Testing equality of parametric semi-linear sets

Engel Lefaucheux

Testing equality of parametric semi-linear sets

Engel Lefaucheux

Université de Lorraine, CNRS, Inria, LORIA

1 Introduction

The first order theory of the integers commonly known as Peano arithmetic equips the integers with addition, multiplication and order. Unfortunately, contrary to his analogous over the reals [12], Peano arithmetic is known to be undecidable. As a consequence, a lot of work has been done to find decidable fragments of Peano arithmetic. The most famous example is Presburger arithmetic [7], which removes multiplication from Peano, and regain decidability this way.

while removing multiplication is an important limitation, Presburger arithmetic has shown a lot of applications. For instance, the Coq proof assistant and the theorem prover Princess [11] are both examples of automated theorem provers that base some or all of their functionality on Presburger arithmetic. Semi-linear sets [6], which are the sets of integers that can be described by formulas in Presburger arithmetic, appear in many fields, from vector addition systems [8] to timed automata [3].

Extensions of Presburger arithmetic often aimed at adding a restricted form of multiplication. Adding the divisibility predicate for instance, while remaining undecidable in general [10] as multiplication can be re-encoded with this predicate, has been shown to be decidable when considering only the purely existential fragment [4]. This direction has been pushed more recently in [2] and [9] where some generalisations are established, with restrictions on the quantifier alternations and shape of the formulas.

In a parallel direction, a line of research (see [1] for instance) has considered parametric Presburger arithmetic, where one or more variables, here called parameters, are considered separately and treated like constants (and thus can be multiplied). While this approach is different in mindset, under some conditions the behaviour of the parameters can be represented with the divisibility predicate and thus fall under the results of [4] or [2].

Similarly to how semi-linear sets arise from Presburger arithmetic, sets can be defined from parametric Presburger arithmetic. The goal of our work is to define formally and to study those parametric semi-linear sets. More precisely, it is known [5] that semi-linear sets can be defined by an union of sets, each of which are described by an initial vector and a set of period vectors that can be added to the initial vector as many times as wanted. Parametric semi-linear sets similarly are an union of sets, each of which are represented by an initial elements and periodic elements to be added to it. The difference with traditional semi-linear sets is that these elements are vectors of intervals which borders are defined by

polynomials in the parameters. Parametric semi-linear sets cannot be defined in Presburger arithmetic in general, and can in fact quickly become complex to analyse. This document focuses mainly on computing for which parameter values, two given parametric semi-linear sets are equal. This is work in progress, some extensions of these results that are currently being written are discussed in conclusion. For now, this paper addresses how to compute those parameters when (1) there is a single parameter, (2) the semi-linear sets defined once a parameter is selected is one dimensional and (3) the shape of the polynomials is restricted. Point (3) should be lifted soon.

2 Preliminaries

Presburger arithmetic [7] is the first order theory of the natural numbers with addition.

Definition 1. *Given a set* VAR *of variables, a Presburger term* t *follows the grammar*

$$t ::= \mathbb{N} \mid \mathsf{VAR} \mid t + t \mid \mathbb{N} * t.$$

Presburger formula are defined as

$$f ::= t \leq t \mid \neg f \mid f \vee f \mid f \wedge f \mid \forall x\ f \mid \exists x\ f$$

where t *represents Presburger terms and* x *is a variable.*

A variable appearing in a Presburger formula is free if it is not bound by any quantifier. We will call a Presburger formula "valid" if it does not have any free variable. If a formula ϕ contains $m \in \mathbb{N}$ free variables $x_1, \ldots, x_m$, then for a tuple $\vec{x} = (y_1, \ldots y_m) \in \mathbb{N}^m$ we note $\phi(\vec{x})$ the valid formula where every occurence of the free variable x_i is replaced by the integer y_i. Such a formula ϕ defines the set of integer tuples $S_\phi = \{\vec{x} \in \mathbb{N}^m \mid \phi(\vec{x})\}$, *i.e.* the set of m dimensional integer vectors ensuring the truth of a formula. Any set defined by a Presburger formula in that way is called semi-linear. It is known that every semi-linear set can be written [6] in the form:

$$\{\vec{x} \in \mathbb{N} \mid \bigvee_{i \in I} \exists k_1, \ldots k_{n_i} \in \mathbb{N}, \vec{x} = \vec{b_0^i} + \sum_{j=1}^{n_i} k_j \vec{b_j^i}\} \tag{1}$$

where I is a finite set and $\vec{b_j^i} \in \mathbb{N}^m$. Intuitively, these sets are a finite union of linear sets that consist in a starting vector, $\vec{b_0^i}$, and a finite number of periods, the $\vec{b_j^i}$.

Parametric Presburger arithmetic [1] extends traditionnal Presburger arithmetic by including one (or more) parameters that are treated like constants, and thus can be multiplied with variables but cannot be quantified.

Definition 2. *Given a set* VAR *of variables and* $\mathbb{P}$ *of parameters, a parametric Presburger term* t *follows the grammar*

$$t ::= \mathbb{N} \cup \mathbb{P} \mid \mathsf{VAR} \mid t + t \mid (\mathbb{N} \cup \mathbb{P}) * t.$$

A parametric Presburger formula then follows the same grammar as a traditionnal Presburger formula, albeit relying on parametric Presburger terms. Imitating Equation 1, one can define parametric semi-linear sets (pSl) as:

$$\begin{aligned} S(\vec{p}) = \{\vec{x} \in \mathbb{N}^m \mid \bigvee_{i \in I} \exists \vec{x_0}, \dots \vec{x_{n_i}} \in \mathbb{N}^m, & k_1, \dots k_{n_i} \in \mathbb{N}, \vec{x} = \sum_{j=0}^{n_i} \vec{x_j} \\ & \wedge \vec{b_0^i}(\vec{p}) \leq \vec{x_0} \leq \vec{c_0^i}(\vec{p}) \bigwedge_{j=1}^{n_i} k_j \vec{b_j^i}(\vec{p}) \leq \vec{x_j} \leq k_j \vec{c_j^i}(\vec{p})\} \end{aligned} \tag{2}$$

where I is a finite set, $\vec{p}$ is the vector of parameters and the b_j^i and c_j^i are vectors of polynomials with integer coefficients. Linear-parametric semi-linear sets (LpSl sets) are the restriction of parametric semi-linear sets to linear polynomials. We write (n, m)-pSl (resp. (n, m)-LpSl) for the pSl sets (resp. LpSl sets) of m dimensional vectors depending on a set of n parameters. Note that pSl sets are technically functions as the set they define depend on an argument (the set of parameters). We rely on the "set" terminology by analogy with traditional semi-linear sets.

As for traditionnal semi-linear sets, pSl sets can be interpreted as a finite union of sets, each of which consist in a starting point and a finite number of periods. An important difference however is that the starting point and periods are not a single integer vector but taken in an interval with bounds represented by polynomials in the parameter. Those intervals could easily be replaced by finitely many vectors in Presburger arithmetic, but this cannot be done here in the presence of a parameter. pSl sets are among the simplest one can define that go beyond the scope of Presburger arithmetic.

It is important to note that, while any set defined by a Presburger formula can be represented in the form of Equation 1, this property is lost with parametric Presburger. Consider the set $\{x \in \mathbb{N} \mid \neg\, \exists k, x = kp\}$. This set contains all the integers x that are not a multiple of p. Removing the negation through usual techniques, one obtain the equivalent formula $\bigvee_{i=1}^{p-1} \exists k, x = kp + i$. While this is a semi-linear set when p is a constant, it is not when p is a parameter as the disjunction does not range over a fixed finite set.

In order to simplify future notations, we call the formula

$$\begin{aligned} \exists \vec{x_0}, \dots \vec{x_{n_i}} \in \mathbb{N}^m, k_1, \dots k_{n_i} \in \mathbb{N}, & \vec{x} = \sum_{j=0}^{n_i} \vec{x_j} \wedge \vec{b_0}(\vec{p}) \leq \vec{x_0} \leq \vec{c_0}(\vec{p}) \\ & \bigwedge_{j=1}^{n_i} k_j \vec{b_j}(\vec{p}) \leq \vec{x_j} \leq k_j \vec{c_j}(\vec{p}) \end{aligned}$$

the formula associated to the couple $(\vec{b_0}, \vec{c_0})$ and the periods $(\vec{b_1}, \vec{c_1}), \dots (\vec{b_n}, \vec{c_n})$ and denote it $\Lambda((\vec{b_0}, \vec{c_0}) \mid (\vec{b_1}, \vec{c_1}), \dots (\vec{b_n}, \vec{c_n}))$. In this context, we will also write (a) for the couple (a, a). Also, we will say that a is an integer period if it is a constant polynomial, and in this case we will confuse the polynomial a and the constant it is equal to.

In this paper, we are interested in analysing pSl sets and in particular in detecting when two pSl sets are equal:

Definition 3. *Let $\mathbb{P}$ be a set of n parameters. Given two (n, m)-pSl sets S_1 and S_2, the (n, m)-pSl equality problem consists in computing the sets of parameter values $\vec{p} = (p_1, \dots, p_n) \in \mathbb{N}^n$ such that $S_1(\vec{p}) = S_2(\vec{p})$.*

The existential (n, m)-pSl equality problem consists in deciding whether there exists a set of parameter values such that S_1 and S_2 are equal and the universal (n, m)-pSl equality problem consists in deciding if the equality holds for every set of parameter values.

The above problems can also be defined for pSl sets.

Let S_1 and S_2 be two (n, m)-pSl sets and let ϕ_1 and ϕ_2 be the two parametric Presburger formulas that defines them. Then solving the existential (n, m)-pSl equality problem of S_1 and S_2 is equivalent to deciding whether the formula

$$\exists p_1, \dots p_n, \forall x_1, \dots, x_m, (\phi_1(p_1, \dots, p_n)(x_1, \dots, x_m) \leftrightarrow \phi_2(p_1, \dots, p_n)(x_1, \dots, x_m))$$

is true. If S_1 and S_2 are in linear-parametric, this formula can be written into the first order theory of the integers with divisibility. This theory is known to be decidable in some cases[2,7,9], for instance when the formula is purely existential. To our knowledge, the formula above goes beyond any known decidable fragment of this theory.

In the rest of this paper, we will focus on $(1, 1)$-LpSl sets.

3 A simple formula for $(1, 1)$-LpSl sets

Let S be a $(1, 1)$-LpSl sets and ϕ be the parametric Presburger formula describing it. We have the following simplification:

Lemma 1. *We can compute $N_0 \in \mathbb{N}$ and formulas $\psi_1, \dots, \psi_m$ such that for $p \geq N_0$, we have*

$$\{x \in \mathbb{N} \mid \phi(p)(x)\} = \{x \in \mathbb{N} \mid \bigvee_{i=1}^{m} \psi_i(p)(x)\}.$$

Moreover, there exists $f \in \mathbb{N}$ such that for all $i \leq m$ we have that $\psi_i = \Lambda((b_0^i, c_0^i) \mid (b_1^i), \dots (b_{n_i}^i))$ and satisfies one of the following condition:

First type: $b_0^i = c_0^i$ *and either every b_j^i with $j \geq 1$ is an integer period, or none are.*

Second type: $n_i = 2$, $c_0^i(p) - b_0^i(p) = dp + e$ *with $d > 0$ and there exists $g_1, g_2 \in \mathbb{N}$ integers such that $b_1^i = fp + g_1$ and $b_2^i = fp + g_2$.*

Proof. This proof consists in three simplification steps, each of which starts with a formula ϕ_0 associated to a couple (b_0, c_0) and periods $(b_1, c_1), \dots (b_n, c_n)$ and turns it into a disjunction of formulas closer to the form stated in the lemma. As ϕ is a disjunction of formulas of the above form and the simplifications can be applied independently on all of them, we will obtain the equivalent disjunction claimed in the lemma.

• Let ϕ_0 be a formula associated to a couple (b_0, c_0) and periods $(b_1, c_1), \dots (b_n, c_n)$. As a first step, we show how to build a formula ψ and an integer N_1 such that if $p \geq N_1$, ψ is equivalent to ϕ_0 and such that $\psi(p)(x) = \bigvee_{i \in I} \psi_i(p)(x)$ where every ψ_i is of the form $\Lambda(b_0^i, c_0^i) \mid (b_1^i), \dots (b_{n_i}^i))$ (thus limiting periods to singletons). The idea of the transformation here is to show that either these interval periods can be transformed into finitely many singleton periods or that after finitely many steps, every integer is accepted (which can be represented with a period (1).

We consider first the period (b_1, c_1). we denote $c_1(p) - b_1(p) = dp + e$. We consider two cases depending on d.

- if $d = 0$, then the formula obtained by replacing in ϕ_0 the period (b_1, c_1) by the finite number of periods $(b_1), (b_1 + 1) \dots (c_1)$ is trivially equivalent to ϕ_0 (*i.e.* ϕ_0 is equivalent to $\Lambda((b_0, c_0) \mid (b_2, c_2), \dots, (b_n, c_n), (b_1), (b_1{+}1) \dots (c_1)))$.
- If $d > 0$, denote $b_1(p) = f_1 p + g_1$, then setting $m_1 = 2(f_1 + g_1)$, we have that for all $p \geq \max(1, \frac{2e}{1-2d})$, $m_1(c_1(p) - b_1(p)) \geq m_1 p/2 = f_1 p + g_1 p \geq b_1(p)$. As a consequence, any integer greater than $b_0(p) + m_1 b_1(p)$ is accepted by ϕ_0. Indeed, let $y \geq b_0(p) + m_1 b_1(p)$ and k such that $k b_1(p) \leq y - b_0(p) \leq (k{+}1) b_1(p)$. Then as $k \geq m_1$, $k c_1(p) \geq (k{+}1) b_1(p)$ hence $k b_1(p) \leq y - b_0(p) \leq k c_1(p)$ and thus y is accepted by ϕ_0.
  We can thus rewrite ϕ_0 by taking into account how many times the period (b_1, c_1) was taken. More precisely, ϕ_0 is equivalent to

$$\Big(\bigvee_{i=1}^{m_j - 1} \Lambda((b_0 + i b_j, c_0 + i c_j) \mid (b_2, c_2), \dots, (b_n, c_n))\Big) \vee \Lambda((b_0 + m_j b_j) \mid (1)).$$

In either cases, the number of non-singleton period was reduced by 1. Thus repeating this process on every period eventually terminates and produces an equivalent formula where every period is a singleton. [1] Let $N_1 \in \mathbb{N}$ be such that the above process is correct for all $p \geq N_1$.

• We now consider a formula ϕ_0 associated to a couple (b_0, c_0) and periods $(b_1), \dots (b_n)$. Note quickly that if $c_0 - b_0 \in \mathbb{N}$, similarly to the previous reduction step, ϕ_0 is equivalent to

$$\bigvee_{i=b_0}^{c_0} \Lambda((i) \mid (b_1), \dots, (b_n)).$$

[1] Note that the naive approach described here could lead to an exponential blow-up in the size of the formula. We believe this blow-up could be avoided, but complexity is not the main focus of this paper.

We will thus assume here that $b_0 = c_0$ and show how to obtain a formula corresponding to the first type mentioned in the lemma, while in the next step we will consider the case where $c_0(p) - b_0(p) = dp + e$ with $d > 0$ and obtain a formula of the second type.

If there exists i, j such that b_i is an integer period and b_j is not. Then every time we would need to take the period (b_j) b_i times, we can instead take the period (b_i) $b_j(p)$ times. As such, the period (b_j) can be removed, by considering multiple starting vectors depending on how many times (b_j) was taken modulo b_i. In other words, and assuming $j = 1$ for simplicity, ϕ_0 is equivalent to

$$\bigvee_{k=0}^{b_i-1} \Lambda((b_0 + kb_1) \mid (b_2), \dots, (b_n)).$$

We thus removed one non-integer period, and by repeating this process we obtain a disjunction of formula, each of which satisfying the first type of the lemma.

• We finally consider a formula ϕ_0 associated to a couple (b_0, c_0) and periods $(b_1), \dots (b_n)$ and assume that $c_0(p) - b_0(p) = dp + e$ is a linear polynomial in p with $d > 0$. As in the case $d > 0$ of the first step, we will use this interval to simplify the formula. However, as the initial couple cannot be repeated contrary to the periods, the simplification is more limited.

For all $i = 1, \dots, n$, we denote $b_i(p) = f_i p + g_i$.

- if there exists $i \leq n$ such that $f_i = 0$, then if $dp + e \geq g_i$ (which can be ensured with p large enough), ϕ_0 is trivially equivalent to $\Lambda((b_0) \mid (1))$.
- Otherwise, one can assume that there exists $f > 0$ such that $f_i = f$. Indeed, let f be the lowest common multiple of the f_i. For all i, let $h_i \in \mathbb{N}$ be such that $f_i h_i = f$, then ϕ_0 can be decomposed in multiple formula depending on whether the period b_i is taken a number of time equal to some m_i modulo h_i and is thus equivalent to

$$\bigvee_{m_i < h_i, i=1,\dots,n} \Lambda((b_0 + \sum_{i=1}^{n} m_i b_i, c_0 + \sum_{i=1}^{n} m_i b_i) \mid (h_1 b_1), \dots, (h_n b_n)).$$

As $h_i b_i(p) = fp + h_i g_i$, this new formula has the desired shape.

The last step is to limit the formula to two periods. Assuming thanks to the above that for all i, $f_i = f > 0$, and ordering the periods so that $g_1 < g_2 < \dots < g_n$, then if $dp + e \geq g_n - g_1$ (which again is ensured when p is large enough), ϕ_0 is equivalent to the formula $\Lambda((b_0, c_0) \mid (b_1), (b_n))$. In other words, the formula ϕ_0 accepts the integers belonging to every interval $[b_0 + kb_1; c_0 + kb_n]$ for some $k \in \mathbb{N}$. Every removed period would fall within these intervals.

Let $N_2 \in \mathbb{N}$ such that for $p \geq N_2$ the conditions on p encountered in the above process are satisfied. Then denoting $N_0 = \max(N_1, N_2)$, we have that if $p \geq N_0$ the cascade of simplifications of this proof transform the initial formula in an equivalent disjunction of formulas as stated in the lemma. □

Unfortunately this simplification cannot be extended to higher number of dimensions or parameters. For instance, consider the two dimensional period $((1,1),(1,p))$ (*i.e.* the period increasing by 1 the first coordinate and by a value between 1 and p the second). Component wise, the first one is already a singleton while the second one can be used to accept every integer above 1. However, a relation exists between the two coordinates which stops us from turning the second component into a singleton.

4 Solving the $(1,1)$-pS-equality problem

We are interested in showing the $(1,1)$-pS-equality problem. As the current paper represents work in progress, we will focus on a simpler case first.

Let ϕ_1 and ϕ_2 be two parametric Presburger formulas describing two $(1,1)$-LpSl sets S_1 and S_2. We assume that these formulas were simplified thanks to Lemma 1. Let N_0 be the constant generated by this lemma beyond which the simplification holds. We say that such formulas are of type 1 if its subformulas are of type 1 as defined in Lemma 1.

Let us show the following result

Proposition 1. *If ϕ_1 and ϕ_2 are of type 1, then we can decide for which value of the parameter p $S_1(p) = S_2(p)$.*

Dans le cas formule de type 1, pointer vers le papier en commentaire Sauf erreur de ma part, la section 4 permet de traiter ces cas, même de façon plus générale (plusieurs paramètres). Moyen de traduire ces formules sans inégalités en presburger avec divisibilité ? L'élimination des quantificateurs sans inégalités passent mieux ? Non : la formule ici a une négation (le point n'appartient pas à la seconde formule). Ce qu'on ne peut pas faire. Donc pointer vers le papier en commentaire et expliquer pourquoi on ne peut pas l'utiliser.

Overview of the proof

Given a parametric Presburger formula ϕ describing a $(1,1)$-LpSl set with parameter p, we build a (non-parametric) two-dimensional semi-linear set L_ϕ that represents the integers that are accepted by the formula such that the term $(x,y) \in L_\phi$ corresponds to the integer $xp + y$.

We will show that, for large enough values of p (when the two coordinates do not can be considered separately), two $(1,1)$-LpSl sets differ iff their associated two-dimensional semi-linear sets differ. Therefore, in order to compute the parameter values such that the equality of two $(1,1)$-LpSl sets one can first test a finite number of initial values for p (until p is "large enough") then test the equality of the two semi-linear sets.

The exact construction of the set L_ϕ depends on some properties of p, however as there is finitely many options in those properties, we will be able to repeat the construction for each possibility. This result imply that the set of parameter

values for which the $(1,1)$-LpSl sets are equal is ultimately periodic. This is in agreement with the results of [1]. The contribution of this proposition can thus be seen as putting a bound on when the periodic behaviour starts.

Preliminary assumptions

Let ϕ be a parametric Presburger formula of type 1 describing a $(1,1)$-LpSl set with parameter p that has the simple form described in Lemma 1 and let N_0 be the constant produced by the lemma. More precisely, ϕ is of the form $\bigvee_{i\in I}\phi_i$ where for all i, ϕ_i is a formula of the first type as described in Lemma 1. For every formula ϕ_i, we will build a two dimensional semi-linear set L_{ϕ_i} and set $L_\phi = \cup_{i\in I} L_{\phi_i}$.

We focus on one sub-formula ϕ_i of ϕ which is of the form $\Lambda((b_0) \mid (b_1), \dots (b_n))$ as ϕ is assumed to be of type 1. We denote for all $j \leq n$, $b_j = f_j p + g_j$. If every f_j is equal to 0 (integer period case of type 1), assuming b_n is the largest period, we guess the set $G_i \subseteq \{(r,s) \in \mathbb{N}^2 \mid r \leq nb_n \wedge s \leq b_n\}$: $(r,s) \in G_i$ means that there exists $k_1, \dots, k_n$ such that $g_0 + \sum_{i=1}^n k_i b_i = rp + s$. In particular $(0, g_0) \in G_i$. There is finitely many such guesses where $r \leq nb_n$ and $s \leq b_n$, and thus the following construction and analysis can be made in each possible case. If the periods are not integers (integer period case of type 2), then let M be such that $f_j M \geq g_j$ for all $j = 0, \dots, n$.

Constructing the semi-linear set for a formula of type 1

If every period of ϕ_i is not an integer period, we build the following set

$$L_{\phi_i} = \{(x,y) \in \mathbb{N}^2 \mid \exists k_1, \dots, k_n, \\ (x,y) = (f_0, g_0) + \sum_{j=1}^{n} k_j (f_j, g_j) \\ \}.$$

Lemma 2. *Let $x, y \in \mathbb{N}$. If $(x,y) \in L_{\phi_i}$, we have that $\phi_i(p)(xp+y)$ is true for all $p \geq N_0$.*

Conversely, for $p \geq N_0$, if $\phi_i(p)(k)$ is true for some $k \in \mathbb{N}$, then there exists $x, y \in \mathbb{N}$ such that $k = xp + y$, $y \leq Mx$ and $(x,y) \in L_{\phi_i}$.

Proof. Let $x, y \in \mathbb{N}$ such that $(x,y) \in L_{\phi_i}$. By construction of L_{ϕ_i}, there exists $k_1, \dots, k_n \in \mathbb{N}$ such that

$$(x,y) = (f_0, g_0) + \sum_{j=1}^{n} k_j (f_j, g_j).$$

Starting from $f_0 p + g_0$ and adding k_j times the period $b_j(p)$ we obtain the value

$$f_0 p + g_0 + \sum_{j=1}^{n} k_j (f_j p + g_j) = xp + y.$$

Thus $\psi_i(p)(xp+y)$ is true.

Now let $p \geq N_0$ and $k \in \mathbb{N}$ such that $\phi_i(p)(k)$ holds. By definition of ϕ_i, there exists $k_1, \ldots, k_n, x, y \in \mathbb{N}$ such that $k = f_0 p + g_0 + \sum_{j=1}^{n} k_j(f_j p + g_j) = xp + y$. Thus by construction of L_{ϕ_i}, the point (x, y) is in L_{ϕ_i}. Moreover, as for all $j \geq 0$ we have $g_j \leq M f_j$, we have that $y \leq Nx$. □

Now if every period of ϕ_i is an integer period, we build a similar set where the periods can be taken many times at once to go above p and the initial point might be shifted by some periods to go above p as well:

$$L_{\phi_i} = \bigcup_{(r_0,s_0)\in L_i} \{(x,y) \in \mathbb{N}^2 \mid \exists k_1, \ldots, k_n, t_1, \ldots t_n \\ (x,y) = (f_0 + r_0, s_0) + \sum_{j=1}^{n} k_j(0, g_j) + \sum_{j=1}^{n} t_j(b_j, 0)\}.$$

Lemma 3. *Let $x, y \in \mathbb{N}$. If $(x, y) \in L_{\phi_i}$, we have that $\phi_i(p)(xp + y)$ is true for all $p \geq N_0$ satisfying G_i.*

Conversely, for $p \geq N_0$ satisfying G_i, if $\phi_i(p)(k)$ is true for some $k \in \mathbb{N}$, then for all $x, y \in \mathbb{N}$ such that $k = xp + y$ and $x \geq f_0$, we have $(x, y) \in L_{\phi_i}$.

Proof. Let $x, y \in \mathbb{N}$ such that $(x, y) \in L_{\phi_i}$ and $p \geq N_0$ satisfying G_i. By construction of L_{ϕ_i}, there exists $(r_0, s_0) \in G_i, k_1, \ldots, k_n, t_1, \ldots t_n \in \mathbb{N}$ such that

$$(x,y) = (f_0 + r_0, s_0) + \sum_{j=1}^{n} k_j(0, g_j) + \sum_{j=1}^{n} t_j(b_j, 0).$$

By definition of (r_0, s_0), there exists $c_1^0, \ldots c_n^0$ such that $g_0 + \sum_{l=1}^{n} c_l^0 b_l = r_0 p + s_0$.

Starting from $f_0 p + g_0$, and adding $k_j + c_j^0 + t_j p$ times the period b_j we obtain the value

$$f_0 p + g_0 + \sum_{j=1}^{n} (k_j + c_j^0 + t_j p) g_j$$
$$= (f_0 + r_0 + \sum_{j=1}^{n} t_j g_j) p + s_0 + \sum_{j=1}^{n} k_j g_j = xp + y.$$

Thus $\psi_i(p)(xp + y)$ is true.

Now let $p \geq N_0$ satisfying G_i and $k \in \mathbb{N}$ such that $\phi_i(p)(k)$ holds. Let $x, y \in \mathbb{N}$ such that $x \geq f_0$ and $k = xp + y$. By definition of ψ_i, there exists $k_1, \ldots, k_n \in \mathbb{N}$ such that $k = f_0 p + g_0 + \sum_{j=1}^{n} k_j g_j$. These k_j occurrences of the period g_j can be represented in L_{ϕ_i} by either taking periods in groups (through the $(b_j, 0)$ and G_i elements), or individually (by the period $(0, g_j)$). Formally, let us show how to reach (x, y) in L_{ϕ_i} by recurrence over $x - f_0$.

• if $x - f_0 = 0$, then $(x, y) = (f_0, g_0) + \sum_{j=1}^{n} k_j(0, g_j)$ which is in L_{ϕ_i} as $(0, g_0) \in G_i$.

• if $nb_n \geq x - f_0 > 0$, then let $k'_1, \ldots, k'_n$ be such that for all $j = 1, \ldots, n, k'_j \leq k_j$

and $g_0 + \sum_{j=1}^{n} k'_j g_j = rp + s$ where $(r, s) \in G_i$ and $r = x - f_0$. Such a pair exists, as we assume p satisfies G_i and the sum of the periods and g_0 is greater than $(x - f_0)p$. Let $x' = x - r$ and $y' = y - s$. Both are positive as $x' = f_0$ and $y' \geq g_0$. By construction, we have that $\phi_i(p)(x'p + y')$ holds and by recurrence we have that $(x', y') \in L_{\phi_i}$. From (x', y'), we can add (r, s) and reach (x, y) as $(r, s) \in G_i$ and the element of G_i that was previously used is $(0, g_0)$.
• if $x - f_0 > nb_n$, then as b_n is the largest period, there exists $j \leq n$ such that $k_j \geq p$. Then $\phi_i(p)((x - b_j)p + y)$ holds and by recurrence $(x - b_j, y) \in L_{\phi_i}$. By construction, from $(x - b_j, y)$ one can add $(b_j, 0)$ and reach (x, y). □

We set $L_\phi = \cup_{i \in I} L_{\phi_i}$.

Conclusion on Proposition 1

Let ϕ and ψ be two parametric Presburger formulas describing two $(1, 1)$-LpSl sets S_ϕ and S_ψ and let $N_0, N'_0 \in \mathbb{N}$ be given by Lemma 1 we have the following relation:

Lemma 4. *There exists N such that if $p > N$ and p satisfies the guesses G_i, $L_\phi = L_\psi$ iff $S_\phi(p) = S_\psi(p)$.*

Proof. Let $N_1 = \max(N_0, N'_0)$. Let H be the constant computed in Corollary 23 of [5] on L_ϕ and L_ψ[2]. We fix $N = \max(N_1, MH)$ and assume $p > N$ and p satisfies the guesses G_i.

Assume first that $L_\phi \neq L_\psi$. By Corollary 23 of [5], there exists a pair (x, y) that distinguishes the two sets (say, $(x, y) \in L_\phi \setminus L_\psi$) such that the modulus of x and y are smaller than H. As $p > N_0$, by Lemma 2 and Lemma 3 $\phi(xp+y)$ holds. Moreover assume per contradiction that $\phi(xp + y)$ holds. Then either Lemma 2 or Lemma 3 applies. In the first case, there exists x', y' with $x'p + y' = xp + y$, $y' \leq Mx'$ and $(x', y') \in L_\psi$. As $y \leq H$ and $p > MH$, we have $y \leq p$ and thus that $x' \leq x$. As a consequence, $y' \leq Mx' \leq Mx \leq MH \leq p$ and thus $x' \geq x$. Therefore $x = x'$ and thus $y = y'$ which raise a contradiction as $(x', y') \in L_\psi$.

Conversely, assume that $L_\phi = L_\psi$. Let $k \in \mathbb{N}$ such that $\phi(k)$ holds. Then by Lemma 2 and Lemma 3 there exists (x, y) such that $k = xp + y$ and $(x, y) \in L_\phi$. By assumption, we thus have that $(x, y) \in L_\psi$ and by Lemma 2 and Lemma 3 we have that $\psi(xp + y)$. As this holds as well by inverting ϕ and ψ we have that $S_\phi = S_\psi$. □

Proof (Proof of Proposition 1). Let S_1 and S_2 be two $(1, 1)$-LpSl sets described by two type 1 formulas ϕ_1 and ϕ_2.

The integers can be effectively finitely partitioned $\mathbb{N} = \bigcup_{i \in I} K_i$ based on which guesses G_i they satisfy. For any set K_i, applying Lemma 4, one obtains

[2] This Corollary states that if the two semi-linear sets are different, then there exist a vector with coefficients smaller than H that differentiates both sets. The value of H is exponential in the formulas and doubly exponential in the dimension (here 2). In particular, this results give a naïve process on how to distinguish two semi-linear sets.

a constant N_i such that for any $p \in K_i$, if $p > N_i$, then the equality of $S_1(p)$ and $S_2(p)$ can be deduced from the equality of two semi-linear sets. As testing equality of two semi-linear sets is decidable, we define $J \subseteq I$ the set of indices i of sets K_i such that the two semi-linear sets associated to K_i are equal. We also fix $N = \max_{i \in I} N_i$. Let K_0 be the set of integers p smaller than N for which $S_1(p)$ and $S_2(p)$. K_0 can be computed as there are finitely many integers to test and for a fixed p $S_1(p)$ and $S_2(p)$ are semi-linear sets. Denoting $J' = J \cup \{0\}$, we thus have that the set of parameters p such that $S_1(p) = S_2(p)$ is $\bigcup_{i \in J'} K_i$. □

5 Conclusion

In this document, we considered parametric Presburger arithmetic, an extension of Presburger arithmetic with parameters that are treated like constants. Imitating semi-linear sets defined in Presburger arithmetic, we introduced parametric semi-linear sets. While semi-linear sets are easy to manipulate, the same cannot be said for parametric semi-linear sets. We show how to test whether two such sets are equal under important assumptions: the sets must define a space of dimension 1 and the shape of the polynomials defining the sets are restricted, being linear among other things.

This document represents work in progress. The goal is to get a better understanding of when testing the equality of two semi-linear sets is possible before publication. In particular, the following results should be added in the future:

- Undecidability of the equality of two parametric semi-linear sets in the general case
- Removing the restriction to type 1 formula in the previous proof, mainly by introducing a third dimension in the semi-linear sets that are created
- Considering other simple classes of restrictions (when only the lower bound or the upper bound depends on the polynomials for instance).

References

1. T. Bogart, J. Goodrick, and K. Woods. Parametric presburger arithmetic: logic, combinatorics, and quasi-polynomial behavior. *Discrete Analysis*, 2017.
2. M. Bozga and R. Iosif. On decidability within the arithmetic of addition and divisibility. In *FoSSaCS'05*, pages 425–439. Springer, 2005.
3. V. Bruyère, E. Dall'Olio, and J.-F. Raskin. Durations and parametric model-checking in timed automata. *ACM Transactions on Computational Logic*, 9(2):12:1–12:23, 2008.
4. L. CERDA-ROMERO and C. MARTINEZ-RANERO. The diophantine problem for addition and divisibility over subrings of the rationals. *The Journal of Symbolic Logic*, 82(3):1140–1149, 2017.
5. D. Chistikov and C. Haase. The Taming of the Semi-Linear Set. In *ICALP'16*, volume 55 of *LIPIcs*, pages 128:1–128:13. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016.
6. S. Ginsburg and E. H. Spanier. Bounded algol-like languages. *Transactions of the American Mathematical Society*, 113(2):333–368, 1964.

7. C. Haase. A survival guide to presburger arithmetic. *ACM*, 5(3):67–82, jul 2018.
8. Jérôme Leroux. The general vector addition system reachability problem by presburger inductive invariants. *Log. Methods Comput. Sci.*, 6(3), 2010.
9. G. A. Pérez and Ritam R. Revisiting parameter synthesis for one-counter automata. In *CSL'22*, volume 216 of *LIPIcs*, pages 33:1–33:18. Schloss Dagstuhl, 2022.
10. J. Robinson. Definability and decision problems in arithmetic. *The Journal of Symbolic Logic*, 14(2):98–114, 1949.
11. P. Rümmer. A constraint sequent calculus for first-order logic with linear integer arithmetic. In *LPAR'08*, volume 5330 of *LNCS*, pages 274–289. Springer.
12. Alfred Tarski. A decision method for elementary algebra and geometry. In *Quantifier elimination and cylindrical algebraic decomposition*, pages 24–84. Springer, 1951.