# Weak Integer Quantifier Elimination Beyond the Linear Case

Aless Lasaruk[1] and Thomas Sturm[2]

[1] FORWISS, Universität Passau, 94030 Passau, Germany
lasaruk@uni-passau.de
[2] FIM, Universität Passau, 94030 Passau, Germany
sturm@uni-passau.de

**Abstract.** We consider the integers using the language of ordered rings extended by ternary symbols for congruence and incongruence. On the logical side we extend first-order logic by bounded quantifiers. Within this framework we describe a weak quantifier elimination procedure for univariately nonlinear formulas. Weak quantifier elimination means that the results possibly contain bounded quantifiers. For fixed choices of parameters these bounded quantifiers can be expanded into finite disjunctions or conjunctions. In univariately nonlinear formulas all congruences and incongruences are linear and their modulus must not contain any quantified variable. All other atomic formulas are linear or contain only one quantified variable, which then may occur there with an arbitrary degree. Our methods are efficiently implemented and publicly available within the computer logic system REDLOG, which is part of REDUCE. Various application examples demonstrate the applicability of our new method and its implementation.

## 1 Introduction

After the fundamental work of Presburger [1] there has been considerable research on Presburger arithmetic, which is the additive theory of the integers with ordering and congruences. The largest part of this research was concerned with complexity issues and with decidability [2,3,4,5,6,7,8]. Weispfenning [9,10] was the first one who was explicitly interested in quantifier elimination as such in contrast to using it as a technique for decision. His quantifier elimination procedures are triply exponential, which is known to be optimal [3]. He managed, however, to optionally decrease that complexity by one exponential step to doubly exponential using the following technical trick: certain systematic disjunctions occurring during the elimination process are not written down explicitly. Instead one uses big $\bigvee$ (disjunction) and $\bigwedge$ (conjunction) operators with an index variable running over a finite range of integers. It is important to understand that at any time these big operators could be expanded such that one obtains a regular first-order formula at the price of considerably increasing the size of the representation. Independently, Weispfenning and others have developed virtual substitution techniques for quantifier elimination in various theories starting with the reals and including also valued fields and Boolean algebras [11,12,13,14,15].

V.G. Ganzha, E.W. Mayr, and E.V. Vorozhtsov (Eds.): CASC 2007, LNCS 4770, pp. 275–294, 2007.

In a recent publication [16] the authors of the present paper combined the two research areas by presenting integer quantifier elimination within the framework of virtual substitution. Furthermore, they extended that framework in order to cover a considerable generalization of Presburger arithmetic admitting as coefficients arbitrary polynomials in the parameters, i.e., the unquantified variables. This extension is called the *full linear theory of the integers*. It perfectly corresponds to what is referred to as linear quantifier elimination for the reals or for valued fields [12,14]. Recall that in regular Presburger arithmetic, in contrast, all coefficients must be numbers. The difference vanishes when considering decision problems. It is well-known that the full linear theory of the integers does not admit quantifier elimination in the traditional sense [10]. Instead one uses *weak quantifier elimination*. This does not necessarily deliver quantifier-free equivalents but formulas that possibly contain some *bounded quantifiers*. For this, one extends the language of logic by two additional quantifiers $\bigsqcup_{k:\,\beta}$ and $\bigsqcap_{k:\,\beta}$. Here $k$ is a variable, and $\beta$ is a formula not containing any quantifier. The semantics of the new quantifiers are defined as follows:

$$\bigsqcup_{k:\,\beta} \varphi \quad \text{iff} \quad \exists k(\beta \wedge \varphi), \qquad \bigsqcap_{k:\,\beta} \varphi \quad \text{iff} \quad \forall k(\beta \longrightarrow \varphi). \tag{1}$$

The quantifier $\bigsqcup_{k:\,\beta}$ is called an *existential bounded quantifier* if the solution set of $\beta$ wrt. $k$ is finite for all interpretations of all other variables. Under the same condition $\bigsqcap_{k:\,\beta}$ is called a *universal bounded quantifier*. Such formulas $\beta$ are called $k$-bounds. Formulas containing no quantifiers at all are called *strictly quantifier-free*. Formulas containing exclusively bounded quantifiers are called *weakly quantifier-free*. The choice of notation obviously resembles Weispfenning's big disjunction and conjunction operators. In general, however, bounded quantifiers can be explicitly expanded only for fixed choices of all parameters occurring therein.

In this paper, we introduce weak quantifier elimination for a subset of first-order formulas, which considerably extends the full linear theory of the integers discussed in [16]: Our language is

$$L = \{0^{(0)}, 1^{(0)}, -^{(1)}, +^{(2)}, \cdot^{(2)}, \neq^{(2)}, \leq^{(2)}, >^{(2)}, \geq^{(2)}, <^{(2)}, \equiv^{(3)}, \not\equiv^{(3)}\}.$$

Consider a formula $\varphi$ with parameters $a_1, \ldots, a_r$. Let $\varphi$ contain quantifiers $Q_1, \ldots, Q_s$ with quantified variables $x_1, \ldots, x_s$, where each occurrence of any of our new quantifiers is in fact a bounded quantifier. Assume furthermore that all right hand sides of equations, inequalities, congruences, and incongruences in $\varphi$ are 0, which can always be achieved by obvious equivalence transformations. Then we are able to eliminate from $\varphi$ all the regular quantifiers provided that $\varphi$ satisfies the following requirements:

($\mathrm{U}_1$) None of the quantified variables $x_1, \ldots, x_s$ occurs within moduli of congruences or incongruences. Note, however, that the moduli may be arbitrary polynomials in $a_1, \ldots, a_r$.

($U_2$) Considering the left hand side terms of congruences and incongruences as polynomials in $x_1, \ldots, x_s$, over the coefficient ring $\mathbb{Z}[a_1, \ldots, a_r]$ each such term has a total degree less than or equal to 1.

($U_3$) Considering the left hand side terms of equations and inequalities as polynomials in $x_1, \ldots, x_s$, over the coefficient ring $\mathbb{Z}[a_1, \ldots, a_r]$ each such term is either a nonlinear univariate polynomial or has a total degree less than or equal to 1.

We call formulas $\varphi$ satisfying these three conditions *univariately nonlinear*. If especially in ($U_3$) every single left-hand side term matches the second case, then $\varphi$ is a linear formula, and we are in the situation discussed in [16]. Thus note that according to our definition, every linear formula is also univariately nonlinear.

Accordingly, we refer to atomic subformulas of $\varphi$ the left hand sides of which match ($U_2$) or the second case in ($U_3$) as *linear atomic formulas* (wrt. $x_1, \ldots, x_s$). Those matching the first case in ($U_3$) are called *superlinear univariate atomic formulas* (wrt. $x_1, \ldots, x_s$).

As an example, consider the following formula, which is univariately nonlinear:

$$\forall y \exists x (ax - y < 0 \wedge x^2 + x + a > 0). \tag{2}$$

The atomic formula $ax - y < 0$ is linear, and the atomic formula $x^2 + x + a > 0$ is superlinear univariate.

As within the framework of [16], the elimination of regular quantifiers possibly introduces several new bounded quantifiers. It is noteworthy that in contrast to similar elimination procedures for higher degrees over the reals [13], we can positively decide by inspection of the original input that we are able to eliminate *all* present regular quantifiers.

The plan of the paper is as follows: Section 2 recalls some basic definitions and results from [16] and generalizes these to our extended framework here. In Section 3 we formulate and prove our elimination theorem. Section 4 gives an overview of our implementation in REDLOG and discusses various computation examples in order to give an idea about possible applications as well as the practical efficiency and limitations of our method. In Section 5 we summarize and evaluate our results and mention some ideas for future research.

## 2 Extended Virtual Substitution Framework

Our quantifier elimination procedure for univariately nonlinear formulas is going to use distinct substitution procedures for test terms originating from superlinear univariate atomic formulas on the one hand and from linear atomic formulas on the other hand.

This gives rise to two extensions of the existing framework: First, with each test point there must be stored in addition the respective substitution procedure. Second, our new substitution for test terms from linear atomic formulas is going to considerably extend the existing concept of virtual substitution. It is going to be called *constrained virtual substitution*.

### 2.1 Parametric Elimination Sets

Let $\varphi$ be a weakly quantifier-free formula. We recall some definitions and results from [16]. Originally, a parametric pre-elimination set for $\exists x\varphi$ had been defined there as a finite set

$$E = \big\{\, (\gamma_i, t_i, B_i) \;\big|\; 1 \le i \le n \,\big\}, \quad \text{where} \quad B_i = \big(\, (k_{ij}, \beta_{ij}) \;\big|\; 1 \le j \le m_i \,\big). \tag{3}$$

The *guards* $\gamma_i$ are strictly quantifier-free formulas, the *test points* $t_i$ are pseudo-terms possibly involving division, the $k_{ij}$ are variables, and the $\beta_{ij}$ are $k_{ij}$-bounds. Originally a parametric elimination set $E$ for $\exists x\varphi$ is then a parametric pre-elimination set such that for some virtual substitution procedure $\nu$ the set $E$ satisfies

$$\exists x\varphi \longleftrightarrow \bigvee_{(\gamma_i,t_i,B_i)\in E} \bigsqcup_{k_{i1}:\,\beta_{i1}} \cdots \bigsqcup_{k_{im_i}:\,\beta_{im_i}} \big(\gamma_i \wedge \nu(\varphi, t_i, x)\big). \tag{4}$$

With this definition, there is one single virtual substitution procedure used for all pseudo-terms in $E$. For the present paper we generalize this as follows: A *parametric pre-elimination set* for $\exists x\varphi$ is a finite set

$$E = \big\{\, (\gamma_i, t_i, \sigma_i, B_i) \;\big|\; 1 \le i \le n \,\big\}, \tag{5}$$

where the definitions of $\gamma_i$, $t_i$ and $B_i$ are as before. Each $\sigma_i$ is either the regular substitution $[\cdot/\cdot]$ of terms for variables or our new constrained virtual substitution $[\cdot/\!\!/\cdot]$, which we are going to explain in detail in the next subsection. A *parametric elimination set* $E$ for $\exists x\varphi$ is a parametric pre-elimination set such that

$$\exists x\varphi \longleftrightarrow \bigvee_{(\gamma_i,t_i,\sigma_i,B_i)\in E} \bigsqcup_{k_{i1}:\,\beta_{i1}} \cdots \bigsqcup_{k_{im_i}:\,\beta_{im_i}} \big(\gamma_i \wedge \sigma_i(\varphi, t_i, x)\big). \tag{6}$$

Assume that $\varphi$ contains parameters $a_1, \ldots, a_r$. Let $E$ be a parametric pre-elimination set for $\exists x\varphi$. For $z_1, \ldots, z_r \in \mathbb{Z}$, strictly quantifier-free formulas $\psi$, and pseudo-terms $t$ we use for a moment the notational convention

$$\psi' = \psi[z_1/a_1, \ldots, z_r/a_r], \quad t' = t[z_1/a_1, \ldots, z_r/a_r]. \tag{7}$$

Furthermore, for formulas $\beta'$ in at most one variable $k$ we denote the solution set wrt. $k$ by $S^k_{\beta'} = \{\, z \in \mathbb{Z} \mid \beta'(z) \,\}$. The *projection* $\Pi(E, z_1, \ldots, z_r)$ of $E$ is then defined as the finite set

$$\begin{aligned}\big\{\, &\big(\gamma'[y_1/k_1, \ldots, y_m/k_m],\ t'[y_1/k_1, \ldots, y_m/k_m],\ \sigma\big) \;\big|\; (\gamma, t, \sigma, B) \in E,\\ &B = ((k_j, \beta_j) \mid 1 \le j \le m),\ y_1 \in S^{k_1}_{\beta'_1},\ \ldots,\ y_m \in S^{k_m}_{\beta'_m[y_1/k_1,\ldots,y_{m-1}/k_{m-1}]} \,\big\}.\end{aligned} \tag{8}$$

**Lemma 1.** *Let $\varphi$ be a weakly quantifier-free formula with parameters $a_1, \ldots, a_r$. Let $E$ be a parametric pre-elimination set for $\exists x\varphi$ with the following property: For each interpretation $z_1, \ldots, z_r \in \mathbb{Z}$ of the parameters $a_1, \ldots, a_r$, we have*

$$\exists x\varphi' \longleftrightarrow \bigvee_{(\gamma',t',\sigma)\in\Pi(E,z_1,\ldots,z_r)} \big(\gamma' \wedge \sigma(\varphi', t', x)\big),$$

*where $\varphi' = \varphi[z_1/a_1, \ldots, z_r/a_r]$. Then $E$ is a parametric elimination set for the formula $\exists x\varphi$.* □

### 2.2 Constrained Virtual Substitution

The essential idea of a virtual substitution $\nu$ is to be able to substitute for a variable $x$ a parametric test point $t$ that is formally not a term of the underlying language. For instance, in our language $t$ might be a fraction. The virtual substitution maps atomic formulas to strictly quantifier-free formulas in such a way that whenever for some choice of values $z_1, \ldots, z_r \in \mathbb{Z}$ for the parameters $a_1, \ldots, a_r$ the test point evaluates to a number $t(z_1, \ldots, z_r) \in \mathbb{Z}$, then the following equivalence holds:

$$\varphi[t(z_1, \ldots, z_r)/x](z_1, \ldots, z_r) \longleftrightarrow \nu(\varphi, t, x)(z_1, \ldots, z_r). \tag{9}$$

Here, $[\cdot/\cdot]$ denotes regular substitution of terms for variables, where we allow ourselves to identify integers with corresponding sums of 1 or $-1$ representing them.

This notion of virtual substitution was sufficiently general for weak quantifier elimination from linear formulas as discussed in [16]. For our generalized setup here, however, we have to extend the concept of virtual substitution to *constrained virtual substitution*. Before giving a formal definition for this, let us return to our example in (2) in order to get a first idea about how our weak quantifier elimination would proceed for the elimination of $\exists x$:

$$\forall y \exists x (ax - y < 0 \wedge x^2 + x + a > 0). \tag{10}$$

From now on we allow ourselves to use absolute values within formulas as an abbreviated notation. Depending on the context they either stand for suitable case distinctions or for corresponding approximations by squares. The following suitable parametric elimination set for our example contains one entry originating from the first atomic formula and one entry from the second one:

$$E = \big\{\big(a \neq 0 \wedge y + k \equiv_a 0,\ \tfrac{y+k}{a},\ [\cdot /\!\!/ \cdot],\ ((k, |k| \leq |a|))\big),\\ \big(\text{true},\ k,\ [\cdot/\cdot],\ ((k, |k| \leq |a| + 2))\big)\big\}. \tag{11}$$

The pseudo-term $\frac{y+k}{a}$ in the first entry describes a finite set of points around the solution of the equation $ax - y = 0$ corresponding to the first atomic formula. The guard $a \neq 0 \wedge y + k \equiv_a 0$ ensures that the pseudo-term evaluates to an integer. The $k$-bound $|k| \leq |a|$ describes the range of an existential bounded quantifier to be introduced for $k$. The substitution $(ax - y < 0)[\frac{y+k}{a} /\!\!/ x]$ is defined as regular substitution of terms for variables followed by multiplication with the square of the denominator that comes into existence.

Assume for a moment that we define the substitution $(x^2 + x + a > 0)[\frac{y+k}{a} /\!\!/ x]$ in the same fashion: This would yield $(y + k)^2 + a(y + k) + a^3 > 0$. This is neither linear nor superlinear univariate wrt. $y$ and $k$. We thus make the following alternative definition:

$$(x^2 + x + a > 0)[\tfrac{y+k}{a} /\!\!/ x] := |ay + ak| > |a|^3 + 2a^2. \tag{12}$$

Notice that division of the right hand side of the definition by $a^2$ yields $|\frac{y+k}{a}| > |a| + 2$, where $|a| + 2$ is the Cauchy bound plus 1 of $x^2 + x + a$. So the right hand

side of (12) formulates that $\frac{y+k}{a}$ satisfies $x^2 + x + a > 0$ due to the fact that it lies outside the Cauchy bounds of this parabola, which extends to $+\infty$. For the possible case that $\frac{y+k}{a}$ lies in contrast within the Cauchy bounds but still satisfies $x^2 + x + a > 0$ there is something left to do.

This turns us to the other entry in (11). Here $|k| \leq |a| + 2$ is the bound of a bounded quantifier that substituting $k$ within its scope exactly covers every single point within the Cauchy bounds expanded by 1 of $x^2 + x + a$. Recall that the substitution $[\cdot/\cdot]$ is the regular substitution of terms for variables.

The overall elimination result for our example is the following weakly quantifier-free formula:

$$\bigsqcup_{k:\,|k|\leq|a|} \big(a \neq 0 \wedge y + k \equiv_a 0 \wedge k < 0 \wedge |ay + ak| > |a|^3 + 2a^2\big) \vee \bigsqcup_{k:\,|k|\leq|a|+2} \big(ak - y < 0 \wedge k^2 + k + a > 0\big). \tag{13}$$

For understanding why it is important to consider the Cauchy bounds expanded by 1 in contrast to simply the Cauchy bounds themselves, consider the example $\exists x(x^2 - 1 > 0 \wedge x = 1)$.

We now turn to formal definitions for the virtual substitution $[\cdot /\!\!/ \cdot]$. Substitution into linear atomic formulas works exactly as usual [13,16]:

$$\begin{aligned} (ax = b)\big[\tfrac{b'}{a'} /\!\!/ x\big] &:= (ab' = a'b), \\ (ax \leq b)\big[\tfrac{b'}{a'} /\!\!/ x\big] &:= (aa'b' \leq a'^2 b), \\ (ax \equiv_m b)\big[\tfrac{b'}{a'} /\!\!/ x\big] &:= (ab' \equiv_{ma'} a'b). \end{aligned} \tag{14}$$

One easily verifies that these substitutions satisfy Equivalence (9).

As already indicated by our example, we are going to use Cauchy bounds for substitution into superlinear univariate atomic formulas. Consider a parametric integer polynomial $p = c_n x^n + \cdots + c_0 \in \mathbb{Z}[a_1, \ldots, a_r][x]$. We define the *uniform Cauchy bound* of $p$ as $|c_{n-1}| + \cdots + |c_0| + 1$.

**Lemma 2.** *For $p = c_n x^n + \cdots + c_0 \in \mathbb{Z}[x]$ the following hold:*

*(i) For $c_n \neq 0$ the uniform Cauchy bound is always greater than or equal to the regular Cauchy bound:*

$$|c_{n-1}| + \cdots + |c_0| + 1 \geq \max\Big(1, \frac{|c_{n-1}| + \cdots + |c_0|}{|c_n|}\Big).$$

*(ii) Let $c_i \neq 0$ for at least one $i \in \{1, \ldots, n\}$. If $p(\xi) = 0$ for some $\xi \in \mathbb{R}$, then $|\xi| < |\xi| + 1 \leq |c_{n-1}| + \cdots + |c_0| + 1$.*

*Proof.* To start with, note that $|c_n| \geq 1$. If in (i) the regular Cauchy bound equals 1, then our claim is obvious. Else our claim follows from the following observation by division by $|c_n|$:

$$|c_n| \cdot (|c_{n-1}| + \cdots + |c_0| + 1) \geq |c_{n-1}| + \cdots + |c_0| + 1 > |c_{n-1}| + \cdots + |c_0|.$$

In (ii) we have, of course, $|\xi|$ less than or equal to the regular Cauchy bound. If $|c_n| \leq |c_{n-1}|+\cdots+|c_0|$, then we obtain $|\xi| \leq |c_n|\cdot|\xi| \leq |c_{n-1}|+\cdots+|c_0|$, which implies $|\xi|+1 \leq |c_{n-1}|+\cdots+|c_0|+1$. If, in contrast, $|c_n| > |c_{n-1}|+\cdots+|c_0|$, then $|\xi| \leq 1$. In case $|c_{n-1}|+\cdots+|c_0| > 0$ we obtain $|\xi|+1 \leq 2 \leq |c_{n-1}|+\cdots+|c_0|+1$. In case $|c_{n-1}| + \cdots + |c_0| = 0$ we have $p = c_n x^n$ with $c_n \neq 0$, thus $\xi = 0$, and it follows that $|\xi| + 1 = 1 = |c_{n-1}| + \cdots + |c_0| + 1$. □

We adopt from [16] the definition of an interval boundary. For a subset $S \subseteq \mathbb{Z}$ a number $z \in S$ is an *interval boundary* if $z - 1 \notin S$ or $z + 1 \notin S$. In the former case, $z$ is called a *lower* interval boundary. In the latter case, $z$ is called an *upper* interval boundary. Let now $\alpha$ be an atomic formula in at most one variable $x$. The *characteristic points* of $\alpha$ are the interval boundaries of the solution set $S^x_\alpha = \{\, z \in \mathbb{Z} \mid \alpha(z) \,\}$ wrt. $x$ of $\alpha$.

**Lemma 3.** *For $c_0, \ldots, c_n \in \mathbb{Z}$ consider an atomic formula $c_n x^n + \cdots + c_0 \;\varrho\; 0$ in at most one variable $x$, where $\varrho \in \{=, \neq, \leq, <, >, \geq\}$.*

*(i) For all characteristic points $k$ of $c_n x^n + \cdots + c_0 \;\varrho\; 0$ we have*

$$|k| \leq |c_{n-1}| + \cdots + |c_0| + 1.$$

*(ii) The atomic formula $c_n x^n + \cdots + c_0 \;\varrho\; 0$ has a constant truth value for choices $l$ of $x$ with $|c_{n-1}| + \cdots + |c_0| + 1 < l$. The same holds for choices $l$ of $x$ with $l < -(|c_{n-1}| + \cdots + |c_0| + 1)$.*

*Proof.* If we have in part (i) that $c_1 = \cdots = c_n = 0$, then there are no characteristic points at all, and the statement is trivial. Else let $k \in \mathbb{Z}$ be a characteristic point of $c_n x^n + \cdots + c_0 \;\varrho\; 0$. Using the definition above and the intermediate value theorem it is easy to see that there is a real zero $\xi$ of $c_n x^n + \cdots + c_0$ within the interval $[k-1, k+1]$. With Lemma 2(ii) it follows that $|k| \leq |\xi| + 1 \leq |c_{n-1}| + \cdots + |c_0| + 1$.

Part (ii) follows by induction from the observation that if an atomic formula has different truth values at $l$ and $l+1$, then either $l$ or $l+1$ is a characteristic point. □

Let now $p = c_n x^n + \cdots + c_0 \in \mathbb{Z}[a_1, \ldots, a_r][x]$, and let $\varrho$ be any of the relations in our language or equality. We define

$$\begin{aligned}\big(p \;\varrho\; 0\big)\big[\tfrac{b'}{a'} /\!\!/ x\big] := {} & \big(a'b' > a'^2(|c_{n-1}| + \cdots + |c_0| + 1) \wedge (p \;\varrho\; 0)[\infty /\!\!/ x]\big) \vee \\ & \big(a'b' < -a'^2(|c_{n-1}| + \cdots + |c_0| + 1) \wedge (p \;\varrho\; 0)[-\infty /\!\!/ x]\big). \quad (15)\end{aligned}$$

For substituting the nonstandard numbers $\pm\infty$ into atomic formulas we follow ideas by Weispfenning [13]. Substitution into equations and into negated equations is straightforward:

$$(p = 0)[\pm\infty /\!\!/ x] := \bigwedge_{i=0}^{n} c_i = 0, \quad (p \neq 0)[\pm\infty /\!\!/ x] := \bigvee_{i=0}^{n} c_i \neq 0. \qquad (16)$$

For ordering inequalities $\omega$ the definition is recursive. Denote by $q = c_{n-1}x^{n-1} + \cdots + c_0$ the formal reductum of $p$, and let $\omega^s := (\omega \setminus =)$ be the strict part of $\omega$. We first give the substitution for $+\infty$:

$$\begin{aligned}(p\ \omega\ 0)[\infty /\!\!/ x] &:= c_n\ \omega^s\ 0 \vee \big(c_n = 0 \wedge (q\ \omega\ 0)[\infty /\!\!/ x]\big) \quad \text{for} \quad n > 0,\\ (c_0\ \omega\ 0)[\infty /\!\!/ x] &:= c_0\ \omega\ 0.\end{aligned} \tag{17}$$

For the substitution of $-\infty$ one has to consider in addition the parities of the degrees during recursion:

$$\begin{aligned}(p\ \omega\ 0)[-\infty /\!\!/ x] &:= (-1)^n c_n\ \omega^s\ 0 \vee \big(c_n = 0 \wedge (q\ \omega\ 0)[-\infty /\!\!/ x]\big) \quad \text{for} \quad n > 0,\\ (c_0\ \omega\ 0)[-\infty /\!\!/ x] &:= c_0\ \omega\ 0.\end{aligned} \tag{18}$$

In contrast to our substitution (14) into linear atomic formulas our substitution (15) into superlinear univariate atomic formulas does not necessarily satisfy Equivalence (9). As a counterexample consider $((x-1)^2(x-2)^2 > 0)[0/\!\!/x]$. It satisfies, however, a weaker condition, which is made precise in the following lemma. It is one crucial technical observation of our paper that the weaker condition can still be exploited to establish an elimination theorem.

**Lemma 4 (Constrained Virtual Substitution).** *Consider a pseudo term $t = b'/a'$ and a superlinear univariate atomic formula $p\ \varrho\ 0$, where $p = c_n x^n + \cdots + c_0$ for $n \geq 2$. Set*

$$\lambda := |a'b'| > a'^2(|c_{n-1}| + \cdots + |c_0| + 1).$$

*Whenever for some choice $\mathbf{z} = (z_1, \ldots, z_r) \in \mathbb{Z}^r$ of the parameters $a_1, \ldots, a_r$ the test point $t$ evaluates to a number $t(\mathbf{z}) \in \mathbb{Z}$, then the following holds:*

$$\lambda(\mathbf{z}) \longrightarrow \big((p\ \varrho\ 0)[t(\mathbf{z})/x](\mathbf{z}) \longleftrightarrow (p\ \varrho\ 0)[t/\!\!/x](\mathbf{z})\big).$$

*Proof.* To start with, it is noteworthy that the premise $\lambda$ is equivalent to the disjunction of the two inequalities on the right hand side of the definition in (15). Furthermore, these inequalities exclude each other. Let $\mathbf{z} = (z_1, \ldots, z_r) \in \mathbb{Z}^r$ such that $t(\mathbf{z}) \in \mathbb{Z}$. Assume that $\lambda(\mathbf{z})$ holds. Then w.l.o.g. the first inequality $(a'b' > a'^2(|c_{n-1}| + \cdots + |c_0| + 1))(\mathbf{z})$ holds. This is equivalent to $t(\mathbf{z}) > (|c_{n-1}| + \cdots + |c_0| + 1)(\mathbf{z})$. By Lemma 3 we have

$$(p\ \varrho\ 0)[t(\mathbf{z})/x](\mathbf{z}) \longleftrightarrow (p\ \varrho\ 0)[l/x](\mathbf{z})$$

for all $l \geq t(\mathbf{z}) > (|c_{n-1}| + \cdots + |c_0| + 1)(\mathbf{z})$. The substitution of $\infty$ exactly simulates such points $l$:

$$(p\ \varrho\ 0)[t(\mathbf{z})/x](\mathbf{z}) \longleftrightarrow (p\ \varrho\ 0)[\infty /\!\!/ x](\mathbf{z}).$$

Since we are already in a situation where $(a'b' > a'^2(|c_{n-1}| + \cdots + |c_0| + 1))(\mathbf{z})$ holds, we finally obtain

$$\begin{aligned}(p\ \varrho\ 0)[\infty /\!\!/ x](\mathbf{z}) &\longleftrightarrow (a'b' > a'^2(|c_{n-1}| + \cdots + |c_0| + 1))(\mathbf{z}) \wedge (p\ \varrho\ 0)[\infty /\!\!/ x](\mathbf{z})\\ &\longleftrightarrow (p\ \varrho\ 0)[t/\!\!/x](\mathbf{z}).\end{aligned}$$

□

For clarity, we refer to virtual substitution procedures satisfying only the weaker condition described by the previous lemma as *constrained virtual substitution* procedures. The idea behind this notion is that $\lambda$ serves as a constraint under which the virtual substitution behaves well. Note, however, that for substitution into linear formulas we still have Equivalence (9) without any constraints.

## 3 Univariate Quantifier Elimination

In this section we present a quantifier elimination procedure for the set of univariately nonlinear formulas.

### 3.1 Elimination of One Quantifier

The following representation lemma implies that characteristic points can generally be expressed by weakly quantifier-free formulas in terms of the coefficients of the input formula.

**Lemma 5 (Representation Lemma).** *Consider the superlinear univariate atomic formula* $c_n x^n + \cdots + c_0 \varrho\ 0$ *wrt.* $x$ *where* $c_0, \ldots, c_n \in \mathbb{Z}[a_1, \ldots, a_r]$. *For a new variable* $k$, *we define the following strictly quantifier-free formula:*

$$\beta := |k| \leq |c_{n-1}| + \cdots + |c_0| + 1.$$

*Then* $\beta$ *is linear in* $k$. *Furthermore* $\beta$ *is a* $k$*-bound. Finally, for each interpretation* $z_1, \ldots, z_r \in \mathbb{Z}$ *of the parameters* $a_1, \ldots, a_r$ *the solution set* $S^k_\beta(z_1, \ldots, z_r)$ *contains all characteristic points of* $(c_n x^n + \cdots + c_0 \varrho\ 0)(z_1, \ldots, z_r)$.

*Proof.* The linearity of $\beta$ and the finiteness of its solution set wrt. $k$ are obvious. Choose interpretations $z_1, \ldots, z_r \in \mathbb{Z}$ of the parameters $a_1, \ldots, a_r$. If $c_i(z_1, \ldots, z_r) = 0$ for all $i \in \{1, \ldots, n\}$, then there are no characteristic points at all. Otherwise let $i \in \{1, \ldots, n\}$ be the largest index such that $c_i(z_1, \ldots, z_r) \neq 0$, and apply Lemma 3(i) to $(c_i x^i + \cdots + c_0 \varrho\ 0)(z_1, \ldots, z_r)$. □

**Lemma 6.** *Let* $\sigma$ *be one of our substitutions* $[\cdot/\cdot]$, $[\cdot/\!\!/\cdot]$. *Let* $\varphi'$ *be a weakly quantifier-free positive formula in one free variable* $x$. *Let* $t'$ *be a variable-free pseudo-term that possibly contains division but describes an integer* $t^* \in \mathbb{Z}$. *Assume that for all atomic subformulas* $\alpha$ *of* $\varphi'$ *and all interpretations* $y_1, \ldots, y_n \in \mathbb{Z}$ *of bound variables* $k_1, \ldots, k_n$ *occurring in* $\alpha$ *the following holds:*

$$\sigma(\alpha, t', x)(y_1, \ldots, y_n) \longrightarrow \alpha[t^*/x](y_1, \ldots, y_n).$$

*Then* $\sigma(\varphi', t', x) \longrightarrow \varphi'[t^*/x]$.

*Proof.* We proceed by induction on the word length of the formula $\varphi'$. If $\varphi'$ is an atomic formula, then it follows from the requirements of the lemma that $\sigma(\varphi', t', x)(y_1, \ldots, y_n) \longrightarrow \varphi'[t^*/x](y_1, \ldots, y_n)$ for all possible interpretations $y_1, \ldots, y_n \in \mathbb{Z}$ of the bound variables $k_1, \ldots, k_n$ occurring in $\varphi'$. Since both

$\sigma(\varphi', t', x)$ and $\varphi'[t^*/x]$ do not contain any other variables besides $k_1, \ldots, k_n$, it follows that $\sigma(\varphi', t', x) \longrightarrow \varphi'[t^*/x]$. Consider now the case that $\varphi'$ not atomic. Since $\varphi'$ is positive, it suffices to consider formulas of the form $\varphi' = \varphi'_1 \vee \varphi'_2$, $\varphi' = \varphi'_1 \wedge \varphi'_2$, $\varphi' = \bigsqcup_{k:\,\beta} \varphi'_1$, and $\varphi' = \bigsqcap_{k:\,\beta} \varphi'_1$.

Consider the case $\varphi' = \varphi'_1 \vee \varphi'_2$. Assume that $\sigma(\varphi', t', x) = \sigma(\varphi'_1 \vee \varphi'_2, t', x)$ holds. By our induction hypothesis we have both

$$\sigma(\varphi'_1, t', x) \longrightarrow \varphi'_1[t^*/x] \quad \text{and} \quad \sigma(\varphi'_2, t', x) \longrightarrow \varphi'_2[t^*/x].$$

Since both our substitutions are defined in terms of substitutions for atomic formulas it follows that $\sigma(\varphi'_1 \vee \varphi'_2, t', x) = \sigma(\varphi'_1, t', x) \vee \sigma(\varphi'_2, t', x)$. Thus at least one of $\sigma(\varphi'_1, t', x)$, $\sigma(\varphi'_2, t', x)$ holds and, accordingly, at least one of $\varphi'_1[t^*/x]$, $\varphi'_2[t^*/x]$ holds. Hence $\varphi'_1[t^*/x] \vee \varphi'_2[t^*/x]$ holds. The case $\varphi' = \varphi'_1 \wedge \varphi'_2$ is similar.

Next, consider the case $\varphi' = \bigsqcup_{k:\,\beta} \varphi'_1$. Assume that the premise of our desired implication holds:

$$\sigma(\varphi', t', x) = \sigma\Big(\bigsqcup_{k:\,\beta} \varphi'_1, t', x\Big) = \bigsqcup_{k:\,\beta} \sigma(\varphi'_1, t', x).$$

Then there is $y \in S^k_\beta$ such that $\sigma(\varphi'_1, t', x)[y/k] = \sigma(\varphi'_1[y/k], t', x)$. By the induction hypothesis it follows that $\varphi'_1[y/k][t^*/x] = \varphi'_1[t^*/x][y/k]$ holds. Hence by our choice of $y$ we obtain that the conclusion of our implication holds:

$$\bigsqcup_{k:\,\beta} (\varphi'_1[t^*/x]) = \Big(\bigsqcup_{k:\,\beta} \varphi'_1\Big)[t^*/x].$$

The case of a bounded universal quantifier is similar. Notice that then the induction hypothesis has to be applied several but finitely many times. □

It is not hard to see that the previous lemma does not hold for non-positive formulas $\varphi'$.

In [16] we have explicitly given a parametric elimination set for the subset of linear formulas in the uniform Presburger arithmetic. We are going to use that very set as a subset of our elimination set for the more general case discussed here. Note that in the following lemma the elimination set $E$ does not depend on the logical structure of $\varphi$ but only on the bounded quantifiers and the set of atomic formulas contained therein.

**Lemma 7 (Elimination of One Quantifier, Linear Case).** *Consider a linear formula $\exists x \varphi$ with parameters $a_1, \ldots, a_r$, where $\varphi$ is weakly quantifier-free, positive, and in prenex normal form:*

$$\varphi = \underset{k_1:\beta_1}{Q_1} \ \ldots \ \underset{k_n:\beta_n}{Q_n} \ \psi.$$

*Let the set of all atomic formulas of $\psi$ that contain $x$ be*

$$\{\, n_i x \ \varrho_i \ s_i + r_i \mid i \in I_1 \dot\cup I_2 \,\}.$$

*Here, the $n_i$ and $r_i$ are polynomials in the parameters $a_1, \ldots, a_r$. The $s_i$ are polynomials in both the parameters $a_1, \ldots, a_r$ and the bound variables $k_1, \ldots, k_n$. For $i \in I_1$, we have $\varrho_i \in \{=, \neq, <, \leq, \geq, >\}$. For $i \in I_2$, we have that $\varrho_i$ is either $\equiv_{m_i}$ or $\not\equiv_{m_i}$, where $m_i$ is a polynomial in $a_1, \ldots, a_r$. Let $k, k_1^*, \ldots, k_n^*$ denote new variables. Define*

$$\beta_1^* = \beta_1[k_1^*/k_1, \ldots, k_n^*/k_n], \quad \ldots, \quad \beta_n^* = \beta_n[k_1^*/k_1, \ldots, k_n^*/k_n].$$

*Define $m = \operatorname{lcm}\{\, m_i^2 + 1 \mid i \in I_2 \,\}$. For $i \in I_1 \cup I_2$ define*

$$s_i^* = s_i[k_1^*/k_1, \ldots, k_n^*/k_n] \quad \text{and} \quad \delta_i = -|n_i|m \leq k - s_i^* \leq |n_i|m.$$

*Then $E = \{\, (\gamma_i, t_i, B_i) \mid i \in I_1 \cup I_2 \,\} \cup \{(\text{true}, 0, \varnothing)\}$, where*

$$\gamma_i = (n_i \neq 0 \wedge r_i + k \equiv_{n_i} 0), \;\; t_i = \frac{r_i + k}{n_i}, \;\; B_i = \big((k_1^*, \beta_1^*), \ldots, (k_n^*, \beta_n^*), (k, \delta_i)\big),$$

*is a parametric elimination set for $\exists x \varphi$.* □

Note that the definition of $\gamma_i$ is such that whenever $\gamma_i$ holds, then the corresponding $t_i$ is defined and evaluates to an integer.

**Lemma 8 (Elimination of One Quantifier).** *Consider a univariately nonlinear formula $\exists x \varphi$ with parameters $a_1, \ldots, a_r$, where*

$$\varphi = \underset{k_1 : \beta_1}{Q_1} \; \ldots \; \underset{k_n : \beta_n}{Q_n} \; \psi.$$

*is weakly quantifier-free, positive, and in prenex normal form. Let $E_0$ be the (regular) parametric elimination set according to Lemma 7 for the subset of linear atomic formulas in $\psi$ and the bounded quantifiers occurring in $\varphi$. Let $\{\, p_i \;\varrho_i\; 0 \mid i \in I \,\}$ be the subset of superlinear univariate atomic formulas of $\psi$. Let $\{\, t_j \;\varrho_j\; 0 \mid j \in J \,\}$ be the set of all congruences and incongruences occurring in $\varphi$, i.e., $\varrho_j$ is either $\equiv_{m_j}$ or $\not\equiv_{m_j}$. Let $m = \operatorname{lcm}\{\, m_j^2 + 1 \mid j \in J \,\}$. For $i \in I$, denote by $u_i$ the uniform Cauchy bound of $p_i$, and define for a new variable $k$ the following strictly quantifier-free formula:*

$$\delta := \bigvee_{i \in I} |k| \leq u_i + m.$$

*Then the following is a parametric elimination set for $\exists x \varphi$:*

$$E = \{\, (\gamma, t, [\cdot /\!\!/ \cdot], B) \mid (\gamma, t, B) \in E_0 \,\} \cup \{(\text{true}, k, [\cdot/\cdot], ((k, \delta))\}.$$

*Proof.* Fix an interpretation $z_1, \ldots, z_r$ of the parameters $a_1, \ldots, a_r$. According to Lemma 1 it is sufficient to show that the projection $\Pi(E, z_1, \ldots, z_r)$ satisfies the following equivalence for $\varphi' := \varphi[z_1/a_1, \ldots, z_r/a_r]$:

$$\exists x \varphi' \longleftrightarrow \bigvee_{(\gamma', t', \sigma) \in \Pi(E, z_1, \ldots, z_r)} \big(\gamma' \wedge \sigma(\varphi', t', x)\big).$$

We first prove the implication from the right to the left. In contrast to the linear case and due to our constrained virtual substitution, this is not trivial. Suppose that the right hand side holds. Then for at least one $(\gamma'_j, t'_j, \sigma_j) \in \mathit{\Pi}(E, z_1, \ldots, z_r)$ the corresponding $\gamma'_j \wedge \sigma_j(\varphi', t'_j, x)$ holds. Recall that $t'_j$ is a pseudo-term possibly containing division. On the other hand, the validity of $\gamma'_j$ guarantees that $t'_j$ corresponds to an integer. Denote that integer by $t^*_j$. We are now going to prove the following, which by Lemma 6 implies that our $t^*_j \in \mathbb{Z}$ is one possible choice for $x$, such that $\exists x \varphi'$ holds: Let $\alpha$ be any atomic sub-formula of $\varphi'$ with bound variables $k_1, \ldots, k_n$. Let $y_1, \ldots, y_n \in \mathbb{Z}$ be an interpretation of $k_1, \ldots, k_n$. Then

$$\sigma_j(\alpha, t'_j, x)(y_1, \ldots, y_n) \longrightarrow \alpha[t^*_j/x](y_1, \ldots, y_n).$$

If $\sigma_j$ is the regular substitution $[\cdot/\cdot]$, then the implication is trivial. Else $\sigma_i$ is our constrained virtual substitution $[\cdot/\!\!/\cdot]$. If $\alpha$ is linear, then $[\cdot/\!\!/\cdot]$ satisfies Equivalence 9, and our implication is just the direction from the right to the left of that equivalence. If, in contrast, $\alpha$ is a superlinear univariate atomic formula $p_i \ \varrho_i \ 0$, where $i \in I$, then we make a case distinction on $t^*_j \in \mathbb{Z}$. If $|t^*_j| > u_i(z_1, \ldots, z_r)$, i.e., it lies outside the uniform Cauchy bound of $\alpha$, then our implication follows from Lemma 4. Otherwise, one verifies by inspection of Definition (15) that $\sigma_i(\alpha, t'_j, x)(k_1, \ldots, k_n) \longleftrightarrow$ false such that the implication holds trivially.

Assume vice versa that $\exists x \varphi'$ holds. Consider first the degenerate case that $S^x_{\varphi'} = \mathbb{Z}$. If $I \neq \varnothing$, then we have $(\mathrm{true}, 0, [\cdot/\cdot]) \in \mathit{\Pi}(E, z_1, \ldots, z_r)$. Otherwise, we have inherited from Lemma 7 $(\mathrm{true}, 0, [\cdot/\!\!/\cdot]) \in \mathit{\Pi}(E, z_1, \ldots, z_r)$, and in the absence of superlinear univariate formulas $[0/\!\!/x] = [0/x]$. Let now $\varnothing \subsetneq S^x_{\varphi'} \subsetneq \mathbb{Z}$.

If $S^x_{\varphi'} \cap S^k_\delta \neq \varnothing$, say, $z \in S^x_{\varphi'} \cap S^k_\delta$, then there is $(\mathrm{true}, z, [\cdot/\cdot]) \in \mathit{\Pi}(E, z_1, \ldots, z_r)$ originating from the test point $(\mathrm{true}, k, [\cdot/\cdot], ((k, \delta)))$.

Assume now that, in contrast, $S^x_{\varphi'} \cap S^k_\delta = \varnothing$. Then we are in a situation, where we can consider instead $\varphi$ the formula $\bar{\varphi}$ as follows: We replace in $\varphi$ each superlinear univariate atomic formula $p_i \ \varrho_i \ 0$ by the following strictly quantifier-free formula:

$$\big(x < -u_i \wedge (p_i \ \varrho_i \ 0)[-\infty/\!\!/x]\big) \vee \big(x > u_i \wedge (p_i \ \varrho_i \ 0)[\infty/\!\!/x]\big).$$

Defining $\bar{\varphi}' := \bar{\varphi}[z_1/a_1, \ldots, z_r/a_r]$ and on our assumption that $S^x_{\varphi'} \cap S^k_\delta = \varnothing$, we have $\bar{\varphi}' \longleftrightarrow \varphi'$, from which it follows that $\exists x \bar{\varphi}'$ holds. Since $\exists x \bar{\varphi}$ obtained this way is a linear formula we know by Lemma 7 a regular parametric elimination set for this:

$$E_0 \cup \big\{ (\mathrm{true}, -u_i + k, ((k, |k| \leq m))), (\mathrm{true}, u_i + k, ((k, |k| \leq m))) \ \big| \ i \in I \big\}.$$

We adapt this set to our constrained virtual substitution framework by adding to each test point the constrained virtual substitution:

$$\bar{E} := \{ (\gamma, t, [\cdot/\!\!/\cdot], B) \mid (\gamma, t, B) \in E_0 \} \cup \big\{ (\mathrm{true}, \pm u_i + k, [\cdot/\!\!/\cdot], ((k, |k| \leq m))) \ \big| \ i \in I \big\}.$$

Recall from the definition of our constrained virtual substitution $[\cdot /\!\!/ \cdot]$ that for linear formulas it equals the virtual substitution used in [16]. It thus follows that $\mathit{\Pi}(\bar{E}, z_1, \ldots, z_r)$ is an elimination set for $\exists x \bar{\varphi}'$. Consequently from the validity of $\exists x \bar{\varphi}'$ it follows that the following formula holds:

$$\bigvee_{(\gamma', t', \sigma) \in \mathit{\Pi}(\bar{E}, z_1, \ldots, z_r)} \big(\gamma' \wedge \sigma(\bar{\varphi}', t', x)\big).$$

Let $(\gamma', t', \sigma) \in \mathit{\Pi}(\bar{E}, z_1, \ldots, z_r)$ such that $\gamma' \wedge \sigma(\bar{\varphi}', t', x)$. We make a case distinction on the origin of $(\gamma', t', \sigma)$. If

$$(\gamma', t', \sigma) \in \mathit{\Pi}(\{\, (\gamma, t, [\cdot /\!\!/ \cdot], B) \mid (\gamma, t, B) \in E_0 \,\}, z_1, \ldots, z_r),$$

then it follows from $\{\, (\gamma, t, [\cdot /\!\!/ \cdot], B) \mid (\gamma, t, B) \in E_0 \,\} \subseteq E$ that $(\gamma', t', \sigma) \in \mathit{\Pi}(E, z_1, \ldots, z_r)$. In the other case, where

$$(\gamma', t', \sigma) \in \mathit{\Pi}(\{\, (\mathrm{true}, \pm u_i + k, [\cdot /\!\!/ \cdot], ((k, |k| \leq m))) \;\big|\; i \in I \,\}, z_1, \ldots, z_r),$$

recall from the formulation of the present lemma the definition of $m$, and observe that the following relation holds for all $y$ with $|y| \leq m(z_1, \ldots, z_r)$:

$$|\pm u_i(z_1, \ldots, z_r) + y| \leq (\pm u_i + m)(z_1, \ldots, z_r).$$

So there is a test point $(\gamma', t', [\cdot / \cdot]) \in \mathit{\Pi}(E, z_1, \ldots, z_r)$, which differs from our considered point only by the substitution procedure. In both cases, we have found a test point $(\gamma', t', \sigma^*) \in \mathit{\Pi}(E, z_1, \ldots, z_r)$, which differs from our considered point at most by the substitution procedure. We are now going to show that

$$\gamma' \wedge \sigma^*(\varphi', t', x).$$

Note that this is not trivial even in the first case where $\sigma = \sigma^* = [\cdot /\!\!/ \cdot]$, because in $\varphi'$ there possibly occur superlinear univariate formulas. Recall that we are in a situation where in particular $\gamma'$ holds, which implies that $t' \in \mathbb{Z}$. This allows to apply Equivalence (9), and it follows that $|t'| > u_i(z_1, \ldots, z_r)$ for all $i \in I$. Hence, using Lemma 4 and Equivalence (9), $\sigma^*(\varphi', t', x')$ equivalently replaces every single atomic formula in $\varphi'$ such that we obtain our desired observation $\gamma' \wedge \sigma^*(\varphi', t', x)$. Hence

$$\bigvee_{(\gamma', t', \sigma) \in \mathit{\Pi}(E, z_1, \ldots, z_r)} \big(\gamma' \wedge \sigma(\varphi', t', x)\big),$$

which is what had to be shown. □

### 3.2 Elimination Theorem

In order to possibly iterate weak quantifier elimination we next have to make sure that the output of our elimination procedure is again univariately nonlinear; in other words, it satisfies the defining conditions ($\mathrm{U}_1$)–($\mathrm{U}_3$) in the introduction. In contrast to the linear case, this observation is not trivial:

**Lemma 9.** *Let $\varphi$ be weakly quantifier-free, positive and prenex. Assume that $\exists x\varphi$ occurs within a univariately nonlinear formula $\hat{\varphi}$. Then replacing $\exists x\varphi$ in $\hat{\varphi}$ with the result of the application of the parametric elimination set $E$ from Lemma 8 is again univariately nonlinear.*

*Proof.* Let $x_1, \ldots, x_s$ be the quantified variables occurring in $\hat{\varphi}$. We have to show, that the formula

$$\varphi' = \bigvee_{(\gamma_i,t_i,\sigma_i,B_i)\in E} \bigsqcup_{k_{i1}:\beta_{i1}} \cdots \bigsqcup_{k_{im_i}:\beta_{im_i}} \big(\gamma_i \wedge \sigma_i(\varphi, t_i, x)\big). \tag{19}$$

satisfies our conditions $(\mathrm{U}_1)$–$(\mathrm{U}_3)$ wrt. $x_1, \ldots, x_s$. The bounds of our newly created bounded quantifiers obtained according to Lemma 5 do not contain any of the variable $x_1, \ldots, x_s$. Since each nontrivial guard originates from a regular elimination set, all guards also satisfy the conditions $(\mathrm{U}_1)$–$(\mathrm{U}_3)$. It is hence sufficient to consider formulas of the form $\sigma_i(\alpha, t_i, x)$ for each atomic formula $\alpha$ occurring in $\varphi$. If $\alpha$ is a linear formula the statement is trivial. For the case $\alpha$ is univariately nonlinear the statement is easily obtained by inspection of the definition in (15). □

**Theorem 10 (Elimination Theorem).** *The ordered ring of the integers with congruences admits weak quantifier elimination for univariately nonlinear formulas.*

*Proof.* Let $\hat{\varphi}$ be a univariately nonlinear formula. We proceed by induction on the number $n$ of regular quantifiers in $\hat{\varphi}$. If $n = 0$, then $\hat{\varphi}$ is already weakly quantifier-free. So there is nothing to do. Consider now the case $n > 0$. There is then a subformula of $\hat{\varphi}$ of one of the forms $\exists x\varphi$ or $\forall x\varphi$, where $\varphi$ is weakly quantifier-free. The latter case can be reduced to the former one by means of the equivalence $\forall x\varphi \longleftrightarrow \neg\exists x\neg\varphi$. We may w.l.o.g. assume that $\varphi$ is in prenex normal form and positive. By Lemma 8, there exists a parametric elimination set $E$ for $\exists x\varphi$. That is, $\exists x\varphi$ is equivalent to

$$\varphi' = \bigvee_{(\gamma_i,t_i,\sigma_i,B_i)\in E} \bigsqcup_{k_{i1}:\beta_{i1}} \cdots \bigsqcup_{k_{im_i}:\beta_{im_i}} \big(\gamma_i \wedge \sigma_i(\varphi, t_i, x)\big),$$

where $B_i = \big((k_{ij}, \beta_{ij}) \mid 1 \le j \le m_i\big)$. We obtain $\hat{\varphi}'$ from $\hat{\varphi}$ by equivalently replacing $\exists x\varphi$ with $\varphi'$. Lemma 9 states that $\hat{\varphi}'$ is again univariately nonlinear. Hence we can eliminate the remaining quantifiers from $\hat{\varphi}'$ by our induction hypothesis. □

**Corollary 11 (Decidability of Sentences).** *In the ordered ring of the integers with congruences univariately nonlinear sentences are decidable.*

*Proof.* Consider a univariately nonlinear sentence. Apply weak quantifier elimination. The result is an equivalent sentence containing only bounded quantifiers. In the absence of parameters these can be expanded into disjunctions and conjunctions. After this, all atomic formulas are variable-free such that we straightforwardly obtain either "true" or "false." □

## 4 Implementation and Computation Examples

The procedure described in this paper has been implemented in REDLOG, which stands for REDUCE *logic system* [17,18]. It provides an extension of the computer algebra system REDUCE to a computer logic system implementing symbolic algorithms on first-order formulas with respect to temporarily fixed first-order languages and theories. Such a choice of language and theory is called a *domain* or, alternatively, a *context*.

Before turning to the integer context relevant for our work here, we briefly summarize the other existing domains together with short names and alternative names, which are supported for backward compatibility:

BOOLEAN, B, IBALP. The class of Boolean algebras with two elements. These algebras are uniquely determined up to isomorphisms. BOOLEAN comprises quantified propositional calculus [15].

COMPLEX, C, ACFSF. The class of algebraically closed fields such as the complex numbers over the language of rings.

DIFFERENTIAL, DCFSF. A domain for computing over differentially closed fields. There is no natural example for such a field, but in special cases the methods can be used for obtaining relevant and interpretable results also for reasonable differential fields [19].

PADICS, DVFSF. One prominent example for discretely valued fields are the $p$-adic numbers for some prime $p$ with abstract divisibility relations encoding order between values. All PADICS algorithms are optionally uniform in $p$ [14].

QUEUES, QQE. A (two-sided) queue is a finite sequence of elements of some basic type. There are two sorts of variables, one for the basic type and one for the queue type. Accordingly, there is first-order quantification possible for both sorts. So far, the implementation is restricted to the reals as basic type [20].

REALS, R, OFSF. The class of real closed fields such as the real numbers with ordering. This context was the original motivation for REDLOG. It is still the most important and most comprehensive one [21].

TERMS, TALP. Free Malcev-type term algebras. The available function symbols and their arity can be freely chosen. [22].

The work discussed here has been integrated into another such domain:

INTEGERS, Z, PASF. The full linear theory of the integers.

This domain had been originally introduced for the methods described in [16]. It now naturally extends to univariately nonlinear formulas without loosing any of its previous features.

The idea of REDLOG is to combine methods from computer algebra with first-order logic thus extending the computer algebra system REDUCE to a computer logic system. In this extended system both the algebraic side and the logic side greatly benefit from each other in numerous ways. The current release REDLOG 3.0 is an integral part of the computer algebra system REDUCE 3.8. The implementation of our methods described here is part of the current development

version of REDLOG. It is going to be distributed with REDUCE 3.9. Until then it is freely available on the REDLOG homepage.[1]

We are now going to discuss various computations with our implementation. The idea is to illustrate the possible application range but also the limits of our method and of the current implementation. All our computations have been performed on a 1.66 GHz Intel Core 2 Duo processor T5500 using only one core and 128 MB RAM.

### 4.1 Optimization

We define a *parametric linear optimization problem with univariately nonlinear constraints* as follows: Minimize a cost function $\gamma_1 x_1 + \cdots + \gamma_n x_n$ subject to

$$A\mathbf{x} \geq \mathbf{b}, \quad p_1 \ \varrho_1 \ 0, \quad \ldots, \quad p_r \ \varrho_r \ 0.$$

As usual, $A = (\alpha_{ij})$ is an $m \times n$-matrix, and $\mathbf{b} = (\beta_1, \ldots, \beta_m)$ is an $m$-vector. For $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ we have $\alpha_{ij}$, $\beta_i$, $\gamma_j \in \mathbb{Z}[a_1, \ldots, a_k]$, i.e., all these coefficients are possibly parametric. For each $s \in \{1, \ldots, r\}$ we have $p_s \in \mathbb{Z}[a_1, \ldots, a_k][x_j]$ for some $j \in \{1, \ldots, n\}$, i.e., the $p_1, \ldots, p_r$ are parametric univariate polynomials. Each corresponding $\varrho_s$ is one of $=$, $\neq$, $\leq$, $>$, $\geq$, or $<$.

Using a new variable $z$ for the minimum such a problem can be straightforwardly translated to our framework as follows:

$$\exists x_1 \ldots \exists x_n \Big( \sum_{j=1}^{n} \gamma_j x_j \leq z \wedge \bigwedge_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} x_j \geq \beta_i \wedge \bigwedge_{s=1}^{r} p_s \ \varrho_s \ 0 \Big).$$

*Example 12.* Minimize $x + y$ subject to the following constraints:

$$x \geq 0, \quad y \geq 0, \quad x + y \geq 0, \quad \text{and} \quad x^2 + a < 0.$$

The formulation as a quantifier elimination problem reads as follows:

$$\exists x \exists y \big( x + y \leq z \wedge x \geq 0 \wedge y \geq 0 \wedge x + y \geq 0 \wedge x^2 + a < 0 \big).$$

For this REDLOG computes within 20 ms a weakly quantifier-free equivalent containing 103 atomic formulas. Setting then $a = 10$ and automatically simplifying yields within 2190 ms the result $z > 3$, i.e., the minimum for $x+y$ is 4. This final simplification step includes in particular expansion of all present bounded quantifiers. If we plug in $a = 10$ before the elimination, then we directly obtain $z > 3$ in only 330 ms. This amazing difference in time, we had already observed for the full linear theory of the integers [16]. It can be explained as follows: In both Lemma 7 and Lemma 8, we compute the least common multiple of the squares of all moduli. For non-parametric moduli we optimize this by using instead the absolute values of the moduli.

Generalizing our method discussed in the present paper to *extended quantifier elimination* [23,13,24,15] would admit to obtain in addition a sample point for the computed optimum. The optimization addressed above with the absolute value instead of squares could be applied in the parametric case as well by adding to the language a symbol for the absolute value.

[1] `www.redlog.eu`

### 4.2 Software Security

*Information flow control* is one important issue in software security [25,26]. The question is whether it is possible to manipulate parameters in such a way that sensitive information can become accessible outside of special code segments. We are going to discuss a modification of an example from [16].

*Example 13.* For the following piece of code there is a security risk if there are choices for `a` and `b` such that `y` is assigned the value of some `A[n^2]`.

```
if (a < b) then
   if (a+b mod 2 = 0) then
      n := (a+b)/2
   else
      n := (a+b+1)/2
   fi
   A[n^2] := get_sensitive_data(x)
   send_sensitive_data(trusted_receiver,A[n^2])
fi
y := A[abs(b-a)].
```

An attacker would be interested in a description of all values of `a` and `b` such that this happens. This can be formulated as follows:

$$\begin{aligned}\exists n\big((a<b\wedge a+b\equiv_2 0\wedge 2n=a+b\wedge\\ ((a<b\wedge b-a=n^2)\vee(a\geq b\wedge a-b=n^2)))\vee\\ (a<b\wedge a+b\not\equiv_2 0\wedge 2n=a+b+1\wedge\\ ((a<b\wedge b-a=n^2)\vee(a\geq b\wedge a-b=n^2)))\big).\end{aligned}$$

Our implementation computes in less than 10 ms the following weakly quantifier-free description:

$$\bigsqcup_{k\colon |k|\leq (a-b)^2+2}(a-b<0\wedge a-b+k^2=0\wedge a+b\not\equiv_2 0\wedge a+b-2k+1=0)\vee$$
$$\bigsqcup_{k\colon |k|\leq (a-b)^2+2}(a-b<0\wedge a-b+k^2=0\wedge a+b\equiv_2 0\wedge a+b-2k=0).$$

### 4.3 Integer Roots

*Example 14.* Consider the generic polynomial $p=ax^2+bx+c$. The question whether $p$ has an integer root can be expressed by a univariately nonlinear formula as follows:

$$\exists x(ax^2+bx+c=0).$$

Our elimination procedure yields after less than 10 ms the following weakly quantifier-free equivalent:

$$\bigsqcup_{k\colon |k|<|b|+|c|+2} ak^2+bk+c=0.$$

This result exactly substitutes all integers inside the uniform Cauchy bounds of $p$ expanded by 1. This expansion is the least common multiple of the (non-existing) moduli in the input.

This result obviously does not provide much mathematical insight. A helpful though imprecise intuition about our method is the following: *Its intelligence works mostly outside of the relevant Cauchy bounds.* Anyway, a slight modification of our previous example yields useful information:

*Example 15.* Given suitable $n_1$, $n_2 \in \mathbb{Z}$ and $d_1$, $d_2 \in \mathbb{N} \setminus \{0\}$ we look for integer zeros of $p(x) = \alpha x^2 + \beta x + \gamma$ within the interval $[n_1/d_1, n_2/d_2]$. This can be formulated as follows:

$$\exists x \big( p = 0 \wedge d_1 x \geq n_1 \wedge d_2 x \leq n_2 \big).$$

Let us consider the polynomial $p = x^5 - 3x^2 + 1$. We want to know whether there is a zero of $p$ in $[1/3, 3]$. This yields the following input:

$$\exists x (x^5 - 3x^2 + 1 = 0 \wedge 3x \geq 1 \wedge x \leq 3).$$

For this, our implementation computes "false" in less than 10 ms. In fact, our chosen $p$ has no integer zeros at all.

This last example illustrates the fact that our method combined with automatic simplification yields a *decision procedure* for univariately nonlinear sentences. So for sentences, we are able to obtain as a result either "true" or "false," which both do not contain any bounded quantifiers. Hence, concerning the decision of sentences, we provide a considerable extension of the original Presburger framework, where the user need not accept any additional syntactic constructs.

## 5 Conclusions

We have considered the integers using the language of ordered rings extended by ternary symbols for congruence and incongruence. On this basis we have given a weak quantifier elimination procedure for the set of univariately nonlinear formulas. The notion of weak quantifier elimination refers to the fact that the result possibly contains bounded quantifiers. For fixed choices of parameters these bounded quantifiers can be expanded into disjunctions or conjunctions. For decision problems they can be completely avoided. Our methods are efficiently implemented and publicly available within the computer logic system REDLOG, which is part of REDUCE. The applicability of our new method and its implementation has been demonstrated by means of various application examples. For the future it is planned to provide also an extended quantifier elimination procedure within the framework considered here. Furthermore, it appears to be a promising idea to extend the language by a symbol for the absolute value. This would allow to considerably reduce the ranges of the bounded quantifiers coming into existence.

## References


1. Presburger, M.: Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In: Comptes Rendus du premier congres de Mathematiciens des Pays Slaves, Warsaw, Poland, pp. 92–101 (1929)
2. Cooper, D.C.: Theorem proving in arithmetic without multiplication. Machine Intelligence 7, 91–99 (1972)
3. Fischer, M., Rabin, M.: Super-exponential complexity of Presburger arithmetic. SIAM-AMS Proceedings 7, 27–41 (1974)
4. Ferrante, J., Rackoff, C.W.: A decision procedure for the first-order theory of real addition with order. SIAM Journal on Computing 4, 69–77 (1975)
5. Ferrante, J., Rackoff, C.W.: The Computational Complexity of Logical Theories. Lecture Notes in Mathematics, vol. 718. Springer, Berlin (1979)
6. von zur Gathen, J., Sieveking, M.: A bound on solutions of linear integer equalities and inequalities. Proceedings of the AMS 72, 155–158 (1978)
7. Berman, L.: Precise bounds for Presburger arithmetic and the reals with addition. In: FOCS 1977. 18th Annual Symposium on Foundations of Computer Science, Providence, RI, USA, October 3–November 2, pp. 95–99. IEEE Press, Los Alamitos (1977)
8. Berman, L.: The complexity of logical theories. Theoretical Computer Science 11, 71–77 (1980)
9. Weispfenning, V.: The complexity of almost linear diophantine problems. Journal of Symbolic Computation 10(5), 395–403 (1990)
10. Weispfenning, V.: Complexity and uniformity of elimination in Presburger Arithmetic. In: Küchlin, W.W. (ed.) ISSAC 97. Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation, Maui, HI, pp. 48–53. ACM Press, New York, NY (1997)
11. Weispfenning, V.: The complexity of linear problems in fields. Journal of Symbolic Computation 5(1&2), 3–27 (1988)
12. Loos, R., Weispfenning, V.: Applying linear quantifier elimination. The Computer Journal 36(5), 450–462 (1993) (special issue on computational quantifier elimination)
13. Weispfenning, V.: Quantifier elimination for real algebra—the quadratic case and beyond. Applicable Algebra in Engineering Communication and Computing 8(2), 85–101 (1997)
14. Sturm, T.: Linear problems in valued fields. Journal of Symbolic Computation 30(2), 207–219 (2000)
15. Seidl, A.M., Sturm, T.: Boolean quantification in a first-order context. In: Ganzha, V.G., Mayr, E.W., Vorozhtsov, E.V. (eds.) Computer Algebra in Scientific Computing. Proceedings of the CASC 2003, Institut für Informatik, Technische Universität München, München, Germany, pp. 329–345 (2003)
16. Lasaruk, A., Sturm, T.: Weak quantifier elimination for the full linear theory of the integers. a uniform generalization of presburger arithmetic. Technical Report MIP-0604, FMI, Universität Passau, D-94030 Passau, Germany (2006) (to appear in the journal AAECC)
17. Dolzmann, A., Sturm, T.: Redlog: Computer algebra meets computer logic. ACM SIGSAM Bulletin 31(2), 2–9 (1997)
18. Dolzmann, A., Sturm, T.: Redlog user manual. Technical Report MIP-9905, FMI, Universität Passau, D-94030 Passau, Germany, Edition 2.0 for Version 2.0 (1999)

19. Dolzmann, A., Sturm, T.: Generalized constraint solving over differential algebras. In: Ganzha, V.G., Mayr, E.W., Vorozhtsov, E.V. (eds.) Computer Algebra in Scientific Computing. Proceedings of the CASC 2004, Institut für Informatik, Technische Universität München, München, Germany, pp. 111–125 (2004)
20. Straßer, C.: Quantifier elimination for queues. In: Draisma, J., Kraft, H. (eds.) Rhine Workshop on Computer Algebra. Proceedings of the RWCA 2006, pp. 239–248. Universität Basel, Basel (2006)
21. Dolzmann, A., Sturm, T., Weispfenning, V.: Real quantifier elimination in practice. In: Matzat, B.H., Greuel, G.M., Hiss, G. (eds.) Algorithmic Algebra and Number Theory, pp. 221–247. Springer, Berlin (1998)
22. Sturm, T., Weispfenning, V.: Quantifier elimination in term algebras. The case of finite languages. In: Ganzha, V.G., Mayr, E.W., Vorozhtsov, E.V. (eds.) Computer Algebra in Scientific Computing. Proceedings of the CASC 2002, Institut für Informatik, Technische Universität München, München, Germany, pp. 285–300 (2002)
23. Weispfenning, V.: Simulation and optimization by quantifier elimination. Journal of Symbolic Computation 24(2), 189–208 (1997) (special issue on applications of quantifier elimination)
24. Dolzmann, A., Sturm, T.: P-adic constraint solving. In: Dooley, S. (ed.) ISSAC 99. Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation, Vancouver, BC, pp. 151–158. ACM Press, New York, NY (1999)
25. Snelting, G.: Quantifier elimination and information flow control for software security. In: Dolzmann, A., Seidl, A., Sturm, T. (eds.) Algorithmic Algebra and Logic. Proceedings of the A3L 2005, BoD, Germany, Norderstedt, pp. 237–242 (2005)
26. Snelting, G., Robschink, T., Krinke, J.: Efficient path conditions in dependence graphs for software safety analysis. ACM Transactions on Software Engineering and Methodology 15(4), 410–457 (2006)