \\
I_{ij}^{ef} &= \begin{cases} \{\mathbf{q} \mid \mathbf{q}_i = r, \mathbf{q}_j = \ell, \mathbf{q}_h \in \{\ell r, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{i, j\}\} & \text{if } i \neq j \\ \{\mathbf{q} \mid \mathbf{q}_i = \ell r, \mathbf{q}_h \in \{\ell r, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{i\}\} & \text{if } i = j \end{cases} \\
F_{ij}^{ef} &= \{r\ell, \perp\}^N \\
I_{ij}^{eb} &= \{\ell r, \perp\}^N \\
F_{ij}^{eb} &= \begin{cases} \{\mathbf{q} \mid \mathbf{q}_i = \ell, \mathbf{q}_j = r, \mathbf{q}_h \in \{\ell r, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{i, j\}\} & \text{if } i \neq j \\ \{\mathbf{q} \mid \mathbf{q}_i = r\ell, \mathbf{q}_h \in \{\ell r, \perp\}, \forall h \in \{1, \dots, N\} \setminus \{i\}\} & \text{if } i = j \end{cases}
\end{aligned}$$

Example 7 Figure 7 (b) shows the zigzag automaton \mathcal{A}_{24}^{ef} of the difference bounds relation $R \Leftrightarrow x_2 - x'_1 \leq -1 \wedge x_3 - x'_2 \leq 0 \wedge x_1 - x'_3 \leq 0 \wedge x'_4 - x_4 \leq 0 \wedge x'_3 - x_4 \leq 0$ from Ex. 4 and Ex. 6. Note that the states that are not both reachable and co-reachable are not shown in this figure, hence the alphabet symbols \mathcal{G}_6 and \mathcal{G}_7 are not used. Fig. 7 (c) shows a run of \mathcal{A}_{24}^{ef} on the word $\gamma = \mathcal{G}_3.(\mathcal{G}_1.\mathcal{G}_2.\mathcal{G}_3)^2.\mathcal{G}_4$, encoding an *ef*-path. ■

The following theorem wraps up the above definition, by relating the language of a zigzag automaton with the weights of the minimal paths in the unfolding \mathcal{G}_R^n of the constraint graph defining a difference bounds relation R .

Theorem 3 Let $R \in \text{DB}_{\mathbf{x}}$ be a $*$ -consistent saturated difference bounds relation, for $\mathbf{x} = \{x_1, \dots, x_N\}$. Then, for each $n > 0$ and all $1 \leq i, j \leq N$, the following hold:

1. each word $w \in \mathcal{L}(\mathcal{A}_{ij}^{of})$ of length n encodes a saturated path $x_i^{(0)} \xrightarrow{*} x_j^{(n)}$ and the weight of a minimal such path is $\min w_{\mathcal{A}_{ij}^{of}}(n)$,
2. each word $w \in \mathcal{L}(\mathcal{A}_{ij}^{ob})$ of length n encodes a saturated path $x_i^{(n)} \xrightarrow{*} x_j^{(0)}$ and the weight of a minimal such path is $\min w_{\mathcal{A}_{ij}^{ob}}(n)$,

3. each word $w \in \mathcal{L}(\mathcal{A}_{ij}^{ef})$ of length n encodes a saturated path $x_i^{(0)} \xrightarrow{*} x_j^{(0)}$ and the weight of a minimal such path is $\minw_{\mathcal{A}_{ij}^{ef}}(n)$,
4. each word $w \in \mathcal{L}(\mathcal{A}_{ij}^{eb})$ of length n encodes a saturated path $x_i^{(n)} \xrightarrow{*} x_j^{(n)}$ and the weight of a minimal such path is $\minw_{\mathcal{A}_{ij}^{eb}}(n)$.

Proof: We prove that each word $w \in \mathcal{L}(\mathcal{A}_{ij}^t)$, for $t \in \{of, ob, ef, eb\}$ encodes a saturated path by contradiction. Suppose that π is a path in $w = \sigma_1 \dots \sigma_n$ which is not saturated. Then there exists two adjacent edges $x_s^{(k)} \rightarrow x_t^{(\ell)} \rightarrow x_u^{(m)}$ on π belonging to the same alphabet symbol $\sigma_i \in \Sigma_R$, for some $1 \leq i \leq n$, $s \sim_R u$ and either (i) $\ell = k + 1$ and $m \in \{k, k + 1\}$, or (ii) $\ell = k - 1$ and $m \in \{k - 1, k\}$. Since $w \in \mathcal{L}(\mathcal{A}_{ij}^t)$, there exists a run $\mathbf{q}_0 \xrightarrow{\sigma_1} \mathbf{q}_1 \dots \xrightarrow{\sigma_n} \mathbf{q}_n$ in \mathcal{A}_{ij}^t . In the first case, we have $(\mathbf{q}_i)_t = r\ell$ and in the second case $(\mathbf{q}_{i-1})_t = \ell r$. In both cases, however, we must have $s \not\sim_R u$, by the definition of the transition relation of \mathcal{A}_{ij}^t , contradiction.

For the characterization of the weights of minimal paths, the proofs of the points (1), (2), (3) and (4) are based on [7, Lemmas 4.6, 4.7, 4.3 and 4.4], respectively. \square

5.3 Weighted Graphs and Periodic Powers of Matrices

This section recalls several results from the theory of weighted graphs, needed to prove the periodicity of the minimal weight functions $\minw_{\mathcal{A}}$ of weighted automata. Based on these facts, we characterize the prefixes and periods of a sequence of powers of a matrix, which sets the ground for the analysis of the periodicity of difference bounds relations (Section 5.4).

Let $\mathcal{G} = \langle V, E, w \rangle$ be a weighted graph for the rest of this section. A path π is of *minimal weight for its length* if, for any path π' such that $\text{src}(\pi') = \text{src}(\pi)$, $\text{dst}(\pi') = \text{dst}(\pi)$ and $|\pi'| = |\pi|$, we have $w(\pi) \leq w(\pi')$. Two paths π and π' in \mathcal{G} are *equivalent* if $\text{src}(\pi) = \text{src}(\pi')$, $\text{dst}(\pi) = \text{dst}(\pi')$, $|\pi| = |\pi'|$ and $w(\pi) = w(\pi')$. The *average weight* of a path π is $\overline{w}(\pi) = \frac{w(\pi)}{|\pi|}$. A cycle is said to be *critical* if it has minimal average weight among all cycles of \mathcal{G} . The *critical graph* \mathcal{G}^c consists of those vertices and edges of \mathcal{G} that belong to a critical cycle. The following theorem states a classical result [1, Theorem 3.96]:

Theorem 4 *For any weighted graph \mathcal{G} , every cycle of the critical graph \mathcal{G}^c is critical.*

If C is a strongly connected component (SCC) of \mathcal{G}^c , we define its *cyclicity* as the greatest common divisor of the lengths of all its elementary cycles. The cyclicity of \mathcal{G}^c is the least common multiple of the cyclicities of its SCCs, and the cyclicity of \mathcal{G} , denoted $c(\mathcal{G})$, is the cyclicity of \mathcal{G}^c .

Weighted graphs are intimately related with the powers of their incidence matrices, defined as follows. For two matrices $A, B \in \mathbb{Z}_{\infty}^{n \times n}$, let $(A \boxtimes B)_{ij} = \min_{1 \leq k \leq n} (A_{ik} + B_{kj})$ and $\mathbf{1}_n$ be the matrix $(\mathbf{1}_n)_{ii} = 0$, for all $1 \leq i \leq n$ and $(\mathbf{1}_n)_{ij} = \infty$, for all $1 \leq i, j \leq n$, where $i \neq j$. The powers of a matrix M are defined as $M^0 = \mathbf{1}_n$ and $M^{k+1} = M \boxtimes M^k$, for all $k \geq 0$. If M is the incidence matrix of a weighted graph \mathcal{G} , the coefficient $(M^k)_{ij}$ is the weight of a minimal path of length k between the vertices

i and j in \mathcal{G} . In this case, we also write $c(M)$ for $c(\mathcal{G})$. The following theorem provides a basic tool for proving periodicity of a sequence of relations, in the following.

Theorem 5 *For a matrix $M \in \mathbb{Z}_{\infty}^{n \times n}$, the sequence $\{M^k\}_{k=0}^{\infty}$ is periodic and its period divides $c(M)$.*

Proof: See [33, Theorem 3.3]. \square

Despite our best efforts, no estimation of the prefix of a power sequence of a matrix could be found in the literature, so far. This gap is filled by the next theorem. We recall that, for a matrix M , $\mu(M)$ stands for the maximum between the absolute values of its coefficients and one.

Theorem 6 *Given a matrix $M \in \mathbb{Z}_{\infty}^{n \times n}$, the prefix of the periodic sequence $\{M^k\}_{k=0}^{\infty}$ is at most $4\mu(M) \cdot n^6$.*

Proof: See Appendix A. \square

Observe that the prefix of a sequence of matrix powers $\{M^k\}_{k=0}^{\infty}$ depends linearly on the maximal coefficient(s) and polynomially on the dimension of M . On the other hand, its period (Theorem 5) depends (exponentially) only on the dimension of M .

Finally, we state the results of Theorems 5 and 6 in terms of weighted automata, instead of weighted graphs. Given a weighted automaton $A = \langle \Sigma, \omega, Q, I, F, \Delta \rangle$, its underlying weighted graph is defined as $\mathcal{G}(A) = \langle Q, \delta, w \rangle$, where for all $q, q' \in Q$: (i) $(q, q') \in \delta$ iff there exists $\sigma \in \Sigma$ such that $(q, \sigma, q') \in \Delta$, and (ii) $w(q, q') = \min \{ \omega(\sigma) \mid \exists \sigma \in \Sigma . (q, \sigma, q') \in \Delta \}$. We write $c(A)$ and $\mu(A)$ for $c(\mathcal{G}(A))$ and $\mu(\mathcal{G}(A))$, respectively.

Corollary 1 *For a weighted automaton $A = \langle \Sigma, \omega, Q, I, F, \Delta \rangle$, the infinite sequence $\{\min w_A(n)\}_{n=0}^{\infty}$ is periodic, with prefix $b = \mathcal{O}(\mu(A) \cdot c(A) \cdot \|Q\|^{10})$ and period c that divides $c(A)$.*

Proof: Let $I = \{q_{i_1}, \dots, q_{i_k}\}$, $F = \{q_{j_1}, \dots, q_{j_\ell}\}$ be the sets of initial and final states of A . Clearly we have $k, \ell \leq \|Q\|$. By denoting $m_{st}(n) = \min \{ \omega(w) \mid q_{i_s} \xrightarrow{w} q_{j_t}, |w| = n \}$, we have:

$$\min \{ \omega(w) \mid w \in \mathcal{L}(A), |w| = n \} = \min_{s=1}^k \min_{t=1}^{\ell} m_{st}(n) .$$

By Theorem 5, each sequence $\{m_{st}(n)\}_{n=0}^{\infty}$ is periodic, with prefix b_{st} and period c_{st} that divides $c(A)$. By Lemma 19, the sequence $\min \{ \omega(w) \mid w \in \mathcal{L}(A), |w| = n \}$ is periodic, with period c that divides $\text{lcm}_{s=1}^k \text{lcm}_{t=1}^{\ell} c_{st}$. Since each c_{st} divides $c(A)$, we have that c divides $c(A)$ as well. For an upper bound on the prefix b of this sequence, let $b_{\max} = \max_{s=1}^k \max_{t=1}^{\ell} b_{st}$. By Lemma 19, we obtain:

$$\begin{aligned} b &\leq b_{\max} + k \cdot \ell \cdot \max_{i=0}^{c-1} \left(\sum_{s=1}^k \sum_{t=1}^{\ell} \text{abs}(m_{s,t}(b_{\max} + i)) \right) \\ b &\leq b_{\max} + k^2 \cdot \ell^2 \cdot (b_{\max} + c - 1) \cdot \mu(A) \end{aligned}$$

By Thm. 6, we have $b_{\max} \leq 4\mu(A) \cdot \|Q\|^6$, hence we obtain, after simplifications $b \leq 4 \cdot \mu(A) \cdot c(A) \cdot \|Q\|^{10}$. \square

5.4 The Periodicity of Difference Bounds Relations

We are now ready to prove that the sequence of matrices $\{\sigma(R^n)\}_{n=0}^\infty$ is periodic, where $R \in \text{DB}_x$ is any difference bounds relation and $x = \{x_1, \dots, x_N\}$ is a set of variables. The coefficients of $\sigma(R^n)$ are the weights of the minimal paths between extremal vertices from the set $x^{(0)} \cup x^{(n)}$ in the unfolding \mathcal{G}_R^n of the constraint graph \mathcal{G}_R — see the constraints (5). By Theorem 3, these weights are given by the functions $\min w_{\mathcal{A}_{ij}^t}(n)$, where $\mathcal{A}_{ij}^t = \langle Q, \omega, I_{ij}^t, F_{ij}^t, \Delta \rangle$ are the zigzag automata for the relation R . Since these functions are periodic, it follows that the sequence $\{\sigma(R^n)\}_{n=0}^\infty$ is periodic (Corollary 1). Moreover, the prefix of this sequence is polynomially bounded by the common cyclicity of zigzag automata and $\|Q\|$ and its period divides this cyclicity. Since $\|Q\| = 2^{O(N)}$ by the construction of zigzag automata, we are essentially left with bounding the cyclicity of zigzag automata.

Let us start by proving a structural property of cycles in a zigzag automaton. A cycle $\mathbf{q} \xrightarrow{\gamma} \mathbf{q}$ in the underlying weighted graph $\mathcal{G}(\mathcal{A}_{ij}^t)$ of \mathcal{A}_{ij}^t is *critical* if it is a critical cycle of $\mathcal{G}(\mathcal{A}_{ij}^t)$ and, moreover, \mathbf{q} is both reachable and co-reachable in \mathcal{A}_{ij}^t .

Lemma 10 *Let \mathcal{A} be a zigzag automaton for a saturated relation $R \in \text{DB}_x$, where $x = \{x_1, \dots, x_N\}$ and $\mathbf{q} \xrightarrow{\gamma} \mathbf{q}$ be one of its a critical cycles, for $|\gamma| > 0$. Then γ is a set of saturated paths $\{\xi_k : x_k^{(p_k)} \xrightarrow{*} x_k^{(q_k)}\}_{k \in K}$, either forward or backward, such that $k \sim_R \ell$ only if $k = \ell$, for all $k, \ell \in K$, where $K \subseteq \{1, \dots, N\}$.*

Proof: We give the proof for the odd-forward case, the other three cases being similar. Since \mathbf{q} is a reachable and co-reachable state of \mathcal{A}_{ij}^{of} , there exists a word $w = \mu \cdot \gamma \cdot \nu \in \mathcal{L}(\mathcal{A}_{ij}^{of})$ and a run $\mathbf{q}_0 \xrightarrow{\mu} \mathbf{q} \xrightarrow{\gamma} \mathbf{q} \xrightarrow{\nu} \mathbf{q}_f$ in \mathcal{A}_{ij}^{of} for some $\mathbf{q}_0 \in I_{ij}^t$ and $\mathbf{q}_f \in F_{ij}^t$. By Theorem 3 (1), w encodes a saturated path $\pi : x_i^{(0)} \xrightarrow{*} x_j^{(|w|)}$. We recall that \mathbf{q} is a N -tuple of elements from the set $\{\ell, r, \ell r, r\ell, \perp\}$ and we denote by \mathbf{q}_k the k -th element of \mathbf{q} , for $1 \leq k \leq N$. Let us consider first the case $\mathbf{q}_k = r$ (the case $\mathbf{q}_k = \ell$ is symmetric). Then either:

- π has a subpath $\xi : x_k^{(|\mu|)} \xrightarrow{*} x_k^{(|\mu|+|\gamma|)}$. Since π is saturated and all variables that occur on ξ are from the same equivalence class of \sim_R , all edges in ξ must have the same direction, forward or backward. But since ξ has strictly positive travel, i.e. $\tau(\xi) > 0$, at least one edge on ξ is forward, thus ξ is forward.
- π has a subpath $\xi : x_k^{(|\mu|+|\gamma|)} \xrightarrow{*} x_k^{(|\mu|)}$ and all edges of ξ must have the same direction, either forward or backward. Since $\tau(\xi) < 0$, at least one edge on ξ must be backward, thus ξ is backward. But since $\mathbf{q}_k = r$, the only outgoing edge from $x_k^{(|\mu|)}$ must be either forward or vertical, contradiction.

The cases $\mathbf{q}_k \in \{\ell r, r\ell\}$ both lead to contradictions, using similar arguments. We have established that γ consists of saturated paths that do not change their direction.

For the second point, assume that there exist two indices $k, \ell \in K$, such that $k \sim_R \ell$ and $k \neq \ell$. Then π has a subpath $\xi : x_k^{(|\mu|)} \xrightarrow{*} x_\ell^{(|\mu|)}$, and since $k \sim_R \ell$, it follows that all variables occurring on ξ must be equivalent. Since π is saturated, the same holds for ξ , thus the edges on ξ are either all forward or all backward. But this contradicts the fact that the endpoints of ξ are both on the same position $|\mu|$. \square

A path $x_k^{(p)} \xrightarrow{*} x_k^{(q)}$ is *essential* if all variables occurring on it are pairwise distinct,

except for the labels of its source and destination vertices. Clearly, the length of an essential path is bounded by the number N of variables occurring on this path. An *essential power* is a path ξ^n obtained from the concatenation of an essential path ξ with itself $n > 0$ times.

The following lemma shows that each critical cycle $\mathbf{q} \xrightarrow{\gamma} \mathbf{q}$ in a zigzag automaton \mathcal{A} is necessarily connected to a critical cycle $\mathbf{q} \xrightarrow{\lambda} \mathbf{q}$, where λ consists of a finite set of essential powers. This allows us to bound the length of λ by a simply exponential value, which divides $\text{lcm}(1, \dots, N)$. It follows that the common cyclicity of these zigzag automata is a divisor of $\text{lcm}(1, \dots, N)$. We use the fact that $\text{lcm}(1, \dots, N) = 2^{\mathcal{O}(N)}$ [30], which occurs as a consequence of the Prime Number Theorem, and bound the cyclicity of zigzag automata by $2^{\mathcal{O}(N)}$.

Lemma 11 *Let \mathcal{A} be a zigzag automaton for a saturated relation $R \in \text{DB}_x$, where $\mathbf{x} = \{x_1, \dots, x_N\}$ and $\mathbf{q} \xrightarrow{\gamma} \mathbf{q}$ be one of its a critical cycles, for $|\gamma| > 0$. Then there exists a critical cycle $\mathbf{q} \xrightarrow{\lambda} \mathbf{q}$ in \mathcal{A} such that λ is a set of essential powers $\{\pi_k^{n_k}\}_{k \in K}$ and $|\lambda| = \text{lcm}_{k \in K} \{|\pi_k|\}$, for some $K \subseteq \{1, \dots, N\}$.*

Proof: By Lemma 10, γ is a set $\{\xi_k\}_{k \in K}$ of forward/backward paths, such that the labels of ξ_k and ξ_ℓ lie in different equivalence classes of the \sim_R relation, for all $k \neq \ell$. We shall build a word λ as a set of essential powers $\{\pi_k^{n_k}\}_{k \in K}$. Clearly, the paths $\pi_k^{n_k}$ and $\pi_\ell^{n_\ell}$ may not intersect (because they cannot share labels), for any $k \neq \ell$, thus $\mathbf{q} \xrightarrow{\lambda} \mathbf{q}$ is a valid cycle w.r.t. the definition of the transition table of \mathcal{A} . Before giving the definition of the set $\{\pi_k^{n_k}\}_{k \in K}$, we prove the following fact:

Fact 1 *For a path $\xi_k : x_k^{(p_1)} \xrightarrow{\alpha_1} \dots \xrightarrow{\alpha_{m-1}} x_k^{(p_m)}$, let \mathcal{G}_k be the restriction of \mathcal{G}_R^f to the vertices and edges on ξ_k . Then $\zeta_k : x_k \xrightarrow{\alpha_1} \dots \xrightarrow{\alpha_{m-1}} x_k$ is a critical cycle of \mathcal{G}_k .*

Proof: Suppose, by contradiction, that there exists $k \in K$ such that ζ_k is not a critical cycle of \mathcal{G}_k , hence there exists a cycle θ in \mathcal{G}_k , such that $\overline{w}(\theta) < \overline{w}(\zeta_k)$. Since \mathcal{G}_k is strongly connected, there exist paths μ and ν such that $\eta_{st} = \mu \cdot \theta^s \cdot \nu \cdot (\mu \cdot \nu)^t$ is a path in \mathcal{G}_k , with source and destination x_k , for all $s, t \geq 0$. It is sufficient to build a path η_{st} such that $|\eta_{st}|$ is a multiple of $|\gamma|$. In this case, there exist a cycle $\mathbf{q} \xrightarrow{\gamma'} \mathbf{q}$, where γ' is the set consisting of $\eta_{s+K|\gamma|, t}$, for a sufficiently large $K > 0$, and powers of ξ_ℓ , for all $\ell \in K \setminus \{k\}$, such that $\overline{w}(\gamma') < \overline{w}(\gamma)$, which contradicts the fact that $\mathbf{q} \xrightarrow{\gamma} \mathbf{q}$ is a critical cycle of \mathcal{A} .

In order to find s and t such that $|\eta_{st}|$ is a multiple of $|\gamma|$, let $n_1 = |\mu| + |\nu|$, $n_2 = |\theta|$, $m_1 = \frac{n_1}{\gcd(n_1, n_2)}$ and $m_2 = \frac{n_2}{\gcd(n_1, n_2)}$. Clearly m_1 and m_2 are coprimes, and let $g(n_1, n_2) = n_1 n_2 - (n_1 + n_2)$ be their Frobenius number. We know that any integer $h > g(n_1, n_2)$ is a conical combination $h = k_1 m_1 + k_2 m_2$, where $k_1, k_2 \geq 0$. Let $n_2 = \ell_1 m_1 + \ell_2 m_2$, for some $\ell_1, \ell_2 \geq 0$, be the smallest multiple of $|\gamma|$ that is greater than $g(n_1, n_2)$, and let $u = \gcd(n_1, n_2) \cdot \frac{n_2}{|\gamma|}$. It is now easy to verify that $|\eta_{s, t}| = s \cdot n_1 + t \cdot n_2 = u \cdot |\gamma|$, where $s = \ell_1 \cdot \gcd(n_1, n_2)$ and $t = \ell_2 \cdot \gcd(n_1, n_2)$. \square

It follows that each graph \mathcal{G}_k is critical, i.e. $\mathcal{G}_k = \mathcal{G}_k^c$, and we can chose, in \mathcal{G}_k , an elementary cycle ρ_k with both source and destinatin labeled by x_k . By Theorem 4, we

have that ρ_k is a critical cycle as well, and consequently $\overline{w}(\rho_k) = \overline{w}(\zeta_k)$. Let π_k be the path in \mathcal{G}_R^n that corresponds to the path ρ_k in \mathcal{G}_R^f , for each $k \in K$. Since all edges of ρ_k correspond to edges of \mathcal{G}_R^n that are either all forward or all backward, we know that such a path exists. The path π_k is essential, because ρ_k is an elementary cycle of \mathcal{G}_R^f and, moreover, $\overline{w}(\pi_k) = \overline{w}(\rho_k) = \overline{w}(\zeta_k) = \overline{w}(\xi_k)$. The word λ consists of the essential powers $\{\pi_k^{n_k}\}_{k \in K}$, where $n_k = \frac{\text{lcm}_{\ell \in K} |\pi_\ell|}{|\pi_k|}$, for all $k \in K$. We have $|\lambda| = \text{lcm}_{k \in K} |\pi_k|$ by the construction of λ . Moreover, $\mathbf{q} \xrightarrow{\lambda} \mathbf{q}$ is a cycle in \mathcal{A} , and since $\frac{w(\lambda)}{|\lambda|} = \frac{w(\gamma)}{|\gamma|}$, it is also a critical cycle. \square

The next theorem concludes the proof of periodicity for the class DB, providing simply exponential upper bounds on the prefix and the period of a difference bounds relation. The next section extends this result to the class OCT and finalizes the proof of Theorem 2, which is the main result of this paper.

Theorem 7 *There exists a constant $d > 0$ such that, for every relation $R \in \text{DB}$, the sequence $\{\sigma(R^k)\}_{k=0}^\infty$ is periodic, with prefix $b = 2^{\mathcal{O}(|R|^d)}$ and period $c = 2^{\mathcal{O}(|R|^d)}$.*

Proof: Let $R \subseteq \mathbb{Z}^x \times \mathbb{Z}^x$, where $\mathbf{x} = \{x_1, \dots, x_N\}$ and assume w.l.o.g. that each variable x_i occurs in a formula ϕ that defines R . Then we have $N \leq |R|$. Let \mathcal{G}_R be the constraint graph of R . We also have that $\mu(\mathcal{G}_R) \leq 2^{\mathcal{O}(|R|)}$.

We distinguish two cases. First, if R is not \ast -consistent, its period is $c = 1$ and its prefix is $b \leq 6N^7 \cdot \mu(\mathcal{G}_R) = 6|R|^6 \cdot 2^{\mathcal{O}(|R|)} = 2^{\mathcal{O}(|R|)}$, by Lemma 7 (2). Second, if R is \ast -consistent, there exists a saturated relation $R_{\text{sat}} \subseteq R$ such that R is periodic if R_{sat} is periodic, and

- $b = b_{\text{sat}} + \mathcal{O}(2^{N \log N}) \cdot \max(\mu(\sigma(R^{N^2})), \max_{0 \leq i < c_{\text{sat}}} \mu(\sigma(R_{\text{sat}}^{b_{\text{sat}}+i})))$,
- c divides c_{sat} ,

where b_{sat} and c_{sat} are the prefix and period of R_{sat} (Lemma 9). By Theorem 3 the sequence $\{\sigma(R_{\text{sat}}^k)\}_{k=0}^\infty$ is periodic, and by Corollary 1, we have $b_{\text{sat}} = \mathcal{O}(\mu(\mathcal{A}) \cdot c(\mathcal{A}) \cdot \|Q\|^{10})$ and c_{sat} is a divisor of $c(\mathcal{A})$, where $\mathcal{A} = \langle \Sigma_R, \omega, Q, I, F, \Delta \rangle$ is any zigzag automaton for R_{sat} . Since Σ_R is a set of subgraphs of \mathcal{G}_R , we have $\mu(\mathcal{A}) \leq \mu(\mathcal{G}_R) = 2^{\mathcal{O}(|R|)}$. Moreover, $\|Q\| = 2^{\mathcal{O}(|R|)}$, by the definition of zigzag automata.

Observe that the choice of \mathcal{A} does not influence $c(\mathcal{A})$, because all zigzag automata share the same transition table. It remains to give a bound for the cyclicity of \mathcal{A} . By Lemma 11 each critical cycle of $\mathcal{G}(\mathcal{A})$ is connected to a critical cycle of length that divides $\text{lcm}(1, \dots, N)$. Thus the cyclicity of each SCC of \mathcal{A} divides $\text{lcm}(1, \dots, N)$ and the same holds for $c(\mathcal{A})$, which is the least common multiple of the cyclicities of the SCCs of $\mathcal{G}(\mathcal{A})$. Since $\text{lcm}(1, \dots, N) = 2^{\mathcal{O}(N)}$ [30], we obtain that $c(\mathcal{A}) = 2^{\mathcal{O}(N)} = 2^{\mathcal{O}(|R|)}$. Summing up, we obtain $b_{\text{sat}} = 2^{\mathcal{O}(|R|)}$ and $c_{\text{sat}} = 2^{\mathcal{O}(|R|)}$. By Lemma 3, for all $0 \leq i < c_{\text{sat}}$, we have $|R_{\text{sat}}^{b_{\text{sat}}+i}| = \mathcal{O}((|R|_{\text{sat}} \cdot \log(b_{\text{sat}} + c_{\text{sat}}))^d)$, for a constant $d > 0$ that does not depend on the choice of R_{sat} . Since $|R_{\text{sat}}| \leq N^2 \cdot |R|$, by the definition of R_{sat} (Lemma 8), we can conclude that $\max_{0 \leq i < c_{\text{sat}}} \mu(\sigma(R_{\text{sat}}^{b_{\text{sat}}+i})) = 2^{\mathcal{O}(|R|^d)}$. A similar reasoning leads to $\mu(\sigma(R^{N^2})) = 2^{\mathcal{O}(|R|^d)}$, and thus $b = 2^{\mathcal{O}(|R|^d)}$. Since $c \leq c_{\text{sat}}$, we also have that $c = 2^{\mathcal{O}(|R|^d)}$, which concludes the proof. \square

5.5 Finalizing the Proof of Theorem 2

We have gathered all the elements necessary to prove the second point of Theorem 2, namely that the octagonal relations are periodic, with simply exponential prefixes and periods. As a consequence, the class of problems REACHFLAT(OCT) is NP-complete. The theorem below is a consequence of the similar result for difference bounds constraints and of the relation between the powers of octagonal relations and their encodings using difference bounds constraints (Lemma 4). We infer that the class OCT is periodic, because for two periodic sequences $\{s_k\}_{k=0}^\infty$ and $\{t_k\}_{k=0}^\infty$, with prefixes $b_s, b_t \geq 0$ and periods $c_s, c_t > 0$, respectively, the following sequences are periodic:

- $\{s_k + t_k\}_{k=0}^\infty$ with prefix at most $\max(b_s, b_t)$ and period which divides $\text{lcm}(c_s, c_t)$ (Lemma 16, Appendix B),
- $\{\lfloor \frac{s_k}{2} \rfloor\}_{k=0}^\infty$ with prefix b_s and period $2c_s$ (Lemma 17, Appendix B),
- $\{\min(s_k, t_k)\}_{k=0}^\infty$ with prefix at most $\max(b_s, b_t) + \sum_{i=0}^{\text{lcm}(c_s, c_t)} (\text{abs}(s_i) + \text{abs}(t_i))$ and period which divides $\text{lcm}(c_s, c_t)$ (Lemma 18, Appendix B).

Theorem 8 *There exists a constant $d > 0$ such that, for every relation $R \in \text{OCT}$, the sequence $\{\sigma(R^k)\}_{k=0}^\infty$ is periodic, with prefix $b = 2^{\mathcal{O}(|R|^d)}$ and period $c = 2^{\mathcal{O}(|R|^d)}$.*

Proof: Let $R \in \text{OCT}_{\mathbf{x}}$, where $\mathbf{x} = \{x_1, \dots, x_N\}$ and $\bar{R} \in \text{DB}_{\mathbf{y}}$ be the difference bounds relation that encodes R , for $\mathbf{y} = \{y_1, \dots, y_{2N}\}$. We have that $|\bar{R}| \leq 2|R|$.

We consider first the case in which R is $*$ -consistent. By Lemma 4 the matrix $\sigma(R^k)$ is octagonal-consistent, for any $k \geq 0$. Then, by Theorem 1, $\sigma(\bar{R}^k)$ is consistent, for all $k \geq 0$, thus \bar{R} is $*$ -consistent. Moreover, by Theorem 7, the sequence $\{\sigma(\bar{R}^k)\}_{k=0}^\infty$ is periodic, with period b and prefix c of the order of $2^{\mathcal{O}(|\bar{R}|^d)}$, for a constant d which does not depend on the choice of R . By Lemma 1, each of the sequences $\{\sigma(\bar{R}^k)_{ij}\}_{k=0}^\infty$ is periodic, with prefix $b_{ij} \leq b$ and period c_{ij} that divides c . By Lemma 17, the sequences $\{\lfloor \frac{\sigma(\bar{R}^k)_{i\bar{i}}}{2} \rfloor\}_{k=0}^\infty$ and $\{\lfloor \frac{\sigma(\bar{R}^k)_{j\bar{j}}}{2} \rfloor\}_{k=0}^\infty$ are periodic with prefixes $b_{i\bar{i}}, b_{j\bar{j}}$ and periods $2c_{i\bar{i}}, 2c_{j\bar{j}}$, respectively. By Lemma 16, the sequence $\{\lfloor \frac{\sigma(\bar{R}^k)_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\sigma(\bar{R}^k)_{j\bar{j}}}{2} \rfloor\}_{k=0}^\infty$ is periodic with prefix at most $\max(b_{i\bar{i}}, b_{j\bar{j}})$ and period which divides $2\text{lcm}(c_{i\bar{i}}, c_{j\bar{j}})$, thus also $2c$. Then $\{\sigma(R^k)_{ij}\}_{k=0}^\infty$ is periodic with period which divides $2c$ and prefix at most:

$$b_m + \max_{\ell=0}^{2c-1} (\text{abs}(\sigma(\bar{R}^{b_m+\ell})_{ij}) + \text{abs}(\sigma(\bar{R}^{b_m+\ell})_{i\bar{i}}) + \text{abs}(\sigma(\bar{R}^{b_m+\ell})_{j\bar{j}}))$$

where $b_m = \max(b_{ij}, b_{i\bar{i}}, b_{j\bar{j}})$. Since $b_{ij}, b_{i\bar{i}}, b_{j\bar{j}}$ and c are of the order of $2^{\mathcal{O}(|R|^d)}$, for all $\ell = 0, \dots, 2c-1$, the coefficients $\sigma(\bar{R}^{b_m+\ell})_{ij}$, $\sigma(\bar{R}^{b_m+\ell})_{i\bar{i}}$ and $\sigma(\bar{R}^{b_m+\ell})_{j\bar{j}}$ are of the order of $2^{\mathcal{O}((|R| \cdot \log(b_m+c))^e)}$, for a constant e that does not depend on R (Lemma 5), thus b is of the order of $2^{\mathcal{O}(|R|^{de})}$. Finally, by Lemma 1, we obtain that the sequence $\{\sigma(R^k)\}_{k=0}^\infty$ is periodic, with prefix and period both of the order of $2^{\mathcal{O}(|R|^{de})}$.

Second, if R is not $*$ -consistent, we have two cases:

- if \bar{R} is not $*$ -consistent, then its period is 1 and its prefix is of the order of $\mathcal{O}(|R|^7 \cdot 2^{|R|})$, by Lemma 7 (2) and the same bounds apply to R .

- otherwise, \bar{R} is $*$ -consistent and there exists $b_0 \geq 0$ and $1 \leq i \leq 2N$ such that $\lfloor \frac{\sigma(\bar{R}^m)_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\sigma(\bar{R}^m)_{\bar{i}i}}{2} \rfloor < 0$ for all $m \geq b_0$ (Theorem 1). Because $\{\sigma(\bar{R}^m)\}_{m=0}^\infty$ is periodic, with prefix b and period c of the order of $2^{\mathcal{O}(|R|^d)}$, we obtain:

$$\lfloor \frac{\sigma(\bar{R}^{b+2kc})_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\sigma(\bar{R}^{b+2kc})_{\bar{i}i}}{2} \rfloor \geq \lfloor \frac{\sigma(\bar{R}^b)_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\sigma(\bar{R}^b)_{\bar{i}i}}{2} \rfloor + k \cdot \left(\lfloor \frac{\lambda_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\lambda_{\bar{i}i}}{2} \rfloor \right)$$

where $\lambda_{i\bar{i}}$ and $\lambda_{\bar{i}i}$ are the rates of the sequences $\{\sigma(\bar{R}^m)_{i\bar{i}}\}_{m=0}^\infty$ and $\{\sigma(\bar{R}^m)_{\bar{i}i}\}_{m=0}^\infty$, respectively. It must be the case that $\lfloor \frac{\lambda_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\lambda_{\bar{i}i}}{2} \rfloor < 0$, otherwise we could not have $\lfloor \frac{\sigma(\bar{R}^m)_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\sigma(\bar{R}^m)_{\bar{i}i}}{2} \rfloor < 0$ for all $m \geq b_0$. Moreover, $\lfloor \frac{\sigma(\bar{R}^{b+2kc})_{i\bar{i}}}{2} \rfloor + \lfloor \frac{\sigma(\bar{R}^{b+2kc})_{\bar{i}i}}{2} \rfloor < 0$ for all $k \geq \text{abs}(\sigma(\bar{R}^b)_{i\bar{i}}) + \text{abs}(\sigma(\bar{R}^b)_{\bar{i}i}) = 2^{\mathcal{O}(|R|^d)}$. It is easy to see that $b_0 \leq b + 2c(\text{abs}(\sigma(\bar{R}^b)_{i\bar{i}}) + \text{abs}(\sigma(\bar{R}^b)_{\bar{i}i})) = 2^{\mathcal{O}(|R|^d)}$, which gives the bound on the prefix of the sequence $\{\sigma(\bar{R}^k)\}_{k=0}^\infty$ in this case. \square

Conclusions

We prove that the class of reachability problems for flat counter machines, with octagonal relations labeling the cycles, is NP-complete. This result is based on the analysis of the periodic behavior of the matrices that encode the power sequences of relations. These sequences of matrices have, moreover, simply exponential prefixes and the periods. The crux of the proof is the reduction from octagonal to a simpler class of difference bounds constraints, who are proved to be periodic by the construction of a weighted automaton. The prefix and period of difference bounds relations are shown to be simply exponential by a detailed analysis of this weighed automaton.

References

- [1] Baccelli, F., Cohen, G., Olsder, G.J., Quadrat, J.P.: Synchronization and linearity : an algebra for discrete event systems. Wiley series in probability and mathematical statistics, J. Wiley & Sons (1992), <http://opac.inria.fr/record=b1082205>
- [2] Bagnara, R., Hill, P.M., Zaffanella, E.: An improved tight closure algorithm for integer octagonal constraints. In: Proc. of VMCAI. LNCS, vol. 4905, pp. 8–21. Springer Verlag, Berlin, Heidelberg (2008)
- [3] Blondin, M., Finkel, A., Göller, S., Haase, C., McKenzie, P.: Reachability in two-dimensional vector addition systems with states is pspace-complete. CoRR abs/1412.4259 (2014), <http://arxiv.org/abs/1412.4259>
- [4] Boigelot, B.: Symbolic Methods for Exploring Infinite State Spaces. PhD, Univ. de Liège (1999)

- [5] Bozga, M., Gîrlea, C., Iosif, R.: Iterating octagons. In: Proc. of TACAS. LNCS, vol. 5505, pp. 337–351. Springer Verlag, Berlin, Heidelberg (2009)
- [6] Bozga, M., Iosif, R., Konečný, F.: Fast acceleration of ultimately periodic relations. In: CAV. LNCS, vol. 6174, pp. 227–242 (2010)
- [7] Bozga, M., Iosif, R., Lakhnech, Y.: Flat parametric counter automata. *Fundamenta Informaticae* 91(2), 275–303 (2009)
- [8] Bozga, M., Iosif, R., Konečný, F.: Deciding conditional termination. *Logical Methods in Computer Science* 10(3) (2014), [http://dx.doi.org/10.2168/LMCS-10\(3:8\)2014](http://dx.doi.org/10.2168/LMCS-10(3:8)2014)
- [9] Bozzelli, L., Pinchinat, S.: Verification of gap-order constraint abstractions of counter systems. In: VMCAI. pp. 88–103 (2012)
- [10] Comon, H., Jurski, Y.: Multiple counters automata, safety analysis and presburger arithmetic. In: CAV. LNCS, vol. 1427, pp. 268–279 (1998)
- [11] Demri, S., Dhar, A.K., Sangnier, A.: On the complexity of verifying regular properties on flat counter systems. In: ICALP (2). pp. 162–173 (2013)
- [12] Demri, S., Dhar, A., Sangnier, A.: Taming past LTL and flat counter systems. In: IJCAR. vol. 7364, pp. 179–193 (2012)
- [13] Finkel, A., Leroux, J.: How to compose presburger-accelerations: Applications to broadcast protocols. In: Proc. of FST TCS. LNCS, vol. 2556, pp. 145–156. Springer Verlag, Berlin, Heidelberg (2002)
- [14] Ganty, P., Iosif, R.: Interprocedural reachability for flat integer programs. In: Fundamentals of Computation Theory - 20th International Symposium, FCT 2015, Gdańsk, Poland, August 17-19, 2015, Proceedings. pp. 133–145 (2015)
- [15] Gaubert, S.: On rational series in one variable over certain dioids. *Rapport de recherche RR-2162, INRIA* (1994), <http://hal.inria.fr/inria-00074510>
- [16] Gawlitza, T.M., Monniaux, D.: Invariant generation through strategy iteration in succinctly represented control flow graphs. *Logical Methods in Computer Science* 8(3) (2012)
- [17] Gurari, E.M., Ibarra, O.H.: The complexity of decision problems for finite-turn multicounter machines. *J. Computer and System Sciences* 22, 220–229 (1981)
- [18] Hojjat, H., Iosif, R., Konečný, F., Kuncak, V., Rümmer, P.: Accelerating Interpolants, pp. 187–202. Springer Berlin Heidelberg (2012)
- [19] Hopcroft, J., Pansiot, J.J.: On the reachability problem for 5-dimensional vector addition systems. *Theoretical Computer Science* 8(2), 135 – 159 (1979)

- [20] Ibarra, O.H.: Reversal-bounded multicounter machines and their decision problems. *J. ACM* 25(1), 116–133 (1978)
- [21] Konečný, F.: PTIME computation of transitive closures of octagonal relations. *CoRR* abs/1402.2102 (2014), <http://arxiv.org/abs/1402.2102>
- [22] Kosaraju, S.R.: Decidability of reachability in vector addition systems (preliminary version). In: *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*. pp. 267–281. *STOC '82*, ACM (1982)
- [23] Leroux, J.: Vector Addition Systems Reachability Problem (A Simpler Solution). In: Voronkov, A. (ed.) *The Alan Turing Centenary Conference. EPiC Series*, vol. 10, pp. 214–228. Andrei Voronkov, Manchester, United Kingdom (2012), <https://hal.archives-ouvertes.fr/hal-00674970>
- [24] Leroux, J., Sutre, G.: *CONCUR 2004 - Concurrency Theory: 15th International Conference*, London, UK, August 31 - September 3, 2004. *Proceedings*, chap. On Flatness for 2-Dimensional Vector Addition Systems with States, pp. 402–416. Springer Berlin Heidelberg (2004)
- [25] Lin, A.W.: *Model Checking Infinite-State Systems: Generic and Specific Approaches*. Ph.D. thesis, School of Informatics, University of Edinburgh (August 2010)
- [26] Lipton, R.J.: The reachability problem is exponential-space-hard. *Tech. Rep. 62*, Yale University, Department of Computer Science (1976)
- [27] Mayr, E.W.: An algorithm for the general petri net reachability problem. In: *Proceedings of the Thirteenth Annual ACM Symposium on Theory of Computing*. pp. 238–246. *STOC '81*, ACM (1981)
- [28] Miné, A.: Weakly relational numerical abstract domains. Ph.D. thesis, École Polytechnique (Dec 2004), <http://www.di.ens.fr/~mine/these/these-color.pdf>
- [29] Minsky, M.: *Computation: Finite and Infinite Machines*. Prentice-Hall (1967)
- [30] Nair, M.: A new method in elementary prime number theory. *Journal of the London Mathematical Society* s2-25(3), 385–391 (1982), <http://jllms.oxfordjournals.org/content/s2-25/3/385.short>
- [31] Rackoff, C.: The covering and boundedness problems for vector addition systems. *Theoretical Computer Science* 6(2), 223 – 231 (1978)
- [32] Revesz, P.Z.: A closed-form evaluation for Datalog queries with integer (gap)-order constraints. *Theor. Comput. Sci.* 116(1&2), 117–149 (1993)
- [33] Schutter, B.D.: On the ultimate behavior of the sequence of consecutive powers of a matrix in the max-plus algebra. *Linear Algebra and Its Applications* 307(1-3), 103–117 (2000)

- [34] Schutzenberger, M.: On the definition of a family of automata. Information and Control 4(23), 245–270 (1961)
- [35] Verma, K.N., Seidl, H., Schwentick, T.: On the complexity of equational horn clauses. In: CADE '05. LNCS, vol. 1831, pp. 337–352 (2005)

A Proof of Theorem 6

For two paths π and π' , such that $\text{dst}(\pi) = \text{src}(\pi')$, we denote by $\pi.\pi'$ their concatenation. The empty path is denoted ε and if π is a cycle, we define $\pi^0 = \varepsilon$, $\pi^{k+1} = \pi.\pi^k$ and $\pi^* = \{\pi^k \mid k \geq 0\}$. For two sets of paths S and S' let $S.S' = \{\sigma.\sigma' \mid \sigma \in S, \sigma' \in S', \text{dst}(\sigma) = \text{src}(\sigma')\}$. If $\sigma_1, \dots, \sigma_{k+1}$ are paths and $\lambda_1, \dots, \lambda_k$ are pairwise distinct elementary cycles, such that $\text{dst}(\sigma_i) = \text{src}(\sigma_{i+1}) = \text{src}(\lambda_i) = \text{dst}(\lambda_i)$, for all $i = 1, \dots, k$, the set $\sigma_1.\lambda_1^*.\sigma_2 \dots \sigma_k.\lambda_k^*.\sigma_{k+1}$ is called a *path scheme*.

First, we show that all paths in $\mathcal{G} = \langle V, E, w \rangle$, that are minimal for their length, are captured by path schemes in which the number of cycles is at most quadratic in the number of vertices.

Proposition 2 *For each minimal path ρ there exists an equivalent path ρ' and a path scheme $\theta = \sigma_1.\lambda_1^* \dots \sigma_k.\lambda_k^*.\sigma_{k+1}$, such that $k \leq \|V\|^2$ and $\rho' \in \theta$.*

Proof: A similar statement, using a slightly different terminology is proved in [25, Lemma 7.3.2]. Namely, for each path ρ (not necessarily minimal) in \mathcal{G} , there exists an equivalent path $\rho' = \sigma_1.\lambda_1^{n_1} \dots \sigma_k.\lambda_k^{n_k}.\sigma_{k+1}$ where $\sigma_1, \dots, \sigma_{k+1}$ are paths, $\lambda_1, \dots, \lambda_k$ are elementary cycles, $n_1, \dots, n_k > 0$ and $|\sigma_1 \dots \sigma_{k+1}| \leq (\|V\| - 1)^2$. \square

A path scheme $\sigma_1.\lambda_1^*.\sigma_2 \dots \sigma_k.\lambda_k^*.\sigma_{k+1}$ is *bi-quadratic* if $|\sigma_1.\sigma_2 \dots \sigma_{k+1}| \leq \|V\|^4$. Next we show that, for every path in the graph, minimal for its length, there exists an equivalent path which is captured by a bi-quadratic path scheme with one cycle:

Lemma 12 *For each path ρ , minimal for its length, there exists an equivalent path ρ' and a bi-quadratic path scheme $\sigma.\lambda^*.\sigma'$, such that $\rho' \in \sigma.\lambda^*.\sigma'$.*

Proof: By Prop. 2, for any path ρ there exists a path scheme $\theta = \sigma_1.\lambda_1^*.\sigma_2 \dots \sigma_k.\lambda_k^*.\sigma_{k+1}$ such that $k \leq \|V\|^2$, and an equivalent path $\rho' = \sigma_1.\lambda_1^{n_1}.\sigma_2 \dots \sigma_k.\lambda_k^{n_k}.\sigma_{k+1}$, for some $n_1, \dots, n_k \geq 0$. Suppose that λ_i is a cycle with minimal average weight among all cycles in the scheme, i.e. $\overline{w}(\lambda_i) = \frac{w(\lambda_i)}{|\lambda_i|} \leq \frac{w(\lambda_j)}{|\lambda_j|} = \overline{w}(\lambda_j)$, for all $1 \leq j \leq k$. For each n_j there exist $p_j \geq 0$ and $0 \leq q_j < |\lambda_i|$, such that $n_j = p_j \cdot |\lambda_i| + q_j$. Let ρ'' be the path:

$$\sigma_1.\lambda_1^{q_1}.\sigma_2 \dots \sigma_{i-1}.\lambda_i^{n_i + \sum_{j=1}^{i-1} p_j \cdot |\lambda_j| + \sum_{j=i+1}^k p_j \cdot |\lambda_j|}.\sigma_{i+1} \dots \sigma_k.\lambda_k^{q_k}.\sigma_{k+1}$$

It is easy to check that $|\rho''| = |\rho'|$ and $w(\rho'') = w(\rho')$, since ρ' is minimal for its length. Clearly ρ'' is captured by the path scheme $\rho_1.\lambda_i^*.\rho_2$, where $\rho_1 = \sigma_1.\lambda_1^{q_1}.\sigma_2 \dots \sigma_{i-1}$ and $\rho_2 = \sigma_{i+1} \dots \sigma_k.\lambda_k^{q_k}.\sigma_{k+1}$. Since $\sigma_1, \dots, \sigma_k, \sigma_{k+1}$ are acyclic elementary paths, by

Prop. 2, $|\sigma_i| < \|V\|$. Also, since $\lambda_1, \dots, \lambda_k$ are elementary cycles, we have $|\lambda_i| \leq \|V\|$. Since $q_i < |\lambda_i| \leq \|V\|$, and $k \leq \|V\|^2$, by Prop. 2, we obtain:

$$\begin{aligned} |\rho_1 \cdot \rho_2| &\leq (k+1) \cdot (\|V\| - 1) + k \cdot (\|V\|) \cdot (\|V\| - 1) \\ &\leq (\|V\|^2 + 1) \cdot (\|V\| - 1) + \|V\|^2 \cdot (\|V\|) \cdot (\|V\| - 1) \\ &= \|V\|^4 - \|V\|^2 + \|V\| - 1 \leq \|V\|^4 \end{aligned}$$

Hence $\rho_1 \cdot \lambda_i^* \cdot \rho_2$ is a bi-quadratic path scheme. \square

For any $\ell \geq 0$ and vertices $u, v \in V$, let $\text{biq}(u, v, \ell)$ denote the set of all bi-quadratic path schemes $\sigma \cdot \lambda^* \cdot \sigma'$, for which there exists a path $\sigma \cdot \lambda^k \cdot \sigma'$ of length ℓ , between u and v and $\text{minbiq}(u, v, \ell) = \{\sigma \cdot \lambda^* \cdot \sigma' \in \text{biq}(u, v, \ell) \mid \forall \tau \cdot \eta^* \cdot \tau' \in \text{biq}(u, v, \ell) \cdot \overline{w}(\lambda) \leq \overline{w}(\eta)\}$ be the subset of $\text{biq}(u, v, \ell)$ consisting of bi-quadratic path schemes of the form $\sigma \cdot \lambda^* \cdot \sigma'$, where λ has minimal average weight. The following lemma shows that, for each sufficiently long path that is minimal for its length, there exists an equivalent path following a bi-quadratic path scheme of the form $\sigma \cdot \lambda^* \cdot \sigma'$, whose cycle λ has minimal average weight, among all path schemes of this form. We recall that $\mu(\mathcal{G})$ is the maximum between the absolute values of the weights of \mathcal{G} and 1.

Lemma 13 *For every path ρ that is minimal for its length $|\rho| > 4\mu(\mathcal{G}) \cdot \|V\|^6$, there exists an equivalent path ρ' and a path scheme $\sigma \cdot \lambda^* \cdot \sigma' \in \text{minbiq}(\text{src}(\rho), \text{dst}(\rho), |\rho|)$, such that $\rho' \in \sigma \cdot \lambda^* \cdot \sigma'$.*

Proof: Let $u = \text{src}(\rho)$ and $v = \text{dst}(\rho)$. By Lemma 12, for every path ρ , minimal for its length $L > 0$, there exists an equivalent path ρ' which is captured by at least one biquadratic path scheme from $\text{biq}(u, v, L)$. We will show that if $L \geq 4\mu(\mathcal{G}) \cdot \|V\|^6$, the cycle in this path scheme must have minimal average weight among the cycles of all path schemes in $\text{biq}(u, v, L)$. Let $\sigma_i \cdot \lambda_i^* \cdot \sigma'_i, \sigma_j \cdot \lambda_j^* \cdot \sigma'_j \in \text{biq}(u, v, L)$ be two path schemes such that $\rho_i = \sigma_i \cdot \lambda_i^{b_i} \cdot \sigma'_i$ and $\rho_j = \sigma_j \cdot \lambda_j^{b_j} \cdot \sigma'_j$ are two paths of length L , between the same vertices, for some $b_i, b_j \geq 0$. First, we compute:

$$\begin{aligned} b_i &= \frac{L - |\sigma_i \cdot \sigma'_i|}{|\lambda_i|} & w(\rho_i) &= w(\sigma_i \cdot \sigma'_i) + \frac{L - |\sigma_i \cdot \sigma'_i|}{|\lambda_i|} \cdot w(\lambda_i) \\ b_j &= \frac{L - |\sigma_j \cdot \sigma'_j|}{|\lambda_j|} & w(\rho_j) &= w(\sigma_j \cdot \sigma'_j) + \frac{L - |\sigma_j \cdot \sigma'_j|}{|\lambda_j|} \cdot w(\lambda_j) \end{aligned}$$

Assume w.l.o.g. that $\overline{w}(\lambda_i) < \overline{w}(\lambda_j)$. We compute:

$$\begin{aligned} w(\rho_i) \leq w(\rho_j) &\Leftrightarrow \\ w(\sigma_i \cdot \sigma'_i) + \frac{L - |\sigma_i \cdot \sigma'_i|}{|\lambda_i|} \cdot w(\lambda_i) &\leq w(\sigma_j \cdot \sigma'_j) + \frac{L - |\sigma_j \cdot \sigma'_j|}{|\lambda_j|} \cdot w(\lambda_j) \Leftrightarrow (7) \\ \frac{|\lambda_i| \cdot |\lambda_j| \cdot (w(\sigma_i \cdot \sigma'_i) - w(\sigma_j \cdot \sigma'_j)) + |\lambda_i| \cdot |\sigma_j \cdot \sigma'_j| \cdot w(\lambda_j) - |\lambda_j| \cdot |\sigma_i \cdot \sigma'_i| \cdot w(\lambda_i)}{w(\lambda_j) \cdot |\lambda_i| - w(\lambda_i) \cdot |\lambda_j|} &\leq L \end{aligned}$$

Since $w(\lambda_j) \cdot |\lambda_i| - w(\lambda_i) \cdot |\lambda_j| > 0$ and since $w(\lambda_i), w(\lambda_j), |\lambda_i|, |\lambda_j| \in \mathbb{Z}$, we have that $w(\lambda_j) \cdot |\lambda_i| - w(\lambda_i) \cdot |\lambda_j| \geq 1$. By Lemma 12, we have $|\sigma_i \cdot \sigma'_i|, |\sigma_j \cdot \sigma'_j| \leq \|V\|^4$, and moreover, for any path π , $w(\pi) \leq |\pi| \cdot \mu(\mathcal{G})$. Since $1 \leq |\lambda_i|, |\lambda_j| \leq \|V\|$, we compute:

$$\begin{aligned} &\frac{|\lambda_i| \cdot |\lambda_j| \cdot (w(\sigma_i \cdot \sigma'_i) - w(\sigma_j \cdot \sigma'_j)) + |\lambda_i| \cdot |\sigma_j \cdot \sigma'_j| \cdot w(\lambda_j) - |\lambda_j| \cdot |\sigma_i \cdot \sigma'_i| \cdot w(\lambda_i)}{w(\lambda_j) \cdot |\lambda_i| - w(\lambda_i) \cdot |\lambda_j|} \\ &\leq |\lambda_i| \cdot |\lambda_j| \cdot (w(\sigma_i \cdot \sigma'_i) - w(\sigma_j \cdot \sigma'_j)) + |\lambda_i| \cdot |\sigma_j \cdot \sigma'_j| \cdot w(\lambda_j) - |\lambda_j| \cdot |\sigma_i \cdot \sigma'_i| \cdot w(\lambda_i) \\ &\leq 4\mu(\mathcal{G}) \cdot \|V\|^6 \end{aligned}$$

Combining this with equation (7), we infer that $\bar{w}(\lambda_i) < \bar{w}(\lambda_j)$ and $L \geq 4\mu(\mathcal{G}) \cdot \|V\|^6$ only if $w(\rho_i) \leq w(\rho_j)$, or, dually, that $w(\rho_i) > w(\rho_j)$ and $L \geq 4\mu(\mathcal{G}) \cdot \|V\|^6$ only if $\bar{w}(\lambda_i) \geq \bar{w}(\lambda_j)$. Therefore, a path, minimal for its length, which is greater than $4\mu(\mathcal{G}) \cdot \|V\|^6$ must belong to a biquadratic path scheme, whose cycle has minimal average weight, among all path schemes, to which the path may belong. \square

The following lemma shows that the minbiq sets are invariant for fixed vertices and lengths that belong to certain arithmetic progressions.

Lemma 14 *Given vertices u and v , for each arithmetic progression $\{\ell_k\}_{k=0}^\infty$ of rate $\text{lcm}(1, \dots, \|V\|)$ and $\ell_0 \geq \|V\|^4$, $\text{minbiq}(u, v, \ell_i) = \text{minbiq}(u, v, \ell_j)$, for all $i, j \geq 0$.*

Proof: Let $C = \text{lcm}(1, \dots, \|V\|)$. It is sufficient to show $\text{biq}(u, v, \ell_k) = \text{biq}(u, v, \ell_{k+1})$, for all $k \geq 0$. Let $\theta = \sigma.\lambda^*.\sigma' \in \text{biq}(u, v, \ell_0 + kC)$ be a path scheme. Clearly, $\ell_0 + kC = |\sigma.\sigma'| + p \cdot |\lambda|$ for some $p \in \mathbb{N}$. Since θ is bi-quadratic, then $|\sigma.\sigma'| \leq \|V\|^4$. Since $\ell_0 \geq \|V\|^4$, we obtain that $\ell_0 \geq |\sigma.\sigma'|$. As a consequence, $p \cdot |\lambda| \geq kC$. Thus, $p \geq \frac{kC}{|\lambda|}$ and hence $p' = p - \frac{kC}{|\lambda|} \geq 0$. Observe that $p' \in \mathbb{Z}$ because $|\lambda| \in \{1, \dots, \|V\|\}$ (λ is an elementary cycle) and $|\lambda|$ divides C . We can define a path $\rho = \sigma.\lambda^{p'}.\sigma'$, such that:

$$|\rho| = |\sigma.\sigma'| + p' \cdot |\lambda| = |\sigma.\sigma'| + p \cdot |\lambda| - kC = \ell_0.$$

Thus, we have $\theta \in \text{biq}(u, v, \ell_0)$ and since $\theta \in \text{biq}(u, v, \ell_0 + kC)$ was an arbitrary path, we have $\text{biq}(u, v, \ell_0 + kC) \subseteq \text{biq}(u, v, \ell_0)$, for all $k \in \mathbb{N}$. The other direction is trivial, by taking $k = 0$. \square

We denote by $\text{minw}_{\mathcal{G}}(u, v, \ell)$ the minimal weight among the paths of length ℓ , between vertices u and v in \mathcal{G} , or ∞ if no such path exists. The next lemma proves that the minimal weights corresponding to a certain arithmetic progression of lengths form an arithmetic progression as well.

Lemma 15 *Given two vertices u and v , for any $\ell_0 > 4\mu(\mathcal{G}) \cdot \|V\|^6$ there exists an arithmetic progression $\{\ell_k\}_{k=0}^\infty$ such that the sequence $\{\text{minw}_{\mathcal{G}}(u, v, \ell_k)\}_{k=0}^\infty$ forms an arithmetic progression.*

Proof: It is sufficient to show that, there exists an integer $c > 0$ such that, for any $\ell_0 > 4\mu(\mathcal{G}) \cdot \|V\|^6$, there exists $r \in \mathbb{Z}$ such that $\text{minw}(u, v, \ell_0 + (k+1)c) = r + \text{minw}(u, v, \ell_0 + kc)$, for all $k \geq 0$. Let $c = \text{lcm}(1, \dots, \|V\|)$. By Lemma 14 we have that $\text{minbiq}(u, v, \ell_0) = \text{minbiq}(u, v, \ell_0 + kc)$, for all $k \geq 0$.

We distinguish two cases. First, $\text{minw}(u, v, \ell_0 + kc) = \infty$, i.e. $\text{minbiq}(\ell_0 + kc, u, v) = \text{minbiq}(u, v, \ell_0 + (k+1)c) = \emptyset$, and therefore we obtain $\text{minw}(u, v, \ell_0 + (k+1)c) = \infty$ as well. Second, suppose that $\text{minw}(u, v, \ell_0 + kc) < \infty$. Then there exists a path ρ between u and v , minimal for its length $|\rho| = \ell_0 + kc > 4 \cdot \mu(\mathcal{G}) \cdot \|V\|^6$. By Lemma 13, there exists an equivalent path ρ' and a biquadratic path scheme $\sigma.\lambda^*.\sigma' \in \text{minbiq}(u, v, \ell_0 + kc)$ such that $\rho' = \sigma.\lambda^b.\sigma'$ for some $b \geq 0$. Let ρ'' be the path $\sigma.\lambda^{b+\frac{c}{|\lambda|}}.\sigma'$. We will show that ρ'' is minimal for its length. For, if this is the case, then $|\rho''| = |\rho| + c$ and $w(\rho'') = w(\rho) + c \cdot \bar{w}(\lambda)$, i.e. $\text{minw}(u, v, \ell_0 + kc) = \text{minw}(u, v, \ell_0 + (k+1)c) + c \cdot \bar{w}(\lambda)$. Since $\bar{w}(\lambda)$ is the common average weight of all path schemes in

$\text{minbiq}(u, v, \ell_0 + kc) = \text{minbiq}(u, v, \ell_0 + k'c)$, for any $k, k' \geq 0$, the value of the rate $c \cdot \overline{w}(\lambda)$ does not depend on the value k .

To show that ρ'' is indeed minimal for its length, suppose it is not, and let π'' be a path of length $|\rho''| = \ell_0 + (k+1)c > 4\mu(\mathcal{G}) \cdot \|V\|^6$ such that $w(\pi'') < w(\rho'')$. By Lemma 13, there exists an equivalent path π' and a biquadratic path scheme $\tau.\eta^*.\tau' \in \text{minbiq}(u, v, \ell_0 + (k+1)c) = \text{minbiq}(u, v, \ell_0 + kc)$ (by Lemma 14) such that $\pi' = \tau.\eta^d.\tau'$, for some $d \geq 0$. We define the path $\pi = \tau.\eta^{d-\frac{c}{|\eta|}}.\tau'$, of length $\ell_0 + kc$. We have the following relations:

$$\begin{array}{llll} \rho &= \sigma.\lambda^b.\sigma' & \rho'' &= \sigma.\lambda^{b+\frac{c}{|\lambda|}}.\sigma' & w(\rho) &\leq w(\pi) & w(\rho'') &> w(\pi'') \\ \pi &= \tau.\eta^{d-\frac{c}{|\eta|}}.\tau' & \pi'' &= \tau.\eta^d.\tau' & |\rho| &= |\pi| & |\rho''| &= |\pi''| \end{array}$$

Since $\overline{w}(\lambda) = \overline{w}(\eta)$, we infer that

$$w(\rho'') - w(\rho) = w(\lambda) \cdot \frac{c}{|\lambda|} = \overline{w}(\lambda) \cdot c = \overline{w}(\eta) \cdot c = w(\eta) \cdot \frac{c}{|\eta|} = w(\pi'') - w(\pi) \quad (8)$$

Also, $w(\rho) \leq w(\pi)$ and $w(\pi'') < w(\rho'')$ implies that $w(\rho) + w(\pi'') < w(\pi) + w(\rho'')$ which contradicts equation (8). \square

Proof of Theorem 6 By Lemma 1, the prefix of the sequence $\{M^k\}_{k=0}^\infty$ is $\max_{1 \leq i, j \leq n} b_{ij}$ where b_{ij} is the prefix of the sequence $\{M_{ij}\}_{k=0}^\infty$. Thus it is sufficient to show that $b_{ij} \leq 4\mu(M) \cdot n^6$, for each pair $1 \leq i, j \leq n$. Let $\mathcal{G}_M = (\{1, \dots, n\}, E, w)$ be the weighted graph whose incidence matrix is M and $i, j \in \{1, \dots, n\}$ be two vertices of \mathcal{G}_M . By Lemma 15, there exists an arithmetic progression $\{\ell_k\}_{k=0}^n$, where $\ell_0 = 4\mu(M) \cdot n^6 + 1$, such that $\{\text{minw}_{\mathcal{G}_M}(i, j, \ell_k)\}_{k=0}^\infty$ is an arithmetic progression. Then it must be the case that $b_{ij} \leq 4\mu(M) \cdot n^6$. \square

B Periodic Sequences of Sums, Minima and Half Terms

Lemma 16 Let $\{s_k\}_{k=0}^\infty$ and $\{t_k\}_{k=0}^\infty$ be two periodic sequences, with prefixes b_s, b_t and periods c_s, c_t , respectively. Then the sequence $\{s_k + t_k\}_{k=0}^\infty$ is periodic, and:

$$\exists \lambda_0, \dots, \lambda_{c-1} \quad \forall k \geq 0 \quad \forall i \in [c] \quad s_{b+(k+1)c+i} + t_{b+(k+1)c+i} = \lambda_i + s_{b+kc+i} + t_{b+kc+i}$$

where $c = \text{lcm}(c_s, c_t)$ and $b = \max(b_s, b_t)$.

Proof: Let $\lambda_0^s, \dots, \lambda_{c_s-1}^s$ be the rates of $\{s_k\}_{k=0}^\infty$ and $\lambda_0^t, \dots, \lambda_{c_t-1}^t$ be the rates of $\{t_k\}_{k=0}^\infty$, respectively. For all $i \in [c]$, we define $\lambda_i = \frac{c}{c_s} \lambda_{i \bmod c_s}^s + \frac{c}{c_t} \lambda_{i \bmod c_t}^t$. With these definitions, the required equality is an easy check. \square

Lemma 17 Let $\{s_n\}_{n=0}^\infty$ be a periodic sequence with prefix $b \geq 0$ and period $c > 0$. Then the sequence $\{\lfloor \frac{s_n}{2} \rfloor\}_{n=0}^\infty$ is periodic, and:

$$\exists \lambda_0, \dots, \lambda_{2c-1} \quad \forall k \geq 0 \quad \forall i \in [2c] \quad \left\lfloor \frac{s_{b+(k+1)2c+i}}{2} \right\rfloor = \lambda_i + \left\lfloor \frac{s_{b+k2c+i}}{2} \right\rfloor.$$

Proof: Let $\lambda_0^s, \dots, \lambda_{c-1}^s$ be the rates of the sequence $\{s_n\}_{n=0}^\infty$. We define $\lambda_i = \lambda_{i+c} = \lambda_i^s$, for all $i \in [c]$. Then, we compute:

$$\begin{aligned} \left\lfloor \frac{s_{b+(k+1)2c+i}}{2} \right\rfloor &= \left\lfloor \frac{2\lambda_i^s + s_{b+k \cdot 2c+i}}{2} \right\rfloor = \lambda_i + \left\lfloor \frac{s_{b+k \cdot 2c+i}}{2} \right\rfloor \\ \left\lfloor \frac{s_{b+(k+1)2c+c+i}}{2} \right\rfloor &= \left\lfloor \frac{2\lambda_i^s + s_{b+k \cdot 2c+c+i}}{2} \right\rfloor = \lambda_i + \left\lfloor \frac{s_{b+k \cdot 2c+c+i}}{2} \right\rfloor. \end{aligned}$$

□

Lemma 18 *Let $\{s_k\}_{k=0}^\infty$ and $\{t_k\}_{k=0}^\infty$ be two periodic sequences, with prefixes b_s, b_t and periods c_s, c_t , respectively. Then the sequence $\{\min(s_k, t_k)\}_{k=0}^\infty$ is periodic, and:*

$$\begin{aligned} &\exists b \leq b_{\max} + \max_{i=0}^{c-1} (\text{abs}(s_{b_{\max}+(i \bmod c_s)}) + \text{abs}(t_{b_{\max}+(i \bmod c_t)})) \\ &\exists \lambda_0, \dots, \lambda_{c-1} \quad \forall k \geq 0 \quad \forall i \in [c]. \\ &\min(s_{b+(k+1)c+i}, t_{b+(k+1)c+i}) = \lambda_i + \min(s_{b+kc+i}, t_{b+kc+i}) \end{aligned}$$

where $c = \text{lcm}(c_s, c_t)$ and $b_{\max} = \max(b_s, b_t)$.

Proof: We prove first the following facts, for all

$$k \geq \left\lceil \frac{\text{abs}(s_{b_{\max}+(i \bmod c_s)}) + \text{abs}(t_{b_{\max}+(i \bmod c_t)})}{c} \right\rceil$$

and each $i \in [c]$:

1. if $\frac{\lambda_i^s \bmod c_s}{c_s} < \frac{\lambda_i^t \bmod c_t}{c_t}$ then $s_{b_{\max}+kc+i} \leq t_{b_{\max}+kc+i}$,
2. if $\frac{\lambda_i^s \bmod c_s}{c_s} > \frac{\lambda_i^t \bmod c_t}{c_t}$ then $s_{b_{\max}+kc+i} \geq t_{b_{\max}+kc+i}$,
3. if $\frac{\lambda_i^s \bmod c_s}{c_s} = \frac{\lambda_i^t \bmod c_t}{c_t}$ then

$$s_{b_{\max}+kc+i} - t_{b_{\max}+kc+i} = s_{b_{\max}+(i \bmod c_s)} - t_{b_{\max}+(i \bmod c_t)} + (i \div c_s) \lambda_i^s \bmod c_s - (i \div c_t) \lambda_i^t \bmod c_t.$$

where \div denotes integer division.

Observe that $s_{b_{\max}+kc+i} = s_{b_{\max}+(i \bmod c_s)} + (k \frac{c}{c_s} + i \div c_s) \lambda_i^s \bmod c_s$ and similar for $t_{b_{\max}+kc+i}$. We have thus the following equivalences:

$$\begin{aligned} s_{b_{\max}+kc+i} &\leq t_{b_{\max}+kc+i} \\ s_{b_{\max}+(i \bmod c_s)} + (k \frac{c}{c_s} + i \div c_s) \lambda_i^s \bmod c_s &\leq t_{b_{\max}+(i \bmod c_s)} + (k \frac{c}{c_t} + i \div c_t) \lambda_i^t \bmod c_t \\ kc \left(\frac{\lambda_i^s \bmod c_s}{c_s} - \frac{\lambda_i^t \bmod c_t}{c_t} \right) &\leq t_{b_{\max}+(i \bmod c_s)} - s_{b_{\max}+(i \bmod c_s)} + (i \div c_t) \lambda_i^t \bmod c_t - (i \div c_s) \lambda_i^s \bmod c_s. \end{aligned}$$

Under the assumption of this first point, we have:

$$kc \geq \text{abs}(s_{b_{\max}+(i \bmod c_s)}) + \text{abs}(t_{b_{\max}+(i \bmod c_t)}) \Rightarrow s_{b_{\max}+kc+i} \leq t_{b_{\max}+kc+i}$$

The second point is symmetric. We obtain the last point by a similar argument. The statement of the lemma follows, with the definition below. For all $i \in [c]$:

$$\lambda_i = \begin{cases} c \frac{\lambda_i^s \bmod c_s}{c_s} & \text{if } \frac{\lambda_i^s \bmod c_s}{c_s} < \frac{\lambda_i^t \bmod c_t}{c_t} \\ c \frac{\lambda_i^t \bmod c_t}{c_t} & \text{if } \frac{\lambda_i^s \bmod c_s}{c_s} > \frac{\lambda_i^t \bmod c_t}{c_t} \\ 0 & \text{otherwise} \end{cases}$$

Observe that, if $b = b_{\max} + \max_{i=0}^{c-1} (\text{abs}(s_{b_{\max}+(i \bmod c_s)}) + \text{abs}(t_{b_{\max}+(i \bmod c_t)}))$, then $b + kC + i = b_{\max} + \left(\left\lceil \frac{\max_{i=0}^{c-1} (\text{abs}(s_{b_{\max}+(i \bmod c_s)}) + \text{abs}(t_{b_{\max}+(i \bmod c_t)}))}{c} \right\rceil + k \right) c + i$, and a case split based on the above three facts can be applied. \square

Lemma 19 *Let $\{s_k^1\}_{k=0}^\infty, \dots, \{s_k^n\}_{k=0}^\infty$ be periodic sequences with prefixes b_1, \dots, b_n , periods c_1, \dots, c_n and rates $\lambda_0^1, \dots, \lambda_{c_1-1}^1, \dots, \lambda_0^n, \dots, \lambda_{c_n-1}^n$, respectively. Let m_k^1, \dots, m_k^ℓ be linear combinations of s_k^1, \dots, s_k^n , respectively. Then the sequence $\{\min(m_k^1, \dots, m_k^\ell)\}_{k=0}^\infty$ is periodic, with prefix at most b and period that divides c , where:*

- $b \leq \max_{i=1}^n (b_i) + n \cdot \max_{i=0}^{c-1} (\sum_{j=1}^\ell \text{abs}(m_j(s_{b_{\max}+i}^1, \dots, s_{b_{\max}+i}^n)))$ and
- $c = \text{lcm}_{i=1}^n (c_i)$.

Proof: Applying Lemma 16, we obtain that, each sequence $\{m_k^i\}_{k=0}^\infty$, for $i = 1, \dots, \ell$, is periodic, with prefix at most $\max_{i=1}^n (b_i)$ and period which divides c . The upper bound on the prefix and period of $\{\min(m_k^1, \dots, m_k^\ell)\}_{k=0}^\infty$ is obtained by applying n times Lemma 18. \square