Safety Problems are NP-complete for Flat Integer Programs with Octagonal Loops

Marius Bozga[1], Radu Iosif[1], and Filip Konecny[2]

[1] VERIMAG/CNRS, Grenoble, France
[2] École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

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 $\{R^i\}_{i=1}^{\infty}$ of such relations can be encoded as a periodic sequence of matrices, and that both the prefix and the period of this sequence are $2^{O(\|R\|_2)}$ in the size of the binary encoding $\|R\|_2$ 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 [8, 11], 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 [19], hence, in theory, any program can be viewed as a counter machine. In practice, effective reductions to counter systems have been designed for programs with dynamic heap data structures [4], arrays [6], dynamic thread creation and shared memory [2], etc. Since counter machines with only two variables are Turing-complete [19], all their decision problems (reachability, termination) are undecidable. This early negative result motivated researchers to find classes of systems with decidable problems, such as: (branching) vector addition systems [15, 17], reversal-bounded counter machines [16], Datalog programs with gap-order constraints [21], and flat counter machines [3, 11, 8]. Despite the fact that reachability of a set of configurations is decidable for these classes, few of them are actually supported by tools, and used for real-life verification purposes. The main reason is that the complexities of the reachability problems for these systems are, in general, prohibitive. Thus, most software verifiers rely on incomplete algorithms, which, due to loss of precision, raise large numbers of false alarms.

We study the complexity of the reachability problems for a class of *flat counter machines* (i.e., the control structure forbids nested loops), in which the transitions occurring inside loops are all labeled with *difference bounds constraints*, i.e. conjunctions of linear inequalities of the form $x - y \leq c$ where $x, y \in \mathbf{x} \cup \mathbf{x}'$ and $c \in \mathbb{Z}$ is a constant. Furthermore, we extend the result to the case of octagonal relations, which are conjunctions of the form $\pm x \pm y \leq c$.

The decidability of the reachability problem for these classes relies on the fact that the transitive closures R^+ of relations R, defined by difference bounds and octagonal constraints, are expressible in Presburger arithmetic [11]. In [8], we presented a concise proof of this fact, based on the observation that any sequence of powers $\{R^i\}_{i=1}$, can be encoded as a *periodic sequence* of matrices, which can be defined by a quantifier-free Presburger formula whose size depends on the prefix and the period of the matrix sequence. In this paper we show primarily that both the prefix and period and this sequence are of the order of $2^{O(\|R\|_2)}$, where $\|R\|_2$ is the size of the binary encoding of the relation. More precisely, the quantifier-free Presburger formula defining a transitive closure (and, implicitly, the reachability problem for the counter machine) has $2^{O(\|R\|_2)}$ many disjuncts of polynomial size. A non-deterministic Turing machine that solves the reachability problem can guess, for each loop relation R, the needed disjunct of R^+, and validate its guess in PTIME($\|R\|_2$).

The main outcome of this result is the definition of a non-trivial class of counter machines, for which the safety problems can be decided relatively easy e.g., by using powerful Satisfiability Modulo Theories (SMT) solvers.

Related Work The complexity of safety, and, more generally, temporal logic properties of integer counter machines has received relatively little attention. For instance, the exact complexity of reachability for vector addition systems (VAS) is an open problem (the only known upper bound is non-primitive recursive), while the coverage and boundedness problems are EXPSPACE-complete for VAS [20], and 2EXPTIME-complete for branching VAS [15]. On what concerns counter machines with gap-order constraints (a restriction of difference bounds constraints $x-y \leq c$ to the case $c \geq 0$), reachability is PSPACE-complete [10], even in the absence of the flatness restriction on the control structure. Our result is incomparable to [10], as we show NP-completeness for flat counter machines with more general[3], difference bounds relations on loops. The results which are probably closest to ours are the ones in [14, 13], where flat counter machines with deterministic transitions of the form $\sum_{i=1}^{n} a_i \cdot x_i + b \leq 0 \wedge \bigwedge_{i=1}^{n} x_i' = x_i + c_i$ are considered. In [14] it is shown that model-checking LTL 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 [13], matching the complexity of CTL* model checking for gap-order constraints [10]. These results are again incomparable with ours, since (i) the linear 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).

2 Preliminary Definitions

We denote by $\mathbb{Z}$, $\mathbb{N}$ and $\mathbb{N}_+$ the sets of integers, positive (including zero) and strictly positive integers, respectively. We denote by $\mathbb{Z}_\infty$ and $\mathbb{Z}_{-\infty}$ the sets $\mathbb{Z} \cup \{\infty\}$ and $\mathbb{Z} \cup \{-\infty\}$, respectively.We write $[n]$ for the interval $\{0,\ldots,n-1\}$, $\mathrm{abs}(n)$ for the absolute value of the integer $n \in \mathbb{Z}$, and $lcm(n_1,\ldots,n_k)$ for the least common multiple of $n_1,\ldots,n_k \in \mathbb{N}$. Let $\mathbf{x}$ denote a nonempty set of variables, and $\mathbf{x}' = \{x' \mid x \in \mathbf{x}\}$. A *valuation* of $\mathbf{x}$ is a function $\nu : \mathbf{x} \to \mathbb{Z}$. The set of all such valuations is denoted by $\mathbb{Z}^{\mathbf{x}}$, and we denote by $\mathbb{Z}^N$ the N-times cartesian product $\mathbb{Z} \times \ldots \times \mathbb{Z}$, for some $N > 0$. A *linear term* t over a set of variables $\mathbf{x} = \{x_1,\ldots,x_N\}$ is a linear combination $a_0 + \sum_{i=1}^{N} a_i x_i$, where $a_0, a_1, \ldots, a_N \in \mathbb{Z}$. An *atomic proposition* is a predicate of the form $t \leq 0$ where t is a linear term, $c \in \mathbb{N}_+$ is a constant, and $\equiv_c$ denotes equality modulo c. *Quantifier-free*

[3] The generalization of gap-order to difference bound constraints suffices to show undecidability of non-flat counter machines, hence the restriction to flat control structures is crucial.

Presburger Arithmetic (QFPA) is the set of boolean combinations of atomic propositions of the above form. For a QFPA formula ϕ, let $Atom(\phi)$ denote the set of atomic propositions in ϕ, and $\varphi[t/x]$ denote the formula obtained by substituting the variable x with the term t in φ.

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 $\nu(x)$ is valid. A formula φ is said to be *consistent* if and only if there exists ν such that $\nu \models \varphi$. For two formulae $\phi_1(\mathbf{x})$ and $\phi_2(\mathbf{x})$, we write $\phi_1 \Leftrightarrow \phi_2$ if, for all $\nu \in \mathbb{Z}^{\mathbf{x}}$, $\nu \models \phi_1$ if and only if $\nu \models \phi_2$.

A formula $\phi(\mathbf{x}, \mathbf{x}')$ is evaluated with respect to two valuations $\nu_1, \nu_2 \in \mathbb{Z}^{\mathbf{x}}$, by replacing each occurrence of $x \in \mathbf{x}$ with $\nu_1(x)$ and each occurrence of $x' \in \mathbf{x}'$ with $\nu_2(x)$ in ϕ. The satisfaction relation is denoted by $(\nu_1, \nu_2) \models \phi(\mathbf{x}, \mathbf{x}')$. A formula $\phi_R(\mathbf{x}, \mathbf{x}')$ is said to *define* a relation $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$ whenever for all $\nu_1, \nu_2 \in \mathbb{Z}^{\mathbf{x}}$, $(\nu_1, \nu_2) \in R$ if and only if $(\nu_1, \nu_2) \models \phi_R$. 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 $R_1 \circ R_2$, defined by the formula $\exists \mathbf{y} . \varphi_1(\mathbf{x}, \mathbf{y}) \wedge \varphi_2(\mathbf{y}, \mathbf{x}')$. The *identity relation* $Id_{\mathbf{x}}$ is defined by the formula $\bigwedge_{x \in \mathbf{x}} x' = x$.

Definition 1. *A* class of relations *is a set $\mathcal{R}$ of QFPA formulae $\phi_R(\mathbf{x}, \mathbf{x}')$ defining relations $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$, such that, for any two $\mathcal{R}$-definable relations $R_1, R_2 \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$, there exists a formula $\phi(\mathbf{x}, \mathbf{x}') \in \mathcal{R}$ defining $R_1 \circ R_2$.*

Notice that any set $\mathcal{R}$ of formulae $\varphi(\mathbf{x}, \mathbf{x}')$ that has quantifier elimination is a class of relations. If the class of a relation is not specified a-priori, we consider it to be the set of all QFPA formulae. For any relation $R \subseteq \mathbb{Z}^{\mathbf{x}}$, we define $R^0 = Id_{\mathbf{x}}$ and $R^{i+1} = R^i \circ R$, for all $i \geq 0$. R^i is called the *i-th power* of R in the sequel. With these notations, $R^+ = \bigcup_{i=1}^{\infty} R^i$ denotes the *transitive closure* of R, and $R^* = R^+ \cup Id_{\mathbf{x}}$ denotes the *reflexive and transitive closure* of R.

For a constant $c \in \mathbb{Z}$, we denote by $\|c\|_2 = \lceil \log_2(\mathrm{abs}(c)) \rceil$, if $\mathrm{abs}(c) > 2$ and $\|c\|_2 = 2$, otherwise, the *size of its binary encoding*[4]. Notice that $\|c\|_2 \geq 2$, for every integer $c \in \mathbb{Z}$. The *binary size* of an atomic proposition is defined as $\|a_0 + \sum_{i=1}^N a_i x_i \leq 0\|_2 = \sum_{i=0}^N \|a_i\|_2$ and $\|a_0 + \sum_{i=1}^N a_i x_i \equiv_c 0\|_2 = \sum_{i=0}^N \|a_i\|_2 + \|c\|_2$. The binary size of a QFPA formula φ is defined as $\|\varphi\|_2 = \sum_{p \in Atom(\varphi)} \|p\|_2$, and gives (a faithful underapproximation of) the number of bits needed to represent φ[5]. It is known that the satisfiability problem for QFPA is NP-complete in the binary size of the formula [23]. The binary size of an $\mathcal{R}$-definable[6] relation R is $\|R\|_2^{\mathcal{R}} = \min\{\|\phi_R\|_2 \mid \phi_R \in \mathcal{R},\ \phi_R \text{ defines } R\}$. When the class of a relation is obvious from the context, it will be omitted. If φ is a QFPA formula, let $\nabla(\varphi)$ denote the sum of the absolute values of its coefficients, formally $\nabla(a_0 + \sum_{i=1}^N a_i x_i \leq 0) = \sum_{j=0}^N \mathrm{abs}(a_j)$, $\nabla(a_0 + \sum_{i=1}^N a_i x_i \equiv_c 0) = \sum_{j=0}^N \mathrm{abs}(a_j) + c$ and $\nabla(\varphi) = \sum_{p \in Atom(\varphi)} \nabla(p)$. Since every formula φ has at least one non-null coefficient, we have $\nabla(\varphi) > 0$. The following relates the sum of absolute values to the binary size of a formula:

Proposition 1. *For every QFPA formula φ, we have $\|\varphi\|_2 \geq \log_2(\nabla(\varphi))$.*

---

[4] Abstracting from particular machine representations, we assume that at least 2 bits are needed to encode each integer.

[5] We consider classical encodings of formulae as strings, and do not deal with issues related to data compression.

[6] The class $\mathcal{R}$ is relevant here, because the same relation can be defined by a smaller formula not in $\mathcal{R}$

Proof: Let $c_0, \ldots, c_n \in \mathbb{Z}$ be the coefficients of φ. We show the following inequality:

$$\|\varphi\|_2 = \sum_{i=0}^{n} \|c_i\|_2 \geq \log_2(\sum_{i=0}^{n} \mathrm{abs}(c_i)) = \log_2(\nabla(\varphi)) \tag{1}$$

by induction on $n \geq 0$. The case $n = 0$ is trivial. For $n > 0$, we have, by the induction hypothesis:

$$\begin{array}{lll} \sum_{i=0}^{n} \|c_i\|_2 & \geq \log_2(\sum_{i=0}^{n-1} \mathrm{abs}(c_i)) + \|c_n\|_2 & \\ & \geq \log_2(\sum_{i=0}^{n-1} \mathrm{abs}(c_i)) + \log_2(\mathrm{abs}(c_n)) & \text{if } \mathrm{abs}(c_n) > 2 \\ & \geq \log_2(\sum_{i=0}^{n} \mathrm{abs}(c_i)) & \text{if } \sum_{i=0}^{n-1} \mathrm{abs}(c_i) \geq 2 \end{array}$$

If $0 \leq \mathrm{abs}(c_n) \leq 2$ we have $\|c_n\|_2 = 2$ and hence:

$$\begin{array}{rll} 4 \cdot (\sum_{i=0}^{n-1} \mathrm{abs}(c_i)) \geq & \sum_{i=0}^{n-1} \mathrm{abs}(c_i) + 2 & \geq \sum_{i=0}^{n} \mathrm{abs}(c_i) \\ \log_2(\sum_{i=0}^{n-1} \mathrm{abs}(c_i)) + \|c_n\|_2 = & \log_2(\sum_{i=0}^{n-1} \mathrm{abs}(c_i)) + 2 & \geq \log_2(\sum_{i=0}^{n} \mathrm{abs}(c_i)) \end{array}$$

Otherwise, if $0 \leq \sum_{i=0}^{n-1} \mathrm{abs}(c_i) \leq 1$, we have $0 \leq \mathrm{abs}(c_i) \leq 1$, hence $\|c_i\|_2 = 2$ for all $i = 0, \ldots, n-1$. It suffices to show that:

$$\sum_{i=0}^{n} \|c_i\|_2 = 2n + \|c_n\|_2 \geq \log_2(\mathrm{abs}(c_n) + 1) \geq \log_2(\sum_{i=0}^{n} \mathrm{abs}(c_i))$$

For $\|c_n\|_2 = 0, 1, 2$ we have $\|c_n\|_2 \geq \log_2(\mathrm{abs}(c_n) + 1)$. For $\|c_n\|_2 > 2$, we have $2 + \log_2(\mathrm{abs}(c_n)) > \log_2(\mathrm{abs}(c_n) + 1)$. $\square$

3 The Reachability Problem for Flat Counter Machines

In this section we define *counter machines*, which are essentially a generalization of integer programs, by allowing non-determinism, and the possibility of describing the program steps by Presburger formulae. Since the class of counter machines with only two counters with increment, decrement and zero test is already Turing-complete [19], we consider a decidable class of *flat counter machines* [3, **?**,**?**], by forbidding nested loops in the control structure of the machine. By further restricting the relations on the loops to several classes of conjunctive formulae (e.g. difference bounds, octagons, finite monoid affine relations) we obtain that reachability is decidable [3, 9, 8]. In this paper we strengthen the decidability results by showing that the reachability problem is in fact NP-complete for flat counter machines with loops labeled by difference bounds and octagonal relations. Formally, a counter machine is a tuple $M = \langle \mathbf{x}, \mathcal{L}, \ell_{init}, \ell_{fin}, \Rightarrow, \Lambda \rangle$, where $\mathbf{x}$ is a set of first-order variables ranging over $\mathbb{Z}$, $\mathcal{L}$ is a set of *control locations*, $\ell_{init}, \ell_{fin} \in \mathcal{L}$ are *initial* and *final* control locations, $\Rightarrow$ is a set of *transition rules* of the form $\ell \overset{R}{\Rightarrow} \ell'$, where $\ell, \ell' \in \mathcal{L}$ are control locations, and $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$ is a relation, and $\Lambda(\ell \overset{R}{\Rightarrow} \ell')$ gives the class of R. A *loop* is a path in the control graph $\langle \mathcal{L}, \Rightarrow \rangle$ of M, where the source and the destination locations are the same, and every transition rule appears only once. A counter machine is said to be *flat* if and only if every control location is the source/destination of at most one loop. The *binary size* of a counter machine M is $\|M\|_2 = \sum_{\ell \overset{R}{\Rightarrow} \ell'} \|R\|_2^{\Lambda(\ell \overset{R}{\Rightarrow} \ell')}$.

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 counters. A *run of M to ℓ* is a sequence of configurations $(\ell_0, \nu_0), \ldots, (\ell_k, \nu_k)$, of length $k \geq 0$, where $\ell_0 = \ell_{init}$, $\ell_k = \ell$, and for each $i = 0, \ldots, k-1$, there exists a transition rule $\ell_i \stackrel{R_i}{\Rightarrow} \ell_{i+1}$ such that $(\nu_i, \nu_{i+1}) \in R_i$. If ℓ is not specified, we assume $\ell = \ell_{fin}$, and say that the sequence is a *run of M*.

The *reachability problem* asks, given a counter machine M, whether *there exists a run in M?* This problem is, in general, undecidable [19], and it is decidable for flat counter machines whose loops are labeled only with certain, restricted, classes of QFPA relations, such as difference bounds (Def. 7) or octagons (Def. 10). The crux of the decidability proofs in these cases is that the transitive closure of any relation of the above type can be defined in QFPA, and is, moreover, effectively computable (see [8] for an algorithm). The goal of this paper is to provide tight bounds on the complexity of the reachability problem in these decidable cases. The parameter of the decision problem is the binary size of the input counter machine M, i.e. $\|M\|_2$. The following theorem proves decidability of the reachability problem for flat counter machines, under the assumption that the composition L of the relations on every loop in a counter machine has a *QFPA-definable transitive closure*.

Theorem 1 ([9, 8, 3]). *The reachability problem is decidable for any class of counter machines $\mathcal{M} = \{M$ flat counter machine $|$ for each loop $q_1 \stackrel{R_1}{\Rightarrow} \ldots \stackrel{R_n}{\Rightarrow} q_1$ in M, the transitive closure $(R_1 \circ \ldots \circ R_n)^+$ is QFPA-definable$\}$*

Proof: Let $M = \langle \mathbf{x}, \mathcal{L}, \ell_1, \ell_n, \Rightarrow, \Lambda \rangle$ be a flat counter machine, where $\mathcal{L} = \{\ell_1, \ldots, \ell_n\}$. First, we reduce the control graph $\langle \mathcal{L}, \Rightarrow \rangle$ of M to a dag (and several self-loops), by replacing each non-trivial loop of M:

$$\ell_{i_0} \stackrel{R_0}{\Rightarrow} \ell_{i_1} \stackrel{R_1}{\Rightarrow} \ell_{i_2} \ \ldots \ \ell_{i_{k-2}} \stackrel{R_{k-2}}{\Rightarrow} \ell_{i_{k-1}} \stackrel{R_{k-1}}{\Rightarrow} \ell_{i_0}$$

where $k > 1$, with the following:

$$\overset{L_0(\mathbf{x},\mathbf{x}')}{\overset{\curvearrowright}{\ell_{i_0}}} \stackrel{R_0}{\Rightarrow} \overset{L_1(\mathbf{x},\mathbf{x}')}{\overset{\curvearrowright}{\ell_{i_1}}} \quad \ldots \quad \overset{L_{k-1}(\mathbf{x},\mathbf{x}')}{\overset{\curvearrowright}{\ell_{i_{k-1}}}} \stackrel{R_{k-1}}{\Rightarrow} \ell'_{i_0} \stackrel{R_0}{\Rightarrow} \ldots \stackrel{R_{k-1}}{\Rightarrow} \ell'_{i_{k-1}} \tag{2}$$

where $L_j = R_j \circ \ldots \circ R_{k-1} \circ R_0 \circ \ldots \circ R_{j-1}$, and $\ell'_{i_1}, \ldots, \ell'_{i_{k-1}}$ are fresh control locations not in $\mathcal{L}$, and for each rule $\ell_{i_j} \Rightarrow \ell_m$ of M, where $m \neq i_{(j+1) \bmod k}$, we add a rule $\ell'_{i_j} \stackrel{\phi}{\Rightarrow} \ell_m$, for each $j = 0, \ldots, k-1$. This operation doubles at most the number of control locations in $\mathcal{L}$. Without loss of generality, we can consider henceforth that each control location ℓ_i belongs to at most one self loop labeled by a formula $L_i(\mathbf{x}, \mathbf{x}')$, whose transitive closure L_i^* is QFPA-definable.

The second phase of the reduction uses a simple breadth-first dag traversal algorithm to label each control location in $\ell_i \in \mathcal{L}$ with a QFPA formula $\sigma_i(\mathbf{x}, \mathbf{x}')$ 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 $\ell_i \in \mathcal{L}$ there exists a path in M from ℓ_1 to ℓ_i, and (ii) there are no rules of the form $\ell_j \Rightarrow \ell_1$ i.e, no self-loop involving ℓ_1 in M. We define:

$$\begin{aligned} \sigma_1 &\equiv Id_{\mathbf{x}} \\ \sigma_j &\equiv \Big(\bigvee_{\ell_i \stackrel{R_{ij}}{\Rightarrow} \ell_j} \sigma_i \circ R_{ij}\Big) \circ L_j^*, \text{ for } j = 2, \ldots, n \end{aligned} \tag{3}$$

Since for every location in $\mathcal{L}$ 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 σ_i before ℓ_j is visited by the algorithm, ensuring that (3) is a proper definition. Moreover, the fact that the structure is essentially a dag guarantees that it is sufficient to visit each node only once in order to label each location with a summary.

Claim. Let $M = \langle \mathbf{x}, \mathcal{L}, \ell_1, \ell_n, \Lambda \Rightarrow \rangle$ be a flat counter machine, and $\sigma_1, \ldots, \sigma_n$ be the labeling of the control locations $\ell_1, \ldots, \ell_n \in \mathcal{L}$, respectively, as defined by (3). Then, for all $\nu, \nu' \in \mathbb{Z}^{\mathbf{x}}$ and $\ell_i \in \mathcal{L}$:

$$(\nu, \nu') \models \sigma_i \text{ if and only if } M \text{ has a run } (\ell_1, \nu), \ldots, (\ell_i, \nu')$$

Proof: "$\Rightarrow$" By induction on the maximum number $m > 0$ of control locations on each path between ℓ_1 and ℓ_i in $\langle \mathcal{L}, \Rightarrow \rangle$. If $m = 1$ then $i = 1$ is the only possibility, and consequently, we have $\sigma_i \equiv Id_{\mathbf{x}}$, by (3). But then $\nu = \nu'$, and (ℓ_1, ν) is a run of M. If $m > 1$, then σ_i is defined according to (3), and there exists $\ell_j \in \mathcal{L}$, $j \neq i$, such that $\ell_j \overset{R_{ji}}{\Rightarrow} \ell_i$ is a transition rule in M, and $(\nu, \nu') \models \sigma_j \circ R_{ji} \circ L_i^*$. Since $\langle \mathcal{L}, \Rightarrow \rangle$ is a dag, the maximum number of control locations on each path from ℓ_1 to ℓ_j is less than m, and, by the induction hypothesis, for each $\overline{\nu} \in \mathbb{Z}^{\mathbf{x}}$, there exists a run $(\ell_1, \nu), \ldots, (\ell_j, \overline{\nu})$ in M if and only if $(\nu, \overline{\nu}) \models \sigma_j$. As $(\nu, \nu') \models \sigma_j \circ R_{ji} \circ L_i^*$, there exist valuations $\nu'', \nu''' \in \mathbb{Z}^{\mathbf{x}}$, such that $(\nu, \nu'') \models \sigma_j$, $(\nu'', \nu''') \models R_{ji}$ and $(\nu''', \nu') \models L_j^*$. Then M has a run $(\ell_1, \nu), \ldots, (\ell_j, \nu''), (\ell_i, \nu'''), \ldots, (\ell_i, \nu')$.

"$\Leftarrow$" By induction on the number $m > 0$ of control locations that occur on the run. If $m = 1$, the only possibility is that the run consists of one configuration (ℓ_1, ν), and $\sigma_1 \equiv Id_{\mathbf{x}}$, by (3). But then $(\nu, \nu) \models \sigma_1$, for every valuation $\nu \in \mathbb{Z}^{\mathbf{x}}$. If $m > 1$, the run is of the form $(\ell_1, \nu), \ldots, (\ell_j, \nu_{k-1}), (\ell_i, \nu_k), \ldots$
(ℓ_i, ν'), where (ℓ_i, ν_k) is the first occurrence of ℓ_i on the run, and in between $(\ell_i, \nu_k), \ldots, (\ell_i, \nu')$, all control locations are ℓ_i. By the induction hypothesis, $(\nu, \nu_{k-1}) \models \sigma_j$. Since $\ell_j \overset{R_{ji}}{\Rightarrow} \ell_i$ is a transition rule of M, we have $(\nu_{k-1}, \nu_k) \models R_{ji}$, and, moreover, $(\nu_k, \nu') \models L_i^*$. Hence $(\nu, \nu') \models \sigma_i$, by (3). □

It is now manifest that M has a run if and only if the summary σ_n corresponding to its final control location ℓ_n is satisfiable. If all transitive closures L_i^* occurring in (3) are QFPA-definable, the reachability problem for a flat counter machine M is decidable.

□

4 Periodic Relations

We introduce a notion of periodicity on classes of relations that can be naturally represented as matrices. In general, an infinite sequence of integers is said to be *periodic* if the elements of the sequence beyond a certain threshold (prefix), and which are situated at equal distance (period) one from another, differ by the same quantity (rate). This notion of periodicity is lifted to matrices of integers, entry-wise. Assuming that each power R^k of a relation R is represented by a matrix M_k, R is said to be periodic if the infinite sequence $\{M_k\}_{k=0}^{\infty}$ of matrix representations of powers of R is periodic. Periodicity guarantees that the sequence has an infinite subsequence which can be captured by a QFPA formula, which thus defines infinitely many powers of the relation. Then, the remaining powers can be computed by composing this formula with only finitely many (i.e., the size of the period) powers of the relation.

Example 1. For instance, consider the relation $R\ :\ x' = y+1 \wedge y' = x$. This relation is periodic, and we have $R^{2k+1}\ :\ x' = y+k+1 \wedge y' = x+k$ and $R^{2k+2}\ :\ x' = x+k \wedge y' = y+k$, for all $k \geq 0$.

For two matrices $A, B \in \mathbb{Z}_\infty^{m\times m}$, we define the sum $(A+B)_{ij} = A_{ij} + B_{ij}$.

Definition 2. *An infinite sequence of matrices* $\{A_k\}_{k=1}^{\infty} \in \mathbb{Z}_\infty^{m\times m}$ *is said to be* periodic *if and only if there exist integers* $b, c > 0$ *and matrices* $\Lambda_0, \ldots, \Lambda_{c-1} \in \mathbb{Z}_\infty^{m\times m}$ *such that* $A_{b+(k+1)c+i} = \Lambda_i + A_{b+kc+i}$*, for all* $k \geq 0$ *and* $i \in [c]$.

The smallest integers b, c are called the *prefix* and the *period* of the sequence. The matrices Λ_i, corresponding to the prefix-period pair (b, c), are called the *rates* of the sequence. A relation R is said to be $*$-*consistent* if and only if $R^n \neq \emptyset$, for all $n > 0$.

Definition 3. *A class of relations* $\mathcal{R}$ *is said to be* periodic *iff there exist two functions* $\sigma : \mathcal{R} \to \bigcup_{m>0} \mathbb{Z}_\infty^{m\times m}$ *and* $\rho : \bigcup_{m>0} \mathbb{Z}_\infty^{m\times m} \to \mathcal{R}$*, such that* $\rho(\sigma(\phi)) \Leftrightarrow \phi$*, for each formula* $\phi \in \mathcal{R}$*, and for any* $*$*-consistent relation R defined by a formula from* $\mathcal{R}$*, the sequence of matrices* $\{\sigma(R^i)\}_{i=1}^{\infty}$ *is periodic.*

If R is a $*$-consistent relation, the prefix, period $b, c > 0$ and rates $\Lambda_0, \ldots, \Lambda_{c-1} \in \mathbb{Z}^{m\times m}$ of the $\{\sigma(R^i)\}_{i=1}^{\infty}$ sequence are called the *prefix*, *period* and *rates* of R, respectively. Otherwise, if R is not $*$-consistent, we convene that its prefix is the smallest $b \geq 0$ such that $R^b = \emptyset$, and its period is one.

Definition 4. *Let* $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$ *be a relation. The* closed form *of R is the formula* $\widehat{R}(k, \mathbf{x}, \mathbf{x}')$*, where* $k \notin \mathbf{x}$*, such that the formula* $\widehat{R}[n/k]$ *defines* R^n*, for all* $n > 0$*.*

If $\mathcal{R}$ is a class of relations, let $\mathcal{R}[k]$ denote the set of closed forms of relations defined by formulae in $\mathcal{R}$[7]. Let $\mathbb{Z}[k]_\infty^{m\times m}$ be the set of matrices $M[k]$ of univariate linear terms, i.e. $M_{ij} \equiv a_{ij} \cdot k + b_{ij}$, where $a_{ij}, b_{ij} \in \mathbb{Z}$, for all $1 \leq i, j \leq m$ or $M_{ij} = \infty$. In addition to the σ and ρ functions from Def. 3, we consider a function $\pi : \bigcup_{m>0} \mathbb{Z}[k]_\infty^{m\times m} \to \mathcal{R}[k]$, mapping matrices into formulae $\phi(k, \mathbf{x}, \mathbf{x}')$ such that $\pi(M)[n/k] \Leftrightarrow \rho(M[n/k])$, for all $n > 0$. The following lemma characterizes the closed form of a periodic relation, by defining an infinite periodic subsequence of powers of the form $\{R^{kc+b+i}\}_{k\geq 0}$, for some $b, c > 0$ and $i \in [c]$.

Lemma 1. *Let R be a* $*$*-consistent periodic relation,* $b, c > 0$ *be integers, and* Λ_i *be matrices such that* $\sigma(R^{b+c+i}) = \Lambda_i + \sigma(R^{b+i})$*, for all* $i \in [c]$*. Then the following are equivalent, for all* $i \in [c]$*:*

1. $\forall k \geq 0\ .\ \widehat{R}(k \cdot c + b + i, \mathbf{x}, \mathbf{x}') \ \Leftrightarrow\ \pi(k \cdot \Lambda_i + \sigma(R^{b+i}))$
2. $\forall k \geq 0\ \exists \mathbf{y}\ .\ \pi(k \cdot \Lambda_i + \sigma(R^{b+i}))(\mathbf{x}, \mathbf{y}) \wedge \sigma(\rho(R^c))(\mathbf{y}, \mathbf{x}') \ \Leftrightarrow\ \pi((k+1) \cdot \Lambda_i + \sigma(R^{b+i}))(\mathbf{x}, \mathbf{x}')$

[7] The closed form of a QFPA-definable relation can always be defined in first-order arithmetic, using Gödel's encoding of integer sequences, and is not, in general, equivalent to a QFPA formula.

Proof: "(1) ⇒ (2)" Let $n \geq 0$ be an arbitrary integer. We compute:

$$\begin{aligned}\exists \mathbf{y} \,.\, \pi(n \cdot \Lambda_i + \sigma(R^{b+i}))(\mathbf{x},\mathbf{y}) \wedge \widehat{R}(c,\mathbf{y},\mathbf{x}') &\Leftrightarrow \widehat{R}(nc+b+i,\mathbf{x},\mathbf{y}) \wedge \sigma(\rho(R^c))(\mathbf{y},\mathbf{x}') \\ &\Leftrightarrow \widehat{R}((n+1)c+b+i,\mathbf{x},\mathbf{x}') \\ &\Leftrightarrow \pi((n+1)\cdot\Lambda_i + \sigma(R^{b+i}))(\mathbf{x},\mathbf{x}')\end{aligned}$$

"(2) ⇒ (1)" We prove, by induction on $n \geq 0$, that:

$$\widehat{R}(nc+b+i,\mathbf{x},\mathbf{x}') \Leftrightarrow \pi(n \cdot \Lambda_i + \sigma(R^{b+i}))(\mathbf{x},\mathbf{x}')$$

The base case $n = 0$ follows from Def. 3. For the induction step, we compute:

$$\begin{aligned}\widehat{R}((n+1)c+b+i,\mathbf{x},\mathbf{x}') &\Leftrightarrow \exists \mathbf{y} \,.\, \widehat{R}(nc+b+i,\mathbf{x},\mathbf{y}) \wedge \sigma(\rho(R^c))(\mathbf{y},\mathbf{x}') \\ &\Leftrightarrow \exists \mathbf{y} \,.\, \pi(n\cdot\Lambda_i + \sigma(R^{b+i}))(\mathbf{x},\mathbf{y}) \wedge \sigma(\rho(R^c))(\mathbf{y},\mathbf{x}') \quad \text{by the induction hypothesis} \\ &\Leftrightarrow \pi((n+1)\cdot\Lambda_i + \sigma(R^{b+i}))(\mathbf{x},\mathbf{x}') \quad \text{by point (2).}\end{aligned}$$

□

Notice that b and c in Lemma 1 are not necessarily the prefix and period of R: b can be an arbitrary integer larger than the prefix, and c may be a multiple of the period.

5 Flat Counter Machines with Periodic Loops

For simplicity's sake, consider first the counter machines with the structure below:

$$\ell_{init} \overset{I(\mathbf{x}')}{\Longrightarrow} \overset{R(\mathbf{x},\mathbf{x}')}{\overset{\curvearrowright}{\ell}} \overset{F(\mathbf{x})}{\Longrightarrow} \ell_{fin} \tag{4}$$

where $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$ is a periodic relation (Def. 3), and $I, F \subseteq \mathbb{Z}^{\mathbf{x}}$ are QFPA-definable sets of valuations. In the following, we give sufficient conditions (Def. 6) under which the reachability problem for the counter machines (4) is NP-complete.

Definition 5. *A class of relations $\mathcal{R}$ is said to be* poly-logarithmic *if and only if there exist integer constants $p,q,r > 0$, depending on $\mathcal{R}$, such that, for all $P,Q,R \in \mathcal{R}$:*

1. $\|R^n\|_2 = O(\|R\|_2^p \cdot (\log_2 n)^q)$*, for all $n > 0$*
2. *the composition $P \circ Q$ can be computed in time $O((\|P\|_2 + \|Q\|_2)^r)$*

If $\mathcal{R}$ is a poly-logarithmic class of relations, it is not difficult to see that there exists a constant $d > 0$, depending of $\mathcal{R}$, such that, for any $\mathcal{R}$-definable relation R, the n-th power R^n can be computed by the fast exponentiation algorithm (Alg. 1) in time $O((\|R\|_2 \cdot \log_2 n)^d)$.

Definition 6. *A class of periodic relations $\mathcal{R}$ is said to be* exponential *if and only if (A) $\mathcal{R}$ is poly-logarithmic, (B) the mappings* σ, ρ *and* π *(Def. 3) are computable in PTIME, and (C) for each $\mathcal{R}$-definable relation $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$:*

1. *there exist integer constants $p,q > 0$, depending on $\mathcal{R}$, such that the prefix and period of R are $b = 2^{O(\|R\|_2^p)}$ and $c = 2^{O(\|R\|_2^q)}$, respectively*
2. *for all $i \in [c]$ and Λ_i such that $\sigma(R^{b+c+i}) = \Lambda_i + \sigma(R^{b+i})$, the second point of Lemma 1 can be checked in NPTIME($\|R\|_2$)*

Algorithm 1 Fast Exponentiation Algorithm

function FASTPOWER(R,n)
  $Q \leftarrow R$
  $P \leftarrow Id_N$
  for $i = 1,\ldots,\lceil \log_2 n \rceil$ **do**
    if the i-th bit of n is 1 **then** [2^i occurs in the binary decomposition of n]
      $P \leftarrow P \circ Q$
    $Q \leftarrow Q \circ Q$ [at this point $Q = R^{2^i}$]
  return P

The idea of the reduction is to show the existence of a non-deterministic Turing machine that computes, in polynomial time, a QFPA formula, which encodes the reachability question. Since the size of this formula is also polynomial (NP $\subseteq$ PSPACE), and the satisfiability of a QFPA formula is an NP-complete problem, it turns out that the reachability problem for the counter machines (4) is in NP. Since I and F can be any QFPA-definable sets, the reachability problem for such counter machines is also NP-hard, by reduction from the satisfiability problem for QFPA.

To start with, observe that the reachability problem for (4) can be stated as the satisfiability of the following formula: $I(\mathbf{x}) \wedge k \geq 0 \wedge \widehat{R}(k,\mathbf{x},\mathbf{x}') \wedge F(\mathbf{x}')$. Since, in general, the closed form $\widehat{R}(k,\mathbf{x},\mathbf{x}')$ is not QFPA-definable, we focus on the case where R is a periodic relation (Def. 3). We distinguish two cases. First, if R is not $*$-consistent i.e., $R^i = \emptyset$ if and only if i is greater or equal than the prefix b of R, the reachability problem for (4) is equivalent to the satisfiability of the formula $I(\mathbf{x}) \wedge \left[\bigvee_{i=0}^{b-1} \rho(\sigma(R^i))\right] \wedge F(\mathbf{x}')$. Second, if R is $*$-consistent, the reachability problem for (4) is equivalent to the satisfiability of the following formula:

$$I(\mathbf{x}) \wedge \big[\underbrace{\bigvee_{i=0}^{b-1} \rho(\sigma(R^i))}_{\text{prefix}} \vee \underbrace{\bigvee_{j=0}^{c-1} k \geq 0 \wedge \pi(k \cdot \Lambda_j + \sigma(R^{b+j}))}_{\text{period}}\big] \wedge F(\mathbf{x}') \tag{5}$$

where $b,c > 0$ are integers, and $\Lambda_0,\ldots,\Lambda_{c-1}$ are matrices meeting the conditions of the second point of Lemma 1. The first disjunct above takes care of the case when the number of iterations of the loop is smaller than the prefix b, and the second one deals with the other case, when $kc+b+j$ iterations of the loop are needed, for some $k \geq 0$ and $j \in [c]$.

To prove that the reachability problem for the counter machines (4), whose loops are labeled by relations from a periodic exponential class $\mathcal{R}$, is in NP, we define a nondeterministic Turing machine $\mathcal{T}$, that decides the reachability problem in time polynomial in $\|R\|_2 + \|I\|_2 + \|F\|_2$. The first guess of $\mathcal{T}$ is whether R is $*$-consistent or not. If the guess was that R is not $*$-consistent, $\mathcal{T}$ guesses further a constant $B = 2^{O(\|R\|_2^p)}$, where $p > 0$ depends on the class $\mathcal{R}$. Then it checks that B is the prefix of R, by computing R^{B-1} and R^B, and checking that $R^{B-1} \neq \emptyset$ and $R^B = \emptyset$. This check can be carried out in time $O((\|R\|_2 \cdot \log_2 B)^d)$, for some $d > 0$, using Alg. 1. Since $B = 2^{O(\|R\|_2^p)}$, the prefix check is polynomial in $\|R\|_2$. The reachability problem can be encoded in QFPA by further guessing $i \in [B]$, and computing the formula $I(\mathbf{x}) \wedge \rho(\sigma(R^i))(\mathbf{x},\mathbf{x}') \wedge F(\mathbf{x}')$. Since $\mathcal{R}$ is a periodic exponential class, $\|R^i\|_2 = O(\|R\|_2^r \cdot (\log_2 i)^s) = O(\|R\|_2^{r+s})$, for some $r,s > 0$, depending on $\mathcal{R}$. Moreover, the binary size of this formula is polynomial in $\|I\|_2 + \|R\|_2 + \|F\|_2$, and the reachability problem, can be answered by $\mathcal{T}$ in NPTIME($\|R\|_2 + \|I\|_2 + \|F\|_2$), in this case.

If, on the other hand, the first guess of $\mathcal{T}$ was that R is $*$-consistent, then $\mathcal{T}$ will further guess constants $B = 2^{O(\|R\|_2^p)}$ and $C = 2^{O(\|R\|_2^q)}$, for $p,q > 0$ depending on $\mathcal{R}$, $0 \leq i \leq B$ and $0 \leq j < C$. Next, it computes the powers R^i, R^{B+j} and R^{B+C+j} in time polynomial in $\|R\|_2$, using Alg. 1, and lets $\Lambda_j = \sigma(R^{B+C+j}) - \sigma(R^{B+j})$. $\mathcal{T}$ establishes further whether the choices of B,C,j and Λ_j are adequate for defining the closed form of the infinite sequence of powers $\{R^{C\cdot k+B+j}\}_{k>0}$, using Lemma 1 (first point). To this end, it must check the condition of the second point of Lemma 1, which by Def. 6 (point C.2) can be done in NPTIME($\|R\|_2$). Next, $\mathcal{T}$ outputs a QFPA formula, by chosing the i-th and j-th disjuncts from (5), and substituting the computed formulae $\rho(\sigma(R^i))$, $\rho(\sigma(R^{B+j}))$ and the matrix Λ_j, which yields a QFPA formula of size polynomial in $\|I\|_2 + \|R\|_2 + \|F\|_2$. The satisfiability problem for this formula, and thus the reachability for counter machines (4), can be solved in NPTIME($\|I\|_2 + \|R\|_2 + \|F\|_2$) by $\mathcal{T}$.

It is not difficult to see that the reachability problem for (4) is NP-hard, by reduction from the satisfiability problem for QFPA [23]: let $I(\mathbf{x})$ be any QFPA formula over $\mathbf{x}$, $R = Id_{\mathbf{x}}$ and $F = true$. Then q_f is reachable from q_i if and only if $I(\mathbf{x})$ is satisfiable. The following theorem generalizes the proof from (4) to general flat counter machines.

Theorem 2. *If $\mathcal{R}$ is a periodic exponential class of relations, the reachability problem for the class $\mathcal{M}_{\mathcal{R}} = \{M$ flat counter machine $\mid$ for all rules $q \xrightarrow{R} q'$ on a loop of M, R is $\mathcal{R}$-definable$\}$ is NP-complete.*

Proof: NP-hardness is by reduction from the satisfiability problem for QFPA. To show that the problem is in NP, let $M \in \mathcal{M}_{\mathcal{R}}$ be a flat counter machine. The reduction builds from M a QFPA formula Φ_M of size $\|\Phi_M\|_2 = O(\|M\|_2^k)$, for some constant $k > 0$ depending on $\mathcal{R}$, such that M has a run from the initial to the final state if and only if Φ_M is satisfiable – the latter condition can be checked by an NP algorithm which guesses a solution of polynomial size in Φ_M and verifies the correctness of the guess. The construction of Φ_M is done along the same lines as the proof of decidability for flat counter machines whose loops have QFPA-definable transitive closures. First, we reduce each loop of M to a path with single self loops, following the idea of (2). Since the class $\mathcal{R}_{DB}$ is poly-logarithmic (Lemma 6), each relation labeling a self-loop can be computed in time $O(\|M\|_2^r)$, for some $r > 0$ depending on $\mathcal{R}$, as the compositions of all relations on that loop. Consequently, the size of each relation R_λ, labeling a self-loop λ in M, is $\|R_\lambda\|_2 = O(\|M\|_2^r)$. A non-deterministic Turing machine will guess first, for each self loop λ, whether R_λ is $*$-consistent or not:

1. In the case R_λ is not $*$-consistent, the Turing machine guesses $b_\lambda = 2^{O(\|M\|_2^p)}$, for some $p > 0$ depending on $\mathcal{R}$, and then validates in PTIME($\|R_\lambda\|_2$) the guess that R_λ was not $*$-consistent, i.e. it checks that $R_\lambda^{b_\lambda} \neq \emptyset$ and $R_\lambda^{b_\lambda+1} = \emptyset$.
2. Otherwise, if R_λ is $*$-consistent, the Turing machine guesses constants $b_\lambda = 2^{O(\|M\|_2^p)}$, $c_\lambda = 2^{O(\|M\|_2^q)}$, $0 \leq i_\lambda \leq b_\lambda$ and $0 \leq j_\lambda < c_\lambda$, for some $p,q > 0$ depending on $\mathcal{R}$. It then validates the guess, by checking the second condition of Lemma 1, which can also be done in NPTIME($\|R_\lambda\|_2$).

In each case, the Turing machine outputs, for each self-loop λ a QFPA formula Ψ_λ of size polynomial in $\|R_\lambda\|_2$. The last step of the reduction is labeling each control state of M by the summary formulae (3), where the Ψ_λ formulae are used instead of the transitive closures:

$$\sigma_j \equiv \Big(\bigvee_{\ell_i \xRightarrow{R_{ij}} \ell_j} \sigma_i \circ R_{ij} \Big) \circ \Psi_{\lambda_j} \tag{6}$$

where $\lambda_j : \ell_j \xRightarrow{R_\lambda} \ell_j$ is the self-loop around control location ℓ_j. The labeling is achieved by visiting each control label of M exactly once, hence its size is polynomial in $\|M\|_2$. By an argument similar to the one used in the proof of Thm. 1 (Claim 3), the reachability problem is reduced to the satisfiability of the summary formula σ_{fin}, corresponding to the final location of M. The latter is of size polynomial in $\|M\|_2$. Hence the reachability problem for the class $\mathcal{M}_{\mathcal{R}}$ is in NP. $\square$

6 The Periodicity of Tropical Matrix Powers

6.1 Weighted Graphs

Weighted graphs are central to the upcoming developments. The main intuition is that the sequence of matrices representing the powers of a difference bounds relation captures *minimal weight paths* of lengths $1,2,3\ldots$ in a weighted graph. Formally a *weighted digraph* 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 $w : E \to \mathbb{Z}$ is a weight function. When G is clear from the context, we denote by $u \xrightarrow{n} v$ the fact that $(u,v) \in E$ and $w(u,v) = n$. Let $\mu(G) = \max\{\mathrm{abs}(n) \mid u \xrightarrow{n} v \text{ in } G\}$ be the maximum absolute value of all weights in G. A matrix $A \in \mathbb{Z}_\infty^{m \times m}$ is the *incidence matrix* of a weighted digraph $G = \langle V, E, w \rangle$ with vertices $V = \{1, \ldots, m\}$ if and only if $A_{ij} = n$, for each edge $i \xrightarrow{n} j$, and $A_{ij} = \infty$ if there is no edge from i to j. A *path* in G is a sequence $\pi : v_0 \xrightarrow{n_1} v_1 \xrightarrow{n_2} v_2 \ldots v_{p-1} \xrightarrow{n_p} v_p$, where $v_{i-1} \xrightarrow{n_i} v_i$ is an edge in E, for each $1 \le i \le p$. A path is *elementary* if for all $1 \le i < j \le p$, we have $v_i = v_j$ only if $i = 1$ and $j = p$. A *cycle* is a path of length greater than zero, whose source and destination vertices are the same. For two paths π and π', such that the final vertex of π coincides with the initial vertex of π', let $\pi.\pi'$ denote their concatenation. For a path π, we denote its length by $|\pi|$, and its weight (the sum of the weights of all edges on π) by $w(\pi)$. Two paths π and π' are said to be *equivalent* if and only if (i) they start and end in the same vertices, (ii) $w(\rho) = w(\rho')$ and (iii) $|\rho| = |\rho'|$. Notice that two equivalent paths may visit different vertices. A path π is *minimal* if and only if, for any path π' between the same vertices, such that $|\pi| = |\pi'|$, we have $w(\pi) \le w(\pi')$. The *average weight* of π is defined as $\overline{w}(\pi) = \frac{w(\pi)}{|\pi|}$. A cycle is said to be *critical* if it has minimal average weight among all cycles of G. For a subset of vertices $W \subseteq V$, we denote by $G_{[W]} = \langle W, E \cap (W \times W), w \cap (W \times \mathbb{Z}) \rangle$ the subgraph of G induced by W. A subgraph $G_{[W]}$ is *strongly connected* if there exists a path between any two distinct vertices $u, v \in W$. $G_{[W]}$ is a *strongly connected component* (SCC) if it is a maximal strongly connected subgraph of G. Each graph can be partitioned in a set of disjoint strongly connected components. The *cyclicity* of a strongly connected component $G_{[W]}$ of G is the greatest common divisor of the lengths of all its elementary critical cycles, or 1, if $G_{[W]}$ contains no cycles.

For two matrices $A, B \in \mathbb{Z}_\infty^{m\times m}$, we define the *tropical product* as $(A \boxtimes B)_{ij} = \min_{k=1}^{m}(a_{ik} + b_{kj})$. We define $A^{\boxtimes^1} = A$, and $A^{\boxtimes^{k+1}} = A^{\boxtimes^k} \boxtimes A$, for all $A \in \mathbb{Z}_\infty^{m\times m}$ and $k > 0$. Let $A \in \mathbb{Z}_\infty^{m\times m}$ be a square matrix, and G be any weighted graph, such that A is the incidence matrix of G. The sequence $\{A^{\boxtimes^k}\}_{k=1}^{\infty}$ of tropical powers of A gives the minimal weights of the paths of lengths $k = 1, 2, \ldots$ between any two vertices in G. The following theorem shows that any sequence of tropical matrix powers is periodic, and provides an accurate characterization of its period.

Theorem 3 ([22]). *Let $A \in \mathbb{Z}_\infty^{m\times m}$ be a matrix, $G = \langle V, E, w\rangle$ be a weighted graph whose incidence matrix is A, and $W_1, \ldots, W_n$ be the partition of G in strongly connected components. The sequence $\{A^{\boxtimes^k}\}_{k=1}^{\infty}$ is periodic, and its period is $lcm(c_1, \ldots, c_n)$, where $c_1, \ldots, c_n$ are the cyclicities of $W_1, \ldots, W_n$, respectively.*

The above theorem does not give an estimate on the prefix of the sequence. Computing an upper bound on the prefix of a sequence of tropical matrix powers is the goal of Section 6.2.

6.2 Bounding the Prefix of a Sequence of Tropical Matrix Powers

Let $G = \langle V, E, w\rangle$ be a weighted digraph. If $\sigma_1, \ldots, \sigma_{k+1}$ are paths, and $\lambda_1, \ldots, \lambda_k$ are pairwise distinct elementary cycles in G, the expression $\theta = \sigma_1.\lambda_1^*.\sigma_2 \ldots \sigma_k.\lambda_k^*.\sigma_{k+1}$ is called a *path scheme*. If $\sum_{i=1}^{k+1} |\sigma_i| \le \mathrm{card}(V)^4$, we say that θ is *biquadratic*. A path scheme encodes the infinite set of paths $[\![\theta]\!] = \{\sigma_1.\lambda_1^{n_1}.\sigma_2 \ldots \sigma_k.\lambda_k^{n_k}.\sigma_{k+1} \mid n_1, \ldots, n_k \in \mathbb{N}\}$. First, we show that all minimal paths are captured by path schemes with numbers of loops which are at most quadratic in the size of the graph:

Lemma 2. *Let $G = \langle V, E, w\rangle$ be a weighted digraph and ρ be a minimal path in G. Then there exists an equivalent path ρ' and a path scheme $\theta = \sigma_1.\lambda_1^* \ldots \sigma_k.\lambda_k^*.\sigma_{k+1}$ in G, such that $\sigma_1, \ldots, \sigma_{k+1}$ are elementary acyclic paths, $k \le card(V)^2$, and $\rho' \in [\![\theta]\!]$.*

Proof: For each vertex $v \in V$, we partition the set of elementary cycles that start and end in v, according to their length. The representative of each equivalence class is chosen to be a cycle of minimal weight in the class. Since the length of each elementary cycle is at most $\mathrm{card}(V)$, there are at most $\mathrm{card}(V)^2$ such equivalence classes.

Let ρ be any minimal path in G. First, notice that ρ can be factorized as:

$$\rho = \sigma_1.\lambda_1 \ldots \sigma_k.\lambda_k.\sigma_{k+1}$$

where $\sigma_1, \ldots, \sigma_{k+1}$ are elementary acyclic paths, and $\lambda_1, \ldots, \lambda_k$ are elementary cycles. This factorization can be achieved by a traversal of ρ while collecting the vertices along the way in a set. The first vertex which is already in the set marks the first elementary cycle. Then we empty the set and continue until the entire path is traversed.

Next, we repeat the following two steps until nothing changes:

1. For all $i = 1, \ldots, k-1$ move all cycles λ_j, $j > i$, starting and ending with the same vertex as λ_i, next to λ_i, in the ascending order of their lengths. The result is a path ρ' of the same length and weight as ρ.
2. Factorize any remaining non-elementary acyclic path $\sigma_i.\sigma_{i+1} \ldots \sigma_{i+j}$ as in the previous.

The loop above is shown to terminate, since the sum of the lengths of the remaining acyclic paths decreases with every iteration. The result is a path of the same length and weight as ρ, which starts and ends in the same vertices as ρ, in which all elementary cycles of the same length are grouped together. Since ρ was supposed to be a minimal path, so is ρ', and moreover, all elementary cycles can be replaced by their equivalence class representatives, without changing neither the length, nor the weight of the path. The result is a path which belongs to a scheme with at most $\mathrm{card}(V)^2$ cycles. $\square$

Second, for every minimal path in the graph, there exists an equivalent path which is captured by a biquadratic path scheme with one loop:

Lemma 3. *Let $G = \langle V, E, w\rangle$ be a weighted digraph and ρ be a minimal path. Then there exists an equivalent path ρ' and a biquadratic path scheme $\sigma.\lambda^*.\sigma'$, such that $\rho' \in [\![\sigma.\lambda^*.\sigma']\!]$.*

Proof: By Lemma 2, for any path ρ in G there exists a path scheme $\theta = \sigma_1.\lambda_1^*.\sigma_2\ldots\sigma_k.\lambda_k^*.\sigma_{k+1}$, such that $\sigma_1,\ldots,\sigma_{k+1}$ are acyclic and $k \leq \mathrm{card}(V)^2$, and a path ρ', starting and ending in the same vertices as ρ, of the same weight and length as ρ, such that $\rho' = \sigma_1.\lambda_1^{n_1}.\sigma_2\ldots\sigma_k.\lambda_k^{n_k}.\sigma_{k+1}$ for some $n_1,\ldots,n_k \geq 0$. Suppose that λ_i is a cycle with minimal average weight among all cycles in the scheme, i.e. $\frac{w(\lambda_i)}{|\lambda_i|} \leq \frac{w(\lambda_j)}{|\lambda_j|}$, 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:

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

It is easy to check that $|\rho''| = |\rho'|$ and $w(\rho'') = w(\rho')$, since ρ' is minimal.

Clearly ρ'' is captured by the path scheme $\rho_1.\lambda_i^*.\rho_2$, where $\rho_1 = \sigma_1.\lambda_1^{q_1}.\sigma_2\ldots\sigma_{i-1}$ and $\rho_2 = \sigma_{i+1}\ldots.\sigma_k.\lambda_k^{q_k}.\sigma_{k+1}$. Since $\sigma_1,\ldots,\sigma_k,\sigma_{k+1}$ are acyclic elementary paths, by Lemma 2, $|\sigma_i| < \mathrm{card}(V)$. Also, since $\lambda_1,\ldots,\lambda_k$ are elementary cycles, we have $|\lambda_i| \leq \mathrm{card}(V)$. Since $q_i < |\lambda_i| \leq \mathrm{card}(V)$, and $k \leq \mathrm{card}(V)^2$, by Lemma 2, we have that

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

Hence $\rho_1.\lambda_i^*.\rho_2$ is a biquadratic path scheme. $\square$

For any $\ell \geq 0$ and vertices $u, v \in V$, let $biqs(\ell, u, v)$ denote the set of all biquadratic path schemes $\sigma.\lambda^*.\sigma'$, for which there exists a path $\rho \in [\![\sigma.\lambda^*.\sigma']\!]$ of length $|\rho| = \ell$ between u and v. Also, let $min\_biqs(\ell, u, v)$ be the subset of $biqs(\ell, u, v)$ consisting of minimal average weight path schemes i.e., $min\_biqs(\ell, u, v) = \{\sigma.\lambda^*.\sigma' \in biqs(\ell, u, v) \mid \forall \tau.\eta^*.\tau' \in biqs(\ell, u, v)\ .\ \overline{w}(\lambda) \leq \overline{w}(\eta)\}$.

Proposition 2. *Given a weighted graph $G = \langle V, E, w\rangle$, for any integer $\ell \geq card(V)^4$ and vertices $u, v \in V$, we have $biqs(\ell, u, v) = biqs(\ell + k\cdot lcm(1,\ldots,card(V)), u, v)$, for all $k \geq 0$. Moreover, we have $min\_biqs(\ell, u, v) = min\_biqs(\ell + k\cdot lcm(1,\ldots,card(V)), u, v)$, for all $k \geq 0$.*

Proof: Let $C = \mathrm{lcm}(1,\ldots,\mathrm{card}(V))$ in rest of the proof. We prove that, for an arbitrary path scheme θ:

$$\theta \in \mathrm{biqs}(\ell + k\cdot C, u, v) \text{ implies } \theta \in \mathrm{biqs}(\ell, u, v)$$

for all $k \geq 0$ (the other direction is trivial). Let $\theta = \sigma.\lambda^*.\sigma' \in \mathrm{biqs}(\ell + k \cdot C, u, v)$ be a path scheme. Clearly,

$$\ell + k \cdot C = |\sigma.\sigma'| + p \cdot |\lambda|$$

for some $p \geq 0$. Since θ is biquadratic, then $|\sigma.\sigma'| \leq \mathrm{card}(V)^4$. Since $\ell \geq \mathrm{card}(V)^4$, we obtain that:

$$\ell \geq |\sigma.\sigma'|$$

As a consequence, $p \cdot |\lambda| \geq k \cdot C$. Thus, $p \geq \frac{k \cdot C}{|\lambda|}$ and hence $p' = p - \frac{k \cdot C}{|\lambda|} \geq 0$. Hence we can define a path $\rho = \sigma.\lambda^{p'}.\sigma'$. We compute:

$$\begin{aligned} |\rho| &= |\sigma.\sigma'| + p' \cdot |\lambda| \\ &= |\sigma.\sigma'| + p \cdot |\lambda| - kC \\ &= \ell \end{aligned}$$

Thus, we have $\theta \in \mathrm{biqs}(\ell, u, v)$. For the second point, let $L = lcm(1, \ldots, \mathrm{card}(V))$. We compute:

$$\begin{aligned} min\_biqs(\ell, u, v) &= \{\sigma.\lambda^*.\sigma' \in biqs(\ell, u, v) \mid \forall \tau.\eta^*.\tau' \in biqs(\ell, u, v)\ .\ \overline{w}(\lambda) \leq \overline{w}(\eta)\} \\ &= \{\sigma.\lambda^*.\sigma' \in biqs(\ell + k \cdot L, u, v) \mid \forall \tau.\eta^*.\tau' \in biqs(\ell + k \cdot L, u, v)\ .\ \overline{w}(\lambda) \leq \overline{w}(\eta)\} \\ &= min\_biqs(\ell + k \cdot L, u, v) \end{aligned}$$

□

The following lemma shows that, for a sufficiently long minimal path, there exists an equivalent path which follows a biquadratic path scheme which moreover, has minimal average weight among all possible path schemes for that length.

Lemma 4. *Let $G = \langle V, E, w \rangle$ be a weighted digraph, and $u, v \in V$ be two vertices. Then for every minimal path ρ from u to v, such that $|\rho| > \max(card(V), 4 \cdot \mu(G) \cdot card(V)^6)$, there exists an equivalent path ρ', and a minimal average weight biquadratic path scheme $\sigma.\lambda^*.\sigma' \in min\_biqs(\ell, u, v)$, such that $\rho' \in [\![\sigma.\lambda^*.\sigma']\!]$.*

Proof: First we consider the case $\mu(G) = 0$. In this case $\max(\mathrm{card}(V), 4 \cdot \mu(G) \cdot \mathrm{card}(V)^6) = \mathrm{card}(V) > 0$, and any path ρ of length $|\rho| > \mathrm{card}(V)$ has a cyclic subpath. Since $\mu(G) = 0$, all paths in G have zero weight, hence ρ is minimal. By Lemma 3, there exists an equivalent path ρ' which is captured by a biquadratic path scheme $\sigma.\lambda^*.\sigma'$ of zero weight.

Back to the case $\mu(G) > 0$, we have $\max(\mathrm{card}(V), 4 \cdot \mu(G) \cdot \mathrm{card}(V)^6) = 4 \cdot \mu(G) \cdot \mathrm{card}(V)^6$. By Lemma 3, for every minimal path ρ of length $L > 0$, there exists an equivalent path ρ' which is captured by at least one biquadratic path scheme from $biqs(L, u, v)$. We will show that if $L \geq 4 \cdot \mu(G) \cdot \mathrm{card}(V)^6$, the cycle in this path scheme must have minimal average weight among cycles of all path schemes in $biqs(L, u, v)$.

Let $\sigma_i.\lambda_i^*.\sigma_i', \sigma_j.\lambda_j^*.\sigma_j' \in biqs(L, u, v)$ be two path schemes such that $\rho_i = \sigma_i.\lambda_i^{b_i}.\sigma_i'$ and $\rho_j = \sigma_j.\lambda_j^{b_j}.\sigma_j'$ are two paths of length L, between the same vertices, for some

$b_i, b_j \geq 0$. First, we compute:

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

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

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

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 3, we have $|\sigma_i.\sigma'_i|, |\sigma_j.\sigma'_j| \leq \text{card}(V)^4$, and moreover, for any path π, $w(\pi) \leq |\pi| \cdot \mu(G)$. Since $1 \leq |\lambda_i|, |\lambda_j| \leq \text{card}(V)$, we compute:

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

Combining this with Equation (7), we infer that if $\overline{w}(\lambda_i) < \overline{w}(\lambda_j)$ and $4 \cdot \mu(G) \cdot \text{card}(V)^6 \leq L$, then $w(\rho_i) \leq w(\rho_j)$. Therefore, a minimal path of length greater than $4 \cdot \mu(G) \cdot \text{card}(V)^6$ must follow a biquadratic path scheme, whose cycle has minimal average weight, among all possible path schemes, which could be followed by that path. □

Let $min\_weight(\ell, u, v) \in \mathbb{Z}_\infty$ denote the minimal weight among all paths of length ℓ between u and v, or ∞ if no such path exists. The following lemma is crucial in proving the main result of this section:

Lemma 5. *Let $G = \langle V, E, w \rangle$ be a weighted graph and $u, v \in V$ be two vertices. Then the sequence $\{min\_weight(\ell, u, v)\}_{\ell=1}^{\infty}$ is periodic, with prefix at most* $\max(card(V)^4, 4 \cdot \mu(G) \cdot card(V)^6)$.

Proof: It is sufficient to show that, there exists an integer $c > 0$ such that, for any $\ell > \max(\text{card}(V)^4, 4 \cdot \mu(G) \cdot \text{card}(V)^6)$, there exists $\Lambda \in \mathbb{Z}_\infty$ such that $min\_weight(\ell + (k+1)c, u, v) = \Lambda + min\_weight(\ell + kc, u, v)$, for all $k \geq 0$. Let $c = lcm(1, \ldots, \text{card}(V))$. By Prop. 2 is that $min\_biqs(\ell, u, v) = min\_biqs(\ell + kc, u, v)$, for all $k \geq 0$.

We distinguish two cases. First, $min\_weight(\ell + kc, u, v) = \infty$, i.e. $min\_biqs(\ell + kc, u, v) = min\_biqs(\ell + (k+1)c, u, v) = \emptyset$, and therefore we obtain $min\_weight(\ell + (k+1)c, u, v) = \infty$ as well. Second, suppose that $min\_weight(\ell + kc, u, v) < \infty$. Then there exists a minimal path ρ between u and v such that $|\rho| = \ell + kc > \max(\text{card}(V)^4, 4 \cdot \mu(G) \cdot \text{card}(V)^6)$. By Lemma 4, there exists an equivalent path ρ' and a biquadratic path scheme $\sigma.\lambda^*.\sigma' \in min\_biqs(\ell + kc, u, v)$ 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, if this is the case, then $|\rho''| = |\rho| + c$ and $w(\rho'') = w(\rho) + c \cdot \overline{w}(\lambda)$ i.e., $min\_weight(\ell + kc, u, v) =$

$min\_weight(\ell+(k+1)c,u,v)+c\cdot\overline{w}(\lambda)$. Since $\overline{w}(\lambda)$ is the common average weight of all path schemes in $min\_biqs(\ell+kc,u,v) = min\_biqs(\ell+k'c,u,v)$, for any $k,k' \geq 0$, the choice of $\Lambda = c\cdot\overline{w}(\lambda)$ does not depend on the particular value of k.

To show that ρ'' is indeed minimal, suppose it is not, and let π'' be a minimal path of length $|\rho''| = \ell+(k+1)c > \max(\mathrm{card}(V)^4, 4\cdot\mu(G)\cdot\mathrm{card}(V)^6)$. By Lemma 4, there exists an equivalent path π' and a biquadratic path scheme $\tau.\eta^*.\tau' \in min\_biqs(\ell+(k+1)c,u,v) = min\_biqs(\ell+kc,u,v)$ (by Prop. 2) 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 $\ell+kc$. 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). □

The following theorem summarizes the main result of this section, completing the evaluation of the period of a sequence of tropical powers (Theorem 3) with an upper bound on its prefix:

Theorem 4. *Given a matrix $A \in \mathbb{Z}_\infty^{m\times m}$, the sequence $\{A^{\boxtimes k}\}_{k=1}^\infty$ is periodic with prefix at most* $\max(m^4, 4\cdot M\cdot m^6)$*, where* $M = \max\{abs(A_{ij}) \mid 1 \leq i,j \leq m, A_{ij} < \infty\}$.

Proof: Let G be the weighted graph whose incidence matrix is A. Clearly, $M = \mu(G)$, and $(A^{\boxtimes \ell})_{ij} = min\_weight(\ell,i,j)$, for each $\ell > 0$ and $1 \leq i,j \leq m$. Since the prefix of the sequence $\{A^{\boxtimes k}\}_{k=1}^\infty$ is the maximum of the prefixes of $\{(A^{\boxtimes k})_{ij}\}_{k=1}^\infty$, and each of the latter prefixes is at most $\max(m^4, 4\cdot M\cdot m^6)$ (by Lemma 5), the conclusion follows. □

7 Difference Bounds Relations

In the rest of this section, let $\mathbf{x} = \{x_1, x_2, ..., x_N\}$ be a set of variables ranging over $\mathbb{Z}$.

Definition 7. *A formula $\phi(\mathbf{x})$ is a* difference bounds constraint *if it is a finite conjunction of atomic propositions of the form $x_i - x_j \leq \alpha_{ij}$, $1 \leq i,j \leq N, i \neq j$, 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 can be defined by a difference bounds constraint $\phi_R(\mathbf{x},\mathbf{x}')$. The class of difference bounds relations is denoted by $\mathcal{R}_{DB}$.*

Difference bounds constraints are represented either as matrices or as graphs. If $\phi(\mathbf{x})$ is a difference bounds constraint, then a *difference bounds matrix* (DBM) representing ϕ is an $N\times N$ matrix M_ϕ such that $(M_\phi)_{ij} = \alpha_{ij}$ if $x_i - x_j \leq \alpha_{ij} \in Atom(\phi)$, $(M_\phi)_{ii} = 0$, and $(M_\phi)_{ij} = \infty$, otherwise (Fig.1b). The *constraint graph* $\mathcal{G}_\phi = \langle \mathbf{x}, \rightarrow \rangle$ is a weighted graph, where each vertex corresponds to a variable, and there is an edge $x_i \xrightarrow{\alpha_{ij}} x_j$ in $\mathcal{G}_\phi$ if and only if there exists a constraint $x_i - x_j \leq \alpha_{ij}$ in ϕ (Fig. 1a). If R is a difference bounds relation defined by the difference bounds constraint $\phi_R(\mathbf{x},\mathbf{x}')$, the *folded graph* of R is the graph $\mathcal{G}_R^f = \langle \mathbf{x}, \xrightarrow{f} \rangle$, which has an edge $x_i \xrightarrow{f} x_j$ whenever either $x_i \xrightarrow{\alpha} x_j$, $x_i \xrightarrow{\alpha} x'_j$, $x'_i \xrightarrow{\alpha} x_j$ or $x'_i \xrightarrow{\alpha} x'_j$ in $\mathcal{G}_R$ (Fig. 1c). For any two variables $x_i, x_j \in \mathbf{x}$, we write $x_i \sim_R x_j$ whenever x_i and x_j belong to the same SCC of $\mathcal{G}_R^f$. Clearly, M_ϕ is the incidence matrix of $\mathcal{G}_\phi$. If $M \in \mathbb{Z}_\infty^{N\times N}$ is a DBM, we define[8]:

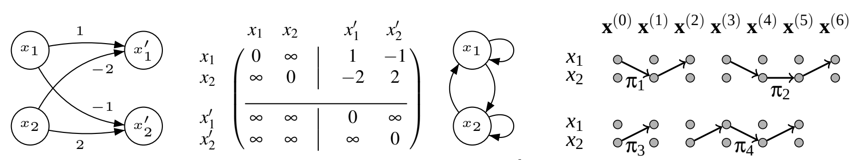
$$\Phi_M^{uu} \equiv \bigwedge_{M_{ij}<\infty} x_i - x_j \leq M_{ij} \qquad \Phi_M^{pu} \equiv \bigwedge_{M_{ij}<\infty} x'_i - x_j \leq M_{ij}$$
$$\Phi_M^{up} \equiv \bigwedge_{M_{ij}<\infty} x_i - x'_j \leq M_{ij} \qquad \Phi_M^{pp} \equiv \bigwedge_{M_{ij}<\infty} x'_i - x'_j \leq M_{ij}$$

A DBM M is said to be *consistent* if and only if Φ_M^{uu} is consistent. For a consistent difference bounds constraint ϕ, let ϕ^* denote its *closure* i.e., the unique difference bounds constraint containing explicitly all the implied constraints of ϕ. It is well known that difference bounds constraints have quantifier elimination[9], and are thus closed under relational composition.

Proposition 3. *Let ϕ and ϕ_1, ϕ_2 be difference bounds constraints, ϕ_1 and ϕ_2 are consistent, $\mathcal{G}_\phi$ be the constraint graph of ϕ and $\mathcal{G}_{\phi_1^*}$, $\mathcal{G}_{\phi_2^*}$ be the constraint graphs of the closures of ϕ_1, ϕ_2, respectively. Then, the following hold:*

- *ϕ is consistent if and only if $\mathcal{G}_\phi$ does not contain an elementary negative weight cycle*
- *$\phi_1 \Leftrightarrow \phi_2$ if and only if $\mathcal{G}_{\phi_1^*} = \mathcal{G}_{\phi_2^*}$.*

Proof: See e.g. [12], §25.5. □



$$\begin{array}{c|c} & \begin{array}{cccc} x_1 & x_2 & x'_1 & x'_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x'_1 \\ x'_2 \end{array} & \left(\begin{array}{cc|cc} 0 & \infty & 1 & -1 \\ \infty & 0 & -2 & 2 \\ \hline \infty & \infty & 0 & \infty \\ \infty & \infty & \infty & 0 \end{array}\right) \end{array}$$

(a) $\mathcal{G}_R$ (b) M_R^* (c) $\mathcal{G}_R^f$ (d) z-paths in $\mathcal{G}_R^\omega$

Fig. 1: Let $R(x_1,x_2,x'_1,x'_2)\ :\ x_1 - x'_1 \leq 1 \wedge x_1 - x'_2 \leq -1 \wedge x_2 - x'_1 \leq -2 \wedge x_2 - x'_2 \leq 2$ be a difference bounds relation. **(a)** shows the graph representation $\mathcal{G}_R$, **(b)** the closed DBM representation of R, and **(c)** the folded graph of $\mathcal{G}_R$, where $x_1 \sim_R x_2$. **(d)** shows several odd forward z-paths: π_1 (essential and repeating), π_2 (repeating), π_3 (essential) and $\pi_4 = \pi_3.\pi_1$ (neither essential nor repeating).

Lemma 6. *The class $\mathcal{R}_{DB}$ is poly-logarithmic.*

Proof: Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathbf{x} = \{x_1, \ldots, x_N\}$ be the set of variables in its arithmetic representation, and let $\mathcal{G}_R$ be its corresponding constraint

[8] The superscripts u and p stand for *unprimed* and *primed*, respectively.
[9] The quantifier elimination procedure relies on the classical Floyd-Warshall closure algorithm.

graph. We assume w.l.o.g. that each variable in $\mathbf{x}$ occurs in at least one atomic proposition of the form $x - y \leq c$, in each arithmetic formula defining R (otherwise we need not consider that variable in $\mathbf{x}$). Hence we have:

$$N \leq 2 \cdot \|R\|_2 \tag{9}$$

We denote by $\mathcal{G}_R^m$ the m-times unfolding of $\mathcal{G}_R$, formally the graph with vertices $\mathbf{x}^{(0)}, \ldots, \mathbf{x}^{(k)}$, where $\mathbf{x}^{(i)} = \{x^{(i)} \mid x \in \mathbf{x}\}$, and the subgraph composed of the edges between $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(i+1)}$ is an isomorphic copy of $\mathcal{G}_R$, for all $i = 0, \ldots, m-1$. Then the m-th power of R is the difference bounds relation:

$$R^m \Leftrightarrow \bigwedge_{1 \leq i,j \leq N} \begin{array}{l} x_i - x_j \leq \min\{x_i^{(0)} \to x_j^{(0)}\} \wedge x'_i - x'_j \leq \min\{x_i^{(m)} \to x_j^{(m)}\} \\ \wedge\, x_i - x'_j \leq \min\{x_i^{(0)} \to x_j^{(m)}\} \wedge x'_i - x_j \leq \min\{x_i^{(m)} \to x_j^{(0)}\} \end{array}$$

where $\min\{x_i^{(p)} \to x_j^{(q)}\}$ denotes the minimal weight among all paths between the extremal vertices $x_i^{(p)}$ and $x_j^{(q)}$ in $\mathcal{G}_R^m$, for $p, q \in \{0, m\}$. Since any such minimal path does not visit any vertex twice, we have $\min\{x_i^{(p)} \to x_j^{(q)}\} \leq N \cdot (m+1) \cdot \nabla(R)$, for all $i, j \in \{1, \ldots, N\}$ and $p, q \in \{0, m\}$. We compute:

$$\begin{array}{rll} \|R^m\|_2 & \leq 4N^2 \cdot log_2(N \cdot (m+1) \cdot \nabla(R)) & \\ & \leq 16\|R\|_2^2 \cdot (\log_2 \|R\|_2 + \log_2(m+1) + \log_2 \nabla(R)) & \text{by (9)} \\ & = O(\|R\|_2^3 \cdot \log_2 m) & \text{by Prop. 1} \end{array}$$

The second point of Def. 5 follows by observing that the composition of two difference bounds relations $P, Q \in \mathbb{Z}^N \times \mathbb{Z}^N$ is computed by the Floyd-Warshall algorithm in $(3N)^3$ steps. Since the maximal values occurring during this computation are less than $3N \cdot (\nabla(P) + \nabla(Q))$, and the operations at each step can be performed in $\log_2 N + \log_2(\nabla(P) + \nabla(Q)) \leq \log_2 N + \log_2 \nabla(P) + \log_2 \nabla(Q)$, the entire computation takes $O((\|P\|_2 + \|Q\|_2)^4)$ time. □

7.1 Zigzag Automata

Zigzag automata have been used in the proof of Presburger definability of transitive closures [9], and of periodicity [8], for difference bounds and octagonal relations. They are needed here for showing that difference bounds relations are exponential (Def. 6). Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathcal{G}_R$ be the constraint graph of R, and $\Sigma_R = 2^{\mathcal{G}_R}$ be the set of subgraphs of $\mathcal{G}_R$. A word of length $n \geq 0$ over Σ_R is a mapping $\gamma : [n] \to \Sigma_R$. The notion of finite words over Σ_R extends naturally to infinite words $\gamma : \mathbb{N} \to \Sigma_R$, and to bi-infinite words $\gamma : \mathbb{Z} \to \Sigma_R$. The concatenation of two finite words $\gamma : [n] \to \Sigma_R$ and $\gamma' : [m] \to \Sigma_R$ is a word $\gamma \cdot \gamma' : [n+m] \to \Sigma_R$, defined as $(\gamma \cdot \gamma')(i) = \gamma(i)$, for all $0 \leq i < n$ and $(\gamma \cdot \gamma')(i) = \gamma'(i-n)$, for all $n \leq i < n+m$. The set of finite words is denoted Σ_R^*. For a finite word $\gamma : [n] \to \Sigma_R$, we denote by γ^ω its infinite iteration, and by ${}^\omega\gamma^\omega$ its bi-infinite iteration, i.e. $\gamma^\omega[i] = \gamma[i \mod n]$, for all $i \in \mathbb{N}$, and ${}^\omega\gamma^\omega[i] = \gamma[i \mod n]$

for all $i \in \mathbb{Z}$. Alternatively, a word γ is represented as a graph[10] with vertices $\bigcup_{i=0}^{n} \mathbf{x}^{(i)}$, where $\mathbf{x}^{(i)} = \{x^{(i)} \mid x \in \mathbf{x}\}$ and edges:

- $x_k^{(i)} \xrightarrow{\alpha} x_\ell^{(i+1)}$ if and only if $x_k \xrightarrow{\alpha} x'_\ell$ in $\gamma(i)$
- $x_k^{(i+1)} \xrightarrow{\alpha} x_\ell^{(i)}$ if and only if $x'_k \xrightarrow{\alpha} x_\ell$ in $\gamma(i)$

for all $1 \leq k, \ell \leq N$ and for all $0 \leq i < n$.

Definition 8. *A finite word $\gamma : [n] \to \Sigma_R$ is said to be* valid *if and only if, for all $1 \leq k \leq N$:*

- *each vertex $x_k^{(i)}$ in γ has in-degree and out-degree at most one, for all $i \in [n]$*
- *each vertex $x_k^{(i)}$ in γ has equal in-degree and has out-degree, for all $i \in [n-1] \setminus \{0\}$*

This notion of validity extends from finite to infinite and bi-infinite words.

Given a difference bounds relation $R \subseteq \mathbb{Z}^{N \times N}$, the set of valid finite words in Σ_R^* is recognizable by a finite weighted automaton, called a *zigzag automaton* in the following. Let $T_R = \langle Q, \Delta, \omega \rangle$ be a weighted graph[11], called the *transition table* of the zigzag automata over Σ_R, where $Q = \{\ell, r, \ell r, r\ell, \perp\}^N$ is a set of states, $\Delta : Q \times \Sigma_R \to Q$ is a transition mapping, and $\omega : \Sigma_R \to \mathbb{Z}_\infty$ is a weight function. Intuitively, a state $\mathbf{q} = \langle \mathbf{q}_{\langle 1 \rangle}, \ldots, \mathbf{q}_{\langle N \rangle} \rangle \in Q$ describes a vertical cut in a word, as follows: $\mathbf{q}_{\langle i \rangle} = \ell$ ($\mathbf{q}_{\langle i \rangle} = r$) if there is a path in the word which traverses the cut at position i form right to left (left to right), $\mathbf{q}_{\langle i \rangle} = \ell r$ ($\mathbf{q}_{\langle i \rangle} = r\ell$) if there is a path from the left (right), which bounces to the left (right) at position i, and $\mathbf{q}_{\langle i \rangle} = \perp$ if the word does not intersect with the cut at position i, for each $i = 1, \ldots, N$ (see Fig. 2 (c) for an intuitive example). The transition function Δ ensures that the (local) validity condition is met. More precisely, each path $\rho : \mathbf{q}_0 \xrightarrow{\gamma_1} \mathbf{q}_1 \xrightarrow{\gamma_2} \ldots \xrightarrow{\gamma_k} \mathbf{q}_k$ in T_R, between two arbitrary states $\mathbf{q}_0, \mathbf{q}_k \in Q$, recognizes a valid word denoted as $\mathcal{G}_\rho = \gamma_1 \cdot \ldots \cdot \gamma_k$. The weight $\omega(G)$ of a graph $G \in \Sigma_R$ is the sum of the weights of its edges, and the weight of a path is $\omega(\rho) = \sum_{i=1}^{k} \omega(G_i)$. Finally, a *zigzag automaton* is a tuple $A = \langle T_R, I, F \rangle$, where $I, F \subseteq Q$ are sets of initial and final states, respectively. We denote the *language* of A as $\mathcal{L}(A) = \{\mathcal{G}_\rho \mid q_i \overset{\rho}{\Rightarrow} q_f, q_i \in I, q_f \in F\}$. A detailed definition of zigzag automata can be found in [9]. For the purposes of the upcoming developments, we rely on the example in Fig. 2 to give the necessary intuition.

Remark 1. The transition table $T_R = \langle Q, \Delta, \omega \rangle$ of a difference bounds relation $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ has at most 5^N vertices, since $Q = \{\ell, r, \ell r, r\ell, \perp\}^N$ is a possible representation of the set of states [9].

[10] We assume w.l.o.g. that $\mathcal{G}_R$ is a bipartite graph, obtained by replacing all constraints of the form $x - y \leq \alpha$ by $x - t' \leq \alpha \wedge t' - y \leq 0$, and all constraints of the form $x' - y' \leq \alpha$ by $x' - t \leq \alpha \wedge t - y' \leq 0$, for some fresh variables $t, t' \notin FV(R)$.

[11] For reasons of presentation, we differ slightly from the definition of a weighted graph given in the previous section – here the weight of an edge is associated with the symbol labeling that edge.

7.2 Paths Recognizable by Zigzag Automata

This section studies the paths that occur within the words recognizable by zigzag automata. Consider the bi-infinite unfolding ${}^{\omega}\mathcal{G}_R^{\omega}$ of $\mathcal{G}_R$. A finite path $\rho : x_{i_1}^{(j_1)} \xrightarrow{\alpha_1} x_{i_2}^{(j_2)} \xrightarrow{\alpha_2} \ldots x_{i_{k-1}}^{(j_{k-1})} \xrightarrow{\alpha_{k-1}} x_{i_k}^{(j_k)}$ in ${}^{\omega}\mathcal{G}_R^{\omega}$, for $j_1,\ldots,j_k \in \mathbb{Z}$ is said to be a *z-path* (see Fig. 1d or Fig. 2 (c) for examples of z-paths) whenever, for all $1 \le p < q \le k$, $i_p = i_q$ and $j_p = j_q$ only if $p = 1$ and $q = k$. We say that a variable x_{i_s} *occurs* on ρ at position j_s, for all $1 \le s \le k$. A z-path is called a *z-cycle* if $i_1 = i_k$ and $j_1 = j_k$. A z-path is said to be *odd* if $j_1 \ne j_k$ and *even* otherwise. For instance, in Fig. 2 (c), the z-path $x_1^{(1)} \xrightarrow{0} x_2^{(2)} \xrightarrow{0} x_3^{(3)} \xrightarrow{0} x_4^{(2)} \xrightarrow{0} x_5^{(1)} \xrightarrow{-1} x_1^{(2)}$ is an odd z-path, while $x_1^{(1)} \xrightarrow{0} x_2^{(2)} \xrightarrow{0} x_3^{(3)} \xrightarrow{0} x_4^{(2)} \xrightarrow{0} x_5^{(1)}$ is an even z-path.

We denote by $\|\rho\| = \mathrm{abs}(j_k - j_1)$ its *relative* length, by $w(\rho) = \sum_{i=1}^{k-1} \alpha_i$ its *weight*, and by $\overline{\overline{w}}(\rho) = \frac{w(\rho)}{\|\rho\|}$ its *relative weight*. We write $vars(\rho)$ for the set $\{x_{i_1},\ldots,x_{i_k}\}$ of variables occurring within ρ, called the *support set* of ρ.

An even z-path is said to be *forward* if $j_1 = j_k = \min(j_1,\ldots,j_k)$ and *backward* if $j_1 = j_k = \max(j_1,\ldots,j_k)$. An even z-path is said to be *fitting* if it is either forward or backward. An odd z-path is said to be *forward* if $j_1 < j_k$ and *backward* if $j_1 > j_k$. An odd forward (backward) z-path is said to be *fitting* if $j_1 = \min(j_1,\ldots,j_k)$ and $j_k = \max(j_1,\ldots,j_k)$ ($j_1 = \max(j_1,\ldots,j_k)$ and $j_k = \min(j_1,\ldots,j_k)$). For instance, in Fig. 2 (c), the odd forward z-path $x_1^{(1)} \xrightarrow{0} x_2^{(2)} \xrightarrow{0} x_3^{(3)} \xrightarrow{0} x_4^{(2)} \xrightarrow{0} x_5^{(1)} \xrightarrow{-1} x_1^{(2)}$ is not fitting, while the odd forward z-path $x_1^{(0)} \to \ldots \to x_7^{(17)}$ is fitting.

We say that a fitting z-path ρ is *encoded* by a word G, if and only if G consists of nothing but ρ and several z-cycles not intersecting with ρ. Observe that every fitting even (odd) z-path is encoded by a valid word $G_1 \cdot \ldots \cdot G_k \in \Sigma_R^*$, such that the z-path traverses each G_i an even (odd) number of times.Let $Enc(G)$ be the set consisting of the single acyclic z-path encoded by G, or the empty set, if G does not contain exactly one acyclic path. Let $Enc(\mathcal{L}) = \bigcup_{G \in \mathcal{L}} Enc(G)$ for any set of words $\mathcal{L} \subseteq \Sigma_R^*$. For instance, in Fig. 2 (c), the valid word $\gamma_0.\gamma_1^2.\gamma_2.\gamma_3.\gamma_4.\gamma_5^3.\gamma_6.\gamma_7.\gamma_8^2.\gamma_9.\gamma_2.\gamma_3.\gamma_4$ encodes the z-path $x_1^{(0)} \to \ldots \to x_7^{(17)}$.

Theorem 5 **([9]).** *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a $*$-consistent difference bounds relation, $\mathbf{x} = \{x_1,\ldots,x_N\}$ be the set of variables used in its definition, and $\mathcal{G}_R$ be its corresponding constraint graph. Then, for every $x_i, x_j \in \mathbf{x}$, there exist zigzag automata*[12] *$A_{ij}^{\bullet} = \langle T_R, I_{ij}^{\bullet}, F_{ij}^{\bullet} \rangle$, $\bullet \in \{ef, eb, of, ob\}$, where $T_R = \langle Q, \Delta, \omega \rangle$, such that $Enc(\mathcal{L}(A_{ij}^{\bullet}))$ are the sets of fitting even/odd, forward/backward z-paths, starting with $x_i^{(k)}$ and ending with $x_j^{(\ell)}$, respectively, for some $k, \ell \in \mathbb{Z}$. Moreover, for each fitting z-path ρ, $\omega(\rho) = \min\{\omega(\gamma) \mid \gamma \in \mathcal{L}(A_{ij}^{ef}) \cup \mathcal{L}(A_{ij}^{eb}) \cup \mathcal{L}(A_{ij}^{of}) \cup \mathcal{L}(A_{ij}^{ob}), \rho \in Enc(\gamma)\}$.*

---

[12] Superscripts ef, eb, of and ob stand for *even forward, even backward, odd forward* and *odd backward*, respectively.

7.3 The Complexity of Acceleration for Difference Bounds Relations

In this section, we prove that difference bounds constraints induce a periodic exponential class of relations (Def. 6). First, we recall that difference bounds relations are periodic (Def. 3) [8]. If $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ is a difference bounds relation, let $\sigma(R) \equiv M_R$ and, for each $M \in \mathbb{Z}_\infty^{2N\times 2N}$, let ${}^{\blacksquare}M$, ${}_{\blacksquare}M$, $M^{\blacksquare}$, $M_{\blacksquare} \in \mathbb{Z}^{N\times N}$ denote its top-left, bottom-left, top-right and bottom-right corners, respectively. Intuitivelly, ${}^{\blacksquare}M$, ${}_{\blacksquare}M$, $M^{\blacksquare}$, $M_{\blacksquare}$ capture constraints of the forms $x_i - x_j \leq c$, $x'_i - x_j \leq c$, $x_i - x'_j \leq c$ and $x'_i - x'_j \leq c$, respectivelly (see Fig. 1b). We define $\rho(M) \equiv \Phi^{uu}_{{}^{\blacksquare}M} \wedge \Phi^{up}_{M^{\blacksquare}} \wedge \Phi^{pu}_{{}_{\blacksquare}M} \wedge \Phi^{pp}_{M_{\blacksquare}}$. Analogously, if $M \in \mathbb{Z}[k]_\infty^{2N\times 2N}$ is a matrix of univariate linear terms in k, $\pi(M)(k,\mathbf{x},\mathbf{x}')$ is defined in the same way as ρ above.

With these definitions, it was shown in [8], that the class of difference bounds relations is periodic (Def. 3). The reason is that the sequence of difference bounds matrices $\{M_{R^i}\}_{i=1}^{\infty}$ corresponding to the powers of a relation R is a pointwise projection of the sequence of tropical powers $\{\mathcal{M}_R^{\boxtimes^i}\}_{i=1}^{\infty}$ of the incidence matrix $\mathcal{M}_R$ of the transition table T_R. By Thm. 3, any sequence of tropical powers of a matrix is periodic, which entails the periodicity of the difference bounds relation R. Recall that the number of vertices in T_R is $5^N = 2^{O(N)}$. Consequently, the prefix of a difference bounds relation can be bounded using Thm. 4:

Lemma 7. *The prefix of a difference bounds relation* $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ *is* $\nabla(R) \cdot 2^{O(N)}$.

Proof: We consider first the case where R is a $*$-consistent relation. In this case, the prefix of R is bounded by the prefix of the sequence $\{\mathcal{M}_R^{\boxtimes^i}\}_{i=0}^{\infty}$, where $\mathcal{M}_R$ is the incidence matrix of the transition table T_R of the zigzag automata for R. The size of this table is $m \leq 5^N$. Let $\mathcal{M}_R \in \mathbb{Z}_\infty^{m\times m}$ be the incidence matrix of T_R. Let $M = \max\{\mathrm{abs}((\mathcal{M}_R)_{ij}) \mid 1 \leq i,j \leq m\}$ be the constant from Thm. 4. Then the prefix of R is bounded by $\max(m^4, 4\cdot M\cdot m^6) \leq \nabla(R)\cdot 5^{6N} = \nabla(R)\cdot 2^{O(N)}$, by Thm. 4.

Assume now that R is not $*$-consistent, i.e. there exist $b > 0$ such that R^b is consistent and, for any $\ell > b$, R^ℓ is not consistent. By Prop. 3, for each $\ell > b$, $\mathcal{G}_{R^\ell}$, contains a z-cycle of negative weight. By Thm. 5, this z-cycle is encoded by a minimal run ρ in a zigzag automaton A_{ii}^{ef}, where $|\rho| = \ell$. By Lemma 3, there exists a biquadratic path scheme $\sigma.\lambda^*.\sigma'$ (i.e. $|\sigma.\sigma'| \leq 5^{4N}$) in T_R such that $\rho = \sigma.\lambda^k.\sigma'$, for some $k \geq 0$. It must be the case that $w(\lambda) < 0$, or else $\sigma.\lambda^k.\sigma'$ could not encode a negative weight z-cycle, for infinitely many $k \geq 0$. Moreover, $|\lambda| \leq 5^N$, since λ is an elementary cycle. Since $w(\rho) = w(\sigma.\sigma') + k\cdot w(\lambda) < 0$, we have:

$$k > \frac{w(\sigma.\sigma')}{-w(\lambda)} \geq w(\sigma.\sigma')$$

since $-w(\lambda) > 0$ and $w(\lambda) \in \mathbb{Z}$, we have $-w(\lambda) \geq 1$. Hence for each $k > \nabla(R)\cdot 5^{4N}$, we have $w(\sigma.\lambda^k.\sigma') < 0$. Since $\ell = |\sigma.\sigma'| + k\cdot|\lambda|$, we obtain that for each $\ell > 5^{4N} + \nabla(R)\cdot 5^{4N}\cdot 5^N$, $w(\rho) < 0$, i.e. R^ℓ is inconsistent. Consequently, it must be the case that $b \leq \nabla(R)\cdot 5^{6N} = \nabla(R)\cdot 2^{O(N)}$. $\square$

A preliminary estimation of the upper bound of the period of a difference bounds relation $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ can be already done using Thm. 3. Since the size of the transition table T_R of the zigzag automata for R is bounded by 5^N, by definition, the cyclicity of any SCC of T_R is at most 5^N, hence, by Thm. 3, the period is bounded by $lcm(1,\ldots,5^N)$. Applying the following lemma, one shows immediately that the period is $2^{2^{O(N)}}$.

Lemma 8. *For each $n \geq 1$, $lcm(1,\ldots,n)$ is $2^{O(n)}$.*

Proof: We know that $lcm(1,\ldots,n) = \prod_{p \leq n} p^{\lfloor log_p(n) \rfloor}$ where the product is taken only over primes p. Obviously, for every prime p we have that $p^{\lfloor log_p(n) \rfloor} \leq p^{log_p(n)} = n$. Hence, $lcm(1,\ldots,n) \leq \prod_{p \leq n} n = n^{\pi(n)}$, where $\pi(n)$ denotes the prime-counting function (which gives the number of primes less than or equal to n, for every natural number n). Using the prime number theorem which states that $\lim_{n\to\infty} \frac{\pi(n)}{n/ln(n)} = 1$ we can effectively bound $\pi(n)$. That is, for any $\varepsilon > 0$, there exists n_ε such that $\frac{\pi(n)}{n/ln(n)} \leq (1+\varepsilon)$ for all $n \geq n_\varepsilon$. Consequently, $n^{\pi(n)} \leq n^{(1+\varepsilon)n/ln(n)} = e^{(1+\varepsilon)n} = 2^{log_2(e)(1+\varepsilon)n} = 2^{O(n)}$ for all $n \geq n_\varepsilon$, which completes the proof. □

We next improve the bound on periods to simply exponential (Thm. 6).

Theorem 6. *The period of a difference bounds relation $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ is $2^{O(N)}$.*

This leads to one of the main results of the paper:

Theorem 7. *The class $\mathcal{R}_{DB}$ is exponential, and the reachability problem for the class $\mathcal{M}_{DB} = \{M$ flat counter machine $|$ for all rules $q \xRightarrow{R} q'$ on a loop of M, R is $\mathcal{R}_{DB}$-definable$\}$ is NP-complete.*

Proof: To show that $\mathcal{R}_{DB}$ is exponential, we consider the four points of Def. 6. Point (A) of Def. 6 is by Lemma 6. Point (B) is trivial, by the definitions of the σ, ρ and π mappings for difference bounds relations. For point (C.1) we use the fact that $N \leq 2 \cdot \|R\|_2$ (9) and $\log_2(\nabla(R)) \leq \|R\|_2$ (Prop. 1) to infer that $b = 2^{O(\|R\|_2)}$ (by Lemma 7) and $c = 2^{O(\|R\|_2)}$ (by Thm. 6). For the last point (C.2), observe that the condition (2) of Lemma 1 states the equivalence of two difference bounds constraints $\phi_\ell(k)$ (for the left hand side of the equivalence) and $\phi_r(k)$ (for the right hand side of the equivalence), for each value of $k > 0$. Since, by the previous point (C.1), $b = 2^{O(\|R\|_2)}$ and $c = 2^{O(\|R\|_2)}$, it follows that the binary size of the equivalence is polynomial in $\|R\|_2$. By Prop. 3 (point 2), it must be that, for each value $n > 0$ we have $\mathcal{G}_{\phi^*_\ell[n/k]} = \mathcal{G}_{\phi^*_r[n/k]}$. Since both $\phi_\ell(k)$ and $\phi_r(k)$ can be represented by constraint graphs with weights of the form $a \cdot k + b$, for $a, b \in \mathbb{Z}$. The closures $\phi^*_\ell(k)$ and $\phi^*_r(k)$ can be computed in polynomial time by a variant of the Floyd-Warshall algorithm that constructs linear min-terms built form min and $+$ operators, k and integer constants. These terms can be represented by dags of polynomial size (by sharing common subterms). In the light of Prop. 3 (point 2), the second point of Lemma 1 is equivalent to the validity of the conjunction of at most N^2 equalities between univariate linear min-terms in k, of polynomial size. This conjunction can be written as a QFPA formula of size polynomial in $\|R\|_2$, and solved in NPTIME($\|R\|_2$).

The NP-completness of the reachability problem for the $\mathcal{M}_{DB}$ class follows from Thm. 2. □

The Period of Difference Bounds Relations (proof idea) Before proceeding with the technical developments, we summarize the proof idea of Thm 6. Let $T_R = \langle Q, \Delta, \omega \rangle$ be the transition table for the difference bounds relation $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$, and let $\mathcal{M}_R$ be its incidence matrix. The main idea of the proof is that each non-trivial SCC of T_R, which is on a path between an initial and a final state, contains a *critical elementary cycle* λ, which consists of nothing but odd z-paths of the form $\pi_1^{n_1},\ldots,\pi_k^{n_k}$, where $\|\pi_i\| \leq N$, for all $i = 1,\ldots,k$ (Lemma 17). Indeed, suppose that this is true. Then the length of the critical elementary cycle λ is $|\lambda| = lcm(\|\pi_1\|,\ldots,\|\pi_k\|)$, which divides $lcm(1,\ldots,N)$. The cyclicity of the SCC containing λ is, by definition, the greatest common divisor of the lengths of all critical elementary cycles of the SCC, and consequently, a divisor of $lcm(1,\ldots,N)$ as well. Since this holds for any non-trivial SCC in T_R, by Thm. 3, the period of the sequence $\{\mathcal{M}_R^{\boxtimes k}\}_{k=1}^{\infty}$ of tropical powers of $\mathcal{M}_R$ is also a divisor of $lcm(1,\ldots,N)$, which is of the order of $2^{O(N)}$ (by Lemma 8).

It remains to prove the existence, in each non-trivial SCC of T_R, of an elementary critical cycle labeled by a set of essential powers. Let $q \in Q$ be a vertex of T_R, on a path from an initial to a final state of the zigzag automaton, and $q \xrightarrow{\gamma} q$ be a critical cycle (in its corresponding SCC) of T_R. The proof is organized in three steps:

- First, we build a word $Z \in \Sigma_R^*$, consisting of several repeating z-paths, such that $\overline{\overline{w}}(Z) = \overline{w}(\gamma)$, and for any $n > 0$, there exists $m > n$, such that Z^n is a subword of $\mathcal{G}_\gamma^m$.
- Second, we prove the existence of a word $\Lambda \in \Sigma_R^*$ *consisting of nothing but essential powers*, such that (i) $\overline{\overline{w}}(\Lambda) \leq \overline{\overline{w}}(Z)$ and that (ii) there exist valid words $V, W \in \Sigma_R^*$ such that $Z^n \cdot V \cdot \Lambda^m \cdot W \cdot Z^p$ is a valid word, for any $n,m,p \geq 0$. In other words, any iteration of Z can be concatenated with any iteration of Λ, and viceversa, via the words V and W, respectively.
- Finally, we prove the existence of another state $q' \in Q$ and of three paths $q \to q'$, $q' \to q$, and $q' \xrightarrow{\Lambda} q'$ in T_R. Since there is a path from q to q' and back, the w cycle is in the same SCC as γ, and moreover, it is a critical cycle of the SCC, because $\overline{w}(\lambda) = \overline{\overline{w}}(\Lambda) \leq \overline{\overline{w}}(Z) = \overline{w}(\gamma)$, and γ was initially assumed to be critical.

This proves the statement of Thm. 6.

Repeating z-Paths This and the next section introduce several technical lemmas which are needed in the proof. If π is the z-path $x_{i_1}^{(j_1)} \to \ldots \to x_{i_n}^{(j_n)}$, we denote by $\overrightarrow{\pi}^k : x_{i_1}^{(j_1+k)} \to \ldots \to x_{i_n}^{(j_n+k)}$ the z-path obtained by *shifting* π by k, where $k \in \mathbb{Z}$. A z-path π is said to be *isomorphic* with another z-path ρ if and only if $\rho = \overrightarrow{\pi}^k$, for some $k \in \mathbb{Z}$. In the following, we will sometimes silently denote by z-paths their equivalence classes w.r.t. the isomorphism relation. The *concatenation* of π with a z-path ρ is the path $\pi.\overrightarrow{\rho}^{j_n}$. The concatenation operation is however undefined if the above is not a valid z-path – this may happen when the π and $\overrightarrow{\rho}^{j_n}$ intersect in some vertex which occurs in the middle of one of the two paths. A z-path π is said to be *repeating* if and only if the i-times concatenation of π with itself, denoted π^i, is defined, for any $i > 0$. If π is repeating, then it clearly starts and ends with the same variable ($i_1 = i_n$), and is necessarily odd ($j_1 \neq j_n$). A repeating z-path π is said to be *essential* if all variables $x_{j_1},\ldots,x_{j_n}$ occurring on the path are distinct, with the exception of x_{j_1} and x_{j_n}, which might be equal. The concatenation of an essential repeating z-path with itself several times is called an *essential power*. The *left spacing* of a z-path represents the number of steps from the leftmost position on the path to the starting position of the path. The *right spacing* is defined symmetrically. Formally, given a z-path $\pi : x_{i_1}^{(j_1)} \to \ldots \to x_{i_n}^{(j_n)}$, we denote by

$[\pi = j_1 - \min\{j_1, \ldots, j_n\}$ and $\pi] = \max\{j_1, \ldots, j_n\} - j_n$ the *left* and *right spacing* of π, respectively. By ε we denote the empty z-path.

Lemma 9. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a $*$-consistent difference bounds relation, and $\mathbf{x} = \{x_1, \ldots, x_N\}$ be the set of variables from its arithmetic encoding. For any two variables $x_i \sim_R x_j$ from $\mathbf{x}$, if $\rho_i : x_i^{(0)} \to \ldots \to x_i^{(n)}$ and $\rho_j : x_j^{(m)} \to \ldots \to x_j^{(0)}$, for some $m, n > 0$, are two forward and backward z-paths in ${}^{\omega}\mathcal{G}_R{}^{\omega}$, respectively, then we have $\overline{\overline{w}}(\rho_i) + \overline{\overline{w}}(\rho_j) \geq 0$.*

Proof: Suppose that $\overline{\overline{w}}(\rho_i) + \overline{\overline{w}}(\rho_j) < 0$. Let us define:

$$p = \mathrm{lcm}(\|\rho_i\|, \|\rho_j\|),\ \ d_i = \frac{p}{\|\rho_i\|},\ \ d_j = \frac{p}{\|\rho_j\|},\ \ \gamma_i = (\rho_i)^{d_i},\ \ \gamma_j = (\rho_j)^{d_j}.$$

Notice that, since ρ_i and ρ_j are not assumed to be repeating, the powers γ_i and γ_j are not necessarily valid z-paths – we shall however abuse notation in the following, and use the usual symbols for the weight and relative length of γ_i and γ_j. We have $\overline{\overline{w}}(\rho_i) = \overline{\overline{w}}(\gamma_i)$ and $\overline{\overline{w}}(\rho_j) = \overline{\overline{w}}(\gamma_j)$. Thus, $\overline{\overline{w}}(\gamma_i) + \overline{\overline{w}}(\gamma_j) < 0$. Furthermore, since $\|\gamma_i\| = \|\gamma_j\| = p$, then $p \cdot \overline{w}(\gamma_i) + p \cdot \overline{w}(\gamma_j) = w(\gamma_i) + w(\gamma_j) < 0$. Since $x_i \sim_R x_j$, there exist essential paths

$$\theta_{ij} = x_i^{(0)} \to \ldots \to x_j^{(q)} \text{ and } \theta_{ji} = x_j^{(0)} \to \ldots \to x_i^{(r)}$$

where $0 \leq \mathrm{abs}(q), \mathrm{abs}(r) < N$. Let $n \geq 0$ be a an arbitrary constant. We build (Fig. 3)

$$\xi = \gamma_i^n . \theta_{ij} . \gamma_j^{2n} . \theta_{ji}$$

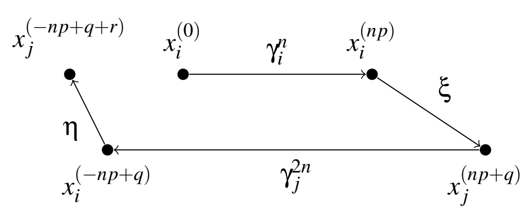


Fig. 3

Clearly, ξ is of the form $\xi : x_i^{(0)} \rightsquigarrow x_i^{(-np+q+r)}$. By choosing $n > \lceil \frac{q+r}{p} \rceil$, we make sure that $-np + r + s < 0$. We repeat the path p-times and obtain $\xi^p : x_i^{(0)} \rightsquigarrow x_i^{(p(-np+q+r))}$. Since $|\gamma_i| = p$ and p divides $p(-np+q+r)$, we build $\zeta = \gamma_i^{(np-r-s)}$ which is of the form $\zeta : x_i^{(p(-np+q+r))} \rightsquigarrow x_i^{(0)}$. Clearly, $\xi^p.\zeta$ forms a cycle with weight

$$np \cdot w(\gamma_i) + p \cdot w(\theta_{ij}) + 2np \cdot w(\gamma_j) + p \cdot w(\theta_{ji}) + (np - q - r) \cdot w(\gamma_i)$$

which simplifies to

$$2np \cdot (w(\gamma_i) + w(\gamma_j)) - (q+r) \cdot w(\gamma_i) + p \cdot (w(\theta_{ij}) + w(\theta_{ji})).$$

Since we assumed that $w(\gamma_i) + w(\gamma_j) < 0$, by choosing a sufficiently large n, we obtain a negative cycle in $\mathcal{G}_R^\omega$. Thus, R is not $*$-consistent, contradiction. □

A z-path π is said to be a *subpath* of ρ if and only if there exists factorizations $\pi = \pi_1. \ldots .\pi_k$ and $\rho = \rho_1. \ldots .\rho_\ell$, with $k < \ell$ and an injective mapping $h : \{1, \ldots, k\} \mapsto \{1, \ldots, \ell\}$ such that (i) for all $1 \le i < j \le k$, $h(i) < h(j)$, and (ii) for all $1 \le i \le k$, π_i is isomorphic with $\rho_{h(i)}$. Notice that the subpath relation is a well-founded preorder.

The following lemma proves, for each repeating z-path, the existence of an essential repeating subpath of smaller or equal average weight, such that arbitrarily many powers of the original z-path can be substituted by powers of its subpath.

Proposition 4. *Any essential z-path $\pi : x_{i_1}^{(j_1)} \to \ldots \to x_{i_n}^{(j_n)}$ such that $i_1 = i_n$ is repeating.*

Proof: Let us prove that $\pi.\pi$ is a valid z-path. The proof for the general case π^k is by induction on $k > 1$. Assume, by contradiction, that $\pi.\pi$ visits a vertex $x_j^{(i)}$ twice. Then, the first occurrence of the vertex must be in the first occurrence of π, and the second one must be in the second occurrence of π, since π is essential. But then π traverses the same variable twice, contradiction. □

Lemma 10. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a $*$-consistent difference bounds relation and $\mathcal{G}_R$ be its constraint graph. Then any forward (backward) repeating z-path π in ${}^\omega\mathcal{G}_R^\omega$ has an essential repeating forward (backward) subpath ρ such that $\overline{\overline{w}}(\rho) \le \overline{\overline{w}}(\pi)$.*

Proof: We give the proof for the case where π is forward, the backward case being symmetric. It is sufficient to prove the existence of a forward subpath ρ of π, starting and ending with the same variable, and such that $\overline{\overline{w}}(\rho) \le \overline{\overline{w}}(\pi)$. The existence of λ_1 and λ_2 follows as a consequence of the fact that ρ is a subpath of π. If ρ is not essential, then one can iterate the construction, taking ρ for π, until an essential subpath is found. By Prop. 4, this subpath is also repeating, and, since the subpath relation is well-founded, such a subpath is bound to exist. The existence of the paths λ_1, λ_2 and the fact that $\lambda_1.\rho^m.\lambda_2$ and π^n start and end with the same vertices is by induction on the number of steps of this derivation.

Let π be the repeating z-path $x_{i_1}^{(j_1)} \to \ldots\ldots \to x_{i_n}^{(j_n)}$, where $i_1 = i_n$. Since π is forward, we have $j_1 < j_n$. Let $P(\pi) = \{s \mid \exists\, t > s \,.\, i_s = i_t\}$ be the set of positions of π labeled with variables that occur more than once. Since π is not essential, we have that $P(\pi) \neq \emptyset$. Let $k = \min(P)$, and ℓ be the last occurrence of x_{i_m} on π, i.e. $i_m = i_p$. Let $\tau_1 : x_{i_1}^{(j_1)} \to \ldots x_{i_k}^{(j_k)}$, $\tau_2 : x_{i_\ell}^{(j_\ell)} \to \ldots\ldots \to x_{i_n}^{(j_n)}$, $\pi_1 = \tau_1.\tau_2$ and $\pi_2 : x_{i_k}^{(j_k)} \to \ldots \to x_{i_\ell}^{(j_\ell)}$. Observe that $\|\pi\| = j_n - j_1$, $\|\pi_1\| = j_n - j_\ell + j_k - j_1$ and $\|\pi_2\| = j_\ell - j_k$. Because of the choice of i_k and i_ℓ, it follows that π_1 is a valid essential z-path. Moreover, by Prop. 4, π_1 is also repeating. We distinguish two cases:

1. both π_1 and π_2 are forward i.e., $j_n - j_\ell + j_k > j_1$ and $j_\ell > j_k$. In this case either $\overline{\overline{w}}(\pi_1) \le \overline{\overline{w}}(\pi)$ or $\overline{\overline{w}}(\pi_2) \le \overline{\overline{w}}(\pi)$ must be the case. To see that this is indeed the case,

suppose that $\overline{\overline{w}}(\pi_1) > \overline{\overline{w}}(\pi)$ and $\overline{\overline{w}}(\pi_2) > \overline{\overline{w}}(\pi)$. We compute:

$$\begin{array}{rclcrcl} \frac{w(\pi_1)}{j_n - j_\ell + j_k - j_1} & > & \frac{w(\pi_1)+w(\pi_2)}{j_n - j_1} & \quad\quad & \frac{w(\pi_2)}{j_\ell - j_k} & > & \frac{w(\pi_1)+w(\pi_2)}{j_n - j_1} \\ \frac{w(\pi_1)}{j_n - j_\ell + j_k - j_1} & > & \frac{w(\pi_2)}{j_\ell - j_k} & & \frac{w(\pi_1)}{j_n - j_\ell + j_k - j_1} & < & \frac{w(\pi_2)}{j_\ell - j_k} \end{array}$$

contradiction.

2. π_1 and π_2 have opposite directions. We consider the case where π_1 is forward, the other case being symmetric. Since R is $*$-consistent, by Lemma 9, we have $\overline{\overline{w}}(\pi_1) + \overline{\overline{w}}(\pi_2) \geq 0$:

$$\begin{array}{rcl} \frac{w(\pi_1)}{j_n - j_\ell + j_k - j_1} + \frac{w(\pi_2)}{j_k - j_\ell} & \geq & 0 \\ \frac{w(\pi_1)+w(\pi_2)}{j_n - j_1} & \geq & \frac{w(\pi_1)}{j_n - j_\ell + j_k - j_1} \\ \overline{\overline{w}}(\pi) & \geq & \overline{\overline{w}}(\pi_1) \end{array}$$

Consequently, either $\overline{\overline{w}}(\pi_1) \leq \overline{\overline{w}}(\pi)$ or $\overline{\overline{w}}(\pi_2) \leq \overline{\overline{w}}(\pi)$ must be the case. If $\overline{\overline{w}}(\pi_1) \leq \overline{\overline{w}}(\pi)$, we let $\rho = \pi_1$ and we are done. Otherwise, if $\overline{\overline{w}}(\pi_2) \leq \overline{\overline{w}}(\pi)$ let $\rho = \pi_2$. The choice of λ_1 and λ_2 depends on the choice of ρ. If $\rho = \pi_1$, then $\lambda_1 = \lambda_2 = \varepsilon$, else if $\rho = \pi_2$ then $\lambda_1 = \tau_1$ and $\lambda_2 = \tau_2$, and finally, if $\rho = \pi_2'$, where $\pi_2 = \sigma_1.\pi_2'.\sigma_2$, then $\lambda_1 = \tau_1.\sigma_1$ and $\lambda_2 = \sigma_2.\tau_2$. $\square$

The next lemma gives a connectivity property of repeating z-paths. Intuitively, any two repeating z-paths traversing only variables that belong to the same SCC of the folded graph can be concatenated via a third path, which traverses variables from the same SCC.

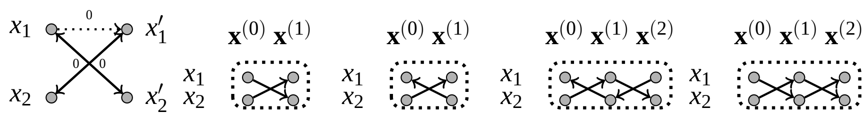
Lemma 11. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathbf{x} = \{x_1, \ldots, x_N\}$ be the variables used in its first-order arithmetic encoding, and let $S \in \mathbf{x}_{/\sim_R}$ be a set of variables belonging to the same SCC of the folded graph of R. For any two repeating forward (backward) z-paths π_1 and π_2 such that $vars(\pi_1) \cup vars(\pi_2) \subseteq S$, there exists a z-path ρ such that $vars(\rho) \subseteq S$ and, for any $n, m > 0$, $\pi_1^n.\rho.\pi_2^m$ is a valid z-path.*

Proof: Suppose that $\pi_1 : x_{i_1}^{(j_1)} \to \ldots \to x_{i_1}^{(j_n)}$ and $\pi_2 : x_{i_2}^{(\ell_1)} \to \ldots \to x_{i_2}^{(\ell_m)}$ are the given paths. Since $x_{i_1} \sim_R x_{i_2}$, there exists a path v from x_{i_1} to x_{i_2} in the folded graph $\mathcal{G}_R^f$ of R. Let v be the shortest path between x_{i_1} to x_{i_2} in $\mathcal{G}_R^f$. Since v visits each variable from S at most once, it induces an essential z-path $\xi : x_{i_1}^{(k_1)} \to \ldots \to x_{i_2}^{(k_p)}$, for some $k_1, \ldots, k_p \in \mathbb{Z}$. We define:

$$L = \max\left(1, \lceil \frac{[\xi}{\|\pi_1\|} \rceil + \lceil \frac{\pi_1]}{\|\pi_1\|} \rceil \right) \quad\quad\quad R = \max\left(1, \lceil \frac{\xi]}{\|\pi_2\|} \rceil + \lceil \frac{[\pi_2}{\|\pi_2\|} \rceil \right)$$

and let μ be the path obtained by concatenating the z-paths π_1^L, ξ and π_2^R. Notice that μ is not necessarily a z-path, because this concatenation does not necessarily result in a valid z-path. Therefore we define ρ to be the z-path obtained by eliminating all cycles from μ. Since π_1 and π_2 are repeating z-paths, the concatenation $\pi_1.\rho.\pi_2$ is defined. The same holds for $\pi_1^n.\rho.\pi_2^m$, for $n, m > 0$, by Prop. 4. Clearly ρ traverses only variables from S. $\square$

Multipaths and Reducts A *multipath* is a (possibly empty) set of z-paths from ${}^{\omega}\mathcal{G}_R{}^{\omega}$, which all start and end on the same positions (see Fig. 4). Formally, a multipath $\mu = \{\pi_1, \ldots, \pi_n\}$ is a set of z-paths such that there exist integers $k < \ell$ such that, for all $i = 1, \ldots, n$, either (i) π_i is a forward (backward) odd z-path from k to ℓ (from ℓ to k), (ii) π_i is an even z-path from k to k (ℓ to ℓ), or (iii) π_i is a z-cycle whose set of positions of variable occurrences is included in the interval $[k, \ell]$, and (iv) no two z-paths in μ intersect each other. The relative length of a multipath μ, is defined as $\|\mu\| = \ell - k$ if $\mu \neq \emptyset$, or $\|\mu\| = 0$ if $\mu = \emptyset$.



(a) $\mathcal{G}_R$ (b) μ_1 (c) μ_2 (d) μ_3 (e) μ_4

Fig. 4: Examples of multipaths. R is $x_1 = x_2' \wedge x_2 = x_1'$ and $\mathcal{G}_R$ is shown in **(a)**. μ_1 is iterable but not repeating, μ_2 is not iterable. Both μ_3 and μ_4 are fitting, iterable, repeating, and they consist of two balanced sc-multipaths each. If R is $x_1 = x_2' \wedge x_2 = x_1' \wedge x_1 \leq x_1'$ instead (the dotted edge $x_1 \xrightarrow{0} x_1'$), then μ_3 is a balanced sc-multipath and μ_4 is an unbalanced sc-multipath, since $\tau_1 \bowtie_R \tau_2$ for the two forward repeating z-paths $\tau_1, \tau_2 \in \mu_4$.

For a multipath μ, we denote by μ^{ac} the set of acyclic z-paths in μ. The weight of μ is defined as $w(\mu) = \sum_{\pi \in \mu}^{n} w(\pi)$, and its average weight is $\overline{\overline{w}}(\mu) = \frac{w(\mu)}{\|\mu\|}$ if $\|\mu\| \neq 0$, or $\overline{\overline{w}}(\mu) = 0$ if $\|\mu\| = 0$. The support set of a multipath is denoted as $vars(\mu) = \bigcup_{\pi \in \mu} vars(\pi)$. The set of variables occurring on the start (end) position k (ℓ) of a multipath μ is called the *left (right) frontier* of μ. The paths in μ starting and ending on k (ℓ) are called *left (right) corners*. The *left* and *right spacing* of μ are defined as $\lfloor\mu = \max(\lfloor\pi_1, \ldots, \lfloor\pi_n)$ and $\mu\rfloor = \max(\pi_1\rfloor, \ldots, \pi_n\rfloor)$, respectively. The concatenation of two multipaths μ_1 and μ_2 is defined if and only if (i) there exists a bijective function $\beta : \mu_1^{ac} \to \mu_2^{ac}$, such that, for all acyclic z-paths $\pi \in \mu_1^{ac}$, $\pi.\beta(\pi)$ is a valid z-path[13], and (ii) the set $\mu_1.\mu_2 = \{\pi.\beta(\pi) \mid \pi \in \mu_1^{ac}\}$ is a valid multipath. A multipath μ is *iterable* if it can be concatenated with itself any number of times, i.e. μ^i is a valid multipath, for all $i > 0$ (Fig. 4 (b)). A *repeating multipath* is an iterable multipath in which all acyclic z-paths are repeating (Fig. 4 (d,e)) – an empty multipath is repeating, by convention. A repeating multipath is said to be *essential* if every acyclic z-path is an essential power. A multipath μ is said to be *fitting* if every acyclic z-path in μ is fitting (Fig. 4 (b-e)).

Definition 9. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathbf{x}$ be the set of variables in its defining formulae, and $\mathcal{G}_R$ be its constraint graph. Let π_1 and π_2 be repeating z-paths in ${}^{\omega}\mathcal{G}_R^{\omega}$. We say that π_1* may join *π_2, denoted $\pi_1 \bowtie_R \pi_2$, if and only if (i) there exists an equivalence class $S \in \mathbf{x}_{/\sim_R}$ such that $vars(\pi_1) \cup vars(\pi_2) \subseteq S$ and (ii) there exists a path in ${}^{\omega}\mathcal{G}_R^{\omega}$ from some vertex in ${}^{\omega}\pi_1{}^{\omega}$ to some vertex in ${}^{\omega}\pi_2{}^{\omega}$.*

[13] An equivalent condition is that set of right frontier variables of μ_1 is the same as the set of left frontier variables of μ_2. Then β maps acyclic z-paths of μ_1 ending in x_i to acyclic z-paths of μ_2 starting with x_i, for all $1 \leq i \leq N$.

Proposition 5. *If $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ is a difference bounds relation, $\bowtie_R$ is an equivalence relation.*

Proof: Let $\mathbf{x}$ be the set of variables in the formulae defining R, and $\mathcal{G}_R$ be the constraint graph of R. $\bowtie_R$ is clearly reflexive, and transitivity is shown using the fact that the z-paths are repeating, hence the connecting paths occur periodically. To prove symmetry, let $\pi_1 \bowtie_R \pi_2$, where π_1 and π_2 are two repeating z-paths. We give the proof for the case where π_1 and π_2 are both forward, the other cases being symmetric. Let $x_i^{(k)}$ and $x_j^{(\ell)}$ be two positions in ${}^{\omega}\pi_1{}^{\omega}$ and ${}^{\omega}\pi_2{}^{\omega}$, respectively, such that there is a path ξ in ${}^{\omega}\mathcal{G}_R{}^{\omega}$ between them. Since π_1 and π_2 are repeating z-paths, there exists sub-paths ρ_1 and ρ_2 of π_1 and π_2, respectively, such that ρ_1 (ρ_2) starts with $x_i^{(k)}$ ($x_j^{(\ell)}$) and ends with a future occurrence of x_i (x_j). Since $x_i \sim_R x_j$, by the definition of $\bowtie_R$, there exists a path η in ${}^{\omega}\mathcal{G}_R{}^{\omega}$ between $x_j^{(\ell)}$ and some occurrence of x_i. We have built two paths, ρ_1 and $\rho_2' = \xi.\rho_2^n.\eta$, between $x_i^{(k)}$ and two occurrences of x_i, call them $x_i^{(m)}$ and $x_i^{(p)}$, respectively. We can assume w.l.o.g. that $m > k$ and $p > k$, for a sufficiently large $n \geq 0$. Let $M = lcm(m-k, p-k)$, $m_1 = \frac{M}{m-k}$ and $m_2 = \frac{M}{p-k}$. Then $x_i^{(k+m_1\cdot(m-k))}$ is a vertex on ${}^{\omega}\pi_1{}^{\omega}$, and $\rho_2^n.(\eta.\xi.\rho_2^n)^{m_2-1}.\eta$ is a path from $x_j^{(\ell)}$ to $x_i^{(k+m_1\cdot(m-k))}$. $\square$

For a repeating multipath μ, we denote by $\mu^{ac}_{/\bowtie_R}$ the partition of the set of acyclic paths μ^{ac} in equivalence classes of the $\bowtie_R$ relation. An *sc-multipath* (for strongly connected multipath) is a repeating multipath whose repeating z-paths belong to the same equivalence class of the $\bowtie_R$ relation (Fig 4 (d) and (e) for $R \equiv x_1 = x_2' \wedge x_2 = x_1' \wedge x_1 \leq x_1'$). A repeating multipath ν is said to be a *reduct* of a repeating multipath μ if and only if $\nu \subseteq \mu$ and, for each equivalence class $C \in \mu^{ac}_{/\bowtie_R}$: if the difference between the number of repeating forward (backward) z-paths and the number of repeating backward (forward) z-paths in C equals $k \geq 0$, then $\nu \cap C$ contains exactly k repeating forward (backward) z-paths and no repeating backward (forward) z-path.

Lemma 12. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a $*$-consistent difference bounds relation and $\mathcal{G}_R$ be its constraint graph. Let μ be a sc-multipath in ${}^{\omega}\mathcal{G}_R^{\omega}$ and ν be a reduct of μ. Then $\overline{\overline{w}}(\nu) \leq \overline{\overline{w}}(\mu)$.*

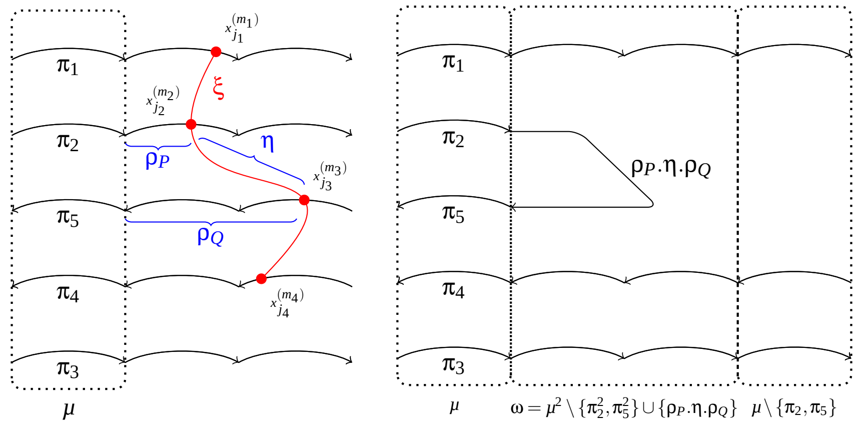
Proof: Let $\pi_1, \pi_2 \in \mu$ be two repeating odd forward and backward z-paths, respectively, such that $\pi_1 \bowtie_R \pi_2$. By Lemma 9, we have that $\overline{\overline{w}}(\pi_1) + \overline{\overline{w}}(\pi_2) \geq 0$, and since $\overline{\overline{w}}(\mu) = \overline{\overline{w}}(\pi_1) + \overline{\overline{w}}(\pi_2) + \overline{\overline{w}}(\mu \setminus \{\pi_1, \pi_2\})$, we obtain $\overline{\overline{w}}(\mu) \geq \overline{\overline{w}}(\mu \setminus \{\pi_1, \pi_2\})$. The statement of the lemma is proved by repeatedly eliminating two opposite z-paths $\pi_1, \pi_2 \in C$, for each $C \in \mu^{ac}_{/\bowtie_R}$, until no more z-paths can be eliminated. The result is a reduct ν of μ, of smaller average weight. Moreover, each reduct of μ can be obtained from μ in this way. $\square$

Consider for instance the sc-multipath $\mu = \{x_1^{(1)} \xrightarrow{0} x_2^{(2)} \xrightarrow{0} x_3^{(3)} \xrightarrow{0} x_4^{(2)} \xrightarrow{0} x_5^{(1)} \xrightarrow{-1} x_1^{(2)}, x_6^{(2)} \xrightarrow{1} x_6^{(1)}, x_7^{(1)} \xrightarrow{1} x_7^{(2)}\}$ from Fig. 2 (c). Then $\nu_1 = \{x_1^{(1)} \xrightarrow{0} x_2^{(2)} \xrightarrow{0} x_3^{(3)} \xrightarrow{0} x_4^{(2)} \xrightarrow{0} x_5^{(1)} \xrightarrow{-1} x_1^{(2)}\}$ and $\nu_2 = \{x_7^{(1)} \xrightarrow{1} x_7^{(2)}\}$ are the reducts of μ, and $1 = \overline{\overline{w}}(\mu) = \overline{\overline{w}}(\nu_2) > \overline{\overline{w}}(\nu_1) = -1$. Furthermore, any repeating multipath can be concatenated left and right with some of its reducts. For an example see Fig. 2 (c) – here μ can be connected with both ν_1 and ν_2, and back.

Lemma 13. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a $*$-consistent difference bounds relation, $\mathcal{G}_R$ be its constraint graph and μ be a sc-multipath in ${}^{\omega}\mathcal{G}_R^{\omega}$. Then there exists two reducts ν and ρ of μ, and multipaths ξ and η, such that $\mu^n.\xi.\nu^m$ and $\rho^n.\eta.\mu^m$ are valid multipaths, for all $n,m \geq 0$.*

Proof: We give the proof for the first point, the second being symmetrical. We assume w.l.o.g. that $\mu^{ac}_{/\bowtie_R}$ contains at least two repeating z-paths of opposite directions – otherwise the only reduct of μ is μ itself and the conclusion follows trivially. We first prove that there exists two z-paths $\pi_1, \pi_2 \in C$, of opposite directions, and a multipath ω, such that the composition $\mu^n.\omega.(\mu \setminus \{\pi_1,\pi_2\})^m$ is defined, for all $n,m \geq 0$. Symmetrically, one can also prove the existence of π_1, π_2 such that $(\mu \setminus \{\pi_1,\pi_2\})^n.\omega.\mu^m$ is defined, for all $n,m \geq 0$ – this later point is left to the reader.

Let $\mu^{ac} = \{\pi_1,\ldots,\pi_i,\pi_{i+1},\ldots,\pi_{i+h}\}$ be an ordering such that $\pi_1,\ldots,\pi_i$ are forward, and $\pi_{i+1},\ldots,\pi_{i+h}$ are backward repeating z-paths, for some $i,h > 0$. We assume w.l.o.g. that π_j is a z-path from $x_j^{(0)}$ to $x_j^{(k)}$ if $1 \leq j \leq i$, and from $x_j^{(k)}$ to $x_j^{(0)}$, if $i < j \leq i+h$, where $k = \|\mu\|$ is the relative length of μ. Since $\pi_1 \bowtie_R \pi_{i+h}$, there exist infinitely many paths in ${}^{\omega}\mathcal{G}_R{}^{\omega}$, from some vertex of ${}^{\omega}\pi_1{}^{\omega}$ to some vertex of ${}^{\omega}\pi_{i+h}{}^{\omega}$, and we can choose one such path ξ, which contains only vertices $x_\ell^{(m)}$, for $m > k$. Let $\pi_{i_1},\ldots,\pi_{i_s}$ be the paths from μ^{ac} which are intersected by ξ. Without loss of generality, we assume that ξ intersects $\pi_{i_1},\ldots,\pi_{i_s}$ exactly in this order. Clearly, there exists $P \in \{1,\ldots,s\}$ and $Q = P+1$ such that π_{i_P} is forward and π_{i_Q} is backward. Let $x_{j_P}^{(m_P)}$ ($x_{j_Q}^{(m_Q)}$) be a vertex where ξ and π_P (π_Q) intersect. Let η be the segment of ξ between $x_{j_P}^{(m_P)}$ and $x_{j_Q}^{(m_Q)}$. Further, let ρ_P and ρ_Q the segments of ${}^{\omega}\pi_P{}^{\omega}$ and ${}^{\omega}\pi_Q{}^{\omega}$ between $x_P^{(k)}$ and $x_{j_P}^{(m_P)}$, and between $x_Q^{(k)}$ and $x_{j_Q}^{(m_Q)}$, respectively. Let $L = \max\left(1, \lceil \frac{\rho_P.\eta.\rho_Q}{k} \rceil\right)$ and $\omega = (\mu^L \setminus \{\pi_h^L, \pi_\ell^L\}) \cup \{\rho_P.\eta.\rho_Q\}$. It is easy to verify that indeed $\mu^n.\omega.(\mu \setminus \{\pi_P,\pi_Q\})^m$ is a valid multipath, for all $n,m > 0$.



(a) finding $\rho_P.\eta.\rho_Q$ (b) connecting μ with $\mu \setminus \{\pi_2, \pi_5\}$

Fig. 5: Connecting sc-multipath with its reduct. Note that $i = 3, h = 2$, and ξ connects π_1^ω with π_4^ω. ξ intersects $\pi_1^\omega, \pi_2^\omega, \pi_5^\omega \pi_4^\omega$ in this order, hence $s = 4, i_1 = 1, i_2 = 2, i_3 = 5, i_4 = 4$. We choose $P = 2, Q = 3$ since $\pi_{i_2}^\omega$ is forward and $\pi_{i_Q}^\omega$ is backward. Since $\rho_P.\eta.\rho_Q] < 2k$ in the figure, $L = 2$ and $\omega = \mu^2 \setminus \{\pi_2^2, \pi_5^2\} \cup \{\rho_P.\eta.\rho_Q\}$.

Applying this reduction for all $C \in \mu^{ac}_{/\bowtie_R}$, until no z-paths can be eliminated, one obtains a reduct ν of μ and a path ξ such that $\mu^n.\xi.\nu^m$ is a valid multipath, for all $n, m \geq 0$. $\square$

Balanced SC-Multipaths A sc-multipath μ is said to be *balanced* if and only if the difference between the number of forward repeating and backward repeating z-paths in μ is either 1, 0, or -1. Let us observe that each reduct of a balanced sc-multipath contains at most one repeating z-path.

Lemma 14. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a $*$-consistent difference bounds relation, $\mathcal{G}_R$ be its constraint graph and μ be a balanced sc-multipath in ${}^\omega\mathcal{G}_R^\omega$. Then there exists an essential sc-multipath, $\tau = \{\tau_0\}$, such that τ_0 is an essential repeating z-path, $\overline{\overline{w}}(\tau) \leq \overline{\overline{w}}(\mu)$, and two sc-multipaths ξ and ζ such that $\mu^m.\xi.\tau^n.\zeta.\mu^p$ is a valid sc-multipath for all $m, n, p \geq 0$.*

Proof: Applying Lemma 13 we obtain two reducts ν and π and two multipaths α and γ such that $\mu^m.\alpha.\nu^n$ and $\pi^m.\gamma.\mu^n$ are valid multipaths, for all $m, n \geq 0$. Since μ is a balanced s-multipath, ν^{ac} and π^{ac} are either both empty, or they consist of a single repeating z-path each.

In the case when both ν^{ac} and π^{ac} are empty, they can be clearly concatenated by an empty, hence essential, multipath. Otherwise, ν^{ac} and π^{ac} consist both of a repeating z-path, and by Lemma 11, there exists a z-path β such that $\nu^p.\beta.\pi^q$ is a valid z-path, for all $p, q \geq 0$. Hence $\mu^m.\alpha.\nu^p.\beta.\pi^q.\gamma.\mu^n$ is a valid multipath, for all $m, n, p, q \geq 0$. Let $\nu^{ac} = \{\nu_0\}$. By Lemma 10, there exists an essential repeating subpath τ_0 of ν_0 such that $\overline{\overline{w}}(\tau_0) \leq \overline{\overline{w}}(\nu_0)$. By Lemma 11, there exist z-paths ρ_0, η_0 such that $\nu_0^m.\rho_0.\tau_0^n$ and $\tau_0^m.\eta_0.\nu_0^n$ are valid z-paths for all $m, n \geq 0$. Defining multipaths $\tau = \{\tau_0\}, \rho = \{\rho_0\}, \eta =$

$\{\eta_0\}$, it is easy to see that $\mu^m.\alpha.\nu^n.\rho.\tau^p.\eta.\nu^q.\beta.\pi^r.\gamma.\mu^s$ is a valid multipath for all $m,n,p,q,r,s \geq 0$. Finally, we can define $\xi = \alpha.\rho$ and $\zeta = \eta.\beta.\gamma$ to see that $\nu^m.\xi.\tau^p.\zeta.\nu^s$ is a valid multipath for all $m,p,s \geq 0$ and that the sc-multipath $\tau = \{\tau_0\}$ consists of a single essential repeating z-path. $\square$

The motivation for defining and studying balanced sc-multipaths can be found when examining the words generated by the iterations of a cycle $q \xrightarrow{\gamma} q$ in a zigzag automaton. Without losing generality, we assume that the state q is both *reachable* (from an initial state) and *co-reachable* (a final state is reachable from q). With this assumption, the following lemma proves that sufficiently many iterations of the γ cycle will exhibit a subword which is an arbitrarily large power of a word composed only of balanced sc-multipaths.

Proposition 6. *Let μ be an iterable multipath. If μ has at least one left corner, then it must also have at least a right corner, and viceversa.*

Proof: We prove the first implication, the second one being symmetrical. Let us suppose, by contradiction, that μ has at least a left corner, but no right corner. If μ is an iterable multipath, then its left and right frontiers are the same. We define the following sets:

- S_ℓ (S_r) is the set of variables from the left (right) frontier of μ which are sources of z-paths in μ ending on the right (left) frontier
- D_ℓ (D_r) is the set of variables from the left (right) frontier of μ which are destinations of z-paths in μ starting on the right (left) frontier
- C_ℓ (C_r) is the set of variables from the left frontier of μ which are either source of destination of left (right) corners

Clearly, S_ℓ, D_ℓ and C_ℓ are pairwise disjoint, and the same holds for S_r, D_r and C_r. Since μ has at least one left corner, but no right corners, all z-paths originating on the right frontier must end on the left frontier, and all z-paths that end on the right frontier must have originated on the left frontier i.e., $\mathrm{card}(S_\ell) = \mathrm{card}(D_r)$ and $\mathrm{card}(D_\ell) = \mathrm{card}(S_r)$. But $\mathrm{card}(C_\ell) > 0$ and $\mathrm{card}(C_r) = 0$, which contradicts with the fact that the left and right frontiers of μ are the same. $\square$

Lemma 15. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathbf{x} = \{x_1, \ldots, x_N\}$ be the variables occurring in its defining formulae, $T_R = \langle Q, \Delta, \omega\rangle$ be its transition table, and $A = \langle T_R, I, F\rangle$ be one of the zigzag automata from Thm. 5. If $q \in Q$ is a reachable and co-reachable state of A, and $q \xrightarrow{\gamma} q$ is a cycle in T_R, then there exist multipaths V, W and Z, such that:*

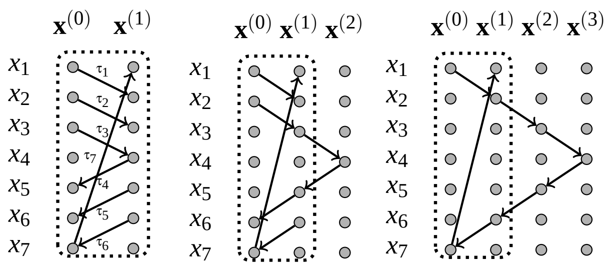
1. $\overline{\overline{w}}(Z) = \overline{w}(\gamma)$
2. *$Z = \mu_1 \cup \ldots \cup \mu_k$, where μ_i are balanced sc-multipaths, and, for all $1 \leq i < j \leq k$, there exist distinct equivalence classes $S_i, S_j \in \mathbf{x}_{/\sim_R}$, such that $vars(\mu_i) \subseteq S_i$ and $vars(\mu_j) \subseteq S_j$*
3. *for all $n \geq 0$ there exists $m > 0$ such that $\gamma^m = V.Z^n.W$*

Proof: The first part of the proof is concerned with the elimination of corners from γ. To this end, we define the following sequences of multipaths. Let $\nu_0 = \gamma$, and for each

$i \geq 0$, we have:

$$
\begin{aligned}
\pi_i &= \begin{cases} x^{(0)}_{p_i} \rightarrow \ldots \rightarrow x^{(0)}_{q_i} & \text{for some } 1 \leq p_i, q_i \leq N \text{, if } \nu_i \text{ has a right corner} \\ \varepsilon & \text{otherwise} \end{cases} \\
\rho_i &= \nu_i \setminus \{\pi_i\} \\
\nu_{i+1} &= \rho_i . \overrightarrow{\pi_i}^{|\gamma|}
\end{aligned}
$$

Since $\mathrm{card}(\nu_i) < \mathrm{card}(\nu_{i+1})$, for all $i \geq 0$, such that $\nu_i \neq \nu_{i+1}$, the sequence eventually reaches its limit for the least index $\ell \geq 0$, such that $\nu_\ell = \nu_{\ell+1}$. Let $Z = \nu_\ell$ denote the limit of this sequence.



(a) ν_0 (b) ν_1 (c) ν_2

$$
\begin{aligned}
\nu_0 &= \{\tau_1, \tau_2, \tau_3.\tau_4, \tau_5, \tau_6.\tau_7\} \\
\pi_0 &= \tau_3.\tau_4 \\
\rho_0 &= \{\tau_1, \tau_2, \tau_5, \tau_6.\tau_7\} \\
\nu_1 &= \{\tau_1, \tau_2.\tau_3.\tau_4.\tau_5, \tau_6.\tau_7\} \\
\pi_1 &= \tau_2.\tau_3.\tau_4.\tau_5 \\
\nu_1 &= \{\tau_1, \tau_6.\tau_7\} \\
\rho_2 &= \{\tau_1.\tau_2.\tau_3.\tau_4.\tau_5.\tau_6.\tau_7\} \\
\pi_2 &= \varepsilon \\
\rho_2 &= \emptyset
\end{aligned}
$$

Fig. 6: Elimination of corners from an sc-multipath ν_0.

First, let us observe that the limit does not depend on the choices of the right corners π_i at each step of the sequence. Second, one can show that Z is an iterable multipath, by induction on $\ell \geq 0$. Moreover, we have:

$$\nu_i.\pi_i.\nu^n_{i+1}.\rho_i = \nu^{n+2}_i \text{ , for all } 0 \leq i < \ell \text{ and } n > 0 \tag{10}$$

Using (10) it is not hard to prove, by induction on $\ell \geq 0$ that:

$$\nu^{n+2\ell}_0 = \nu_0.\pi_0.\cdots.\nu_{\ell-1}.\pi_{\ell-1}.\nu^n_\ell.\rho_{\ell-1}.\cdots.\rho_0 \text{ , for all } n \geq 0$$

Let $V = \nu_0.\pi_0.\cdots.\nu_{\ell-1}.\pi_{\ell-1}$ and $W = \rho_{\ell-1}.\cdots.\rho_0$. We have $\mathcal{G}^{n+2\ell}_\gamma = V.Z^n.W$, for all $n \geq 0$.

Third, we prove that Z has no corners. Clearly, it can have no right corners, since any right corner was eliminated before the $\{\nu_i\}_{i\geq 0}$ reached its fixpoint. But, since Z is an iterable multipath, it cannot have left corners either, by Prop. 6. Finally, we have that $\|Z\| = |\gamma|$ and $\overline{\overline{w}}(Z) = \overline{w}(\gamma)$, also by induction on $\ell \geq 0$.

Since Z is an iterable multipath without corners, each forward (backward) z-path in Z is of the form $x^{(0)}_{p_i} \rightarrow \ldots \rightarrow x^{(m)}_{q_i}$ ($x^{(m)}_{p_i} \rightarrow \ldots \rightarrow x^{(0)}_{q_i}$). Since Z is iterable, the left and right frontier of Z must contain the same set of variables $\mathbf{x}_\gamma$. Then Z induces a permutation Π on the set of indices in $\mathbf{x}_\gamma$: $\Pi(p) = q$ if and only if either $x^{(0)}_p \rightarrow \ldots \rightarrow$

$x_q^{(m)} \in Z$, or $x_p^{(m)} \to \ldots \to x_q^{(0)} \in Z$. Therefore, there exists a constant $s > 0$ such that Π^s is the identity function, and consequently, Z^s is a repeating multipath. Moreover, we clearly have $\overline{\overline{w}}(Z^s) = \overline{w}(\gamma)$.

It remains to be shown that Z^s is composed only of balanced sc-multipaths, whose variables pertain to different equivalence classes of the $\sim_R$ relation. Since $q \xrightarrow{\gamma} q$ is a cycle in A, and q is on a path from an initial to a final state of A, by Thm. 5, there exist initial and final states, $q_i \in I$ and $q_f \in F$, respectively, and paths $q_i \xrightarrow{\sigma} q$ and $q \xrightarrow{\tau} q_f$, in A, such that the word $\sigma.\gamma^m.\tau$ encodes an odd (even) forward (backward) z-path π, for some $m \geq 0$, such that $\mathcal{G}_\gamma^m = V.Z^s.W$ (by the previous argument, such m always exists). Let us consider the case when π is odd forward – the other cases, being symmetric, are left to the reader. Let $\pi_1, \pi_2, \ldots, \pi_h$ be the subpaths of π corresponding to the traversals of Z^s by π, in this order. Then h is odd, and moreover, for all $1 \leq i \leq h$, π_i is forward (backward) repeating z-path if i is odd (even). Thus, each equivalence class $C \in {Z^s}_{/\bowtie_R}$ is a balanced sc-multipath. Since all π_i are repeating z-paths, for each $1 \leq i \leq h$, there exists $S_i \in \mathbf{x}_{/\sim_R}$ such that $vars(\pi_i) \in S_i$. Similarly, there exists $T_i \in Z^s_{/\bowtie_R}$ such that $\pi_i \in T_i$. For all $1 \leq i < j \leq h$, π_i can be connected with π_j with some subpath of π. Therefore, if $S_i = S_j$, it follows from Prop. 5 that $\pi_i \bowtie_R \pi_t \bowtie_R \pi_j$ for all $i < t < j$. It thus follows that $T_i \neq T_j$ implies $S_i \neq S_j$, for all $1 \leq i < j \leq h$. $\square$

For instance, for the γ_1 zigzag cycle in Fig. 2 (b), the balanced sc-multipath is $Z = \{x_1^{(2)} \xrightarrow{0} x_2^{(3)} \xrightarrow{0} x_3^{(4)} \xrightarrow{0} x_4^{(3)} \xrightarrow{0} x_5^{(2)} \xrightarrow{-1} x_1^{(3)},\ x_6^{(3)} \xrightarrow{1} x_6^{(2)},\ x_7^{(2)} \xrightarrow{1} x_7^{(3)}\}$ (highlighted in Fig. 2 (c)), and the connecting multipaths are $V = \{x_2^{(2)} \xrightarrow{0} x_3^{(3)} \xrightarrow{0} x_4^{(2)}\}$ and $W = \{x_1^{(2)} \xrightarrow{0} x_2^{(3)},\ x_4^{(3)} \xrightarrow{0} x_5^{(2)} \xrightarrow{-1} x_1^{(3)},\ x_6^{(3)} \xrightarrow{1} x_6^{(2)},\ x_7^{(2)} \xrightarrow{1} x_7^{(3)}\}$, $\overline{w}(\gamma) = \overline{\overline{w}}(Z) = 1$ and $\gamma_1^n = V.Z^{n-2}.W$, for all $n \geq 2$.

Strongly Connected Zigzag Cycles The next lemma maps Z back into another critical elementary loop $q' \xrightarrow{\lambda} q'$ of the zigzag automaton, belonging to the same SCC as γ, such that λ is composed of essential powers, and $\overline{w}(\lambda) = \overline{\overline{w}}(Z) = \overline{w}(\gamma)$. This is the final step needed to conclude the proof of Thm. 6.

Lemma 16. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathbf{x} = \{x_1, \ldots, x_N\}$ be the variables occurring in its first defining formulae, $T_R = \langle Q, \Delta, \omega \rangle$ be its transition table, and $A = \langle T_R, I, F \rangle$ be one of the zigzag automata from Thm. 5. Let σ and τ be two multipaths such that $\sigma.\tau \in Enc(\mathcal{L}(A))$, and μ be a multipath such that $\sigma.\mu.\tau$ is a valid fitting multipath. Then $\sigma.\mu.\tau \in Enc(\mathcal{L}(A))$.*

Proof: We give the proof only in the case where A encodes all odd forward z-paths between two variables $x_i, x_j \in \mathbf{x}$ i.e., A is the zigzag automaton A_{ij}^{of} from Thm. 5. The other cases are symmetric. For simplicity, we assume that $\sigma.\tau$ consists only of a fitting z-path $\pi : x_i^{(0)} \to \ldots \to x_j^{(k)}$, where $k = |\sigma.\tau|$. The case where $\sigma.\tau$ contains also several cycles is dealt with as an easy generalization.

If $\|\sigma\| = h < k$, let $\nu_0, \ldots, \nu_{2n-1}$ be the vertices of π of the form $x_i^{(h)}$, such that the predecessor and the successor of $x_i^{(h)}$ on π do not lie both in σ, nor in τ. Then, the separation of π between σ and τ induces a factorization $\pi = \pi_1. \cdots .\pi_{2n}$, where each π_ℓ is a subpath of π between two vertices $v_\ell, v_{\ell+1}$, for all $\ell = 0,\ldots,2n-1$, and $x_i^{(0)} = v_0$, $x_j^{(k)} = v_{2n}$. We define the graph $G = \langle \{v_0,\ldots,v_{2n}\}, E\rangle$ in the following way:

- $(v_{2n}, v_0) \in E$
- $(v_\ell, v_{\ell+1}) \in E$, for all $\ell = 0,\ldots,2n-1$
- $(v_\ell, v_m) \in E$ iff there exists a z-path from v_ℓ to v_m in $\sigma.\mu.\tau$

It is easy to prove that each vertex in G has exactly one incoming and one outgoing edge if and only if $\sigma.\mu.\tau$ is a valid multipath. Then the graph $G' = \langle \{v_0,\ldots,v_{2n}\}, E \setminus \{(v_{2n},v_0)\}\rangle$ defines a z-path from $x_i^{(0)}$ to $x_j^{(k)}$, which traverses each vertex $v_1,\ldots,v_{2n-1}$ at most once, and possibly several cycles involving the remaining vertices, not on this z-path. Then $\sigma.\mu.\tau$ encodes a z-path from $x_i^{(0)}$ to $x_j^{(k)}$, which moreover is fitting, by the hypothesis. Hence $\sigma.\mu.\tau \in Enc(\mathcal{L}(A))$. □

Let $\pi : x_{i_1}^{(j_1)} \to \ldots \to x_{i_n}^{(j_n)}$ is a repeating z-path ($i_1 = i_n$) and $1 \le k \le n$ is an integer, the z-path $\overset{\frown}{\pi}^k : (x_{i_k}^{(j_k)} \to \ldots \to x_{i_n}^{(j_n)}).(x_{i_n=i_1}^{(j_1)} \to \ldots \to x_{i_k}^{(j_k)})$ is called the *rotation of* π *by* k. The *left* and *right remainders* of a rotation $\overset{\frown}{\pi}^k$ are the unique z-paths π_l and π_r, respectively, such that $\pi^2 = \pi_l. \overset{\frown}{\pi}^k .\pi_r$. Given a z-path π of the above form and an integer $0 \le s < \|\pi\|$, let $minpos_\pi(s) = min\{k \mid j_k = s + j_1\}$.

Proposition 7. *Any rotation* ρ *of a repeating z-path* π *is repeating,* $\|\rho\| = \|\pi\|$ *and* $\overline{\overline{w}}(\rho) = \overline{\overline{w}}(\pi)$*. Moreover, if* π *is elementary, so is* ρ.

Proof: Let $\pi = \pi_1.\pi_2$, where $\pi_1 : x_{i_1}^{(j_1)} \to \ldots \to x_{i_k}^{(j_k)}$ and $\pi_2 : x_{i_k}^{(j_k)} \to \ldots \to x_{i_n}^{(j_n)}$, and $\rho = \overset{\frown}{\pi}^k = \pi_2.\pi_1$, for some $1 \le k \le n$. Then $\rho^m = \pi_2.(\pi_1.\pi_2)^{m-1}.\pi_1$, which is clearly a valid z-path, for all $m > 0$. Since $\|\pi\| = \|\pi_1\| + \|\pi_2\| = \|\rho\|$, we obtain that $\overline{\overline{w}}(\rho) = \overline{\overline{w}}(\pi)$. The last point can be easily proved by contradiction. □

Proposition 8. *Let* π *be an essential repeating z-path. Then there exists a word* G *such that (i)* ${}^\omega\pi^\omega = {}^\omega G^\omega$*, (ii)* $\|\pi\| = |G|$ *and (iii)* $\overline{w}(G) = \overline{\overline{w}}(\pi)$.

Proof: Let $k \in \mathbb{Z}$ be an integer, and $x_{i_1}^{(k)},\ldots,x_{i_n}^{(k)}$ be the variables that occur at position k in ${}^\omega\pi^\omega$. Then there exists integers $m_1,\ldots,m_n \in \{1,\ldots,|\pi|\}$ and $\ell_1,\ldots,\ell_n \in \mathbb{Z}$, such that each $x_{i_j}^{(k)}$ occurs as the m_j-th variable on the ℓ_j-th copy of π in ${}^\omega\pi^\omega$. Notice that each $x_{i_j}^{(k)}$ has exactly one incoming and one outgoing edge. Since π is an elementary repeating z-path, $\overset{\frown}{\pi}^{m_j}$ is also an elementary repeating z-path, and $\|\overset{\frown}{\pi}^{m_j}\| = \|\pi\|$. for each $j = 1,\ldots,n$ (by Prop. 7). Moreover, $\overset{\frown}{\pi}^{m_j}$ starts and ends with x_{i_j}, and there is no other occurrence of x_{i_j} on it, apart from the initial and final vertex. Hence the "cut" through ${}^\omega\pi^\omega$ at k is identical to the one at position $k + \|\pi\|$: the same variables occur on both positions, and they are traversed by ${}^\omega\pi^\omega$ in the same ways (left-to-right, right-to-left, left-to-left or right-to-right). Let G be the graph consisting of all edges of ${}^\omega\pi^\omega$ situated between k and $k + \|\pi\|$. Clearly ${}^\omega\pi^\omega = {}^\omega G^\omega$ and $|G| = \|\pi\|$. Since every edge in π must occur in G as well, we have $w(G) = w(\pi)$, hence $\overline{w}G = \overline{\overline{w}}(\pi)$ as well. □

Lemma 17. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathbf{x} = \{x_1, \ldots, x_N\}$ be the variables occurring in its defining formulae, $T_R = \langle Q, \Delta, \omega \rangle$ be its transition table, and $A = \langle T_R, I, F \rangle$ be one of the zigzag automata from Thm. 5. If $q \in Q$ is a reachable and co-reachable state of A, and $q \xrightarrow{\gamma} q$ is a cycle, then there exists a state $q' \in Q$, a cycle $q' \xrightarrow{\lambda} q'$, and paths $q \to q'$ and $q' \to q$ in T_R, such that (i) $\overline{w}(\lambda) \leq \overline{w}(\gamma)$, and (ii) $|\lambda| \mid lcm(1, \ldots, N)$.*

Proof: If q is a reachable and co-reachable state of A, there exist initial and final states $q_i \in I$ and $q_f \in F$, respectively, and paths $q_i \xrightarrow{\iota} q$, $q \xrightarrow{\phi} q_f$ and $q \xrightarrow{\gamma} q$ in T_R. By Lemma 15, there exists a multipath Z, composed of balanced sc-multipaths $\mu_1, \ldots, \mu_k$, such that:

- $\overline{\overline{w}}(Z) = \overline{w}(\gamma)$
- for all $1 \leq i < j \leq k$, there exist $S_i, S_j \in \mathbf{x}_{/\sim_R}$, such that $S_i \neq S_j$, $vars(\mu_i) \subseteq S_i$, and $vars(\mu_j) \subseteq S_j$
- $\exists V, W \ . \ \forall n \geq 0 \ \exists m > 0 \ . \ \gamma^n = V.Z^n.W$

Since $\|Z\| = \|\mu_1\| = \ldots = \|\mu_k\|$, we have $\overline{\overline{w}}(Z) = \sum_{i=1}^{k} \overline{\overline{w}}(\mu_i)$. By Lemma 14, for all $1 \leq i \leq k$, there exist:

- essential sc-multipaths $\tau_i = \{\nu_i\}$, where ν_i are essential repeating z-paths, and $\overline{\overline{w}}(\tau_i) \leq \overline{\overline{w}}(\mu_i)$
- sc-multipaths ξ_i and η_i such that:

$$\forall n, m, t \geq 0 \ . \ \mu_i^n.\xi_i.\tau_i^t.\eta_i.\mu_i^m \text{ is a valid sc-multipath}$$

Clearly, for all $i = 1, \ldots, k$ we have $vars(\tau_i) \subseteq S_i$, hence $vars(\tau_i) \cap vars(\tau_j) = \emptyset$, for all $1 \leq i < j \leq k$. Let $P = lcm(\|\tau_1\|, \ldots, \|\tau_k\|)$ and $q_i = \frac{P}{\|\tau_i\|}$, for all $i = 1, \ldots, k$. We have:

$$Z = \begin{bmatrix} \mu_1 \\ \ldots \\ \mu_k \end{bmatrix} \quad \text{and} \quad \mathcal{L}_{\langle j_1, \ldots, j_k \rangle} \stackrel{\text{def}}{=} \begin{bmatrix} \overset{\curvearrowright j_1}{\nu_1^{q_1}} \\ \ldots \\ \overset{\curvearrowright j_k}{\nu_k^{q_k}} \end{bmatrix} \quad \text{for any tuple } \langle j_1, \ldots, j_k \rangle \in [P]^k$$

It suffices to prove the existence of integers $j_1, \ldots, j_k \in [P]$, and of multipaths Z_1 and Z_2 such that:

$$\forall m, n, s \geq 0 \ . \ Z^n.Z_1.\mathcal{L}^s_{\langle j_1, \ldots, j_k \rangle}.Z_2.Z^m \text{ is a valid multipath} \tag{11}$$

Assume first that (11) is true. Since ν_i are essential repeating z-paths, so are $\overset{\curvearrowright j_i}{\nu_i^{q_i}}$, and thus $\mathcal{L}_{\langle j_1, \ldots, j_k \rangle}$ is a (repeating) essential multipath. It is not hard to show that, for all $s > 0$:

$$\begin{bmatrix} \mathcal{L}_{\langle j_1, \ldots, j_k \rangle} = \mathcal{L}^s_{\langle j_1, \ldots, j_k \rangle} \\ \mathcal{L}_{\langle j_1, \ldots, j_k \rangle} = \mathcal{L}^s_{\langle j_1, \ldots, j_k \rangle} \end{bmatrix} \text{ which implies } \begin{bmatrix} Z_1.\mathcal{L}_{\langle j_1, \ldots, j_k \rangle} = Z_1.\mathcal{L}^s_{\langle j_1, \ldots, j_k \rangle} \\ \mathcal{L}_{\langle j_1, \ldots, j_k \rangle}.Z_2 = \mathcal{L}^s_{\langle j_1, \ldots, j_k \rangle}.Z_2 \end{bmatrix}$$

Since for all $V.Z^n.W$ is fitting, for all $n \geq 0$, it turns out that, for all $s \geq 0$:

$$\forall n \geq \left\lceil \frac{\left[Z_1.\mathcal{L}^s_{\langle j_1,\ldots,j_k\rangle}\right]}{\|Z\|} \right\rceil \forall m \geq \left\lceil \frac{\left[\mathcal{L}^s_{\langle j_1,\ldots,j_k\rangle}.Z_2\right]}{\|Z\|} \right\rceil . \iota.V.Z^n.Z_1.\mathcal{L}^s_{\langle j_1,\ldots,j_k\rangle}.Z_2.Z^m.W.\phi \text{ is a fitting multipath}$$

Since $\iota.V.Z^n.Z^m.W.\phi \in \mathcal{L}(A)$ for all $n,m \geq 0$, by Lemma 16, we have that $\iota.V.Z^n.Z_1.\mathcal{L}^s_{\langle j_1,\ldots,j_k\rangle}.Z_2.Z^m.W.\phi \in \mathcal{L}(A)$ as well. Moreover, by the definition of the states $q \in Q$ of the zigzag automaton (as tuples $\mathbf{q} \in \{\ell, r, \ell r, r\ell, \perp\}^N$ describing a vertical cut of a word at some position) it is not hard to see that:

$$q \xrightarrow{V.Z^n.Z_1.\mathcal{L}^s_{\langle j_1,\ldots,j_k\rangle}.Z_2.Z^m.W} q \tag{12}$$

is a cycle in T_R for each $s \geq 0$.

Let us consider the bi-infinite iteration ${}^{\omega}\mathcal{L}^{\omega}_{\langle j_1,\ldots,j_k\rangle}$. Since $\tau_i = \{\nu_i\}$ and ν_i are essential repeating z-paths, $\vartheta_i \stackrel{def}{=} \overset{\curvearrowright j_i}{\nu_i^{q_i}}$ are essential repeating z-paths as well (by Prop 7). Moreover, we have:

$$\|\tau_1\| = \ldots = \|\tau_k\|$$

and, by Prop. 8, there exist words $G_1, \ldots, G_k$ such that:

- ${}^{\omega}\vartheta_i{}^{\omega} = {}^{\omega}G_i{}^{\omega}$
- $|G_i| = \|\vartheta_i\| = \|\tau_i\|$
- $\overline{w}(G_i) = \overline{\overline{w}}(\vartheta_i) = \overline{\overline{w}}(\tau_i)$

Hence the word:

$$\lambda = \begin{bmatrix} G_1 \\ \ldots \\ G_k \end{bmatrix}$$

is well defined, since $|G_1| = \ldots = |G_k|$. Moreover, λ^i is a valid word, for all $i > 0$, hence there exists a cycle λ, within (12). Clearly $|\lambda| = |G_i| = \|\tau_i\| \cdot q_i = P \mid lcm(1, \ldots, N)$, since each ν_i is an essential z-path, and hence $\|\nu_i\| \leq N$. Also, since $\overline{w}(G_i) = \overline{\overline{w}}(\tau_i)$, we obtain that $\overline{w}(\lambda) = \overline{\overline{w}}(\mathcal{L}_{\langle j_1,\ldots,j_k\rangle})$, and consequently:

$$\overline{w}(\lambda) = \overline{\overline{w}}(\mathcal{L}_{j_1,\ldots,j_k}) \stackrel{\text{Prop. 7}}{=} \sum_{i=1}^{k} \overline{\overline{w}}(\tau_i) \stackrel{\text{Lemma 12}}{\leq} \sum_{i=1}^{k} \overline{\overline{w}}(\mu_i) = \overline{\overline{w}}(Z) = \overline{w}(\gamma)$$

Turning back to the existence of an integer tuple $\langle j_1, \ldots, j_k\rangle \in [P]^k$ and of multipaths Z_1 and Z_2 (11), we prove the case $k = 2$, the generalization to the case $k \geq 2$ being among the same lines. We distinguish two cases:

1. $\|\xi_1\| + \|\eta_1\| = \|\xi_2\| + \|\eta_2\|$: It is easy to check that $\|\xi_1.\tau_1^{q_1 \cdot k}.\eta_1\| = \|\xi_2.\tau_2^{q_2 \cdot k}.\eta_2\|$ for all $k > 0$. Since $vars(\tau_1) \cap vars(\tau_2) = \emptyset$, the following is a valid multipath for all $k > 0$:

$$\begin{bmatrix} \xi_1 \ . \ \tau_1^{q_1 \cdot k} \ . \ \eta_1 \\ \xi_2 \ . \ \tau_2^{q_2 \cdot k} \ . \ \eta_2 \end{bmatrix}$$

   We further consider three subcases: $\|\xi_1\| > \|\xi_2\|$, $\|\xi_1\| < \|\xi_2\|$ and $\|\xi_1\| = \|\xi_2\|$. We cover here the subcase $\|\xi_1\| > \|\xi_2\|$ (the subcase $\|\xi_1\| < \|\xi_2\|$ is symmetric and the subcase $\|\xi_1\| = \|\xi_2\|$ is trivial). Let $r > 0$, $0 \leq s < P$ be the unique integers such that $\|\xi_1\| - \|\xi_2\| = r \cdot P - s$ and define $j_1 = minpos_{\tau_1^{q_1}}(s)$, $j_1 = 0$. Let π_ℓ and π_r be the left and right remainders of the rotation $\vartheta_1 = \overset{\curvearrowright\, j_1}{\nu_1^{q_1}}$, respectively, i.e. $\pi_\ell.\vartheta_1.\pi_r = \nu_1^2$. We next define the multipath:

$$\begin{bmatrix} \xi_1.\pi_l \\ \xi_2.\nu_2^{q_2 \cdot r} \end{bmatrix} \cdot \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix}^k \cdot \begin{bmatrix} \pi_r.\nu_1^{q_1 \cdot (r-1)}.\eta_1 \\ \eta_2 \end{bmatrix} = Z_1.\mathcal{L}^k_{\langle j_1, j_2 \rangle}.Z_2$$

By observing that $\|\pi_\ell\| = s$, $\|\pi_r\| = P - s$, and $\|\xi_1\| - \|\xi_2\| = \|\eta_2\| - \|\eta_1\|$, we verify that $\|\xi_1.\pi_\ell\| = \|\xi_2.\tau_2^{q_2 \cdot r}\|$ and $\|\pi_r.\tau_1^{q_1 \cdot (r-1)}.\eta_1\| = \|\eta_2\|$:

$$\begin{aligned} \|\xi_1\| + \|\pi_\ell\| &= \|\xi_2\| + \|\tau_2\| \cdot q_2 \cdot r \\ \|\xi_1\| + s &= \|\xi_2\| + P \cdot r \\ \|\xi_1\| + \|\xi_2\| &= P \cdot r - s \end{aligned} \qquad \begin{aligned} \|\pi_r\| + \|\tau_1\| \cdot q_1 \cdot (r-1) + \|\eta_1\| &= \|\eta_2\| \\ P - s + P \cdot r - P &= \|\eta_2\| - \|\eta_1\| \\ P \cdot r - s &= \|\xi_1\| - \|\xi_2\| \end{aligned}$$

Thus, the above multipath $Z_1.\mathcal{L}^k_{\langle j_1, j_2 \rangle}.Z_2$ is valid for all $k \geq 0$.

2. $\|\xi_1\| + \|\eta_1\| \neq \|\xi_2\| + \|\eta_2\|$: We reduce the problem to the first case. We define

$$\begin{aligned} N &= lcm(\|\xi_1\| + q_1 \cdot \|\tau_1\| + \|\eta_1\|, \|\xi_2\| + q_2 \cdot \|\tau_2\| + \|\eta_2\|) \\ n_1 &= \tfrac{N}{\|\xi_1\| + q_1 \cdot \|\tau_1\| + \|\eta_1\|} - 1 \\ n_2 &= \tfrac{N}{\|\xi_2\| + q_2 \cdot \|\tau_2\| + \|\eta_2\|} - 1 \end{aligned}$$

Clearly, now $\|\xi_1.\tau_1^{q_1}.\eta_1.(\xi_1.\tau_1^{q_1}.\eta_1)^{n_1}\| = \|\xi_2.\tau_2^{q_2}.\eta_2.(\xi_2.\tau_2^{q_2}.\eta_2)^{n_2}\|$ and by choosing

$$\xi_1' = \xi_1.\tau_1^{q_1} \qquad \eta_1' = \eta_1.(\xi_1.\tau_1^{q_1}.\eta_1)^{n_1} \qquad \xi_2' = \xi_2.\tau_2^{q_2} \qquad \eta_2' = \eta_2.(\xi_2.\tau_2^{q_2}.\eta_2)^{n_2}$$

the problem reduces to the first case.

□

Consider, for example, the reachable and co-reachable cycle $\mathbf{q}_2 \xrightarrow{\gamma_1} \mathbf{q}_2$ in Fig. 2 (b). Then we have that $\overline{w}(\gamma_5) = \overline{w}(\gamma_1) = 1$, $-1 = \overline{w}(\gamma_8) < \overline{w}(\gamma_1) = 1$, and the cycles $\mathbf{q}_6 \xrightarrow{\gamma_5} \mathbf{q}_6$ and $\mathbf{q}_{10} \xrightarrow{\gamma_8} \mathbf{q}_{10}$ are in the same SCC as $\mathbf{q}_2 \xrightarrow{\gamma_1} \mathbf{q}_2$. Notice that both the unfoldings of γ_5 and γ_8 encode powers of the essential repeating z-paths $x_7^{(2)} \xrightarrow{1} x_7^{(3)}$ and $x_1^{(2)} \xrightarrow{0} x_2^{(3)} \xrightarrow{0} x_3^{(4)} \xrightarrow{0} x_4^{(3)} \xrightarrow{0} x_5^{(2)} \xrightarrow{-1} x_1^{(3)}$, respectivelly. Moreover, $|\gamma_5| = |\gamma_8| = |\gamma_1| = 1$ and $\mathbf{q}_{10} \xrightarrow{\gamma_8} \mathbf{q}_{10}$ is a critical cycle in its SCC.

Proof of Thm. 6: Let $\mathbf{x} = \{x_1, \ldots, x_N\}$ be the variables occurring in the arithmetic representation of a difference bounds relation $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$, $T_R = \langle Q, \Delta, \omega \rangle$ be its transition table, and $A_{ij}^\bullet = \langle T_R, I_{ij}^\bullet, F_{ij}^\bullet \rangle$, for $\bullet \in \{ef, eb, of, ob\}$ and $1 \leq i, j \leq N$ be the zigzag automata from Thm. 5. If $\mathcal{M}_R$ denotes the incidence matrix of T_R, the sequence $\{M_{R^i}\}_{i=1}^\infty$ of DBM encodings of $R^0, R^1, R^2, \ldots$ is the projection of the sequence $\{\mathcal{M}_R^{\boxtimes^i}\}_{i=1}^\infty$ of tropical powers onto the entries corresponding to all pairs of vertices:

$$(q_i,q_f) \in \bigcup_{1\le i,j\le N} I^{ef}_{ij} \times F^{ef}_{ij} \cup \bigcup_{1\le i,j\le N} I^{eb}_{ij} \times F^{eb}_{ij} \cup \bigcup_{1\le i,j\le N} I^{of}_{ij} \times F^{of}_{ij} \cup \bigcup_{1\le i,j\le N} I^{ob}_{ij} \times F^{ob}_{ij}$$

These entries denote the weights of the minimal paths in T_R between these vertices. By Thm. 3, the period of the $\{M_{R^i}\}_{i=1}^{\infty}$ sequence is the least common multiple of the cyclicities of all SCCs of T_R, which contain at least one state which is both reachable from $I^{\bullet}_{ij}$ and co-reachable from $F^{\bullet}_{ij}$, for some $\bullet \in \{ef,eb,of,ob\}$. By Lemma 17, each SCC in each $A^{\bullet}_{ij}$, $\bullet \in \{ef,eb,of,ob\}$ contains an elementary cycle λ of length $|\lambda| \mid lcm(1,\ldots,N)$. Hence the period of the sequence divides $lcm(1,\ldots,N)$. By Lemma 8, the period is bounded by $lcm(1,\ldots,N) = 2^{O(N)}$. $\square$

8 Octagonal Relations

The class of integer octagonal constraints is defined as follows:

Definition 10. *A formula* $\phi(\mathbf{x})$ *is an* octagonal constraint *if it is a finite conjunction of terms of the form* $x_i - x_j \le a_{ij}$, $x_i + x_j \le b_{ij}$ *or* $-x_i - x_j \le c_{ij}$ *where* $a_{ij}, b_{ij}, c_{ij} \in \mathbb{Z}$, *for all* $1 \le i,j \le N$. *A relation* $R \subseteq \mathbb{Z}^{\mathbf{x}} \times \mathbb{Z}^{\mathbf{x}}$ *is an* octagonal relation *if it can be defined by an octagonal constraint* $\phi_R(\mathbf{x},\mathbf{x}')$.

We represent octagons as difference bounds constraints over the dual set of variables $\mathbf{y} = \{y_1, y_2, \ldots, y_{2N}\}$, with the convention that y_{2i-1} stands for x_i and y_{2i} for $-x_i$, respectively. For example, the octagonal constraint $x_1 + x_2 = 3$ is represented as $y_1 - y_4 \le 3 \wedge y_2 - y_3 \le -3$. In order to handle the $\mathbf{y}$ variables in the following, we define $\bar{\imath} = i-1$, if i is even, and $\bar{\imath} = i+1$ if i is odd. Obviously, we have $\bar{\bar{\imath}} = i$, for all $i \in \mathbb{N}$. We denote by $\overline{\phi}(\mathbf{y})$ the difference bounds constraint over $\mathbf{y}$ that represents $\phi(\mathbf{x})$:

Definition 11. *Given an octagonal constraint* $\phi(\mathbf{x})$, $\mathbf{x} = \{x_1,\ldots,x_N\}$, *its* difference bounds representation $\overline{\phi}(\mathbf{y})$, *over* $\mathbf{y} = \{y_1,\ldots,y_{2N}\}$, *is a conjunction of the following difference bounds constraints, where* $1 \le i,j \le N$, $c \in \mathbb{Z}$.

$$\begin{array}{ll}
(x_i - x_j \le c) \in Atom(\phi) & \Leftrightarrow (y_{2i-1} - y_{2j-1} \le c), (y_{2j} - y_{2i} \le c) \in Atom(\overline{\phi}) \\
(-x_i + x_j \le c) \in Atom(\phi) & \Leftrightarrow (y_{2j-1} - y_{2i-1} \le c), (y_{2i} - y_{2j} \le c) \in Atom(\overline{\phi}) \\
(-x_i - x_j \le c) \in Atom(\phi) & \Leftrightarrow (y_{2i} - y_{2j-1} \le c), (y_{2j} - y_{2i-1} \le c) \in Atom(\overline{\phi}) \\
(x_i + x_j \le c) \in Atom(\phi) & \Leftrightarrow (y_{2i-1} - y_{2j} \le c), (y_{2j-1} - y_{2i} \le c) \in Atom(\overline{\phi})
\end{array}$$

The following equivalence relates ϕ and $\overline{\phi}$:

$$\phi(\mathbf{x}) \Leftrightarrow (\exists y_2, y_4, \ldots, y_{2N} \,.\, \overline{\phi} \wedge \bigwedge_{i=1}^{N} y_{2i-1} = -y_{2i})[x_i/y_{2i-1}]_{i=1}^{N} \tag{13}$$

An octagonal constraint ϕ is equivalently represented by the DBM $M_{\overline{\phi}} \in \mathbb{Z}_{\infty}^{2N \times 2N}$, corresponding to $\overline{\phi}$. We say that a DBM $M \in \mathbb{Z}_{\infty}^{2N \times 2N}$ is *coherent*[14] iff $M_{ij} = M_{\bar{\jmath}\bar{\imath}}$ for all

[14] DBM coherence is needed because $x_i - x_j \le c$ can be represented as both $y_{2i-1} - y_{2j-1} \le c$ and $y_{2j} - y_{2i} \le c$.

$1 \leq i,j \leq 2N$. Dually, for a coherent DBM $M \in \mathbb{Z}_\infty^{2N\times 2N}$, we define:

$$\begin{array}{l}
\Psi_M^{uu} \equiv \bigwedge_{1\leq i,j\leq N} x_i - x_j \leq M_{2i-1,2j-1} \wedge x_i + x_j \leq M_{2i-1,2j} \wedge -x_i - x_j \leq M_{2i,2j-1} \\
\Psi_M^{up} \equiv \bigwedge_{1\leq i,j\leq N} x_i - x'_j \leq M_{2i-1,2j-1} \wedge x_i + x'_j \leq M_{2i-1,2j} \wedge -x_i - x'_j \leq M_{2i,2j-1} \\
\Psi_M^{pu} \equiv \bigwedge_{1\leq i,j\leq N} x'_i - x_j \leq M_{2i-1,2j-1} \wedge x'_i + x_j \leq M_{2i-1,2j} \wedge -x'_i - x_j \leq M_{2i,2j-1} \\
\Psi_M^{pp} \equiv \bigwedge_{1\leq i,j\leq N} x'_i - x'_j \leq M_{2i-1,2j-1} \wedge x'_i + x'_j \leq M_{2i-1,2j} \wedge -x'_i - x'_j \leq M_{2i,2j-1}
\end{array}$$

A coherent DBM M is said to be *octagonal-consistent* if and only if Ψ_M^{uu} is consistent.

Definition 12. *An octagonal-consistent coherent DBM $M \in \mathbb{Z}_\infty^{2N\times 2N}$ is said to be* tightly closed *iff it is closed and, for all $1 \leq i,j \leq 2N$, $M_{i\bar{\imath}}$ is even, and $M_{ij} \leq \lfloor \frac{M_{i\bar{\imath}}}{2} \rfloor + \lfloor \frac{M_{\bar{j}j}}{2} \rfloor$.*

Intuitivelly the conditions of Def. 12 ensure that all knowledge induced by the triangle inequality and the $y_{2i-1} = -y_{2i}$ constraints (13) has been propagated in the DBM. Given an octagonal-consistent coherent DBM $M \in \mathbb{Z}^{2N} \times \mathbb{Z}^{2N}$, we denote the (unique) logically equivalent tightly closed DBM by M^t. The following theorem from [1] provides an effective way of testing octagonal-consistency and computing the tight closure of a coherent DBM.

Theorem 8. *[1] Let $M \in \mathbb{Z}_\infty^{2N\times 2N}$ be a coherent DBM. Then M is octagonal-consistent if and only if M is consistent and $\lfloor \frac{M^*_{i\bar{\imath}}}{2} \rfloor + \lfloor \frac{M^*_{\bar{\imath}i}}{2} \rfloor \geq 0$, 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:*

$$M^t_{ij} = \min\left\{ M^*_{ij}, \left\lfloor \frac{M^*_{i\bar{\imath}}}{2} \right\rfloor + \left\lfloor \frac{M^*_{\bar{j}j}}{2} \right\rfloor \right\}$$

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

The tight closure of DBMs is needed for checking equivalence between octagonal constraints.

Proposition 9 ([18]). *Let ϕ_1 and ϕ_2 be octagonal-consistent octagonal constraints. Then, $\phi_1 \Leftrightarrow \phi_2$ if and only if $M^t_{\phi_1} = M^t_{\phi_2}$.*

Octagonal constraints are closed under existential quantification, thus octagonal relations are closed under composition [5]. Tight closure of octagonal-consistent DBMs is needed for quantifier elimination.

Proposition 10 ([5]). *Let $\phi(\mathbf{x})$, $\mathbf{x} = \{x_1,\ldots,x_N\}$, be an octagonal-consistent octagonal constraint. Further, let $1 \leq k \leq 2N$ and M' be the restriction of M^t_ϕ to $\mathbf{y} \setminus \{y_{2k-1}, y_{2k}\}$. Then, M' is tightly closed, and $\Psi^{uu}_{M'} \Leftrightarrow \exists x_k.\phi(\mathbf{x})$.*

The set of octagonal constraints forms therefore a class, denoted further $\mathcal{R}_{OCT}$.

Lemma 18. *The class $\mathcal{R}_{OCT}$ is poly-logarithmic.*

Proof: Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be a difference bounds relation, $\mathbf{x} = \{x_1,\ldots,x_N\}$ be the set of variables in its defining formulae, and $\overline{R}(\mathbf{y},\mathbf{y}')$, for $\mathbf{y} = \{y_1,\ldots,y_{2N}\}$ be its difference bounds representation. Let $\mathcal{G}^m_{\overline{R}}$ be the m-times unfolding of the constraint graph $\mathcal{G}_{\overline{R}}$, and $\widetilde{M}_{\overline{R}^m}$ be its incidence matrix. Clearly, $\widetilde{M}_{\overline{R}^m}$ is the DBM representation of the conjunction:

$$\overline{R}(\mathbf{y}^{(0)},\mathbf{y}^{(1)}) \wedge \ldots \wedge \overline{R}(\mathbf{y}^{(m-1)},\mathbf{y}^{(m)})$$

Let $\widetilde{M}^t_{\overline{R}^m}$ be the tight closure of $\widetilde{M}_{\overline{R}^m}$. By Thm. 8, we have $(\widetilde{M}^t_{\overline{R}^m})_{ij} \leq (\widetilde{M}^*_{\overline{R}^m})_{ij}$ for all $1 \leq i, j \leq 2N$. Moreover, by Prop. 10, $\overline{R}^m(\mathbf{y}, \mathbf{y}')$ is obtained from $\Psi^{uu}_{\widetilde{M}^t_{\overline{R}^m}}$ by eliminating the atomic propositions involving $\mathbf{y}^{(1)}, \ldots \mathbf{y}^{(m-1)}$, and renaming $\mathbf{y}^{(0)}$ and $\mathbf{y}^{(m)}$ by $\mathbf{y}$ and $\mathbf{y}'$, respectively. Since any minimal path in $\mathcal{G}^m_{\overline{R}}$, and hence, any entry in $\widetilde{M}^*_{\overline{R}^m}$, is bounded by $2N \cdot (m+1) \cdot \nabla(R)$, we have:

$$\begin{aligned} \|R^m\|_2 &\leq (4N)^2 \cdot \log_2(2N \cdot (m+1) \cdot \nabla(R)) \\ &\leq 64\|R\|_2^2 \cdot (\log_2 \|R\|_2 + \log_2(m+1) + \log_2 \nabla(R)) \\ &= O(\|R\|_2^3 \cdot \log_2 m) \end{aligned}$$

This proves the first point of Def. 5. The second point of Def. 5 follows from Thm. 8 – the tight closure of a DBM can be computed by the Floyd-Warshall algorithm followed by a polynomial-time tightening step. □

8.1 The Complexity of Acceleration for Octagonal Relations

The proof idea for the periodicity of $\mathcal{R}_{OCT}$ is the following. Since any power R^i of an octagonal relation R is obtained by quantifier elimination, and since quantifier elimination for octagons uses the tight closure of the DBM representation, then the sequence $\{R^i\}_{i>0}$ is defined by the sequence $\{M^t_{\overline{R^i}}\}_{i>0}$ of tightly closed DBMs. In [8] we prove that this sequence of matrices is periodic, using the result from Thm. 9, below. If $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ is an octagonal relation, let $\sigma(R) \equiv M_{\overline{R}}$ be the characteristic DBM of its difference bounds representation, and for a coherent DBM $M \in \mathbb{Z}^{4N\times 4N}_\infty$, we define $\rho(M) \equiv \Psi^{uu}_{\blacksquare M} \wedge \Psi^{up}_{M\blacksquare} \wedge \Psi^{pu}_{\blacksquare M} \wedge \Psi^{pp}_{M_\blacksquare}$. Analogously, $\pi(M)$ is defined in the same way as ρ, for each matrix $M \in \mathbb{Z}[k]^{4N\times 4N}_\infty$ of univariate linear terms. With these definitions, periodicity of $\mathcal{R}_{OCT}$ has been shown in [8], using the periodicity of $\mathcal{R}_{DB}$ and the following theorem [5], establishing the following relation between $M^t_{\overline{R^m}}$ (the tightly closed octagonal DBM corresponding to the m-th iteration of R) and $M^*_{\overline{R}^m}$ (the closed DBM corresponding to the m-th iteration of the difference bounds relation $\overline{R}$), for all $m > 0$:

Theorem 9. *[5] Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$, be a $*$-consistent octagonal relation. Then, for all $m > 0$ and $1 \leq i, j \leq 4N$: $(M^t_{\overline{R^m}})_{ij} = \min\left\{ (M^*_{\overline{R}^m})_{ij}, \left\lfloor \frac{(M^*_{\overline{R}^m})_{i\bar{\imath}}}{2} \right\rfloor + \left\lfloor \frac{(M^*_{\overline{R}^m})_{\bar{\jmath}j}}{2} \right\rfloor \right\}$.*

In the rest of this section, we show that the periodic class $\mathcal{R}_{OCT}$ is also exponential, which proves NP-completness of the reachability problem for flat counter machines with octagonal constraints labeling their loops.

Lemma 19. *Let $\{s_m\}_{m=1}^\infty$ and $\{t_m\}_{m=1}^\infty$ be two periodic sequences. Then the sequences $\{\min(s_m,t_m)\}_{m=1}^\infty$, $\{s_m+t_m\}_{m=1}^\infty$ and $\left\{ \left\lfloor \frac{s_m}{2} \right\rfloor \right\}_{m=1}^\infty$ are periodic as well. Let b_s (c_s) be the prefix (period) of $\{s_m\}_{m=1}^\infty$, let b_t (c_t) be the prefix (period) of $\{t_m\}_{m=1}^\infty$, and let define $b = \max(b_s, b_t)$, $c = \mathrm{lcm}(c_s, c_t)$, and $b_m = b + \max_{i=0}^{c-1} K_i c$, where (1) $K_i = \left\lceil \frac{s_{b+i} - t_{b+i}}{\lambda_i^{(t)} - \lambda_i^{(s)}} \right\rceil$ if $\lambda_i^{(s)} < \lambda_i^{(t)}$ and $t_{b+i} < s_{b+i}$, (2) $K_i = \left\lceil \frac{t_{b+i} - s_{b+i}}{\lambda_i^{(s)} - \lambda_i^{(t)}} \right\rceil$ if $\lambda_i^{(t)} < \lambda_i^{(s)}$ and $s_{b+i} < t_{b+i}$, and (3) $K_i = 0$, otherwise, for each $i = 0, \ldots, c-1$ and where $\lambda_0^{(s)}, \ldots, \lambda_{c-1}^{(s)}$ ($\lambda_0^{(t)}, \ldots, \lambda_{c-1}^{(t)}$) are* *rates of* $\{s_m\}_{m=1}^{\infty}$ *(*$\{t_m\}_{m=1}^{\infty}$*) with respect to the common prefix b and period c. Then, the prefix and the period of the above sequences are:*

	prefix	*period*
$\{s_m + t_m\}_{m=1}^{\infty}$	b	c
$\left\{\left\lfloor \frac{s_m}{2} \right\rfloor\right\}_{m=1}^{\infty}$	b	$2c$
$\{\min(s_m, t_m)\}_{m=1}^{\infty}$	b_m	c

Proof: We can show that the sum sequence $\{s_m + t_m\}_{m=1}^{\infty}$ is periodic as well, with prefix b, period c and rates $\lambda_0^{(s)} + \lambda_0^{(t)}, ..., \lambda_{c-1}^{(s)} + \lambda_{c-1}^{(t)}$. In fact, for every $k \geq 0$ and $i = 0, ..., c-1$ we have successively:

$$
\begin{aligned}
(s+t)_{b+(k+1)c+i} &= s_{b+(k+1)c+i} + t_{b+(k+1)c+i} && (14)\\
&= \lambda_i^{(s)} + s_{b+kc+i} + \lambda_i^{(t)} + t_{b+kc+i} && (15)\\
&= \lambda_i^{(s)} + \lambda_i^{(t)} + s_{b+kc+i} + t_{b+kc+i} && (16)\\
&= (\lambda_i^{(s)} + \lambda_i^{(t)}) + (s+t)_{b+kc+i} && (17)
\end{aligned}
$$

For the min sequence $\{\min(s_m, t_m)\}_{m=1}^{\infty}$, it can be shown that, for each $i = 0, ..., c-1$ precisely one of the following assertions hold:

1. $(\lambda_i^{(s)} < \lambda_i^{(t)}$ or $\lambda_i^{(s)} = \lambda_i^{(t)}$ and $s_{b+i} < t_{b+i})$ and $\forall k \geq 0.\ s_{b+K_ic+kc+i} \leq t_{b+K_ic+kc+i}$
2. $(\lambda_i^{(t)} < \lambda_i^{(s)}$ or $\lambda_i^{(s)} = \lambda_i^{(t)}$ and $t_{b+i} < s_{b+i})$ and $\forall k \geq 0.\ t_{b+K_ic+kc+i} \leq s_{b+K_ic+kc+i}$

Intuitively, starting from the position $b + K_ic$, on every period c, the minimum amongst the two sequences is always defined by the same sequence i.e., the one having the minimal rate on index i, or if the rates are equal, the one having the smaller starting value.

We can show now that the min sequence $\{\min(s_m, t_m)\}_{m=1}^{\infty}$ is periodic starting at $b_m = b + \max_{i=0}^{c-1} K_ic$, with period c and rates $\min(\lambda_0^{(s)}, \lambda_0^{(t)}), ..., \min(\lambda_{c-1}^{(s)}, \lambda_{c-1}^{(t)})$. That is, we have successively, for every $k \geq 0$ and $i = 0, ..., c-1$, and whenever i satisfies the condition (1) above (the case when i satisfies the condition (2) being similar):

$$
\begin{aligned}
\min(s_{b_m+(k+1)c+i}, t_{b_m+(k+1)c+i}) &= s_{b_m+(k+1)c+i}\\
&= \lambda_i^{(s)} + s_{b_m+kc+i}\\
&= \min(\lambda_i^{(s)}, \lambda_i^{(t)}) + \min(s_{b_m+kc+i}, t_{b_m+kc+i})
\end{aligned}
$$

For the sequence $\left\{\left\lfloor \frac{s_m}{2} \right\rfloor\right\}_{m=1}^{\infty}$, assume that the sequence $\{s_m\}_{m=1}^{\infty}$ is periodic with prefix b, period c and rates $\lambda_0, ..., \lambda_{c-1}$. It can be easily shown that the sequence $\left\lfloor \frac{s_m}{2} \right\rfloor$ is periodic as well with prefix b, period $2c$, and rates $\lambda_0, ..., \lambda_{c-1}, \lambda_0, ..., \lambda_{c-1}$.

We have successively for any $k \geq 0$, and for any $i = 0, ..., c-1$:

$$\left\lfloor \frac{s_{b+(k+1)2c+i}}{2} \right\rfloor = \left\lfloor \frac{2\lambda_i + s_{b+k\cdot 2c+i}}{2} \right\rfloor = \lambda_i + \left\lfloor \frac{s_{b+k\cdot 2c+i}}{2} \right\rfloor$$

Similarly, for any $k \geq 0$ and for any $i = 0, ..., c-1$, we have:

$$\left\lfloor \frac{s_{(b+k+1)2c+c+i}}{2} \right\rfloor = \left\lfloor \frac{2\lambda_i + s_{b+k\cdot 2c+c+i}}{2} \right\rfloor = \lambda_i + \left\lfloor \frac{s_{b+k\cdot 2c+c+i}}{2} \right\rfloor$$

□

The following lemma establishes the bounds on the prefix and period of octagonal relations, needed by Thm. 10, which gives one of the main results of this paper.

Lemma 20. *Let $R \subseteq \mathbb{Z}^N \times \mathbb{Z}^N$ be an octagonal relation. The prefix and period of R are $\nabla(R)^3 \cdot 2^{O(N)}$ and $2^{O(N)}$, respectively.*

Proof: Let $\mathbf{x} = \{x_1, \ldots, x_N\}$, $R(\mathbf{x}, \mathbf{x}')$ be the octagonal constraint defining the relation R, $\mathbf{y} = \{y_1, \ldots, y_{2N}\}$ and $\overline{R}(\mathbf{y}, \mathbf{y}')$ be the difference bounds encoding of R. Since $\nabla(R) = \nabla(\overline{R})$, by Lemma 7, the prefix of the periodic sequence $\{M^*_{\overline{R}^m}\}_{m=1}^{\infty}$ is $\nabla(\overline{R}) \cdot 2^{O(N)}$ and, by Lemma 6, its period is $2^{O(N)}$. We will show that the periodic sequence $\{M^t_{\overline{R^m}}\}_{m=1}^{\infty}$ defining the sequence of powers $\{R^m\}_{m=1}^{\infty}$ has prefix $\nabla(R)^3 \cdot 2^{O(N)}$ and period $2^{O(N)}$, respectively. Let us fix $1 \leq i, j \leq 4N$. By Thm. 9 we have, for all $m > 0$:

$$(M^t_{\overline{R^m}})_{ij} = \min \left\{ (M^*_{\overline{R}^m})_{ij}, \left\lfloor \frac{(M^*_{\overline{R}^m})_{i\bar{\imath}}}{2} \right\rfloor + \left\lfloor \frac{(M^*_{\overline{R}^m})_{\bar{\jmath}j}}{2} \right\rfloor \right\}$$

The period of $\{(M^t_{\overline{R^m}})_{ij}\}_{m=1}^{\infty}$ can be shown to be $2^{O(N)}$ by an application of Lemma 19, which establishes the upper bound for the period of R. To compute the upper bound on the prefix of $\{(M^t_{\overline{R^m}})_{ij}\}_{m=1}^{\infty}$, we consider first the case when R is $*$-consistent. Let us define, for all $m > 0$:

$$s_m = (M^*_{\overline{R}^m})_{ij} \qquad\qquad t_m = \left\lfloor \frac{(M^*_{\overline{R}^m})_{i\bar{\imath}}}{2} \right\rfloor + \left\lfloor \frac{(M^*_{\overline{R}^m})_{\bar{\jmath}j}}{2} \right\rfloor$$

By Lemma 19, the periodic sequence $\{t_m\}_{m=1}^{\infty}$ has prefix b and period $c' = 2c$. The sequence $\{s_m\}_{m=1}$ has prefix b and period c, but we can w.l.o.g. assume that its period is $c' = 2c$. Clearly $b = \nabla(R) \cdot 2^{O(N)}$ and $c' = 2^{O(N)}$, by Lemma 7 and 6, respectively. By Lemma 19, the sequence $\{\min(s_m, t_m)\}_{m=1}^{\infty}$ has period c and prefix defined as $b' = b + \max_{i=0}^{c-1} K_i c'$ where:

$$\begin{array}{ll} K_i = \left\lceil \frac{s_{b+i} - t_{b+i}}{\lambda_i^{(t)} - \lambda_i^{(s)}} \right\rceil & \text{if } \lambda_i^{(s)} < \lambda_i^{(t)} \text{ and } t_{b+i} < s_{b+i} \\ K_i = \left\lceil \frac{t_{b+i} - s_{b+i}}{\lambda_i^{(s)} - \lambda_i^{(t)}} \right\rceil & \text{if } \lambda_i^{(t)} < \lambda_i^{(s)} \text{ and } s_{b+i} < t_{b+i} \\ K_i = 0 & \text{otherwise} \end{array}$$

Observe that:

$$\begin{array}{l} s_b \geq -b \cdot \nabla(\overline{R}) \\ t_b \leq \max\{(M^*_{\overline{R}^b})_{i\bar{\imath}}, (M^*_{\overline{R}^b})_{\bar{\jmath}j}\} \leq b \cdot \nabla(\overline{R}) \end{array}$$

Thus, if $\lambda_i^{(s)} > \lambda_i^{(t)}$ and $t_{b+i} > s_{b+i}$, then:

$$K_i = \left\lceil \frac{t_{b+i} - s_{b+i}}{\lambda_i^{(s)} - \lambda_i^{(t)}} \right\rceil \leq t_{b+i} - s_{b+i} \leq 2 \cdot b \cdot \nabla(\overline{R})$$

Similarly, we infer that $K_i \leq 2 \cdot b \cdot \nabla(\overline{R})$ if $\lambda_i^{(s)} < \lambda_i^{(t)}$ and $t_{b+i} < s_{b+i}$. Hence, $b' = b + 2 \cdot b \cdot \nabla(\overline{R}) \cdot c'$ is the prefix of $\{\min(s_m, t_m)\}_{m=1}^{\infty}$ and thus of R. Thus $b' = \nabla(R)^2 \cdot 2^{O(N)}$.

Second, consider the case when R is not $*$-consistent, i.e. there exists $\ell > 0$ such that $R^\ell \neq \emptyset$ and for all $k > \ell$, $R^k = \emptyset$. According to Thm. 8, this can happen if and only if either:

1. there exists $m > \ell$ such that the DBM $M^*_{\overline{R}^m}$ is inconsistent. In this case, $M^*_{\overline{R}^k}$ is inconsistent for all $k \geq m$, and by by Lemma 7, we have $\ell < m = \nabla(R) \cdot 2^{O(N)}$.
2. for all $m > \ell$ the DBM $M^*_{\overline{R}^m}$ is consistent, hence by Thm. 8, for all $m > \ell$ there exist $i_m \in \{1, \ldots, 2N\}$, such that $t^m_{i_m} \stackrel{def}{=} \lfloor \frac{(M^*_{\overline{R}^m})_{i_m \overline{i_m}}}{2} \rfloor + \lfloor \frac{(M^*_{\overline{R}^m})_{\overline{i_m} i_m}}{2} \rfloor < 0$, or else R^m would be consistent, which contradicts with the initial hypothesis. But then there exists $i \in \{1, \ldots, 2N\}$ such that, for infinitely many $m > \ell$, we have $t^m_i < 0$. By Lemma 19, the sequence $\{t^m_i\}_{m=1}^{\infty}$ is periodic, with prefix[15] $B = \nabla(R)^2 \cdot 2^{O(N)}$ and period $C = 2^{O(N)}$. By passing on to a subsequence, we find infinitely many $m > \ell$, all equal J modulo C, such that $t^m_i < 0$. We have $t^m_i = t^{B+J}_i + m \cdot \lambda_J$, since $\{t^m_i\}_{m=1}^{\infty}$ is periodic, with rate λ_J. Since $t^m_i < 0$ for infinitely many $m > \ell$, we have $\lambda_J < 0$. We compute a lower bound on the first m_0 such that $t^{m_0}_i < 0$:

$$\begin{array}{ll} t^{m_0}_i = t^{B+J}_i + m_0 \cdot \lambda_J < 0 & \\ m_0 > \frac{t^{B+J}_i}{-\lambda_J} & \geq t^{B+J}_i = \nabla(R)^3 \cdot 2^{O(N)} \end{array}$$

The last equality follows from the fact that $B = \nabla(R)^2 \cdot 2^{O(N)}$ and $J < C = 2^{O(N)}$. Clearly $\ell \leq m_0$, which gives the upper bound on ℓ.

$\square$

Theorem 10. *The class $\mathcal{R}_{DB}$ is exponential, and the reachability problem for the class $\mathcal{M}_{OCT} = \{M$ flat counter machine $|$ for all rules $q \stackrel{R}{\Rightarrow} q'$ on a loop of M, R is $\mathcal{R}_{OCT}$-definable$\}$ is NP-complete.*

Proof: To show that $\mathcal{R}_{OCT}$ is exponential, we prove the four points of Def. 6. Point (A) is by Lemma 18. Point (B) is by the definition of the σ, ρ and π mappings for $\mathcal{R}_{OCT}$, for the class $\mathcal{R}_{OCT}$. For point (C.1) we use Lemma 20, and the fact that $N \leq 2 \cdot \|R\|_2$ (9) and $\log_2(\nabla(R)) \leq \|R\|_2$ (Prop. 1). For the last point (C.2) we use Prop. 9 to decide the second condition of Lemma 1, as a N^2 conjunction of equalities between univariate terms, of size polynomial in $\|R\|_2$, built using k, integer constants, min, $+$ and $\lfloor \frac{\cdot}{2} \rfloor$. The validity of this conjuction (for all $k > 0$) can be decided in NPTIME($\|R\|_2$), since it is equivalent to a QFPA formula of size polynomial in the size of the equivalence. Finally, NP-completness of the reachability problem for the $\mathcal{M}_{OCT}$ class follows directly from Thm. 2. $\square$

[15] The computation of B is as follows (using the notation of Lemma 19):

$$\begin{array}{l} b = \max(b_s, b_t) = \nabla(R) \cdot 2^{O(N)} \\ C = lcm(c_s, c_t) = 2^{O(N)} \\ B \leq b + \max_{i \in [C]} \max(s_{b+i}, t_{b+i}) \cdot C = \nabla(R) \cdot 2^{O(N)} + \nabla(R)^2 \cdot 2^{O(N)} \cdot 2^{O(N)} = \nabla(R)^2 \cdot 2^{O(N)} \end{array}$$

9 Conclusions and Future Work

We prove that the verification of reachability properties for flat counter machines with difference bounds and octagonal relations on loops is NP-complete. Future work includes the extension of this result to finite monoid affine relations [8], and the investigation of temporal logic properties of flat counter machines with transitions defined using these classes of relations.

References

1. 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)
2. Bansal, K., Koskinen, E., Wies, T., Zufferey, D.: Structural counter abstraction. In: TACAS. vol. 7795, pp. 62–77 (2013)
3. Boigelot, B.: Symbolic Methods for Exploring Infinite State Spaces. PhD, Univ. de Liège (1999)
4. Bouajjani, A., Bozga, M., Habermehl, P., Iosif, R., Moro, P., Vojnar, T.: Programs with lists are counter automata. In: CAV. LNCS, vol. 4144, pp. 517–531 (2006)
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., Habermehl, P., Iosif, R., Konečný, F., Vojnar, T.: Automatic verification of integer array programs. In: CAV. LNCS, vol. 5643, pp. 157–172 (2009)
7. Bozga, M., Iosif, R., Konecny, F.: Safety Problems are NP-complete for Flat Integer Programs with Octagonal Loops. Tech. Rep. arXiv 1307.5321 (2013), http://arxiv.org/abs/1307.5321
8. Bozga, M., Iosif, R., Konečný, F.: Fast acceleration of ultimately periodic relations. In: CAV. LNCS, vol. 6174, pp. 227–242 (2010)
9. Bozga, M., Iosif, R., Lakhnech, Y.: Flat parametric counter automata. Fundamenta Informaticae 91(2), 275–303 (2009)
10. Bozzelli, L., Pinchinat, S.: Verification of gap-order constraint abstractions of counter systems. In: VMCAI. pp. 88–103 (2012)
11. Comon, H., Jurski, Y.: Multiple counters automata, safety analysis and presburger arithmetic. In: CAV. LNCS, vol. 1427, pp. 268–279 (1998)
12. Cormen, T.H., Stein, C., Rivest, R.L., Leiserson, C.E.: Introduction to Algorithms. McGraw-Hill Higher Education, 2nd edn. (2001)
13. 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)
14. Demri, S., Dhar, A., Sangnier, A.: Taming past ltl and flat counter systems. In: IJCAR. vol. 7364, pp. 179–193 (2012)
15. Demri, S., Jurdzinski, M., Lachish, O., Lazic, R.: The covering and boundedness problems for branching vector addition systems. J. Comput. Syst. Sci. 79(1), 23–38 (2013)
16. Ibarra, O.H.: Reversal-bounded multicounter machines and their decision problems. Journal of the ACM 25(1), 116–133 (1978)
17. Leroux, J.: Vector addition system reachability problem: a short self-contained proof. In: POPL. pp. 307–316 (2011)
18. Miné, A.: The octagon abstract domain. Higher-Order and Symbolic Computation 19(1), 31–100 (2006)
19. Minsky, M.: Computation: Finite and Infinite Machines. Prentice-Hall (1967)
20. Rackoff, C.: The covering and boundedness problems for vector addition systems. Theor. Comput. Sci. 6, 223–231 (1978)
21. Revesz, P.Z.: A closed-form evaluation for datalog queries with integer (gap)-order constraints. Theor. Comput. Sci. 116(1&2), 117–149 (1993)
22. 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, 103–117 (2000)
23. Verma, K.N., Seidl, H., Schwentick, T.: On the Complexity of Equational Horn Clauses. In: CADE-20. LNCS, vol. 3632, pp. 337–352 (2005)

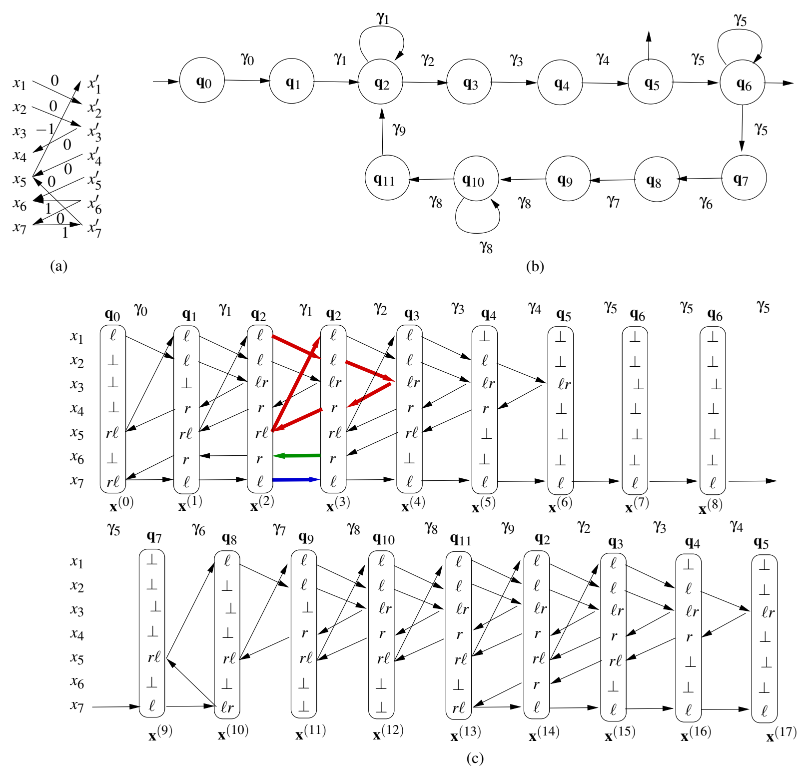


Fig. 2: (a) A difference bounds relation $R \equiv x_1 - x_2' \leq 0 \wedge x_2 - x_3' \leq 0 \wedge x_3' - x_4 \leq 0 \wedge x_4' - x_5 \leq 0 \wedge x_5' - x_6 \leq 0 \wedge x_6' - x_6 \leq 1 \wedge x_6' - x_7 \leq 0 \wedge x_7 - x_7' \leq 1 \wedge x_7' - x_5 \leq 0 \wedge x_5 - x_1' \leq -1$ and its constraint graph $\mathcal{G}_R$ (b) A fragment of the zigzag automaton A_{17}^{of} recognizing odd forward z-paths from x_1 to x_7. Initial states are marked with an incoming arrow, and final states with an outgoing arrow. (c) A run of the zigzag automaton A_{17}^{of} (b) over the valid word $\gamma_0.\gamma_1^2.\gamma_2.\gamma_3.\gamma_4.\gamma_5^3.\gamma_6.\gamma_7.\gamma_8^2.\gamma_9.\gamma_2.\gamma_3.\gamma_4$, encoding a fitting odd forward z-path from $x_1^{(0)}$ to $x_7^{(17)}$.