On the Automata Size for Presburger Arithmetic*

Felix Klaedtke
Albert-Ludwigs-Universität Freiburg, Germany
klaedtke@informatik.uni-freiburg.de

Abstract

Automata provide an effective mechanization of decision procedures for Presburger arithmetic. However, only crude lower and upper bounds are known on the sizes of the automata produced by this approach. In this paper, we prove that the number of states of the minimal deterministic automaton for a Presburger arithmetic formula is triple exponentially bounded in the length of the formula. This upper bound is established by comparing the automata for Presburger arithmetic formulas with the formulas produced by a quantifier elimination method. We also show that this triple exponential bound is tight (even for nondeterministic automata). Moreover, we provide optimal automata constructions for linear equations and inequations.

1. Introduction

Presburger arithmetic (PA) is the first-order theory with addition and the ordering relation over the integers. Relevant decision problems can be expressed in it, such as solvability of (parametric) systems of linear Diophantine equations, integer programming, and various problems in system verification. The decidability of PA was established around 1930 independently by Presburger [23, 24, 34] and Skolem [32, 33] using the method of quantifier elimination. Due to the applicability of PA in various domains, its complexity and the complexity of decision problems for fragments of it have been investigated intensively. For example, Fischer and Rabin [14, 15] gave a lower bound on any decision procedure for PA, namely double exponential in nondeterministic time. Later, Berman [2] showed that the decision problem for PA is complete in the complexity class $LATIME(2^{2^{O(n)}})$, i. e., the class of problems solvable by alternating Turing machines in time $2^{2^{O(n)}}$ with a linear number of alternations. The upper bound for PA is established with a result from Ferrante and Rackoff [13] showing that quantified variables need only to range over a restricted domain of integers. Grädel [19] and Schöning [29] investigated the complexity of decision problems of fragments of PA.

Oppen [22] showed that Cooper's quantifier elimination decision procedure for PA [10] has a triple exponential worst case complexity in deterministic time. Another approach for deciding PA or fragments of it that has recently become popular is to use automata; a point that was already made by Büchi [9]. The idea of the automata-theoretic approach is simple: Integers are represented as words, e. g., using the 2's complement representation, and the word automaton (WA) for a formula accepts precisely the words that represent the integers making the formula true. The WA can be recursively constructed from the formula, where automata constructions handle the logical connectives and quantifiers. Specific algorithms for constructing WAs for linear (in)equations have been developed in [1, 4, 7, 18, 37].

A crude complexity analysis of automata-based decision procedures leads to a non-elementary worst case complexity. Namely, for every quantifier alternation there is an exponential blow-up in the worst case. However, experimental comparisons [1, 18, 31] illustrate that automata-based decision procedures for PA often have good and competitive performance in comparison to other methods. In [7], the authors claimed that the minimal deterministic WA for an arbitrary formula has at most a triple exponential number of states in the length of the formula. Unfortunately, as explained in [37], the argument used in [7] to substantiate this claim is incorrect. Wolper and Boigelot [37] gave an argument why there must be an elementary upper bound on the size of the minimal deterministic WA for a formula. However, their argumentation is rather sketchy and they only indicate that there has to be such an elementary upper bound. [1]

In this paper, we rigorously prove that there is a triple exponential upper bound on the size of the minimal deterministic WA for a formula. This bound on the automata size for

* The author is also affiliated with the ETH Zurich, Switzerland. This work has been partly supported by the German Research Council (DFG) as part of the Transregional Collaborative Research Center "Automatic Verification and Analysis of Complex Systems" (SFB/TR 14 AVACS).

1 Appendix B contains a detailed discussion of Wolper and Boigelot's argumentation.

PA contrasts with the upper bound on the automata size for the monadic second-order logic WS1S, or even WS1S with the ordering relation "$<$" as a primitive but without quantification over monadic second-order variables. There, the number of states of the minimal WA for a formula can be non-elementary larger than the formula's length [26, 35]. In order to establish the upper bound on the automata size for PA, we give a detailed analysis of the deterministic WAs for formulas by comparing the constructed WAs with the quantifier-free formulas produced by the quantifier elimination method in [25], which is an improvement of Cooper's quantifier elimination method [10]. From this analysis, we obtain that the minimal deterministic WA for an arbitrary formula of length n has at most $2^{2^{2^{O(n)}}}$ states.

Furthermore, we show that the triple exponential upper bound on the size of deterministic WAs for formulas is tight. In fact, we show a stronger result. Namely, we give a family of Presburger arithmetic formulas for which even a nondeterministic WA must have at least triple exponentially many states. We also improve the automata constructions in [4, 18, 37] for linear (in)equations. We prove that our automata constructions are optimal in the sense that the constructed deterministic WAs are minimal.

We proceed as follows. Preliminaries are given in §2. In §3, we investigate the WAs for quantifier-free formulas. In §4, we prove the upper bound on the size of the minimal deterministic WA for formulas, and in §5, we give a worst case example. Finally, in §6, we draw conclusions. Appendix A contains additional proof details.

2. Preliminaries

Presburger arithmetic (PA) is the first-order logic over the structure $\mathfrak{Z} := (\mathbb{Z}, <, +)$. We use standard notation. For instance, we write $\mathfrak{Z} \models \varphi[a_1, \ldots, a_r]$ for a formula $\varphi(x_1, \ldots, x_r)$ and $a_1, \ldots, a_r \in \mathbb{Z}$ if φ is true in $\mathfrak{Z}$ when the variable x_i is interpreted as the integer a_i, for $1 \leq i \leq r$. And analogously, $t[a_1, \ldots, a_r]$ denotes the integer of the term $t(x_1, \ldots, x_r)$ when the x_is are interpreted as the a_is. For a formula $\varphi(x_1, \ldots, x_r)$, we define $[\![\varphi]\!] := \{(a_1, \ldots, a_r) \in \mathbb{Z}^r : \mathfrak{Z} \models \varphi[a_1, \ldots, a_r]\}$.

Extended Logical Language. We extend the logical language of PA by (i) constants for the integers 0 and 1, (ii) the unary operation "$-$" for integer negation, and (iii) for all $d \geq 2$, we extend the language by predicates "$d|$" for the divisibility relation. These constructs are definable in PA, e. g., the formula $\exists x(x + \cdots + x = t)$ defines $d|t$, where x is repeated d times in the term $x + \cdots + x$ and x does not appear in the term t. The reason for the extended logical language, where (i), (ii), and (iii) are handled as primitives, is that it admits quantifier elimination (qe), i. e., for a formula $\exists x \varphi(x, \overline{y})$, where φ is quantifier-free, we can construct a logically equivalent quantifier-free formula $\psi(\overline{y})$.

Additionally, we allow the relation symbols $\leq, >, \geq$, and $\neq$ with their obvious meanings. In the following, we assume that terms and formulas are defined in terms of the extended logical language for PA. We denote the set of quantifier-free formulas by QF.

For convenience, we allow the usage of standard symbols when writing terms. For instance, c stands for $1 + \cdots + 1$ (repeated c times) if $c > 0$, and c stands for $-(1 + \cdots + 1)$ if $c < 0$. We say that the term c is a *constant*. We identify the term c with the integer that it represents. Analogously, we write $k \cdot x$ for $x + \cdots + x$ (repeated k times) if $k > 0$, and $-(x + \cdots + x)$ if $k < 0$. Moreover, if $k = 0$ then $k \cdot x$ abbreviates $x + (-x)$. We say that k is the *coefficient* of x. For a term t and $k \in \mathbb{Z}$, $k \cdot t$ denotes the term where the constant and the coefficients in t are multiplied by k.

A term t is *homogeneous* if it is either 0 or of the form $k_1 \cdot x_1 + \cdots + k_r \cdot x_r$, for some $r \geq 1$, where the variables $x_1, \ldots, x_r$ are pairwise distinct and $k_1, \ldots, k_r \in \mathbb{Z}$. Throughout the text, we assume that terms are of the form t or $t + c$, where t is homogeneous and $c \in \mathbb{Z} \setminus \{0\}$. The *normal form* of an (in)equation $t_1 \bowtie t_2$ with $\bowtie \in \{=, \neq, <, \leq, >, \geq\}$ is the logically equivalent (in)equation $t \bowtie c$, where summands of the form $k \cdot x$ in t_1 and t_2 are collected on the left-hand side t and constants in t_1 and t_2 are collected on the right-hand side c according to standard calculation rules.

Length of a Formula. The *length* of a formula is the number of letters used in writing the formula. Note that the length of a formula depends significantly on how we define the length of coefficients and constants. For instance, $x = 10 \cdot y$ contains 6 letters, namely, x, $=$, 1, 0, $\cdot$, and y. The "expanded version" has $2 + 19$ letters since $10 \cdot y$ abbreviates the term $y + y + y + y + y + y + y + y + y + y$. We use the same definition of the length of a formula as in [14, 22, 25]. In particular, the length of a coefficient or constant is the number of letters of the expanded version. However, it is possible to express $k \cdot x$ by a formula of length $\mathrm{O}(\lg |k|)$. The idea is illustrated by $x = 10 \cdot y$: the formula is logically equivalent to $\exists z(x = z + z \wedge \exists x(z = x + x + y \wedge x = y + y))$. Note that we only need a fixed number of variables for any k, see [14]. For the sake of uniformity, we define the length of the formula $d|t$ as the length of the term t plus $d + 1$. Again, there is a logically equivalent formula of length $\mathrm{O}(\lg d)$ plus the length of t. For the results in this paper it does not matter if we define the length of an integer k as $\mathrm{O}(\lg |k|)$ or as $\mathrm{O}(|k|)$.

3. Automata Constructions

In this section, we investigate the automata for quantifier-free PA formulas. We assume that the reader is familiar with the basic notions of automata theory. Recall that a *deterministic word automata* (DWA) is a tu-

ple $\mathcal{A} = (Q, \Sigma, \delta, q_{\mathrm{I}}, F)$, where Q is a finite set of states, Σ is a finite alphabet, $\delta : Q \times \Sigma \to Q$ is the transition function, $q_{\mathrm{I}} \in Q$ is the initial state, and $F \subseteq Q$ is the set of accepting states. The *size* of $\mathcal{A}$ is the cardinality of Q. The *language* of $\mathcal{A}$ is $L(\mathcal{A}) := \{w \in \Sigma^* : \widehat{\delta}(q_{\mathrm{I}}, w) \in F\}$, where $\widehat{\delta}(q, \lambda) := q$ and $\widehat{\delta}(q, wb) := \delta(\widehat{\delta}(q, w), b)$, for $q \in Q$, $b \in \Sigma$, and $w \in \Sigma^*$. Note that λ denotes the empty word.

In §3.1, we define how DWAs recognize sets of integers. In §3.2, we provide optimal automata constructions for linear (in)equations and in §3.3, we give an automata construction for the divisibility relation. Finally, in §3.4, we give an upper bound on the size of the minimal deterministic DWA for a quantifier-free formula.

3.1. Representing Sets of Integers with Automata

We use an idea that goes at least back to Büchi [9] in order to use automata to recognize tuples of numbers by mapping words to tuples of numbers. Our encoding is based on the 2's complement representation of integers, where the most significant bit is the first digit. For $b_n b_{n-1} \dots b_0 \in \{0,1\}^+$, we define

$$\langle b_n b_{n-1} \dots b_0 \rangle := -b_n 2^n + \sum_{0 \le i < n} b_i 2^i$$

and $\langle \lambda \rangle := 0$ for the empty word. Note that the encoding of an integer is not unique since $\langle \lambda \rangle = \langle 0 \rangle = 0$ and $\langle b_n b_{n-1} \dots b_0 \rangle = \langle b_n b_n b_{n-1} \dots b_0 \rangle$. We extend this encoding to tuples of integers as follows: A word $w := \overline{b}_n \dots \overline{b}_0 \in (\{0,1\}^r)^*$ *represents* the tuple $\overline{z} := (z_1, \dots, z_r) \in \mathbb{Z}^r$ of integers, where the ith "track" of the word w encodes the integer z_i. That is, for all $1 \le i \le r$ we have that $z_i = \langle b_{n,i} \dots b_{0,i} \rangle$, where $\overline{b}_j = (b_{j,1}, \dots, b_{j,r})$ for $0 \le j \le n$. The first letter $\overline{b}_n$ of w is the *sign letter* since it determines the signs of the integers $z_1, \dots, z_r$. We abuse notation and write $\langle w \rangle$ to denote the tuple $\overline{z}$ of integers. Moreover, we write $\langle\!\langle \overline{z} \rangle\!\rangle$ for the shortest word in $(\{0,1\}^r)^*$ that represents $\overline{z} \in \mathbb{Z}^r$. Note that $\langle\!\langle \overline{z} \rangle\!\rangle$ is well-defined since (1) there is a word $w \in (\{0,1\}^r)^*$ with $\langle w \rangle = \overline{z}$, and (2) if $\langle v \rangle = \langle v' \rangle$ for $v, v' \in (\{0,1\}^r)^*$, then v and v' have a common suffix $u \in (\{0,1\}^r)^*$ with $\langle u \rangle = \langle v \rangle$.

A natural choice for representing sets of tuples of integers as languages is the following: The set $U \subseteq \mathbb{Z}^r$ is *represented* by the language $L \subseteq (\{0,1\}^r)^*$ if for every $\overline{z} \in \mathbb{Z}^r$ it holds that $\overline{z} \in U$ iff all words that represent $\overline{z}$ are in L. The reason for the requirement that all encodings of the elements in U have to be in the set L is that the representation of $\mathbb{Z}^r \setminus U$ is then the complement of L, i. e., $(\{0,1\}^r)^* \setminus L$. A DWA $\mathcal{A}$ *represents* the set $U \subseteq \mathbb{Z}^r$ if $L(\mathcal{A})$ represents U. Note that under this definition not every language over $\{0,1\}^r$ encodes a set of tuples of integers, and not every DWA with alphabet $\{0,1\}^r$ represents a subset of $\mathbb{Z}^r$. The language $L \subseteq (\{0,1\}^r)^*$ encodes a subset of $\mathbb{Z}^r$ iff L satisfies the two properties $\lambda \in L \Leftrightarrow \overline{0} \in L$, and $\overline{b}u \in L \Leftrightarrow \overline{b}\overline{b}u \in L$, for all $u \in (\{0,1\}^r)^*$ and $\overline{b} \in \{0,1\}^r$.

In the following subsections, we give and analyze the automata constructions for atomic formulas and for Boolean combinations of atomic formulas. For the analysis, we introduce the following notation: We define $\langle \lambda \rangle_{\mathbb{N}} := \overline{0}$ and $\langle \overline{b}_n \dots \overline{b}_0 \rangle_{\mathbb{N}} := \sum_{0 \le i \le n} \overline{b}_i 2^i$ with $\overline{b}_n \dots \overline{b}_0 \in (\{0,1\}^r)^+$. Similar to $\langle\!\langle \overline{z} \rangle\!\rangle$ for $\overline{z} \in \mathbb{Z}^r$, we define $\langle\!\langle \overline{a} \rangle\!\rangle_{\mathbb{N}}$, for $\overline{a} \in \mathbb{N}^r$, as the shortest word $w \in (\{0,1\}^r)^*$ with $\overline{a} = \langle w \rangle_{\mathbb{N}}$.

3.2. Linear Equations and Inequations

In this subsection, we first recall the automata constructions given in [4, 6, 18, 37] for linear (in)equations. Then, we improve these constructions such that they are optimal, i. e., the constructed DWAs are minimal. Assume that the (in)equation $t \bowtie c$ is in normal form, where $t(x_1, \dots, x_r)$ is a homogeneous term, $\bowtie \in \{=, \neq, <, \le, >, \ge\}$, and $c \in \mathbb{Z}$.

First, we make the following observation for a word $w \in (\{0,1\}^r)^*$ and $\overline{b} \in \{0,1\}^r$. If $w \neq \lambda$ then $\langle w\overline{b} \rangle = 2\langle w \rangle + \overline{b}$, and for $w = \lambda$, we have that $\langle \overline{b} \rangle = -\overline{b}$, since we are using 2's complement representation. Given this, it is relatively straightforward to obtain an analog of a DWA with *infinitely* many states for $t \bowtie c$. The set of states is $\{q_{\mathrm{I}}\} \cup \mathbb{Z}$ where q_{I} is the initial state. Note that we identify integers with states. The idea is to keep track of the value of t as successive bits are read. Thus, except for the special initial state, a state in $\mathbb{Z}$ represents the current value of t. Lemma 1 below justifies this intuition. The transition function $\eta : (\{q_{\mathrm{I}}\} \cup \mathbb{Z}) \times \{0,1\}^r \to (\{q_{\mathrm{I}}\} \cup \mathbb{Z})$ is defined as follows for a letter $\overline{b} \in \{0,1\}^r$. For $q \in \mathbb{Z}$, we define $\eta(q, \overline{b}) := 2q + t[\overline{b}]$ and $\eta(q_{\mathrm{I}}, \overline{b}) := -t[\overline{b}]$, for the initial state.

Lemma 1. *For $w \in (\{0,1\}^r)^*$ of length $n \ge 0$,*

(a) $\widehat{\eta}(q, w) = q2^n + t\big[\langle w \rangle_{\mathbb{N}}\big]$, where $q \in \mathbb{Z}$, and

(b) $\widehat{\eta}(q_{\mathrm{I}}, w) = t\big[\langle w \rangle\big]$ if $n \ge 1$.

Proof. (a) is easily proved by induction over n, and (b) follows from (a) and the definition of η. □

Later we make use of the following lemma, which translates the question whether $q \in \mathbb{Z}$ is reachable from $p \in \mathbb{Z}$ via $\widehat{\eta}$ to a number theoretic problem.

Lemma 2. *For $p, q \in \mathbb{Z}$, there are $N, a_1, \dots, a_r \ge 0$ such that $N \ge \lceil \lg(1 + \max\{a_1, \dots, a_r\}) \rceil$ and $p2^N + t[a_1, \dots, a_r] = q$ iff $\widehat{\eta}(p, w) = q$, for some $w \in (\{0,1\}^r)^*$.*

Proof. ($\Rightarrow$) Assume that $\langle\!\langle a_1, \dots, a_r \rangle\!\rangle_{\mathbb{N}}$ has length ℓ. Note that $\ell \le N$. By Lemma 1(a), we have that

$$\widehat{\eta}\big(p, \overline{0}^{N-\ell} \langle\!\langle a_1, \dots, a_r \rangle\!\rangle_{\mathbb{N}}\big) = p2^N + t[a_1, \dots, a_r] = q\,.$$

($\Leftarrow$) Assume that $\widehat{\eta}(p, w) = q$, for some $w \in (\{0,1\}^r)^*$. We have that $N \geq \lceil \lg(1+a) \rceil$, where a is the largest number in the tuple $\langle w \rangle_{\mathbb{N}}$ and N is the length of w. It follows from Lemma 1(a) that $\widehat{\eta}(p, w) = p2^N + t[\langle w \rangle_{\mathbb{N}}]$. □

The automata constructions in [18, 37] are based on the observation that the states $q, q' \in \mathbb{Z}$ can be merged if, intuitively speaking, q and q' are both small or both large. Here, the meaning of "small" and "large" depends on the coefficients of t and on the constant c. More precisely, we say that $q \in \mathbb{Z}$ is *small* if $q < \min\{c, -\|t\|_+\}$, and *large* if $q > \max\{c, \|t\|_-\}$, where

$$\|t\|_- := \sum_{\substack{1 \leq j \leq r \\ \text{and } k_j < 0}} |k_j| \qquad \text{and} \qquad \|t\|_+ := \sum_{\substack{1 \leq j \leq r \\ \text{and } k_j > 0}} k_j$$

assuming that t is of the form $k_1 \cdot x_1 + \cdots + k_r \cdot x_r$. Note that from a small value we can only obtain smaller values and from a large value we can only obtain larger values by η, i. e., for all $\overline{b} \in \{0,1\}^r$, if $q > \|t\|_-$ then $\eta(q, \overline{b}) = 2q + t[\overline{b}] > q$, and if $q < -\|t\|_+$ then $\eta(q, \overline{b}) = 2q + t[\overline{b}] < q$. A difference between the constructions in [37] and [18] are the bounds that determine the meaning of "small" and "large".

For $m < n$, we define $\mathcal{A}^{t \trianglerighteq c}_{(m,n)} := (Q, \{0,1\}^r, \delta, q_{\mathrm{I}}, F)$, where $Q := \{q_{\mathrm{I}}\} \cup \{q \in \mathbb{Z} : m \leq q \leq n\}$ and

$$\delta(q, \overline{b}) := \begin{cases} m & \text{if } \eta(q, \overline{b}) \leq m, \\ n & \text{if } \eta(q, \overline{b}) \geq n, \\ \eta(q, \overline{b}) & \text{otherwise,} \end{cases}$$

for $q \in Q$ and $\overline{b} \in \{0,1\}^r$. Moreover, let $F := \{q \in Q : q \trianglerighteq c\}$, where q_{I} is interpreted as 0.

Fact 3. *The DWA $\mathcal{A}^{t \trianglerighteq c}_{(m,n)}$ represents $[\![t \trianglerighteq c]\!]$ if m is small and n is large. Moreover, $\mathcal{A}^{t \trianglerighteq c}_{(m,n)}$ has $2 + n - m$ states.*

In the following, we optimize the constructions such that the produced DWA for an (in)equation is minimal. Moreover, we give a lower bound on the minimal DWA for an (in)equation. However, these results are not needed for the upper bound on the minimal DWA for a PA formula. In the remainder of this subsection, let $\mathcal{A}^{t \trianglerighteq c}_{(m,n)} = (Q, \{0,1\}^r, \delta, q_{\mathrm{I}}, F)$ for the (in)equation $t \trianglerighteq c$ with $m = \max\{q \in \mathbb{Z} : q \text{ is small}\}$ and $n = \min\{q \in \mathbb{Z} : q \text{ is large}\}$. We restrict ourselves to the cases where $\trianglerighteq \in \{=, <, >\}$. The cases with $\trianglerighteq \in \{\neq, \leq, \geq\}$ reduce to the cases for $=$, $<$, $>$ and complementation of DWAs, since $t \neq c$ is logically equivalent to $\neg t = c$, $t \leq c$ is logically equivalent to $\neg t > c$, and $t \geq c$ is logically equivalent to $\neg t < c$. Note that complementation of a DWA can be done by flipping accepting and non-accepting states. The complemented DWA is minimal iff the original DWA is minimal.

Eliminating Unreachable States. An obvious optimization is to eliminate the states in $Q \cap \mathbb{Z}$ that are not a multiple of the greatest common divisor of the absolute values of the coefficients in the term t, since they are not reachable from the initial state q_{I}. We define the *greatest common divisor* of the term $t(x_1, \ldots, x_r)$ as $\gcd(t) := \gcd(|k_1|, \ldots, |k_r|)$, where k_i is the coefficient of the variable x_i, for $1 \leq i \leq r$.

Lemma 4. *The state $q \in Q \cap \mathbb{Z}$ is reachable from the initial state q_{I} iff q is a multiple of $\gcd(t)$.*

Proof. ($\Rightarrow$) This direction is easy to prove by induction: for all $\overline{b} \in \{0,1\}^r$, it holds that (i) $-t[\overline{b}]$ is a multiple of $\gcd(t)$, and (ii) if $p \in \mathbb{Z}$ is a multiple of $\gcd(t)$ then $2p + t[\overline{b}]$ is a multiple of $\gcd(t)$.
($\Leftarrow$) Assume that q is a multiple of $\gcd(t)$. There are $v_1, \ldots, v_r \in \mathbb{Z}$ with $t[v_1, \ldots, v_r] = q$. From Lemma 1(b) it follows that $\widehat{\delta}\big(q_{\mathrm{I}}, \langle\!\langle v_1, \ldots, v_r \rangle\!\rangle\big) = t[v_1, \ldots, v_r]$. □

Alternatively, instead of filtering out the states $q \in \mathbb{Z}$ that are not a multiple of $\gcd(t)$ we can rewrite the (in)equation $t \trianglerighteq c$ to the logically equivalent atomic formula α and then construct the DWA for α, where α is defined as

$$\alpha := \begin{cases} t' \trianglerighteq \left\lceil \frac{c}{\gcd(t)} \right\rceil & \text{if } \trianglerighteq \text{ is } <, \\ t' \trianglerighteq \left\lfloor \frac{c}{\gcd(t)} \right\rfloor & \text{if } \trianglerighteq \text{ is } >, \\ t' \trianglerighteq \frac{c}{\gcd(t)} & \text{if } \trianglerighteq \text{ is } = \text{ and } c \text{ is a multiple of } \gcd(t), \\ 1 < 0 & \text{otherwise,} \end{cases}$$

where the coefficients in t' are the coefficients of t divided by $\gcd(t)$. In the remainder of this subsection we assume that $\gcd(t) = 1$.

Optimization for Inequations. In the following we assume that the inequation is of the form $t > c$ with $c \geq 0$. The cases where $\trianglerighteq$ is $<$ or $c \geq 0$ are analogous. The following example illustrates that many states of $\mathcal{A}^{t > c}_{(m,n)}$ can be merged if c is significantly larger than $\|t\|_-$.

Example 5. The construction for $x - y > 32$ yields a DWA with the set of states $Q = \{q_{\mathrm{I}}, -2, -1, 0, \ldots, 32, 33\}$; but the minimal DWA for $x - y > 32$ has only 13 states. The reason for this gap is that several states can be merged. First, we merge the states -2 and -1 since from both states only non-accepting states are reachable. Second, we can merge the states in $Q' := \{q \in Q \cap \mathbb{Z} : 2q + a - b > c, \text{ for all } a, b \in \{0,1\}\} = \{17, \ldots, 32\}$ to a single state since all states in Q' are non-accepting and all their transitions go to 33. The state 16 cannot be merged with any other state since if we read the letter $(1,0)$, we end up in the accepting state 33, and if we read the letters $(0,0)$, $(1,1)$, or $(0,1)$ we end up in the non-accepting states 32 or 31, respectively. The states in $\{9, \ldots, 15\}$ can again be merged to a single state since with every transition we reach a state in Q'. Analogously, we can merge the states in $\{5, 6, 7\}$.

We determine the states in $\mathcal{A}^{t>c}_{(m,n)}$ that are equivalent. Recall that $p,q \in Q$ are *equivalent*, $p \sim q$ for short, if it holds that $\widehat{\delta}(p,w) \in F$ iff $\widehat{\delta}(q,w) \in F$, for all $w \in (\{0,1\}^r)^*$. Note that $\sim \subseteq Q \times Q$ is an equivalence relation. We denote the equivalence class of $q \in Q$ by $\widetilde{q}$. Since all states are reachable from q_I, the states $p,q \in Q$ can be merged iff $p \sim q$. This means, the DWA $\mathcal{B} := (\{\widetilde{q} : q \in Q\}, \{0,1\}^r, \delta', \widetilde{q}_\mathrm{I}, \{\widetilde{q} : q \in F\})$ with $\delta'(\widetilde{q},\overline{b}) := \delta(q,\overline{b})$ is minimal and $L(\mathcal{B}) = L(\mathcal{A}^{t>c}_{(m,n)})$.

It holds that $-\|t\|_+ \sim -\|t\|_+ - 1$, since both states are non-accepting and all transitions from these states either go to $-\|t\|_+$ or to $-\|t\|_+ - 1$. In order to identify the other equivalent states in $\mathcal{A}^{t>c}_{(m,n)}$ we define the following strictly monotonically decreasing sequence $d_0 > d_1 > \cdots > d_\ell$, for some $\ell \geq 1$. Let $d_0 := \infty$ and $d_1 := \max\{c+1, \|t\|_-\}$. Assume that $d_0 > d_1 > \cdots > d_i$ are already defined, for some $i \geq 1$.

- If $d_i = \|t\|_-$ then we are done, i. e., $\ell = i$.
- If $d_i > \|t\|_-$ then let $d_{i+1} < d_i$ be the smallest integer greater than $\|t\|_- - 1$ such that for all $\overline{b} \in \{0,1\}^r$, there is an index j with $1 \leq j \leq i$ and

$$2d_{i+1} + t[\overline{b}],\ 2(d_i - 1) + t[\overline{b}] \in [d_j, d_{j-1})\,, \quad (1)$$

  where $[d,d')$ denotes the interval $\{d,\ldots,d'-1\}$ if $d,d' \in \mathbb{Z}$, and $[d,d')$ is $\{z \in \mathbb{Z} : z \geq d\}$ if $d \in \mathbb{Z}$ and $d' = \infty$. Note that d_{i+1} is well-defined since $d_i - 1$ satisfies (1), for all $\overline{b} \in \{0,1\}^r$.

Fact 6. *For all $1 \leq i \leq \ell$, if $p,q \in [d_i, d_{i-1})$ then $p \sim q$.*

The following lemma shows that there are no more equivalent states in $\mathcal{A}^{t>c}_{(m,n)}$.

Lemma 7. *For all $p,q \in Q$, if $p \sim q$ then $p = q$ or $p,q \in \{-\|t\|_+, -\|t\|_+ - 1\}$ or $p,q \in [d_i, d_{i-1})$, for $1 \leq i \leq \ell$.*

Proof. Let $R := \{-\|t\|_+, -\|t\|_+ - 1\}$. We prove the claim by contraposition, i. e., $p \neq q$ and $p \in R \Rightarrow q \notin R$ and for all $1 \leq i \leq \ell$, $p \in [d_i, d_{i-1}) \Rightarrow q \notin [d_i, d_{i-1}$ implies $p \not\sim q$. Assume $p \neq q$. It suffices to distinguish the following three cases.

(I) Assume that $p \in R$ and $q \notin R$. Since we can reach an accepting state from q, we have that $p \not\sim q$.

(II) Assume that $p \in [d_i, d_{i-1})$ and $q \notin [d_i, d_{i-1})$, for some $1 \leq i \leq \ell$. It is straightforward to prove by induction over i that $p \not\sim q$.

(III) Assume that $p \notin R \cup \bigcup_{1 \leq i \leq \ell}[d_i, d_{i-1})$. We have that either $p = q_\mathrm{I}$ or $p \in S$, where $S := \{s \in Q \cap \mathbb{Z} : -\|t\|_+ < s < \|t\|_-\}$.

Assume that $p = q_\mathrm{I}$ and $q \notin R$. Let $\overline{b} \in \{0,1\}^r$ be the letter that has a 1 in its ith coordinate iff the ith coefficient of t is positive. It holds that $q_\mathrm{I} \not\sim q$, since $\widehat{\delta}(q_\mathrm{I}, \overline{b}^n) \notin F$ and $\widehat{\delta}(q, \overline{b}^n) \in F$, for some $n \geq 1$. Note that $\delta(q_\mathrm{I}, \overline{b}) = -t[\overline{b}] = -\|t\|_+$ and $\delta(q, \overline{b}) = 2q + t[\overline{b}] = 2q + \|t\|_+ \geq q$. If $q \in R$ then we conclude similar to (I) that $p \not\sim q$.

Assume that $p \in S$. Note that for every $s \in S$ there is a $\overline{b} \in \{0,1\}^r$ such that $\delta(s,\overline{b}) \in S$. It follows that for every $n \geq 0$ there is a word $u \in (\{0,1\}^r)^*$ of length n such that $\widehat{\delta}(p,u) \in S$. We conclude that there is a word $u \in (\{0,1\}^r)^*$ such that $\widehat{\delta}(p,u) \in S$ and $\widehat{\delta}(q,u) \in R \cup \bigcup_{1 \leq i \leq \ell}[d_i, d_{i-1})$, since $\delta(s,\overline{b}) - \delta(s',\overline{b}) = 2(s-s')$, for all $s,s' \in S$ and all $\overline{b} \in \{0,1\}^r$. Analogously to (I) and (II) we conclude that $p \not\sim q$. ☐

From Lemma 7, it follows that the minimal DWA representing $[\![t > c]\!]$ has at least $\|t\|_- + \|t\|_+$ states. Note that this is in contrast to the number of symbols we need to write the inequation $t > c$ if coefficients are represented as binary numbers. For instance, we need $22 + 7$ letters for $1025 \cdot x - 1024 \cdot y > 0$, since each of the two coefficients can be represented with 11 digits. The same lower bound on the minimal DWA size holds for $t < c$. In the following, we show that a similar lower bound holds for equations.

Optimization for Equations. For an equation $t = c$ we can collapse the states in $\mathcal{A}^{t=c}_{(m,n)}$ from which we cannot reach the accepting state $c \in Q$ to a single non-accepting state. Additionally, if $t = c$ is one of the equations $x_1 = 0$, $-x_1 = 0$, $x_1 - x_2 = 0$, or $-x_1 + x_2 = 0$ we can merge the states q_I and 0. These optimizations produce the minimal DWA for $t = c$. For instance, the case for $p \in Q \cap \mathbb{Z}$ is proved as follows. Assume that we can reach from $p \in Q \cap \mathbb{Z}$ the state c, i. e., there is a $u \in (\{0,1\}^r)^*$, with $\widehat{\delta}(p,u) = c$. Any other states $q \in Q \cap \mathbb{Z}$ with $q \neq p$ from which we can reach c cannot be merged with p, since

$$c \overset{\text{Lemma 1(a)}}{=} p2^n + t\big[\langle u \rangle_\mathbb{N}\big] \neq q2^n + t\big[\langle u \rangle_\mathbb{N}\big] \overset{\text{Lemma 1(a)}}{=} \widehat{\delta}(q,u)\,,$$

where n is the length of u. The other cases are similar.

A lower bound for the minimal DWA representing $[\![t = c]\!]$ is based on the following lemma about the states of the DWA $\mathcal{A}^{t \bowtie c}_{(m,n)} = (Q, \{0,1\}^r, \delta, q_\mathrm{I}, F)$, where $\bowtie \in \{=, \neq, <, \leq, >, \geq\}$. Let $S := \{s \in Q \cap \mathbb{Z} : -\|t\|_+ < s < \|t\|_-\}$ and $[n] := \{0,\ldots,n-1\}$, for $n \geq 0$.

Lemma 8. *Every $q \in Q \cap \mathbb{Z}$ is reachable from every $p \in S$.*

Proof. We need a result from number theory. Let $\gamma > 0$ and let $c_1,\ldots,c_\gamma$ be integers with $0 < c_1 < \cdots < c_\gamma$ and $\gcd(c_1,\ldots,c_\gamma) = 1$. The *Frobenius number* $G(c_1,\ldots,c_\gamma)$ is the greatest integer z for which the linear equation $c_1 \cdot x_1 + \cdots + c_\gamma \cdot x_\gamma = z$ has *no* solution in the natural numbers. For $\gamma = 1$, it trivially holds that $G(c_1) = -1$. For $\gamma > 1$, the upper bound $G(c_1,\ldots,c_\gamma) \leq \frac{c_\gamma^2}{\gamma - 1}$ was proved by Dixmier [11]. It is straightforward to show that for all $\gamma > 0$,

$$G(c_1,\ldots,c_\gamma) < 2^{c_1+\cdots+c_\gamma} - (c_1 + \cdots + c_\gamma)\,. \quad (2)$$

The cases for $r = 0$ and $r = 1$ are trivial since $S = \emptyset$. Assume that $r \geq 2$. By Lemma 2, it suffices to show that there are $a_1, \ldots, a_r \geq 0$ with $p2^N + t[a_1, \ldots, a_r] = q$, for some $N \geq \lceil \lg(1 + \max\{a_1, \ldots, a_r\}) \rceil$.

Case I: $p = 0$. There are positive and negative coefficients in t, since $p \in S$. It follows that the equation $t(x_1, \ldots, x_r) = q$ has infinitely many solutions in the natural numbers. Recall that we assume that $\gcd(t) = 1$. In particular, there are $a_1, \ldots, a_r \geq 0$ with $2^N p + t[a_1, \ldots, a_r] = q$, for some appropriate large enough N.

Case II: $p > 0$ *and* $q \geq 0$. Let $k_{i_1}, \ldots, k_{i_\mu}$ be the positive coefficients in t, and let $k_{j_1}, \ldots, k_{j_\nu}$ be the negative coefficients in t. Let N be the size of the DWA $\mathcal{A}_{(m,n)}^{t \gtrless c}$, i. e., $N = 3 + \max\{|c|, \|t\|_+\} + \max\{c, \|t\|_-\}$. We rewrite the equation $t(x_1, \ldots, x_r) + p2^N = q$ to

$$\zeta + t_1(x_{i_1}, \ldots, x_{i_\mu}) = t_2(x_{j_1}, \ldots, x_{j_\nu}) \,, \tag{3}$$

where $\zeta := p2^N - q$, t_1 is the term $k_{i_1} \cdot x_{i_1} + \cdots + k_{i_\mu} \cdot x_{i_\mu}$, and t_2 is the term $|k_{j_1}| \cdot x_{j_1} + \cdots + |k_{j_\nu}| \cdot x_{j_\nu}$. Note that $\zeta \geq 0$ since $p > 0$ and $2^N \geq q$. Let $D := \gcd(|k_{j_1}|, \ldots, |k_{j_\nu}|)$. In order to show the existence of a solution $a_1, \ldots, a_r \in [2^N]$ of the equation (3), we proceed in two steps:

Step 1: There are $a_{i_1}, \ldots, a_{i_\mu} \in [D]$ such that

$$D \mid \zeta + t_1[a_{i_1}, \ldots, a_{i_\mu}] \,.$$

Step 2: There are $a_{j_1}, \ldots, a_{j_\nu} \in [2^N]$ such that

$$\zeta + t_1[a_{i_1}, \ldots, a_{i_\mu}] = t_2[a_{j_1}, \ldots, a_{j_\nu}] \,.$$

Proof of Step 1: If $\mu = 0$ then there is nothing to prove. Assume that $\mu > 0$. There are $K, R \geq 0$ such that $\zeta = DK + R$ with $R < D$. It suffices to show that there are $0 \leq a_{i_1}, \ldots, a_{i_\mu} < D$ and $K' \geq 0$ with $DK' = R + t_1[a_{i_1}, \ldots, a_{i_\mu}]$, since then

$$\begin{aligned} \zeta + t_1[a_{i_1}, \ldots, a_{i_\mu}] &= DK + R + t_1[a_{i_1}, \ldots, a_{i_\mu}] \\ &= DK + DK' = D(K + K') \,, \end{aligned}$$

and thus, $D | \zeta + t_1[a_{i_1}, \ldots, a_{i_\mu}]$.

First, assume the existence of $a_{i_1}, \ldots, a_{i_\mu} \geq 0$ with $D | R + t_1[a_{i_1}, \ldots, a_{i_\mu}]$, where $a_{i_\xi} \geq D$, for some $1 \leq \xi \leq \mu$. To simplify matters, we assume without loss of generality that $\xi = 1$. There is an $a \geq 0$ with $a_{i_1} = D + a$. Further, assume that there is no $b < a_{i_1}$ with $D | R + t_1[b, a_{i_2}, \ldots, a_{i_\mu}]$. For some $K' \geq 0$, we have that

$$DK' = R + t_1[a_{i_1}, \ldots, a_{i_\mu}] = R + Dk_{i_1} + t_1[a, a_{i_2}, \ldots, a_{i_\mu}] \,.$$

Therefore, $D(K' - k_{i_1}) = R + t_1[a, a_{i_2}, \ldots, a_{i_\mu}]$, i. e., $D | R + t_1[a, a_{i_2}, \ldots, a_{i_\mu}]$. This contradicts the minimality of $D + a$.

It remains to show the existence of $a_{i_1}, \ldots, a_{i_\mu} \geq 0$ with $D | R + t_1[a_{i_1}, \ldots, a_{i_\mu}]$. The existence reduces to the problem of whether the equation

$$D \cdot y - k_{i_1} \cdot x_{i_1} - \cdots - k_{i_\mu} \cdot x_{i_\mu} = R$$

has a solution in the natural numbers. This is the case since $\gcd(D, k_{i_1}, \ldots, k_{i_\mu}) = 1$, by assumption.

Proof of Step 2: Assume that there are $\gamma \geq 1$ distinct coefficients in t_2 of the equation (3). Without loss of generality, assume that $0 < |k_{j_1}| < \cdots < |k_{j_\gamma}|$. Let $W := \frac{\zeta + t_1[a_{i_1}, \ldots, a_{i_\mu}]}{D}$, $L := \frac{\|t\|_-}{D}$, and $\ell_\xi := \frac{|k_{j_\xi}|}{D}$, for $1 \leq \xi \leq \nu$. Note that $\ell_1 < \cdots < \ell_\gamma$ and that $\gcd(\ell_1, \ldots, \ell_\gamma) = 1$. The equation (3) simplifies with the a_is from Step 1 to

$$W = \ell_1 \cdot x_{j_1} + \cdots + \ell_\nu \cdot x_{j_\nu} \,. \tag{4}$$

An upper bound on W is

$$W \leq \tfrac{p2^N - q + (D-1)\|t\|_+}{D} \tag{5}$$

and a lower bound on W is

$$\begin{aligned} W &\geq \tfrac{2^N - q}{D} \geq \tfrac{2^N - \max\{c, \|t\|_-\}}{D} \geq \tfrac{2^{D(\ell_1 + \cdots + \ell_\nu)} - D(\ell_1 + \cdots + \ell_\nu)}{D} \\ &\geq 2^{\ell_1 + \cdots + \ell_\gamma} - (\ell_1 + \cdots + \ell_\gamma) \,. \end{aligned}$$

From the lower bound on W and the upper bound on Frobenius numbers (2), it follows that the equation (4) has a solution in the natural numbers. Let $\kappa \geq 0$ be maximal such that there are $a_1, \ldots, a_\gamma \geq 0$ with

$$W = \ell_1 a_1 + \cdots + \ell_\gamma a_\gamma + \kappa L \,. \tag{6}$$

We show that $a_1, \ldots, a_\gamma < L$. To achieve a contradiction, assume that there is a ξ, $1 \leq \xi \leq \gamma$ with $a_\xi = L + a$, for some $a \geq 0$. Without loss of generality, assume that $\xi = 1$. This contradicts the assumption that κ is maximal:

$$\begin{aligned} W &= \kappa L + \ell_1(L + a) + \ell_2 a_2 + \cdots + \ell_\gamma a_\gamma \\ &= (\kappa + \ell_1)L + \ell_1 a + \ell_2 a_2 + \cdots + \ell_\gamma a_\gamma \,. \end{aligned}$$

From κ and $a_1, \ldots, a_\gamma$, we obtain a solution for the equation (4) in the natural numbers, namely

$$\begin{aligned} W &= \kappa L + \ell_1 a_1 + \cdots + \ell_\gamma a_\gamma \\ &= \kappa(\ell_1 + \cdots + \ell_\nu) + \ell_1 a_1 + \cdots + \ell_\gamma a_\gamma \\ &= \ell_1(\kappa + a_1) + \cdots + \ell_\gamma(\kappa + a_\gamma) + \ell_{\gamma+1}\kappa + \cdots + \ell_\nu \kappa \,. \end{aligned}$$

It suffices to show that $\kappa < 2^N - \max\{a_1, \ldots, a_\gamma\}$. An upper bound on κ is

$$\begin{aligned} \kappa &\overset{(6)}{=} \tfrac{W - (\ell_1 a_1 + \cdots + \ell_\gamma a_\gamma)}{L} \leq \tfrac{W - \max\{a_1, \ldots, a_\gamma\}}{L} \\ &\overset{(5)}{\leq} \tfrac{p2^N - q + (D-1)\|t\|_+}{DL} - \tfrac{\max\{a_1, \ldots, a_\gamma\}}{L} \\ &\leq \tfrac{(\|t\|_- - 1)2^N}{DL} + \tfrac{D-1}{DL}\|t\|_+ - \tfrac{\max\{a_1, \ldots, a_\gamma\}}{L} \\ &\leq 2^N - \tfrac{2^N}{DL} + \tfrac{\|t\|_+ - \max\{a_1, \ldots, a_\gamma\}}{L} \,. \end{aligned}$$

It remains to check whether the inequality

$$2^N - \tfrac{2^N}{DL} + \tfrac{\|t\|_+ - \max\{a_1,\ldots,a_\gamma\}}{L} < 2^N - \max\{a_1,\ldots,a_\gamma\}$$

is valid. The previous inequality rewrites to

$$1 + \tfrac{\|t\|_+ + \max\{a_1,\ldots,a_\gamma\}(L-1)}{L} \leq \tfrac{2^N}{DL}\,.$$

Multiplying with the common denominator DL, the inequality rewrites further to

$$DL + D\|t\|_+ + D\max\{a_1,\ldots,a_\gamma\}(L-1) \leq 2^N\,.$$

Since $\max\{a_1,\ldots,a_\gamma\} \leq L-1$ and $N \geq \|t\|_- + \|t\|_+ = DL + \|t\|_+$, it suffices to show the validity of the inequality

$$DL + D\|t\|_+ + D(L-1)^2 \leq 2^{DL+\|t\|_+}\,. \qquad (7)$$

It is straightforward to show that the inequality (7) is true for all $D, L \geq 1$ and $\|t\|_+ \geq 0$.

Case III: $p < 0$ and $q \leq 0$. It suffices to show the existence of a solution $a_1,\ldots,a_r \in [2^N]$ for

$$t_1(x_{i_1},\ldots,x_{i_\mu}) = |p|2^N - |q| + t_2(x_{j_1},\ldots,x_{j_\nu})\,.$$

where t_1 and t_2 are defined as in Case II. This equation is symmetric to the equation (3) in Case II.

Case IV: $p > 0$ and $q < 0$. This case can be solved with Case I and Case II. Since $p > 0$ and $q < 0$, we have that $0 \in S$. By Case II, the state 0 is reachable from p, and by Case I, q is reachable from state 0.

Case V: $p < 0$ and $q > 0$. Analogously, this case can be solved by Case III and Case I. □

With Lemma 8 at hand it is straightforward to prove for $\mathcal{A}^{t\gtrless c}_{(m,n)}$ that $p \sim q$ iff $p = q$, for all $p, q \in S$. Therefore, we have that the minimal automaton representing $[\![t = c]\!]$ has at least $|S|$ states. Note that Lemma 8 also reveals the interesting fact that S is a strongly connected component in $\mathcal{A}^{t\gtrless c}_{(m,n)}$.

3.3. Divisibility Relation

In this subsection, we give an upper bound of the size of the minimal DWA for a formula $d|t+c$, where $d \geq 2$, $t(x_1,\ldots,x_r)$ is a homogeneous term and $c \in \mathbb{Z}$.

Let $\mathcal{A}^{d|t+c}$ be the DWA with the set of states $Q := \{q_{\mathrm{I}}, 0, 1, \ldots, d-1\}$. A state $q \in Q \cap \mathbb{Z}$ has the intuitive interpretation: if we reach the state q with a word $w \in (\{0,1\}^r)^*$ then the remainder of the division of $t[\langle w\rangle]$ by d equals q. We denote by $\mathrm{rem}(q,d)$ the remainder of $q \in \mathbb{Z}$ divided by d. Let $\mathcal{A}^{d|t+c} := (Q, \{0,1\}^r, \delta, q_{\mathrm{I}}, F)$ with

$$\delta(q,\overline{b}) := \begin{cases} \mathrm{rem}\big(-t[\overline{b}],d\big) & \text{if } q = q_{\mathrm{I}},\\ \mathrm{rem}\big(2q + t[\overline{b}],d\big) & \text{otherwise,}\end{cases}$$

for $q \in Q$ and $\overline{b} \in \{0,1\}^r$, and $F := \{q \in Q : d|q+c\}$, where q_{I} is interpreted as 0. Note that there is exactly one $q \in Q \cap \mathbb{Z}$ with $d|q+c$.

The correctness of our construction follows from the two facts which are straightforward to prove.

(a) For $n \in \mathbb{Z}$, $d|n+c$ iff $d|\,\mathrm{rem}(n,d)+c$.

(b) For $w \in (\{0,1\}^r)^+$, $\widehat{\delta}(q_{\mathrm{I}}, w) = \mathrm{rem}\big(t[\langle w\rangle], d\big)$.

The proof of (a) is straightforward. There are $p, q \in \mathbb{Z}$ such that $pd + q = n$ and $0 \leq q < d$. Note that $q = \mathrm{rem}(n,d)$. By definition, $d|n+c$ iff there is a $k \in \mathbb{Z}$ with $dk = n+c = pd+q+c$. The equality can be rewritten to $d(k-p) = q+c$, i. e., $d|\,\mathrm{rem}(n,d)+c$.

We prove (b) by induction over the length of w. For the base case, let $w = \overline{b} \in \{0,1\}^r$. Since we represent integers using 2's complement, we have that $t[\langle b\rangle] = -t[\overline{b}]$. By definition, $\widehat{\delta}(q_{\mathrm{I}}, \overline{b}) = \mathrm{rem}\big(t[\langle b\rangle], d\big)$. For the step case, assume $\widehat{\delta}(q_{\mathrm{I}}, w) = \mathrm{rem}\big(t[\langle w\rangle], d\big)$ and let $\overline{b} \in \{0,1\}^r$. There are $p, q \in \mathbb{Z}$ with $t[\langle w\rangle] = pd + q$ and $0 \leq q < d$. Note that $q = \mathrm{rem}\big(t[\langle w\rangle], d\big)$ and $t[\langle w\overline{b}\rangle] = 2t[\langle w\rangle] + t[\overline{b}] = 2pd + 2q + t[\overline{b}]$. We have that

$$\begin{aligned}\mathrm{rem}\big(t[\langle w\overline{b}\rangle], d\big) &= \mathrm{rem}(2pd + 2q + t[\overline{b}], d)\\ &= \mathrm{rem}(2q + t[\overline{b}], d) = \delta(q,\overline{b})\\ &\overset{\text{IH}}{=} \delta\big(\widehat{\delta}(q_{\mathrm{I}}, w), \overline{b}\big) = \widehat{\delta}(q_{\mathrm{I}}, w\overline{b})\,.\end{aligned}$$

Fact 9. *The DWA $\mathcal{A}^{d|t+c}$ represents $[\![d|t+c]\!]$ and has $d+1$ states.*

An optimization of the construction is to filter out the states that are not a multiple of $\gcd(\gcd(t), d)$. These states are not reachable from the initial state since $\mathrm{rem}(t[\overline{a}], d)$ is a multiple of $\gcd(\gcd(t), d)$, for every $\overline{a} \in \mathbb{Z}^r$. Additionally, we can merge states from which we cannot reach an accepting state.

3.4. Quantifier-free Formulas

In this subsection, we give an upper bound on the size of the minimal DWA for a quantifier-free formula. This upper bound depends on the maximal absolute value of the constants occurring in the (in)equations of the formula, the homogeneous terms, and the divisibility relations. The upper bound does *not* depend on the Boolean combination of the atomic formulas. This is not obvious since Boolean connectives are handled by the product construction if we construct the DWA recursively over the structure of the quantifier-free formula. The size of the resultant DWA using the product construction is in the worst case the product of the number of states of the two DWAs. Note that we allow the connective $\leftrightarrow$. Eliminating $\leftrightarrow$ can lead to exponentially larger formulas.

Let T be a finite nonempty set of homogeneous terms and let D be a finite set of integers greater than 1. Moreover, let $\ell > \max\{\|t\|_+ : t \in T\} \cup \{\|t\|_- : t \in T\}$.

Theorem 10. *Let ψ be a Boolean combination of atomic formulas $t \mathrel{\Join} c$ and $d|t+c'$, with $t \in T$, $d \in D$, $-\ell < c < \ell$, $c' \in \mathbb{Z}$, and $\Join \in \{=, \neq, <, \leq, >, \geq\}$. The size of the minimal DWA for ψ is at most $(2+2\ell)^{|T|} \cdot (1+\max D)^{|T|\cdot|D|}$.*

Proof. Without loss of generality, we assume that the variables occurring in terms in T are $y_1, \ldots, y_r$. Let $\mathcal{C}$ be the product automaton of all the $\mathcal{A}^{t=0}_{(-\ell,\ell)}$s and $\mathcal{A}^{d|t}$s, for $t \in T$ and $d \in D$. To simplify notation we omit the subscripts $(-\ell, \ell)$ and we assume that $T = \{t_1, \ldots, t_m\}$ and $D = \{d_1, \ldots, d_n\}$. Note that the states of $\mathcal{C}$ are tuples $(q^{t_1}, \ldots, q^{t_m}, q^{d_1,t_1}, \ldots, q^{d_n,t_m})$, where q^{t_i} is a state of $\mathcal{A}^{t_i=0}$ and q^{d_j,t_i} is a state of $\mathcal{A}^{d_j|t_i}$. By Fact 3, $\mathcal{A}^{t_i=0}$ has $2+2\ell$ states, and by Fact 9, $\mathcal{A}^{d_j|t_i}$ has $1+d_j$ states. It follows that the size of $\mathcal{C}$ is at most

$$\prod_{t\in T}(2+2\ell) \cdot \prod_{t\in T}\prod_{d\in D}(1+d) \leq (2+2\ell)^m \cdot (1+\max D)^{mn}.$$

It remains to customize the set of accepting states of $\mathcal{C}$ according to ψ. We identify an initial state of $\mathcal{A}^{t_i=0}$ or $\mathcal{A}^{d_j|t_i}$ with the integer 0. We define the DWA $\mathcal{D}$ as $\mathcal{C}$, except the set E of accepting states is defined as follows. A state $q = (q^{t_1}, \ldots, q^{t_m}, q^{t_1,d_1}, \ldots, q^{t_m,d_n})$ of $\mathcal{D}$ is in E iff $\mathfrak{Z} \models \psi_q$, where ψ_q is the formula obtained by substituting

- the integer q^{t_i} for the term t_i in the atomic formulas of the form $t_i \mathrel{\Join} c$, and
- the integer q^{t_i,d_j} for the term t_i in the atomic formulas of the form $d_j|t_i + c$.

Note that ψ_q is either true or not in $\mathfrak{Z}$ since it is a sentence.

It remains to prove that $\mathcal{D}$ represents $[\![\psi]\!]$. Let $w \in (\{0,1\}^r)^*$ be a word representing $\overline{a} \in \mathbb{Z}^r$. For a term $t \in T$, the value $t[\overline{a}]$ can be replaced by ℓ if $t[\overline{a}] \geq \ell$ and by $-\ell$ if $t[\overline{a}] \leq -\ell$ in every atomic formula of the form $t \mathrel{\Join} c$ without changing its truth value since $-\ell < c < \ell$. This modified value corresponds to the state reached by $\mathcal{A}^{t=0}$ after reading the word w. In an atomic formula of the form $d|t+c$, with $t \in T$ and $d \in D$, we can replace $t[\overline{a}]+c$ by $\mathrm{rem}(t[\overline{a}]+c, d)$ without changing the truth value. This adjusted value corresponds to the state reached by $\mathcal{A}^{d|t}$ after reading the word w. From the definition of E, it follows that $w \in L(\mathcal{D})$ iff $\mathfrak{Z} \models \psi[\overline{a}]$. □

4. Upper Bound on the Automata Size

In this section, we give an upper bound on the size of the minimal DWA for an arbitrary formula. We obtain this bound by examining the quantifier-free formula constructed by Reddy and Loveland's qe method [25], which improves Cooper's qe method [10]. We use Reddy and Loveland's qe method since the produced formulas are "small" with respect to the following parameters on which the upper bound of the minimal DWA in Theorem 10 depends.

Definition 11. *For a formula φ, let*

$$\begin{aligned}
\#_{\mathrm{terms}}(\varphi) &:= |\{t : t \mathrel{\Join} c \in \mathsf{A}(\varphi) \text{ or } d|t \in \mathsf{A}(\varphi)\}| \\
\#_{\mathrm{coefs}}(\varphi) &:= |\{k : k \text{ is a coefficient in } \alpha \in \mathsf{A}(\varphi)\}| \\
\#_{\mathrm{divs}}(\varphi) &:= |\{d : d|t \in \mathsf{A}(\varphi)\}| \\
\max_{\mathrm{const}}(\varphi) &:= \max\{1\} \cup \{|c| : t \mathrel{\Join} c \in \mathsf{A}(\varphi)\} \\
\max_{\mathrm{coef}}(\varphi) &:= \max\{1\} \cup \{|k| : k \text{ is a coefficient in } \alpha \in \mathsf{A}(\varphi)\} \\
\max_{\mathrm{div}}(\varphi) &:= \max\{1\} \cup \{d : d|t \in \mathsf{A}(\varphi)\}
\end{aligned}$$

where $\mathsf{A}(\varphi)$ is the set of atomic formulas of the form $d|t$ and the normal forms of the (in)equations occurring in φ.

4.1. Eliminating Quantifiers

For the sake of completeness, we briefly recall Reddy and Loveland's qe method. Consider the formula $\exists x\varphi$ with $\varphi(x, \overline{y}) \in \mathsf{QF}$. The construction of $\psi(\overline{y}) \in \mathsf{QF}$ proceeds in 2 steps.

Step 1: First, eliminate the connectives $\rightarrow$ and $\leftrightarrow$ in φ using standard rules, e. g., a subformula $\chi \rightarrow \chi'$ is replaced by $\neg\chi \vee \chi'$. Second, push all negation symbols in φ inwards (using De Morgan's laws, etc.) until they only occur directly in front of the atomic formulas. Third, rewrite all atomic formulas and negated atomic formulas in which x occurs such that they are all of one of the forms

$$k \cdot x < t(y_1, \ldots, y_n) \quad \text{(A)}$$
$$t(y_1, \ldots, y_n) < k \cdot x \quad \text{(B)}$$
$$d \mid t(x, y_1, \ldots, y_n) \quad \text{(C)}$$

with $k > 0$. For instance, the negated inequation $\neg 2 \cdot x + 9 \cdot y < 5$ is rewritten to $-9 \cdot y + 5 - 1 < 2 \cdot x$, and the negated equation $\neg 2 \cdot x + 9 \cdot y = 5$ is replaced by the disjunction $-9 \cdot y + 5 < 2 \cdot x \vee 2 \cdot x < -9 \cdot y + 5$. Let $\varphi'(x, \overline{y})$ be the resulting formula.

Step 2: Let $\psi_{-\infty}$ be the formula where all the atomic formulas of type (A) in φ' are replaced by "true", i. e., $0 < 1$, and all atomic formulas of type (B) are replaced by "false", i. e., $1 < 0$. Let B be the set of the atomic formulas in φ' of type (B), and let D be the least common multiple of the ds in the atomic formulas of type (C) and of the coefficients of the variable x in the atomic formulas of type (B). Let ψ be the formula

$$\bigvee_{1\leq j\leq D} \psi_{-\infty}[j/x] \vee \bigvee_{t+c<k\cdot x\in\mathsf{B}} \bigvee_{1\leq j\leq D} \left(k \mid t+c+j \wedge \varphi'[t+c+j/k\cdot x]\right),$$

where $\varphi'[t+c+j/k\cdot x]$ means that every atomic formula α in φ' in which x occurs is first multiplied by k and then

$k \cdot x$ is substituted by $t + c + j$, i. e., for a term t and $k \in \mathbb{Z}$,

$$\alpha[t/k \cdot x] := \begin{cases} k' \cdot t < k \cdot t' & \text{if } \alpha = k' \cdot x < t', \\ k \cdot t' < k' \cdot t & \text{if } \alpha = t' < k' \cdot x, \\ kd | k' \cdot t + k \cdot t' & \text{if } \alpha = d | k' \cdot x + t', \\ \alpha & \text{otherwise.} \end{cases}$$

Fact 12. *The formula ψ is logically equivalent to $\exists x \varphi$.*

4.2. Analysis

In [22], Oppen analyzed the length of the formulas produced by Cooper's qe method by relating the growth in the number of atomic formulas, the maximum of the absolute values of constants and coefficients appearing in these atomic formulas, and the number of distinct coefficients and divisibility predicates that may appear. Similar analysis of improved versions of Cooper's qe method are in [19, 25]. We refine the analysis [25] of Reddy and Loveland's qe method. Such analyses are technical and subtle. For the sake of readability, we have put the proofs of this subsection in the appendix. For a single application of the qe method, we have the following upper bounds.

Lemma 13. *For every formula $Qx\varphi$ with $\varphi \in \mathsf{QF}$ and $Q \in \{\exists, \forall\}$ there is a logically equivalent $\psi \in \mathsf{QF}$ with*

$$\begin{aligned} \#_{\text{terms}}(\psi) &\leq 4\, \#_{\text{terms}}(\varphi)^2 \\ \#_{\text{coefs}}(\psi) &\leq 16\, \#_{\text{coefs}}(\varphi)^4 + 2\, \#_{\text{coefs}}(\varphi) \\ \#_{\text{divs}}(\psi) &\leq \#_{\text{divs}}(\varphi) \cdot \big(\#_{\text{terms}}(\varphi) + 1\big) + \#_{\text{terms}}(\varphi) \\ \max_{\text{const}}(\psi) &\leq 2 \max_{\text{coef}}(\varphi) \cdot \big(\max_{\text{const}}(\varphi) + 1\big) + \\ &\qquad \max_{\text{coef}}(\varphi)^{\#_{\text{coefs}}(\varphi)+1} \cdot \max_{\text{div}}(\varphi)^{\#_{\text{divs}}(\varphi)} \\ \max_{\text{coef}}(\psi) &\leq 2 \max_{\text{coef}}(\varphi)^2 \\ \max_{\text{div}}(\psi) &\leq \max_{\text{coef}}(\varphi) \cdot \max_{\text{div}}(\varphi)\,. \end{aligned}$$

Using Lemma 13 and applying the qe method iteratively, we obtain the following bounds on the parameters.

Theorem 14. *For every formula φ containing $m \geq 0$ quantifiers there is a logically equivalent $\psi \in \mathsf{QF}$ such that*

$$\begin{aligned} \#_{\text{terms}}(\psi), \#_{\text{divs}}(\psi) &\leq a^{2^{3m}} \\ \#_{\text{coefs}}(\psi) &\leq c^{2^{3m}} \\ \max_{\text{coef}}(\psi), \max_{\text{div}}(\psi) &\leq s^{2^{m+1}} \\ \max_{\text{const}}(\psi) &\leq 2s^{(c^{2^{3m}} + a^{2^{3m}} + 2^m)} \end{aligned}$$

with $a \geq \max\{2, |\,\mathsf{A}(\varphi)|\}$, $c \geq \max\{3, \#_{\text{coefs}}(\varphi)\}$, and $s \geq \max\{2, \max_{\text{const}}(\varphi), \max_{\text{coef}}(\varphi), \max_{\text{div}}(\varphi)\}$.

Remark 15. Using the qe method naively to eliminate all quantifiers in a formula is insufficient to prove the upper bounds in Theorem 14. A common way to extend qe methods to arbitrary formulas is (a) to transform the formula into prenex normal form, and (b) to successively iterate the qe method from the innermost quantifier. The formula $(\forall x \varphi) \leftrightarrow \psi$ illustrates that (a) is not a good idea. Quantifiers do not distribute in general over $\rightarrow$ and $\leftrightarrow$. Therefore, we eliminate the connective $\leftrightarrow$ first. We obtain $((\forall x \varphi) \rightarrow \psi) \wedge (\psi \rightarrow \forall x \varphi)$. Eliminating $\rightarrow$ yields $((\neg \forall x \varphi) \vee \psi) \wedge (\neg \psi \vee \forall x \varphi)$. To move the quantifiers to the front, we have to push the first negation inwards. Finally, we obtain $\exists x \forall x' ((\neg \varphi \vee \psi) \wedge (\neg \psi \vee \varphi[x'/x]))$. We have not only doubled the length of the formula but we have also doubled the number of quantifiers. We want to *eliminate* quantifiers and have ended up doubling the work that we have to do. Fortunately, we can do better by successively (i) applying the qe method to subformulas of the form $Qx\varphi$ with $\varphi \in \mathsf{QF}$ and $Q \in \{\forall, \exists\}$, and (ii) replacing the subformula $Qx\varphi$ by the logically equivalent formula produced by the qe method.

4.3. Main Result

We now prove the main result of the paper: The size of the minimal DWA for a formula is at most triple exponential in the length of the formula.

Let $\varphi(y_1, \ldots, y_r)$ be a formula of length n. Note that $n \geq 3$ and that the numbers of quantifiers, atomic formulas, and the maximal absolute integer in φ are at most n. From Theorem 14 it follows that there is a logically equivalent $\psi(y_1, \ldots, y_r) \in \mathsf{QF}$ with

$$\begin{aligned} \#_{\text{terms}}(\psi), \#_{\text{divs}}(\psi) &\leq n^{2^{3n}} = 2^{2^{3n} \lg n} \leq 2^{2^{4n}} \\ \max_{\text{coef}}(\psi), \max_{\text{div}}(\psi) &\leq n^{2^{n+1}} = 2^{2^{n+1} \lg n} \leq 2^{2^{2n}}\,. \end{aligned}$$

Note that $x^y = 2^{y \lg x}$, for $x \geq 1$ and $y \geq 0$. Moreover, a strict upper bound on $\max_{\text{const}}(\psi)$ is

$$\begin{aligned} \max_{\text{const}}(\psi) &\leq 2n^{\left(n^{2^{3n}} + n^{2^{3n}} + 2^n\right)} \leq 2n^{3n^{2^{3n}}} = 2^{1 + 3n^{2^{3n}} \lg n} \\ &< 2^{2^n \cdot n^{2^{3n}}} \leq 2^{2^n \cdot 2^{2^{3n} \lg n}} \leq 2^{2^n \cdot 2^{2^{4n}}} \leq 2^{2^{2^{5n}}}\,. \end{aligned}$$

We have used the fact that $1 + 3x \lg y < 2^y x$, for $x, y \geq 2$.

Note that ψ is a Boolean combination of at most $\#_{\text{terms}}(\psi) \cdot \#_{\text{divs}}(\psi)$ atomic formulas of the form $d|t + c$ and of at most $\#_{\text{terms}}(\psi)$ (in)equations of the form $t \lessgtr c$, where $c \in \mathbb{Z}$. Since every term in ψ contains at most the variables $y_1, \ldots, y_r$, the sum of the absolute values of the coefficients in a term is bounded by $r 2^{2^{2n}} \leq n 2^{2^{2n}} < 2^{2^{2^{5n}}}$. With Theorem 10 at hand, we know that the size of the minimal DWA for ψ is at most

$$\begin{aligned} &\left(2 + 2 \cdot 2^{2^{2^{5n}}}\right)^{2^{2^{4n}}} \cdot \left(1 + 2^{2^{4n}}\right)^{(2^{2^{4n}})^2} \\ &\leq 2^{2^{1 + 2^{5n} + 2^{4n}}} \cdot 2^{2^{4n+1} \cdot 2^{2^{4n+1}}} \\ &\leq 2^{2^{2^{5n+2}}} \cdot 2^{2^{2^{4n+2}}} \leq 2^{2^{2^{5n+3}}}\,. \end{aligned}$$

Theorem 16. *The size of the minimal DWA for a formula of length n is at most $2^{2^{2^{O(n)}}}$.*

The above theorem does not change if we measure the length of integers logarithmically and not linearly. The only

change is that the maximal absolute integer in φ is not bounded by n but by 2^n. We have to adjust the bounds on $\max_{\mathrm{coef}}(\psi)$, $\max_{\mathrm{div}}(\psi)$, and $\max_{\mathrm{const}}(\psi)$. For instance $\max_{\mathrm{coef}}(\psi)$ is now bounded by $(2^n)^{2^{n+1}} = 2^{n2^{n+1}}$. This is still less than $2^{2^{cn}}$, for some $c \geq 1$. Analogously for $\max_{\mathrm{div}}(\psi)$ and $\max_{\mathrm{const}}(\psi)$.

5. Worst Case Example

We show that the upper bound in Theorem 16 is tight. We use the formulas $\mathrm{Prod}_n(x,y,z)$ defined in [14], for $n \geq 0$. In [14], Fischer and Rabin looked at the structure $(\mathbb{N},+)$ and not at $\mathfrak{Z}$, but it is straightforward to adapt the definition of $\mathrm{Prod}_n(x,y,z)$ to $\mathfrak{Z}$. For $n \geq 0$, $\mathrm{Prod}_n(x,y,z)$ has the following properties [14]: its length is in $\mathrm{O}(n)$ and for $a,b,c \in \mathbb{Z}$, we have that

$$\mathfrak{Z} \models \mathrm{Prod}_n[a,b,c] \quad \text{iff} \quad ab = c \text{ and } 0 \leq a,b,c < \prod_{\substack{p \text{ prime and} \\ p < f(n+2)}} p\,,$$

where $f(n) := 2^{2^n}$. It follows from the Prime Number Theorem that $\prod_{\substack{p \text{ prime and} \\ p<f(n+2)}} p \geq 2^{f(n)^2} = 2^{2^{2^{n+1}}} = 2^{f(n+1)}$.

Theorem 17. *Let $n \geq 0$. Every DWA representing $[\![\mathrm{Prod}_n]\!]$ has at least $2^{\frac{f(n+1)}{2}}$ states.*

We first prove the following lemma about the set $\mathrm{MULT}_m := \{(a,b,c) \in \mathbb{Z}^3 : a,b \in [2^m] \text{ and } ab = c\}$, for $m \geq 0$.

Lemma 18. *Let $m \geq 1$. Every DWA representing MULT_m has at least 2^m states.*

We use the following fact to prove Lemma 18.

Fact 19. *Let $\ell \geq 1$. For all $z \in \mathbb{N}$ with $2^{\ell-1} \leq z \leq 2^\ell - 2$, there are $x,y,z' \in [2^\ell]$ such that $xy = 2^\ell z + z'$.*

Proof. If $\ell = 1$ then there is nothing to prove since $2^{\ell-1} > 2^\ell - 2$. In the following, assume that $\ell > 1$ and $2^{\ell-1} \leq z \leq 2^\ell - 2$. Let $x,y \in [2^\ell]$ with $xy \geq 2^\ell z$ and $xy - 2^\ell z$ is minimal. Note that it is always possible to find $x,y \in [2^\ell]$ with $xy \geq 2^\ell z$ since for $x = y = 2^\ell - 1$, we have that

$$xy = 2^{2\ell} - 2^{\ell+1} + 1 \geq 2^{2\ell} - 2^{\ell+1} = 2^\ell(2^\ell - 2) \geq 2^\ell z\,.$$

Let $z' := xy - 2^\ell z$. We have to show that $z' \in [2^\ell]$. Since $xy \geq 2^\ell z$ we have that $z' \geq 0$. For the sake of absurdity, assume that $z' \geq 2^\ell$. It follows that

$$(x-1)y = xy - y = 2^\ell z + z' - y \geq 2^\ell z$$

since $y < 2^\ell$ and $z' \geq 2^\ell$. This contradicts the minimality of $xy - 2^\ell z$ since $xy > (x-1)y \geq 2^\ell z$. □

Proof (Lemma 18). Let $\mathcal{A} = (Q, \{0,1\}^3, \delta, q_{\mathrm{I}}, F)$ be a DWA representing MULT_m, and let K be the set of words $u \in (\{0\}^2 \times \{0,1\})^*$ with $2 \leq |u| \leq m+1$ and u starts with the letters $\overline{0}(0,0,1)$, where $\overline{0} := (0,0,0)$.

We first show that for every $w \in K \setminus \{\overline{0}(0,0,1)^m\}$ there is a $v \in (\{0,1\}^3)^*$ with $wv \in L(\mathcal{A})$. Let

$$\widehat{w} := \begin{cases} w(0,0,0) & \text{if } w \text{ is of the form } \overline{0}(0,0,1)^+, \\ w & \text{otherwise,} \end{cases}$$

and let $1+\ell$ be the length of $\widehat{w}$. Note that $2 \leq \ell \leq m$. The third track of $\widehat{w}$ encodes an integer $z \in \{2^{\ell-1}, \ldots, 2^\ell - 2\}$ since the first letter of $\widehat{w}$ is $\overline{0}$, the second letter is $(0,0,1)$, and $\widehat{w} \neq \overline{0}(0,0,1)^\ell$. From Fact 19, it follows that there are $x,y,z' \in [2^\ell]$ such that $xy = 2^\ell z + z'$. We have that $\widehat{w}\langle\!\langle x,y,z' \rangle\!\rangle_{\mathbb{N}} \in L(\mathcal{A})$. Let

$$v := \begin{cases} \overline{0}\langle\!\langle x,y,z' \rangle\!\rangle_{\mathbb{N}} & \text{if } w \text{ is of the form } \overline{0}(0,0,1)^+, \\ \langle\!\langle x,y,z' \rangle\!\rangle_{\mathbb{N}} & \text{otherwise.} \end{cases}$$

It holds that $\widehat{w}\langle\!\langle x,y,z' \rangle\!\rangle_{\mathbb{N}} = wv$.

Now, let $w' \in K \setminus \{w\}$. Because the first and second track of wv and $w'v$ encode both the pair (x,y), and the third track of $w'v$ does not encode $2^\ell z + z'$. It follows that $\widehat{\delta}(q_{\mathrm{I}}, w) \neq \widehat{\delta}(q_{\mathrm{I}}, w')$, for all $w' \in K \setminus \{w\}$.

Therefore, $\mathcal{A}$ must have at least $|K|$ states. Moreover, since $\lambda \in L(\mathcal{A})$ and $K \cap L(\mathcal{A}) = \emptyset$, these $|K|$ states are all distinct from q_{I}. It follows that $\mathcal{A}$ has at least $|K| + 1 = 2^0 + 2^1 + \cdots + 2^{m-1} + 1 = 2^m$ states. □

Proof (Theorem 17). Assume that for $n \geq 0$, there is a DWA $\mathcal{B}$ with less than $2^{\frac{f(n+1)}{2}}$ states representing the set $[\![\mathrm{Prod}_n]\!]$. Let $m := \frac{f(n+1)}{2}$. It holds that $\mathrm{MULT}_m \subseteq [\![\mathrm{Prod}_n]\!]$ since $(2^m - 1)^2 < 2^{2m} = 2^{f(n+1)}$. It is straightforward to construct from $\mathcal{B}$ a DWA representing the set MULT_m that has as many states as $\mathcal{B}$ by making some of the accepting states in $\mathcal{B}$ non-accepting. This contradicts Lemma 18. □

Note that the proof of Theorem 17 carries over to nondeterministic word automata.

6. Conclusion

We analyzed the automata-theoretic approach for deciding Presburger arithmetic and established a tight upper bound on the automata size. We have used the 2's complement representation to represent integers as words, where the first letter in a word has been interpreted as the most significant bit (big Endian). Similar representations have been used in [4, 18, 37]. Wolper and Boigelot [4, 37] do not restrict themselves to the 2's complement representation of integers. The automata constructions presented in [4, 37] are parameterized by $p \geq 2$ for the p's complement integer representation. The triple exponential upper bound on

the automata size carries over when using the p's complement representation. However, the effect on the automata size of switching from a p's complement representation to a q's complement representation is open.

Another representation for integers as words is to interpret the first letter of a word as the least significant bit (little Endian) [1, 7]. We can switch from one representation to the other by reversing the words and languages. The size of a DWA representing a PA definable set can be exponentially smaller by interpreting the first letter as the least significant bit. There are examples showing that such exponential reductions occur. (see appendix A.3). However, we have not found an example that shows the converse, i. e., where the least significant bit encoding has an exponentially larger minimal DWA than using the most significant bit encoding. It is an open problem how the two representations are precisely related and which representation is superior in practice. We point out that our results rely on interpreting the first letter as the most significant bit. For instance, Theorem 10 does not carry over if the first letter is interpreted as the least significant bit. Using the formulas defined in [14], which are used to describe Turing machines in PA, and using lower bounds on the BDD size for m-bit multiplication [8], it is straightforward to show a similar lower bound for the least significant bit encoding as in Theorem 17. Note that we cannot use lower bounds on the sizes of BDDs (or more generally, branching programs) when using the most significant bit encoding since the reading of the next digits by a DWA involves a left bit shift.

There are also *nonstandard* numeration systems for representing integers, e. g. [17]. In a nonstandard numeration system, the base is an infinite sequence of integers that has certain properties. A standard example is the sequence of Fibonacci numbers $(fib_i)_{i\geq 0}$: a natural number n is represented as a word $b_{n-1}\dots b_0 \in \{0,1\}^*$ with $n = \sum_{0\leq i<n} b_i \cdot fib_i$. The relationship between nonstandard numeration systems and formal language theory has been investigated, e. g., in [16, 20, 30]. It remains as future work to investigate the sizes of DWAs using different nonstandard numeration systems.

The main technique to prove the triple exponential upper bound on the automata size was to relate DWAs with the formulas constructed by a qe method. This technique can also be used to prove upper bounds on the sizes of minimal automata for other logics that admit qe, and where the structures are automata representable [3, 21], i. e., these structures are provided with automata for deciding equality on the domain and the atomic relations of the structure. Examples are the mixed first-order theory over the structure $(\mathbb{R}, \mathbb{Z}, <, +)$ [5, 36] and the first-order theory of queues [27, 28].

References

[1] C. Bartzis and T. Bultan. Efficient symbolic representations for arithmetic constraints in verification. *Int. J. Found. Comput. Sci.*, 14(4):605–624, 2003.

[2] L. Berman. The complexity of logical theories. *Theor. Comput. Sci.*, 11:71–77, 1980.

[3] A. Blumensath and E. Grädel. Automatic structures. In *LICS'00*, pp. 51–62, 2000.

[4] B. Boigelot. *Symbolic Methods for Exploring Infinite State Spaces*. PhD thesis, Faculté des Sciences Appliquées de l'Université de Liège, Liège, Belgium, 1999.

[5] B. Boigelot, S. Jodogne, and P. Wolper. On the use of weak automata for deciding linear arithmetic with integer and real veriables. In *IJCAR'01*, LNCS 2083, pp. 611–625, 2001.

[6] B. Boigelot, S. Rassart, and P. Wolper. On the expressiveness of real and integer arithmetic automata (extended abstract). In *ICALP'98*, LNCS 1443, pp. 152–163, 1998.

[7] A. Boudet and H. Comon. Diophantine equations, Presburger arithmetic and finite automata. In *CAAP'96*, LNCS 1059, pp. 30–43, 1996.

[8] R. Bryant. On the complexity of VLSI implementations and graph representations of Boolean functions with application to integer multiplication. *IEEE Trans. on Comp.*, 40:205–213, 1991.

[9] J. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.

[10] D. Cooper. Theorem proving in arithmetic without multiplication. *Machine Intelligence*, 7:91–99, 1972.

[11] J. Dixmier. Proof of a conjecture by Erdös and Graham concerning the problem of Frobenius. *J. Number Theory*, 34(2):198–209, 1990.

[12] H. Enderton. *A Mathematical Introduction to Mathematical Logic*. Academic Press, 1972.

[13] J. Ferrante and C. Rackoff. *The Computational Complexity of Logical Theories*. LNM 718. 1979.

[14] M. Fischer and M. Rabin. Super-exponential complexity of Presburger arithmetic. In *Symp. appl. Math.*, volume VII of *SIAM-AMS Proc.*, pp. 27–41, 1974.

[15] M. Fischer and M. Rabin. Super-exponential complexity of Presburger arithmetic. In B. Caviness and J. Johnson, editors, *Quantifier elimination and cylindrical algebraic decomposition*, Texts and Monographs in Symbolic Computation, pp. 122–135. 1998. Reprint of the article [14].

[16] C. Frougny. Representations of numbers and finite automata. *Math. Syst. Theory*, 25(1):37–60, 1992.

[17] C. Frougny. Numeration systems. In M. Lothaire, editor, *Algebraic Combinatorics on Words*, ch. 7. 2002.

[18] V. Ganesh, S. Berezin, and D. Dill. Deciding Presburger arithmetic by model checking and comparisons with other methods. In *FMCAD'02*, LNCS 2517, pp. 171–186, 2002.

[19] E. Grädel. Subclasses of Presburger arithmetic and the polynomial-time hierarchy. *Theor. Comput. Sci.*, 56:289–301, 1988.

[20] M. Hollander. Greedy numeration systems and regularity. *Theory Comput. Syst.*, 31(2):111–133, 1998.

[21] B. Khoussainov and A. Nerode. Automatic presentations of structures. In *LCC'94*, LNCS 960, pp. 367–392, 1995.

[22] D. Oppen. A $2^{2^{2^{pn}}}$ upper bound on the complexity of Presburger arithmetic. *J. Comput. Syst. Sci.*, 16:323–332, 1978.

[23] M. Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Sprawozdanie z I Kongresu metematyków slowiańskich, Warszawa 1929*, pp. 92–101, 395, 1930.

[24] M. Presburger. On the completeness of a certain system of arithmetic of whole numbers in which addition occurs as the only operation. *History and Philosophy of Logic*, 12(2):225–233, 1991. English translation of the article [23] by D. Jacquette.

[25] C. Reddy and D. Loveland. Presburger arithmetic with bounded quantifier alternation. In *STOC'78*, pp. 320–325, 1978.

[26] K. Reinhardt. The complexity of translating logic to finite automata. In E. Grädel, W. Thomas, and T. Wilke, editors, *Automata, Logics, and Infinite Games*, LNCS 2500, ch. 13, pp. 231–238. 2002.

[27] T. Rybina and A. Voronkov. A decision procedure for term algebras with queues. *ACM Trans. Comput. Log.*, 2(2):155–181, 2001.

[28] T. Rybina and A. Voronkov. Upper bounds for a theory of queues. In *ICALP'03*, LNCS 2719, pp. 714–724, 2003.

[29] U. Schöning. Complexity of Presburger arithmetic with fixed quantifier dimension. *Theory Comput. Syst.*, 30(4):423–428, 1997.

[30] J. Shallit. Numeration systems, linear recurrences, and regular sets. *Inf. Comput.*, 113(2):331–347, 1994.

[31] T. Shiple, J. Kukula, and R. Ranjan. A comparison of Presburger engines for EFSM reachability. In *CAV'98*, LNCS 1427, pp. 280–292, 1998.

[32] T. Skolem. Über einige Satzfunktionen in der Arithmetik. In *Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo, I. Matematisk naturvidenskapelig klasse*, volume 7, pp. 1–28, Oslo, 1931.

[33] T. Skolem. Über einige Satzfunktionen in der Arithmetik. In J. Fenstad, editor, *Selected Works in Logic*, pp. 281–306. Universitetsforlaget, Oslo, 1970. Reprint of the article [32].

[34] R. Stansifer. Presburger's article on integer arithmetic: Remarks and transl ation. Technical Report TR84-639, Department of Computer Science, Cornell University, Ithaca, NY, USA, 1984.

[35] L. Stockmeyer. *The complexity of decision problems in automata theory and logic*. PhD thesis, Department of Electrical Engineering, MIT, Boston, MA, USA, 1974.

[36] V. Weispfenning. Mixed real-integer linear quantifier elimination. In *ISSAC'99*, pp. 129–136, 1999.

[37] P. Wolper and B. Boigelot. On the construction of automata from linear arithmetic constraints. In *TACAS'00*, LNCS 1785, pp. 1–19, 2000.

A. Additional Proof Details

A.1. Proof Details of Lemma 13

The remainder of this subsection contains the proof of Lemma 13. For the sake of brevity, we write $\#_{\text{terms}}$ instead of $\#_{\text{terms}}(\varphi)$, and analogously for the other parameters.

If $Q = \exists$ then we apply Reddy and Loveland's qe method to the formula $\exists x\varphi$ in order to eliminate the existential quantified variable. If $Q = \forall$ then we rewrite the formula $\forall x\varphi$ to $\neg\exists x\neg\varphi$ and apply Reddy and Loveland's qe method to the formula $\exists x\neg\varphi$. Let ψ be the formula produced by the qe method.

Maximal Absolute Values. We first prove the upper bounds for $\max_{\text{coef}}(\psi)$, $\max_{\text{const}}(\psi)$, and $\max_{\text{div}}(\psi)$.

Step 1 may increase the maximal absolute value of a constant in an (in)equation by 1; the other two parameters are not altered in Step 1. The least common multiple D is at most

$$\max_{\text{coef}}{}^{\#_{\text{coefs}}} \cdot \max_{\text{div}}{}^{\#_{\text{divs}}} .$$

In order to carry out the substitution for an atomic formula $t + c < k \cdot x \in \mathsf{B}$ in Step 2 we multiply each atomic formula α of φ' in which x occurs by k. Recall that

$$\alpha[t{+}c{+}j\,/k{\cdot}x] = \begin{cases} k' \cdot (t + c + j) < k \cdot (t' + c') \\ \qquad\qquad \text{if } \alpha = k' \cdot x < t' + c', \\ k \cdot (t' + c') < k' \cdot (t + c + j) \\ \qquad\qquad \text{if } \alpha = t' + c' < k' \cdot x, \\ kd|k' \cdot (t + c + j) + k \cdot (t' + c') \\ \qquad\qquad \text{if } \alpha = d|k' \cdot x + t' + c. \end{cases}$$

In the worst case, we have that $j = D$, $|c| = |c'| = \max_{\text{const}} +1$, and $d = \max_{\text{div}}$. Moreover, it can be the case that there are coefficients ℓ and ℓ' of a variable y_i in t and t', respectively, with $|\ell| = |\ell'| = \max_{\text{coef}}$.

Since $k, |k'| \leq \max_{\text{coef}}$, we obtain the upper bounds

$$\begin{aligned} \max_{\text{coef}}(\psi) &\leq 2\max_{\text{coef}}{}^2 , \\ \max_{\text{div}}(\psi) &\leq \max_{\text{coef}} \cdot \max_{\text{div}} , \end{aligned}$$

and

$$\begin{aligned} \max_{\text{const}}(\psi) &\leq 2\max_{\text{coef}} \cdot (\max_{\text{const}} +1) + D \cdot \max_{\text{coef}} \\ &\leq 2\max_{\text{coef}} \cdot (\max_{\text{const}} +1) + \\ &\qquad \max_{\text{coef}}{}^{\#_{\text{coefs}}+1} \cdot \max_{\text{div}}{}^{\#_{\text{divs}}} . \end{aligned}$$

Note that the values of the parameters in the subformula $\bigvee_{1 \leq j \leq D} \psi_{-\infty}[j/x]$ of ψ cannot get larger than the bounds given above.

Number of Distinct Objects. For the following, let $\mathsf{B}_0 := \{t < k \cdot x \,:\, t + c \leq k \cdot x \in \mathsf{B}, \text{ for some } c \in \mathbb{Z}\}$. Note that $|\mathsf{B}_0| \leq \#_{\text{terms}}$. It holds that $\#_{\text{divs}}(\psi) \leq \#_{\text{divs}}(\psi_0)$ with

$$\psi_0 := \psi_{-\infty}[0/x] \vee \bigvee_{t < k \cdot x \in \mathsf{B}_0} (k \mid t \wedge \varphi'[t/k \cdot x])$$

since for an atomic formula of the form $d|t + c$ in ψ, we have an atomic formula of the form $d|t + c'$ in ψ_0. Similar upper bounds hold for $\#_{\text{coefs}}(\psi)$ and $\#_{\text{terms}}(\psi)$.

Step 1 does not alter the number of the ds in the atomic formulas of type (C), i. e., $\#_{\text{divs}} = \#_{\text{divs}}(\varphi')$. It holds that

$$\begin{aligned}\#_{\text{divs}}(\psi) &\leq \#_{\text{divs}}(\psi_0) \leq \#_{\text{divs}} + |\mathsf{B}_0| + |\mathsf{B}_0| \cdot \#_{\text{divs}} \\ &\leq \#_{\text{divs}} \cdot (\#_{\text{terms}} + 1) + \#_{\text{terms}} \ .\end{aligned}$$

Although the number of atomic formulas in Step 1 can change, we have that $\#_{\text{terms}} = \#_{\text{terms}}(\varphi')$. It is straightforward to see that the elimination of the connectives $\rightarrow$ and $\leftrightarrow$ does not alter the number of distinct terms. Moreover, this number is also not changed, when the negations are pushed in front of the atomic formulas. Since the atomic formulas in $\mathsf{A}(\varphi)$ and $\mathsf{A}(\varphi')$ are in normal form, the number $\#_{\text{terms}}$ is not altered in the third rewrite step. We have that

$$\begin{aligned}\#_{\text{terms}}(\psi) &\leq \#_{\text{terms}}(\psi_0) \\ &\leq \#_{\text{terms}} + \#_{\text{terms}} + |\mathsf{B}_0| + |\mathsf{B}_0| \cdot \#_{\text{terms}} \\ &\leq 4\, {\#_{\text{terms}}}^2 \ .\end{aligned}$$

It remains to give an upper bound on the number of distinct coefficients occurring in the formula ψ. Step 1 might double the number of distinct coefficients since by rewriting an atomic formula in one of the forms (A), (B), or (C) a coefficient k may become its opposite $-k$. Therefore, the number of distinct coefficients in φ' is at most $2\,\#_{\text{coefs}}$.

Let $t < k \cdot x \in \mathsf{B}_0$. In order to carry out the substitution, we multiply an atomic formula α of φ' in which x occurs by the factor k. α is of one of the forms $k' \cdot x < t' + c'$, $t' + c' < k' \cdot x$, or $d|k' \cdot x + t' + c'$. Let ℓ be a coefficient in the term t for the variable y_i, and let ℓ' be a coefficient in the term t' for the variable y_i. In the worst case, the coefficient $\pm k'\ell \pm k\ell'$ for the variable y_i in the atomic formula $\alpha[t/k \cdot x]$ is different for any possible value of k, k', ℓ, and ℓ'. Here, $\pm$ stands for either $+$ or $-$ depending on the type of α. We do not have to consider if $\pm$ is either $+$ or $-$ since we assume that if $k \in \mathbb{Z}$ is a coefficient in φ' then $-k$ may also occurs as a coefficient in φ'. Since k, k', ℓ, ℓ' are coefficients we get at most $(2\,\#_{\text{coefs}})^4$ distinct coefficients. Moreover, all these coefficients can be different to the coefficients in φ' which may still occur in the formula $\psi_{-\infty}[0/x]$, in the atomic formulas without the variable x, or in the added atomic formulas $k|t$. Therefore, we have that

$$\#_{\text{coefs}}(\psi) \leq 16\, {\#_{\text{coefs}}}^4 + 2\,\#_{\text{coefs}} \ .$$

This completes the proof of Lemma 13.

A.2. Proof Details of Theorem 14

Without loss of generality, we assume that the bound variables in $\varphi(y_1, \ldots, y_n)$ are $x_1, \ldots, x_m$.

Remark 15 illustrates that building the prenex normal form is not a good idea. Fortunately, we can do better by successively (i) applying the qe method to subformulas of the form $Qx\psi$ where ψ is quantifier-free and $Q \in \{\forall, \exists\}$, and (ii) replacing the subformula $Qx\psi$ by the obtained logically equivalent formula.

We define $\varphi_0 := \varphi$, and for $k > 0$, φ_k is the formula *after* the kth iteration of (i) applying the qe method to a subformula of φ_{k-1} that is of the form $Q_k x_k \psi_k$, where ψ_k is quantifier-free, and (ii) replacing $Q_k x_k \psi_k$ by the constructed quantifier-free formula. We denote by χ_k the outcome of the qe method to $Q_k x_k \psi_k$. We iterate (i) and (ii) m times. The sketched algorithm produces a quantifier-free formula φ_m that is logically equivalent to φ. We define $\psi := \varphi_m$.

Number of Distinct Objects. We prove by induction over $k \geq 0$ that the number of distinct terms t_k in φ_k is at most $2^{2^{2k}-1}a^{2^k}$. The base case holds for $k = 0$ due to the choice of a. Let $k \geq 1$. We apply the qe method to the subformula $Q_k x_k \psi_k$ of the formula φ_{k-1}. The worst case is that all distinct terms already occur in ψ_k and the terms in χ_k are all different from the terms occurring in φ_{k-1}. Clearly, it holds that $\#_{\text{terms}}(\psi_k) \leq t_{k-1}$, and by Lemma 13, it holds that $\#_{\text{terms}}(\chi_k) \leq (2\,\#_{\text{terms}}(\psi_k))^2$. Therefore,

$$\begin{aligned}t_k &\leq t_{k-1} + (2t_{k-1})^2 \\ &\overset{\text{IH}}{\leq} 2^{2^{2(k-1)}-1}a^{2^{k-1}} + \left(2 \cdot 2^{2^{2(k-1)}-1}a^{2^{k-1}}\right)^2 \\ &\leq 8 \cdot 2^{2 \cdot 2^{2(k-1)}-2}a^{2^k} = 2^{2^{2k-1}+1}a^{2^k} \\ &\leq 2^{2^{2k}-1}a^{2^k} \ .\end{aligned}$$

The last inequality holds since $2^{2k-1} + 1 \leq 2^{2k} - 1$, for all $k \geq 1$.

From t_m we obtain the claimed upper bound on the distinct number of terms in ψ:

$$\#_{\text{terms}}(\psi) \leq t_m \leq 2^{2^{2m}-1}a^{2^m} \leq a^{2^{2m}-1+2^m} \leq a^{2^{3m}} \ .$$

We prove by induction over $k \geq 0$ that the number d_k of distinct ds in the atomic formula of the form $d|t$ in φ_k is at most $2^{2^{2k}-1}a^{2^k}$. The base case for $k = 0$ is obvious. Let $k \geq 1$. We apply the qe method to the subformula $Q_k x_k \psi_k$ in the formula φ_{k-1}. An upper bound on d_k is

$$\begin{aligned}d_k &\leq d_{k-1} + \#_{\text{divs}}(\chi_k) \\ &\leq d_{k-1} + \#_{\text{divs}}(\psi_k) \cdot (\#_{\text{terms}}(\psi_k) + 1) + \#_{\text{terms}}(\psi_k) \\ &\leq d_{k-1} + d_{k-1}(t_{k-1} + 1) + t_{k-1} \\ &\leq t_{k-1} + (d_{k-1}t_{k-1} + 2d_{k-1} + 1) \\ &\overset{\text{IH}}{\leq} 2^{2^{2(k-1)}-1}a^{2^{k-1}} + (2^{2^{2(k-1)}-1}a^{2^{k-1}} + 1)^2 \\ &\leq 2^{2^{2(k-1)}-1}a^{2^{k-1}} + (2 \cdot 2^{2^{2(k-1)}-1}a^{2^{k-1}})^2 \\ &\leq 2^{2^{2k}-1}a^{2^k} \ .\end{aligned}$$

Analogously to the bound on $\#_{\text{terms}}(\psi)$ we get the bound $\#_{\text{divs}}(\psi) \leq a^{2^{3m}}$.

Let c_k be the number of distinct coefficients in φ_k. We show by induction over $k \geq 0$ that $c_k \leq 3^{4^k-1}c^{4^k}$. This bound obviously holds for $k = 0$. Let $k > 0$. We apply the qe method to the subformula $Q_k x_k \psi_k$ in the formula φ_{k-1}. An upper bound on c_k is

$$\begin{aligned} c_k &\leq c_{k-1} + \#_{\text{coefs}}(\chi_k) \\ &\leq c_{k-1} + 16\,\#_{\text{coefs}}(\psi_k)^4 + 2\,\#_{\text{coefs}}(\psi_k) \\ &\leq 16c_{k-1}^4 + 2c_{k-1} \leq 18c_{k-1}^4 \\ &\overset{\text{IH}}{\leq} 18(3^{4^{k-1}-1}c^{4^{k-1}})^4 \leq 18 \cdot 3^{4^k-4}c^{4^k} \\ &\leq 3^{4^k-1}c^{4^k}. \end{aligned}$$

The claimed bound on $\#_{\text{coefs}}(\psi)$ holds since

$$\begin{aligned} \#_{\text{coefs}}(\psi) \leq c_m &\leq 3^{4^m-1}c^{4^m} \leq c^{4^m-1+4^m} = c^{2^{2m}+2^{2m}-1} \\ &\leq c^{2^{3m}}. \end{aligned}$$

Maximal Absolute Values. The maximal absolute coefficient in the formula φ_k, for every $k \geq 1$, is equal to the maximum of the maximal absolute coefficient of φ_{k-1} and the maximal absolute coefficient that occurs in χ_k. Recall that χ_k is the constructed formula from the qe method applied to the subformula $Q_k x_k \psi_k$ of φ_{k-1}. Analogously for the maximal absolute constant, and the maximal ds in the atomic formulas of the form $d|t$ in of φ_k.

We prove by induction over $k \geq 0$ that $\max_{\text{coef}}(\varphi_k) \leq 2^{2^k-1}s^{2^k}$. The base case $k = 0$ is obvious. For the step case $k \geq 1$, we have that

$$\begin{aligned} \max_{\text{coef}}(\varphi_k) &= \max\{\max_{\text{coef}}(\varphi_{k-1}), \max_{\text{coef}}(\chi_k)\} \\ &\leq \max\{\max_{\text{coef}}(\varphi_{k-1}), 2\max_{\text{coef}}(\psi_k)^2\} \\ &\leq 2\max_{\text{coef}}(\varphi_{k-1})^2 \overset{\text{IH}}{\leq} 2\left(2^{2^{k-1}-1}s^{2^{k-1}}\right)^2 \\ &\leq 2^{2^k-1}s^{2^k}. \end{aligned}$$

Since $s \geq 2$ it holds that

$$\begin{aligned} \max_{\text{coef}}(\psi) = \max_{\text{coef}}(\varphi_m) &\leq 2^{2^m-1}s^{2^m} \leq s^{2^m-1+2^m} \\ &\leq s^{2^{m+1}}. \end{aligned} \tag{8}$$

We prove by induction over $k \geq 0$ that $\max_{\text{div}}(\varphi_k) \leq 2^{2^k-1}s^{2^k}$. The base case $k = 0$ is obvious. For the step case $k \geq 1$, we have that

$$\begin{aligned} \max_{\text{div}}(\varphi_k) &= \max\{\max_{\text{div}}(\varphi_{k-1}), \max_{\text{div}}(\chi_k)\} \\ &\leq \max\{\max_{\text{div}}(\varphi_{k-1}), \max_{\text{coef}}(\psi_k) \cdot \max_{\text{div}}(\psi_k)\} \\ &\leq \max_{\text{coef}}(\varphi_{k-1}) \cdot \max_{\text{div}}(\varphi_{k-1}) \\ &\overset{\text{IH}}{\leq} 2^{2^{k-1}-1}s^{2^{k-1}} \cdot 2^{2^{k-1}-1}s^{2^{k-1}} = 2^{2^k-1}s^{2^k}. \end{aligned}$$

Analogously to (8) we obtain the desired upper bound on $\max_{\text{div}}(\psi)$.

We prove by induction over $k \geq 0$ that $\max_{\text{const}}(\varphi_k) \leq 2s^{\left(c^{2^{3k}}+a^{2^{3k}}+2^k\right)}$. The base case $k = 0$ is obvious. For the step case, assume $k \geq 1$. We first give an upper bound on $N := \max_{\text{coef}}(\psi_k)^{\#_{\text{coefs}}(\psi_k)+1} \cdot \max_{\text{div}}(\psi_k)^{\#_{\text{divs}}(\psi_k)}$. Note that for every subformula ξ of φ_{k-1}, we have that $\max_{\text{coef}}(\xi) \leq \max_{\text{coef}}(\varphi_{k-1})$. Similar inequalities hold for the other parameters. Using our yet established upper bounds, we get that

$$\begin{aligned} N \leq \left(s^{2^k}\right)^{\#_{\text{coefs}}(\psi_k)+\#_{\text{divs}}(\psi_k)+1} &\leq \left(s^{2^k}\right)^{c^{2^{3(k-1)}}+a^{2^{3(k-1)}}+1} \\ = s^{\left(2^k c^{2^{3(k-1)}}+2^k a^{2^{3(k-1)}}+2^k\right)} &\leq s^{\left(c^{k+2^{3(k-1)}}+a^{k+2^{3(k-1)}}+2^k\right)} \\ \leq s^{\left(c^{2^{3k}}+a^{2^{3k}}+2^k\right)}. \end{aligned} \tag{9}$$

An upper bound on $2\max_{\text{coef}}(\psi_k)$ is

$$2\max_{\text{coef}}(\psi_k) \leq 2 \cdot 2^{2^{k-1}-1}s^{2^{k-1}} \leq 2^{2^{k-1}}s^{2^{k-1}} \leq s^{2^k}. \tag{10}$$

Using Lemma 13, we obtain the following upper bound on $\max_{\text{const}}(\varphi_k)$:

$$\begin{aligned} & \max\{\max_{\text{const}}(\varphi_{k-1}), \max_{\text{const}}(\chi_k)\} \\ \leq\ & \max\{\max_{\text{const}}(\varphi_{k-1}), \\ & \qquad 2\max_{\text{coef}}(\psi_k)(\max_{\text{const}}(\psi_k)+1)+N\} \\ \leq\ & 2\max_{\text{coef}}(\psi_k)(\max_{\text{const}}(\varphi_{k-1})+1)+N \\ \overset{(9)\&(10)}{\leq}\ & s^{2^k}\left(2\max_{\text{const}}(\varphi_{k-1})\right) + s^{\left(c^{2^{3k}}+a^{2^{3k}}+2^k\right)} \\ \overset{\text{IH}}{\leq}\ & s^{2^k}\left(4s^{\left(c^{2^{3(k-1)}}+a^{2^{3(k-1)}}+2^{k-1}\right)}\right) + s^{\left(c^{2^{3k}}+a^{2^{3k}}+2^k\right)} \\ \leq\ & s^{\left(c^{2^{3(k-1)}}+a^{2^{3(k-1)}}+2^{k-1}+2^k+2\right)} + s^{\left(c^{2^{3k}}+a^{2^{3k}}+2^k\right)} \\ \leq\ & 2s^{\left(c^{2^{3k}}+a^{2^{3k}}+2^k\right)}. \end{aligned}$$

For $k = m$, we obtain the claimed bound on $\max_{\text{const}}(\psi)$.

A.3. Big Endian Versus Little Endian

The size of a DWA representing a PA definable set can be exponentially smaller by interpreting the first letter as the least significant bit as the following example shows.

For $n \geq 1$, we define the formula $\varphi_n(y)$ as

$$\begin{aligned} & y \geq 0 \wedge \\ & \exists x_0 \ldots \exists x_{n-1} \Big(\textstyle\bigwedge_{0 \leq i < n} (x_i = 0 \vee x_i = 2^i) \wedge \\ & \qquad \textstyle\bigwedge_{0 \leq i < n} 2^{i+1} \mid y - x_0 - \ldots - x_i \wedge \\ & \qquad \neg 2^{n+1} \mid y - x_0 - \ldots - x_{n-1} \Big). \end{aligned}$$

The formula $\varphi_n(y)$ encodes the language of words over $\{0,1\}$, where the $(n+1)$st letter is 1 if we use the encoding where the least significant bit is the first letter. It is straightforward to define a DWA for this language that has $\mathrm{O}(n)$ states. It is well-known that the minimal DWA for the reverse language has $2^{\mathrm{O}(n)}$ states.

We have not found an example that shows the converse, i. e., where the most significant bit encoding leads to an exponentially smaller minimal DWA than using the least significant bit encoding.

B. Wolper and Boigelot's Argument

It has been claimed in [7] that the size of the minimal deterministic WA for a Presburger arithmetic formula is at most 3 exponentials in the length of the formula. Wolper and Boigelot explain in [37] that the proof in [7] is incorrect. In the following quoted paragraph from [37], they argue that there must be an elementary upper bound on the size of the minimal deterministic WA for a Presburger arithmetic formula. They write on page 14:[2]

> That the argument used in [7] is false does not mean that the size of the automaton for a Presburger formula will grow nonelementarily with respect to the number of quantifier alternations. Indeed, an analysis of the traditional quantifier elimination procedure for Presburger arithmetic [12] shows the opposite. Looking at this procedure, one notices that the number of basic formulas that are generated stays elementary in the size of the initial formula. Whatever the quantifier prefix of the formula, the quantifier elimination procedure only generates a Boolean combination of this elementary number of formulas. Hence, the formula obtained by the quantifier elimination procedure is elementary and so will be the corresponding automaton.

Our proof of Theorem 16 is based on an idea similar to that used by Wolper and Boigelot, namely, analyzing the formula produced by a quantifier elimination method. However, Wolper and Boigelot used the quantifier elimination method described in [12] and we used the quantifier elimination method by Reddy and Loveland [25]. The analysis suggested in the paragraph quoted above is not actually spelled out by Wolper and Boigelot. Moreover, their argumentation above is sketchy and unclear, as we explain below.

Reddy and Loveland's method does not require that the formula is put into disjunctive normal form whenever another quantifier is to be eliminated, unlike the method in [12], where it is assumed that when eliminating a quantifier the formula is given in the form

$$\bigvee_{1 \leq i \leq m} \exists x (\alpha_{i,1} \wedge \ldots \wedge \alpha_{i,k_i}) \qquad (11)$$

or possibly with a negation symbol in front of it (the $\alpha_{i,j}$s are atomic formulas or negated atomic formulas[3] and $m, k_1, \ldots, k_m \geq 1$). Putting a formula $\exists x \varphi$, where φ is quantifier-free, in such a disjunctive normal form may cause an exponential growth in the length of φ. The number of distinct atomic formulas remains the same. But since each of the existential quantifiers in (11) is eliminated separately, in spite of what is claimed in the above quoted paragraph, it is unclear why the number of distinct atomic formulas remains elementary bounded when applying this method to formulas with nested and alternating quantifiers.

The meaning of *elementary formula* in the last sentence of the quoted text is also unclear since the length of the obtained formula can be non-elementarily longer than the original formula. A possible meaning is that the obtained formula is logically equivalent to a quantifier-free formula of elementary length. However, even if the number of distinct atomic formulas is elementary bounded then it does not follow that the produced formula is logically equivalent to a quantifier-free formula of elementary length, since the atomic formulas in the produced formula might contain constants that are not elementary bounded.

Moreover, the concrete upper bound remains open in Wolper and Boigelot's argument. A crucial point in establishing the triple exponential upper bound is that deterministic WAs can optimally share the homogeneous terms in the (in)equations and the divisibility relations of Boolean combinations (see Theorem 10). Only taking into account the number of distinct atomic formulas in a Boolean combination will result in a significantly larger upper bound (at least by one exponent). Another reason for a worse upper bound is the following: Wolper and Boigelot have neither given automata constructions for the divisibility predicates nor for the modulo relations (see footnote 3). A crude upper bound on the size of the minimal deterministic WA for the atomic formula $d|t$ can be given by 2^n, where n is the size of the minimal deterministic WA for the equation $d \cdot x = t$, where the variable x does not occur in the term t. Note that $d|t$ is logically equivalent to $\exists x\, d \cdot x = t$. However, as we have seen in Fact 9, the size of the minimal deterministic WA is at most $d + 1$. Using the crude exponential upper bound for divisibility predicates results in an upper bound that is exponentially worse than the upper bound in Theorem 16.

2 The bibliographical references within this quotation have been adapted to correspond to those in the bibliography of this paper.

3 The logical language in [12] contains the modulo relations "$\equiv_d$" for $d \geq 2$ instead of the divisibility predicates "$d|$". Note that $x \equiv_d y$ iff $d|x - y$.