When are two Parametric Semi-linear Sets Equal?

Engel Lefaucheux

When are two Parametric Semi-linear Sets Equal?

Engel Lefaucheux

Université de Lorraine, CNRS, Inria, LORIA

Abstract.

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 [15], Peano arithmetic is known to be undecidable [5]. 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 hence regain decidability.

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 [14] are both examples of automated theorem provers that base some of their functionalities 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 [10] 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 [13] as multiplication can be re-encoded with this predicate, was proven decidable when considering only the purely existential fragment [11,9]. This direction has been pushed more recently in [2] and [12] 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 [1,8] 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 [11,9].

Similarly to how semi-linear sets arise from Presburger arithmetic, sets can be obtained using parametric Presburger arithmetic. The goal of our work is to formally define and to study those parametric semi-linear sets. More precisely, semi-linear sets can be defined [4] by a finite union of linear sets, and linear sets are described by an initial vector and a set of period vectors that can be added to the initial vector as many times as desired. Parametric semi-linear sets

similarly are a finite union of parametric linear sets, which are each represented by an initial element and periodic elements to be added to it. The difference lies in the fact that the building blocks of parametric linear sets are vectors of intervals which borders are defined by polynomials in the parameters. For fixed values of the parameters, a parametric semi-linear set hence become a semi-linear set. Parametric semi-linear sets cannot be defined in Presburger arithmetic in general, and can in fact quickly become complex to analyse. This article focuses mainly on computing for which parameter values two given parametric semi-linear sets are equal.

In this article, we first formally define parametric Presburger arithmetic as well as parametric semi-linear sets (Section 2). Then, we show that parametric semi-linear sets are unfortunately "too expressive": deciding whether there exists parameter values such that two parametric semi-linear sets describe the same object is undecidable. In fact, this remains undecidable even with strong restrictions on the parametric semi-linear sets, such as requiring the polynomials to be linear (Section 3). Finally, we show that if the polynomials describing the sets are (1) linear, (2) depend on a single parameter and (3) span a single dimension, then one can decide the existence of (and compute the) parameter values achieving equality (Section 4)

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 formulas 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* $x \in \mathsf{VAR}$*.*

A variable appearing in a Presburger formula is free if it is not bound by any quantifier. If a formula ϕ contains $m \in \mathbb{N}$ free variables $x_1, \ldots, x_m$, then for a tuple $\vec{y} = (y_1, \ldots y_m) \in \mathbb{N}^m$ we note $\phi(\vec{y})$ the formula where every occurrence 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 ϕ. Any set defined by a Presburger formula this 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 traditional Presburger formula, albeit relying on parametric Presburger terms.
Imitating Equation 1, one can define parametric semi-linear sets (pSl sets) as functions associating a semi-linear set to parameter values[1] in the following way:

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

where I is a finite set, $\vec{p}$ is the vector of parameter values and the b_j^i and c_j^i are vectors of multi-variate polynomials with integer coefficients. Linear parametric semi-linear sets (LpSl sets) are pSl sets where in each dimension, the polynomials b_j^i and c_j^i are linear. We write (n, m)-pSl sets (resp. (n, m)-LpSl sets) for the pSl sets (resp. LpSl sets) of m dimensional vectors depending on a set of n parameters.

As for traditionnal semi-linear sets, pSl sets can be interpreted as a finite union of sets, each of which consists 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 parameters. Those intervals could easily be replaced by finitely many vectors in Presburger arithmetic, but this transformation fails in the presence of a parameter.

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 $\{x \in \mathbb{N} \mid \exists k, k \leq p \wedge x = 2k\}$. This set contains all the even integers smaller than $2p$, which can be represented by the formula $\bigvee_{i=1}^{p} x = 2i$. 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 fact, this set cannot be represented as a pSL set, as, being finite, no period can be repeated, and a pSl set has finitely many starting intervals.

[1] Hence, pSl sets are not technically sets. We still rely on the "set" terminology by analogy with semi-linear sets.

In order to simplify future notations, we call the formula

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

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 period a is an integer period if a is a constant polynomial, and in this case we will confuse the polynomial a and the constant it is equal to. Given such a formula ϕ, we will write S_ϕ for the pSl set implied by this formula.

In this paper, we are interested in analysing pSl sets and in particular in deciding when two pSl sets can be made equal:

Definition 3. *Given two (n, m)-pSl sets S_1 and S_2, the (n, m)-pSl equality problem consists in computing the set 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 parameter values $\vec{p}$ such that $S_1(\vec{p}) = S_2(\vec{p})$.

These two problems can also be defined for (n, m)-LpSl 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. This formula can be rewritten into the first order theory of the integers with divisibility. This theory is known to be decidable in some cases [2,7,12], for instance when the formula is purely existential. To our knowledge, this formula goes beyond any known decidable fragment of this theory, even with a single parameter and linear polynomials.

Example 1. Consider the two pSl sets $S_{\Lambda((4)\mid(2))}$ and $S_{\Lambda((p_1^3-p_2, 2p_3)\mid(p_2-p_3))}$. The first set contains every even integer greater or equal to 4, while the second contains every integer which remains modulo $p_2 - p_3$ lies between $p_1^3 - p_2$ and $2p_3$. In particular, the second set contains the interval $[p_1^3 - p_2, 2p_3]$, for both sets to be equal, and as the first one does not contain consecutive integers, we have that $p_1^3 - p_2 = 2p_3$. Moreover, the smallest integer accepted by each set are respectively 4 and $p_1^3 - p_2$. Hence, equality of the sets imply that $p_1^3 - p_2 = 4$. Combining those two equations, we have that equality of the sets imply $p_3 = 2$. In this example, for equality to hold, the periods of both sets should be equal, so we need $p_2 - p_3 = 2$, hence $p_2 = 4$. Therefore, the only parameter values achieving equality of the two pSl sets are $p_1 = 2, p_2 = 4$ and $p_3 = 2$.

3 pSl sets are too expressive

pSl sets go beyond Presburger arithmetic, but fails to encompass all of Peano arithmetic. They thus fall within an area of arithmetic where the border of decidability and undecidability is unclear. However, pSl sets lie in general are on the wrong side of the border, even in a single dimension, at least when it comes to testing equality of sets.

3.1 Testing equality of pSl sets is undecidable

Given a polynomial P with n variables and coefficients in $\mathbb{Z}$, Hilbert's tenth problem consists in deciding whether there exists integer values $x_1, \ldots, x_n$ such that $P(x_1, \ldots, x_n) = 0$. This problem is famously known to be undecidable [5]. In this section, we show by reduction to Hilbert's tenth problem that the existential $(n, 1)$-pSl equality problem is undecidable. This is not entirely surprising as, in the general case, pSl sets involve multi-variate polynomials over the parameters with coefficients and valuations in $\mathbb{N}$.

Theorem 1. *The existential $(n, 1)$-pSl equality problem is undecidable.*

Proof. Let P be a polynomial with n variables $x_1, \ldots, x_n$ and coefficients in $\mathbb{Z}$. We build the polynomial Q with $2n$ variables $y_1, \ldots y_{2n}$ where for $i \leq n$ every occurrence of x_i in P was replaced by $y_{2i-1} - y_{2i}$. It is clear that P has a zero in $\mathbb{Z}^n$ iff Q has a zero in $\mathbb{N}^{2n}$.

Q is of the form $\sum_{i=1}^{m} a_i g_i$ where the g_i are monomials with coefficient 1 and $a_i \in \mathbb{Z}$. We reorder the terms if needed so that there exists $k \in \mathbb{N}$ such that $a_i \geq 0$ iff $i \leq k$. We build the $(n, 1)$-pSl sets S_1 and S_2 where $S_1 = S_{\Lambda((\sum_{i=1}^{k} a_i g_i)|)}$. and $S_2 = S_{\Lambda((\sum_{i=1}^{k} a_i g_i, \sum_{i=k+1}^{n} -a_i g_i)|)}$. Let us show that given $y_1, \ldots y_{2n} \in \mathbb{N}$, $Q(y_1, \ldots y_{2n}) = 0$ iff $S_1(y_1, \ldots y_{2n}) = S_2(y_1, \ldots y_{2n})$.

For a chosen set of parameter valuations $y_1, \ldots y_{2n} \in \mathbb{N}$ S_1 contains a single element: $\sum_{i=1}^{k} a_i g_i(y_1, \ldots y_{2n})$. On the contrary, S_2 contains an interval of integers ranging from $\sum_{i=1}^{k} a_i g_i(y_1, \ldots y_{2n})$ to $\sum_{i=k+1}^{n} -a_i g_i(y_1, \ldots y_{2n})$. This interval is equal to $\sum_{i=1}^{k} a_i g_i(y_1, \ldots y_{2n})$ iff $\sum_{i=1}^{k} a_i g_i(y_1, \ldots y_{2n}) = \sum_{i=k+1}^{n} -a_i g_i(y_1, \ldots y_{2n})$. In other words, $S_1(y_1, \ldots y_{2n}) = S_2(y_1, \ldots y_{2n})$ iff $\sum_{i=1}^{m} a_i g_i(y_1, \ldots y_{2n}) = 0$ and thus iff $Q(y_1, \ldots y_{2n}) = 0$.

Therefore there exists parameter valuations achieving equality of S_1 and S_2 iff the Hilbert's tenth problem for P has a solution. □

3.2 Testing equality of LpSl sets is undecidable

Due to Theorem 1, some restrictions must be added to pSl sets in order to regain decidability. The two main options are to limit the number of parameters allowed, or to restrict the form of the polynomials involved. We first consider the latter, and show that the problem remains undecidable for $(n, 1)$-LpSl sets.

To do so, we will reduce Hilbert's tenth problem again. Though we need to first modify the statement slightly thanks to the following result which decompose the given polynomial into a conjunction of multiple simpler polynomials.

Lemma 1. *Let P be a multi-variate polynomial with variables ranging in $\mathbb{N}$ and coefficients in $\mathbb{Z}$. Then one can compute polynomials Q_i that are of one of the four following forms (1) $uv \pm z$, (2) $u \pm v \pm z$, (3) $u + kv$ or (4) u where $k \in \mathbb{Z}$ and u, v, z are variables such that there exists $\vec{x}$ with $P(\vec{x}) = 0$ iff there exists $\vec{y}$ with $\bigwedge_i Q_i(\vec{y}) = 0$.*

This lemma directly follows from the classic idea of introducing additional variables which represent addition or multiplication of variables in order to put the entire initial polynomial within a single variable. For instance, to represent the polynomial $P = x_1x_2x_3 + 4x_4$, we introduce four new variables u, v, w and z and the polynomials $Q_1 = x_1x_2 - u$, $Q_2 = ux_3 - v$, $Q_3 = w - 4x_4$, $Q_4 = v + w - z$ and $Q_5 = z$. If Q_1 to Q_4 are equal to 0, we have that $u = x_1x_2$, $v = ux_3 = x_1x_2x_3$, $w = 4x_4$ and $z = v + w = x_1x_2x_3 + 4x_4$. Thus, there exists values of $x_1, x_2, x_3, x_4, u, v, w$ satisfying the above relations and setting z to 0 iff there exist values of x_1, x_2, x_3, x_4 setting P to 0.

Theorem 2. *The existential $(n, 1)$-LpSl equality problem is undecidable.*

For pedagogical reasons, we first show that the existential (n, m)-LpSl equality problem is undecidable. The proof of Theorem 2 follows the same structure, though is more involved in order to gather everything on a single dimension.

app:undecann

Proposition 1. *The existential (n, m)-LpSl equality problem is undecidable.*

Proof. As we are only allowed linear polynomials, we cannot directly apply the simple reduction of Theorem 1. Moreover, the decomposition given by Lemma 1 still contains quadratic polynomials. We will show that through the use of periods and "sliding windows", we can generate tests on quadratic polynomials. Moreover, as we currently allow multiple dimensions, each of the tests can be restricted to its own dimension.

Indeed, note that we can trivially separate the different dimensions in a (n, m)-LpSl set: in the definition of Eq. (2), for a given dimension k and set $J \subseteq I$, if we force that every vector associated to an element of J is equal to 0 outside k, and conversely every vector associated to an element of $I \setminus J$ is 0 on k, the vectors included in the set either have a 0 in the k'th dimension, or have a 0 everywhere else. This allows us to be sure each of the tests we need to do can be confined to their dimension and thus will not affect the other tests. This separation is what we will lack, when tackling the $(n, 1)$-LpSl equality problem.

Let P, a multi-variate polynomial with variables $\vec{x}$ ranging in $\mathbb{N}$ and coefficients in $\mathbb{Z}$, be an instance of the Hilbert's tenth problem. Through Lemma 1, we decompose P into the simplified polynomials Q_i over $\vec{y}$. As stated in the lemma, there is a solution to $P(\vec{x}) = 0$ iff there is a valuation of the variables $\vec{y}$ zeroing every Q_i. We define two (n, m)-LpSl sets S_1 and S_2 such that for each polynomial Q_i, we dedicate one or more dimensions to testing whether the variables (converted into parameters of the sets) ensure zeroness of Q_i.

The parameters of S_1 and S_2 is the set $\vec{p}$ which includes the variables appearing in $\vec{y}$. We thus overload the notations by denoting a variable and the

corresponding parameter by the same name. For each Q_i, we detail the construction of the dimensions and parameters associated to this polynomial in S_1 and S_2.

• Case $Q_i(\vec{y}) = u$ with $u \in \vec{y}$. We dedicate a dimension of the system to Q_i and, on this dimension, we define S_1 to be the set $\{0\}$ and S_2 to be the set $\{0, u\}$. Hence, on this dimension, $S_1(\vec{p}) = S_2(\vec{p})$ iff $Q_i(\vec{y}) = u = 0$.

• Case $Q_i(\vec{y}) = u + kv$ with $u, v \in \vec{y}$ and $k \in \mathbb{Z}$. We dedicate a dimension of the system to Q_i and, on this dimension, if $k \leq 0$ $S_1 = \{0, u\}$ and $S_2 = \{0, -kv\}$ and otherwise $S_1 = \{0\}$ and $S_2 = \{0, u + kv\}$. Hence, we trivially have that on this dimension, $S_1(\vec{p}) = S_2(\vec{p})$ iff $u = -kv$.

• Case $Q_i(\vec{y}) = u \pm v \pm z$ with $u, v, z \in \vec{y}$. We partition $\{u, v, z\}$ into I_+ and I_- depending on whether their coefficient in Q_i is positive or negative (in particular, $u \in I_+$). We then dedicate a dimension of the system to Q_i and, on this dimension, $S_1 = \{0, \sum_{w \in I_+} w\}$ and $S_2 = \{0, \sum_{w \in I_-} w\}$. Again, we immediately have that on this dimension, $S_1(\vec{p}) = S_2(\vec{p})$ iff $\sum_{w \in I_+} w = \sum_{w \in I_-} w$, thus iff $u \pm v \pm z = 0$.

• Case $Q_i(\vec{y}) = uv \pm z$ with $u, v, z \in \vec{y}$. We introduce three new parameters k_i, t_i, w_i and dedicate four dimensions to this formula. k_i will be a high value, used to separate different parts of the test, w_i will be made equal to uv and t_i is used to correct a surplus unfortunately added with the current approach.

In the first dimension, S_1 is the set defined by the formula

$$\Lambda((0, k_i)|) \vee \Lambda((0)|(v + u)) \vee \Lambda((0)|(v + u + 1))$$

and S_2 is the set defined by the formula

$$\Lambda((0, k_i)|) \vee \Lambda((k_i)|(v + u)) \vee \Lambda((k_i + v + u)|(v + u + 1)).$$

Here, on the first k_i values, the sets are equal due to the first part of the formula. Then, for higher integers, due to the use of different periods, the points generated by the second parts of the formulas must be eventually identical, and the same holds for the third parts of the formulas. As a consequence, for the second parts to be equal, k_i must be a multiple of $v + u$, and, for the third part to be equal, $k_i + v + u$ must be a multiple of $v + u + 1$. The smallest value for k_i satisfying these two conditions is $(v+u)^2$. Thus, we can deduce that $S_1(\vec{p}) = S_2(\vec{p})$ implies $k_i \geq (v + u)^2$.

On the second dimension, consider the formulas:

$$\begin{aligned}
\phi_1 &= \Lambda((0) \mid (u + k_i)) \\
\phi_2 &= \Lambda((0, v - 1) \mid (u + k_i - 1)) \\
\phi_3 &= \Lambda((v + 1) \mid (u + k_i - 1, u + k_i)) \\
\phi_4 &= \Lambda((w_i + t_i) \mid)
\end{aligned}$$

S_1 and S_2 are the sets defined respectively by the formulas $\phi_1 \vee \phi_2 \vee \phi_3$ and $\phi_2 \vee \phi_3 \vee \phi_4$.

Initially, the integers accepted by ϕ_1 are also accepted by ϕ_2. However, as the period of ϕ_1 is greater by 1 compared to the period of ϕ_2, after exactly v steps ϕ_1

exits ϕ_2, *i.e.* the number $v(u+k_i)$, is accepted by ϕ_1 but not by ϕ_2 (and neither by ϕ_3). Indeed, note that as $k_i \geq (v+u)^2$, for all $j \leq v$, the points accepted by ϕ_1, ϕ_2 or ϕ_3 after j period lie in the interval $[jk_i; (j+1)k_i]$, thus there is no possible interference between the sets accepted with different number of periods used (at least when the number of periods remain low) and in particular $v(u+k_i)$ cannot be accepted by ϕ_2 or ϕ_3 after $v-1$ or $v+1$ periods.

Moreover, from $v+1$ period added onward (*i.e.* with the number $(v+1)(u+k_i)$), the numbers accepted by ϕ_1 are also accepted by ϕ_3. Thus ϕ_1 only adds to S_1 the number $v(u+k_i)$. For S_1 and S_2 to be equal, ϕ_4 must add to S_2 the very same number, therefore $S_1(\vec{p}) = S_2(\vec{p})$ on this dimension and the previous one iff $w_i + t_i = v(u+k_i)$.

We use the third dimension to ensure $t_i = vk_i$ so that $w_i = vu$ as announced. The method is similar to what we did in the previous dimension but with a smaller period:

$$\begin{aligned}
\phi_1 &= \Lambda((0) \mid (k_i)) \\
\phi_2 &= \Lambda((0, v-1) \mid (k_i - 1)) \\
\phi_3 &= \Lambda((v+1) \mid (k_i - 1, k_i)) \\
\phi_4 &= \Lambda((t-i) \mid)
\end{aligned}$$

and S_1 and S_2 correspond respectively to the formulas $\phi_1 \vee \phi_2 \vee \phi_3$ and $\phi_2 \vee \phi_3 \vee \phi_4$. For similar reasons to before, we have that $S_1(\vec{p}) = S_2(\vec{p})$ on this dimension and the previous two iff $t_i = vk_i$ and thus $w_i = vu$.

The only thing left is to build a dimension (the fourth) that ensures $w_i \pm z = 0$, which can be done with the same method as for the case where $Q_i(\vec{y}) = u \pm v \pm z$. And thus, $S_1(\vec{p}) = S_2(\vec{p})$ on those four dimension iff $Q_i(\vec{y}) = 0$.

By the construction above, we have that there exists parameters $\vec{p}$ such that $S_1(\vec{p}) = S_2(\vec{p})$ (on every dimension) iff there is a valuation such that for all i, $Q_i(\vec{y}) = 0$. By Lemma1 and the undecidability of Hilbert's tenth problem, this implies that the existential (n, m)-LpSl equality problem is undecidable. □

4 Solving the (1, 1)-LpSl-equality problem

Due to the undecidability results established in the previous section, we focus on $(1, 1)$-LpSl sets. We have:

Theorem 3. *Given two $(1, 1)$-LpSl sets S_1 and S_2, one can compute the set of parameter values p such that $S_1(p) = S_2(p)$.*

4.1 Overview of the proof

Given a parametric Presburger formula ϕ describing a $(1, 1)$-LpSl set with parameter p, our first step, in Section 4.2, is to build a new formula ϕ' describing the same set, but enjoying some additional properties. In particular, every

period appearing in ϕ' is a singleton, and if the initial couple is not a singleton, it is associated to at most two periods that have the same coefficient on the parameter.

In a second time, in Section 4.3, thanks to the previously made simplifications of ϕ, we build a (non-parametric) two-dimensional semi-linear set $L_{\phi'}$ that represents the integers that are accepted by the formula such that the term $(x, y) \in L_\phi$ corresponds to the integer $xp + y$.

This representation is not entirely correct, however we will show that, for large enough values of p (when the two coordinates can be considered separately for large enough values of the coordinates), 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 (such as its remainder modulo some constants of the formula), however as there is finitely many options in those properties, we will be able to repeat the construction for each possibility. This result implies 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]. Comparing to [1], the contribution of this section can thus be interpreted as putting a bound on when the periodic behaviour starts.

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

Let S be a $(1, 1)$-LpSl sets and ϕ be the associated parametric Presburger formula.

We have the following simplification:

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

The proof is postponed to Appendix B.

4.3 Solving a simple case of the $(1, 1)$-LpSl-equality problem

We will say that a parametric Presburger formula is of the first type if, after applying the simplification of Lemma 2, all of its subformulas are of the first type.

For pedagogy, as the LpSl sets defined by formulas of first type are easier to handle, we start by proving a restricted version of Theorem 3.

Proposition 2. *Given two parametric Presburger formulas ϕ and ψ of first type describing $(1,1)$-LpSl sets S_ϕ and S_ψ, we can compute the set of values of the parameter p such that $S_\phi(p) = S_\psi(p)$.*

The general case is postponed to appendix C.

Preliminary assumptions

Let ϕ be a parametric Presburger formula of first type describing a $(1,1)$-LpSl set with parameter p that has the simple form described in Lemma 2 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 2. For every formula ϕ_i, we will build a two dimensional semi-linear set L_{ϕ_i} and set $L_\phi = \cup_{i\in I} L_{\phi_i}$.

Considering one sub-formula ϕ_i which is thus of the form $\Lambda((b_0) \mid (b_1), \dots (b_n))$. We denote for all $j \leq n$, $b_j = f_j p + g_j$. By definition of formulas of the first type, either for all $1 \leq j \leq n$, $f_j = 0$, or for all $1 \leq j \leq n$, $f_j \neq 0$.

• If for all $1 \leq j \leq n$, $f_j = 0$. We order the periods so that g_n is the largest integer. Then we guess the set $G_i = \{(r,s) \in \mathbb{N}^2 \mid r \leq n g_n \wedge s \leq g_n \wedge \exists k_1, \dots, k_n \in \mathbb{N}, g_0 + \sum_{i=1}^n k_i g_i = rp + s\}$. In particular $(0, g_0) \in G_i$.

Thanks to the conditions that $r \leq n b_n$ and $s \leq b_n$, there is finitely many different guesses that can be made for the set G_i. The construction (and analysis) which follows apply for every parameter p for which the guess is correct. Hence, our guess produces a partition of $\mathbb{N}$ and we study the equality of sets separately for each element of this partition. Note that the set of integers corresponding to a specific guess G_i can be defined in (traditional) Presburger arithmetic and can thus be represented by a semi-linear set.

• If for all $1 \leq j \leq n$, $f_j \neq 0$. We set M_i to be such that $f_j M_i \geq g_j$ for all $j = 0, \dots, n$.

Constructing a semi-linear set representing ϕ_i

Intuitively, we will build a two dimensional semi-linear set L_{ϕ_i} such that $(x,y) \in L_{\phi_i}$ if $xp + y$ is accepted by ϕ_i, and such that the reverse implication partially holds. We separate the construction, depending on whether the periods of ϕ_i are integer periods.

• If every period of ϕ_i is not an integer period, we build the semi-linear 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 3. *Let $x, y \in \mathbb{N}$. If $(x, y) \in L_{\phi_i}$, $\phi_i(p)(xp + y)$ is true.*

Conversely, if $\phi_i(p)(k)$ is true for some $k \in \mathbb{N}$, then there exists $x, y \in \mathbb{N}$ such that $k = xp + y$, $y \leq M_i x$ 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 $k_1, \ldots, k_n \in \mathbb{N}$ 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 $\phi_i(p)(xp + y)$ 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 $g_j \leq M_i f_j$, we have that $y \leq M_i x$. □

• Conversely, if every period of ϕ_i is an integer period, we build a semi-linear set where the guess G_i is used to allow periods to be taken many times at once, producing a number greater than p and thus increasing the first component of our semi-linear set:

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

Lemma 4. *Let $x, y \in \mathbb{N}$. If $(x, y) \in L_{\phi_i}$, $\phi_i(p)(xp+y)$ is true for all p satisfying the guess G_i.*

Conversely, for p 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 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

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

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

Starting from f_0p+g_0, and adding $k_j+c_j^0+t_jp$ times the period g_j we obtain the value

$$\begin{aligned} &f_0p + g_0 + \sum_{j=1}^{n}(k_j + c_j^0 + t_jp)g_j \\ =&(f_0 + r_0 + \sum_{j=1}^{n} t_jg_j)p + s_0 + \sum_{j=1}^{n} k_jg_j = xp + y. \end{aligned}$$

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

Now let p 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 ϕ_i, there exists $k_1, \ldots, k_n \in \mathbb{N}$ such that $k = f_0p + g_0 + \sum_{j=1}^{n} k_jg_j$. 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 $ng_n \geq x - f_0 > 0$, then let $k'_1, \ldots, k'_n$ be such that for all $j = 1, \ldots, n, k'_j \leq k_j$ and $g_0 + \sum_{j=1}^{n} k'_jg_j = rp + s$ where $r = x - f_0$ and $(r, s) \in G_i$. Such a pair exists, as we assumed p satisfies G_i and the sum of the periods and g_0 is greater or equal to $(x - f_0)p$. Let $x' = x - r$ and $y' = y - s$. Both are positive as $x' = f_0$ and $y' \geq g_0$. As $x' = f_0$, we can reuse the first point and have that $(x, y) = (r, s) + (x', y') = (r, s) + (f_0, g_0) + \sum_{j=1}^{n}(k_j - k'_j)(0, g_j) \in L_{\phi_i}$
- if $x - f_0 > ng_n$, then as g_n is the largest period, there exists $j \leq n$ such that $k_j \geq p$. Then $\phi_i(p)((x - g_j)p + y)$ holds and by recurrence $(x - g_j, y) \in L_{\phi_i}$. By construction, from $(x - g_j, y)$ one can add $(g_j, 0)$ and reach (x, y), while remaining in L_{ϕ_i}. □

As mentioned earlier, we set $L_\phi = \cup_{i \in I} L_{\phi_i}$.

Conclusion on Proposition 2

Let ϕ and ψ be two parametric Presburger formulas describing two $(1, 1)$-LpSl sets S_ϕ and S_ψ and let $N_\phi, N_\psi \in \mathbb{N}$, and $\bigvee_{i \in I_\phi} \phi_i$ and $\bigvee_{i \in I_\psi} \phi_i$ be given by applying Lemma 2 on ϕ and ψ. Assuming every subformula is of first type, we have the following relation:

Lemma 5. *There exists N such that if $p > N$ and p satisfies all the guesses G_i for $i \in I_\phi \cup I_\psi$, then $L_\phi = L_\psi$ iff $S_\phi(p) = S_\psi(p)$.*

Proof. Let $N_1 = \max(N_\phi, N_\psi)$, $M = \max_{i \in I_\phi \cup I_\psi} M_i$. Let H be the constant computed in Corollary 23 of [4] on L_ϕ and L_ψ[2]. We fix $N = \max(N_1, MH)$ and assume that $p > N$ and that p satisfies the guesses G_i for $i \in I_\phi \cup I_\psi$.

[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

Assume first that $L_\phi \neq L_\psi$. By Corollary 23 of [4], 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 $p > N_\phi$, by Lemma 2, Lemma 3 and Lemma 4 $\phi(xp+y)$ holds. Moreover assume per contradiction that $\psi(xp+y)$ holds. Then, as $p > N_\psi$ either Lemma 3 or Lemma 4 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 $(x', y') \in L_\psi$ and $(x, y) \notin L_\psi$. In the second case, there exists x', y' with $x'p + y' = xp + y$, and as $y \leq p$, $x' \leq x$. By Lemma 4, as $x'p + y' = xp + y$, $x' \leq x$ and $(x', y') \in L_\psi$, we also have $(x, y) \in L_\psi$, which raises a contradiction.

Conversely, assume that $L_\phi = L_\psi$. Let $k \in \mathbb{N}$ such that $\phi(p)(k)$ holds. Then by Lemma 3 and Lemma 4 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 3 and Lemma 4 we have that $\psi(p)(xp + y)$. As this holds as well by inverting ϕ and ψ we have that $S_\phi = S_\psi$. □

Proof (Proof of Proposition 2). Let S_ϕ and S_ψ be two $(1, 1)$-LpSl sets described by two first type formulas ϕ and ψ.

The integers can be effectively finitely partitioned $\mathbb{N} = \bigcup_{i \in I} K_i$ based on which set of guesses G_i they satisfy. For any set K_i, applying Lemma 5, one obtains a constant N_i such that for any $p \in K_i$, if $p > N_i$, then the equality of $S_\phi(p)$ and $S_\psi(p)$ 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 built for 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 $S_\phi(p)$ and $S_\psi(p)$. The subset K'_0 of parameter values from K_0 achieving equality of the two $(1, 1)$-LpSl sets can be computed as there are finitely many integers within K_0 and for a fixed value of p $S_\phi(p)$ and $S_\psi(p)$ are semi-linear sets. We thus have that the set of parameter values of p such that $S_\phi(p) = S_\psi(p)$ is $K'_0 \bigcup_{i \in J} K_i$. □

5 Conclusion

In this article, 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 their parametric counterparts. In particular, we show that testing the equality of two pSl sets is undecidable in general. More positively, we show how to test whether two pSl sets are equal under important assumptions: the sets must

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 by testing every vector below this threshold.

define a space of dimension 1, depend on a single parameter and the polynomials defining the sets must be linear.

This line of research was initially motivated by the occurrence of pSl sets within an ongoing work on the opacity of parametric timed automata. We are therefore interested in discovering other problems for which parametric Presburger arithmetic would be an efficient tool. Through those applications, we can identify which classes of pSl sets needs to be studied and for which problems beyond equality. In particular, parametric timed automata motivate the study of upper and lower pSl sets, that is pSl sets where parameters are divided into "upper" and "lower" parameters and such that for any couple (b, c) of the formula defining the set, b depends only on the lower parameters, while c depends only on the upper parameters. Other natural extensions include removing the linear restriction for polynomials over a single parameter, and considering linear polynomials with a small number of parameters.

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. 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.
5. Martin Davis. Hilbert's tenth problem is unsolvable. *The American Mathematical Monthly*, 80(3):233–269, 1973.
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. A. Lasaruk and T. Sturm. Weak quantifier elimination for the full linear theory of the integers. *Appl. Algebra Eng. Commun. Comput.*, 18(6):545–574, 2007.
9. Antonia Lechner, Joël Ouaknine, and James Worrell. On the complexity of linear arithmetic with divisibility. In *30th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS'15*, pages 667–676. IEEE Computer Society, 2015.
10. Jérôme Leroux. The general vector addition system reachability problem by presburger inductive invariants. *Log. Methods Comput. Sci.*, 6(3), 2010.
11. L. Lipshitz. The diophantine problem for addition and divisibility. *Transactions of the American Mathematical Society*, 235:271–283, 1978.
12. 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.
13. J. Robinson. Definability and decision problems in arithmetic. *The Journal of Symbolic Logic*, 14(2):98–114, 1949.
14. 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, 2008.

15. Alfred Tarski. A decision method for elementary algebra and geometry. In *Quantifier elimination and cylindrical algebraic decomposition*, pages 24–84. Springer, 1951.

A Proof of Theorem 2

Compared to the proof of Proposition 1, we now wish to remove the use of the extra dimensions. To do so, the idea is to expand on the use of the k_i parameter that was needed to distance each period in the last case of the previous proof so that every test can be done in parallel, each occurring within their own k_i interval.

Note first that the tests we built for polynomials of the form $Q_i(\vec{y}) = u$, $Q_i(\vec{y}) = u + kv$ or $Q_i(\vec{y}) = u \pm v \pm z$ (the linear polynomials) do not rely on periodical behaviour. As such, we can realise them at the start without impacting any of the other tests. For example, in order to implement the test $Q_i(\vec{y}) = u$ we build the sets S_1 and S_2 such that S_1 includes the set $\{0\}$ and S_2 includes the set $\{0, u\}$ (as was done to ensure that $u = 0$ in the previous proof), then shift every other test by $u + 1$. This way, the formulas created for zeroing Q_i are the only ones adding integers within the interval $[0, u]$, moreover the formulas created for the test of Q_i do not produce any integer farther than u, ensuring independence of each test. Doing this for every linear polynomial, we can build M_0 such that every linear test is done independently within the M_1 first integers.

We now have to handle the tests of the form $Q_i(\vec{y}) = uv \pm z$. Those tests relying on periodical behaviour, it is harder to separate them. Let $P_1, \dots P_n$ be the polynomials of this form. Imitating what we did in the undecidability proof of the (n, m)-LpSl equality problem, we will build a k large enough that can be used to keep each test within intervals of length k so that interactions among tests only occur after the properties we seek have been ensured.

To zero one of the $P_i(\vec{y}) = uv \pm z$, when relying on multiple dimensions, we needed to build k_i such that $k_i \geq (u + v)^2$. Let K be the sum of all variables appearing in the P_i. By ensuring $k \geq (12K)^2$, we will have that $k \geq (u+v)^2$ for each P_i. This is achieved in the same way as before, with an initial shift of M_1: we add to S_1 the set defined by the formula

$$\Lambda((M_1, M_1 + k)|) \vee \Lambda((M_1)|(12K)) \vee \Lambda((M_1)|(12K + 1))$$

and to S_2 the set defined by the formula

$$\Lambda((M_1, M_1 + k)|) \vee \Lambda((M_1 + k)|(12K)) \vee \Lambda((M_1 + k + 12K)|(12K + 1)).$$

which as shown before, as there is no interference from the previous tests, ensures that $k \geq (12K)^2$.

Let $M_2 = M_1 + k + 12K + 2$. As the first two numbers produced by the above formulas outside the interval $[M_1; M_1 + k]$ are $M_1 + k + 12K$ and $M_1 + k + 12K + 1$, and the fact that they are sufficient to ensure the property we seek on k, we can start new tests creating integers after M_2 without affecting the previous statement.

We decompose the integer line into $2n$ repeated intervals of width k and will do each of our test within one of those intervals. Those repeated intervals play the role of the different dimensions that we used previously. One important difference from the previous approach however is that the tests we made to ensure that k is large can affect the following tests. We thus need to slightly modify our previous constructions to make them robust. In particular, note that as the periods used in the test are either $12K$ or $12K+1$, there cannot be more than two intruding integers, generated by the test on k, in each interval of width $12K$.

The constant k being fixed large enough, let us explain precisely how we modify the constructions corresponding to what we did in the second, third and fourth dimensions of the proof for (n,m)-LpSl sets, for the polynomials of the form $Q_i(\vec{y}) = uv \pm z$.

Let $i \leq n$. We introduce the variable t_i and w_i and, for $j \in \{0,1,2\}$, we consider the formulas

$$\phi_1^j = \Lambda((M_2 + 2(i-1)k + j(3v+4), M_2 + 2(i-1)k + 2 + j(3v+4)) \mid (u + 2nk))$$
$$\phi_2^j = \Lambda((M_2 + 2(i-1)k + j(3v+4), M_2 + 2(i-1)k + j(3v+4) + 3v - 1) \mid (u + 2nk - 3))$$
$$\phi_3^j = \Lambda((M_2 + 2(i-1)k + 3v + 3 + j(3v+4)) \mid (u + 2nk - 3, u + 2nk))$$
$$\phi_4^j = \Lambda((M_2 + 2(i-1)k + w_i + t_i + j(3v+4), M_2 + 2(i-1)k + w_i + t_i + 2 + j(3v+4)) \mid)$$

And add to S_1 and S_2 the LpSl sets defined respectively by the formulas $\bigvee_{j\in\{0,1,2\}} \phi_1^j \vee \phi_2^j \vee \phi_3^j$ and $\bigvee_{j\in\{0,1,2\}} \phi_2^j \vee \phi_3^j \vee \phi_4^j$.

In the undecidability proof of the (n,m)-LpSl equality problem, this step ensured that $w_i + t_i = v(u + k_i)$. This was done using a sliding window, where the formulas ϕ_2 and ϕ_3 accepted an interval of integers with a single hole in it, which the formula ϕ_1 was exiting once, forcing ϕ_4 to be equal to the value reached at that moment. Here, the main idea remains the same, though with a few differences. First, we can see that the starting points of the formulas are shifted by $M_2 + 2(i-1)k + j(3v+4)$. This is done to (1) ensure no interaction with the tests that are done before M_2, (2) as we divided the integer line into $2n$ repeated intervals of width k, we wish for this test to occur within the $2(i-1)k$'th such interval and (3) we add $+j(3v+4)$ so that the three sets of formulas do not interact on the first v periods. Second, the period contains now $2nk$ instead of k_i, again, this is done to remain within the allocated interval. Third, for each value of j, we are building with ϕ_2^j and ϕ_3^j a sliding window with a hole of size 3 which formula ϕ_1^j exits exactly once to produce the integers $M_2 + 2(i-1)k + j(3v+4) + v(u+2nk)$ to $M_2 + 2(i-1)k + j(3v+4) + v(u+2nk) + 2$. The formulas ϕ_4^j must thus fill this gap and produce those integers exactly as before. However, due to the possible interference from the test ensuring $k \geq (12K)^2$, two of the integers that need to be produced may be filled by formulas other than ϕ_4^j. This is why we tripled the number of sliding windows and increased their width: as during the first v steps, all the sliding windows are contained within an interval of size at most $12K$, there are at most two "intruding" integers, thus (1) one of the sliding window must be unaffected, and (2) the first sliding window

is not entirely hidden by the intruding integers. This is sufficient to make the construction robust and to ensure that $w_i + t_i = v(u + 2nk)$.

We now use the $2(i-1)k+1$'th interval to ensure that $t_i = 2nkv$ by applying the above modifications to how we handled the third dimension in the previous proof. That is, we consider the formulas

$$\phi_1^j = \Lambda((M_2 + 2(i-1)k + j(3v+4), M_2 + 2(i-1)k + 2 + j(3v+4)) \mid (2nk))$$
$$\phi_2^j = \Lambda((M_2 + 2(i-1)k + j(3v+4), M_2 + 2(i-1)k + j(3v+4) + 3v - 1) \mid (2nk-3))$$
$$\phi_3^j = \Lambda((M_2 + 2(i-1)k + 3v + 3 + j(3v+4)) \mid (2nk-3, 2nk))$$
$$\phi_4^j = \Lambda((M_2 + 2(i-1)k + t_i + j(3v+4), M_2 + 2(i-1)k + t_i + 2 + j(3v+4)) \mid)$$

Adding to S_1 and S_2 the LpSl sets defined respectively by the formulas $\bigvee_{j\in\{0,1,2\}} \phi_1^j \vee \phi_2^j \vee \phi_3^j$ and $\bigvee_{j\in\{0,1,2\}} \phi_2^j \vee \phi_3^j \vee \phi_4^j$, ensures that $t_i = 2nkv$.

The last step to test the polynomial Q_i (ensuring that $w_i \pm z = 0$), corresponding to the fourth dimension of the previous proof, was achieved through a simple test without a periodical behaviour. We can thus assume it was done within the M_1 first integers.

By the construction above, we have that there exists parameter values $\vec{p}$ such that $S_1(\vec{p}) = S_2(\vec{p})$ iff there is a valuation of the variables $\vec{y}$ such that for all i, $Q_i(\vec{y}) = 0$. By Lemma1 and the undecidability of Hilbert's tenth problem, as S_1 and S_2 are $(n,1)$-LpSl sets, this implies that the existential $(n,1)$-LpSl equality problem is undecidable. □

B Proof of Lemma 2

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 $(b_1, c_1), \ldots (b_n, c_n)$ 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 $(b_1, c_1), \ldots (b_n, c_n)$. 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 $\Lambda(b_0^i, c_0^i) \mid (b_1^i), \ldots (b_{n_i}^i))$ (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 $(b_1), (b_1 + 1) \ldots (c_1)$ is trivially equivalent to ϕ_0 (*i.e.* ϕ_0 is equivalent to $\Lambda((b_0, c_0) \mid (b_2, c_2), \ldots, (b_n, c_n), (b_1), (b_1{+}1) \ldots (c_1)))$.

– 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 $kb_1(p) \leq y - b_0(p) \leq (k+1)b_1(p)$. Then as $k \geq m_1$, $kc_1(p) \geq (k+1)b_1(p)$ hence $kb_1(p) \leq y - b_0(p) \leq kc_1(p)$ 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

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

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. [3] 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 $(b_1), \ldots (b_n)$. Note quickly that if $c_0 - b_0 \in \mathbb{N}$, similarly to the previous reduction step, ϕ_0 is equivalent to

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

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

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

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 $(b_1), \ldots (b_n)$ 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

[3] 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.

simplify the formula. However, as the initial couple cannot be repeated contrary to the periods, the simplification is more limited.

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

- if $n = 0$, then the formula is already of the second type.
- 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

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

  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 at most two periods. Assuming that $n \geq 2$ and, thanks to the above, that for all i $f_i = f > 0$, and ordering the periods so that $g_1 < g_2 < \cdots < g_n$, 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. □

C Proof of Theorem 3

The structure of the proof of Theorem 3 is similar to the one of Proposition 2. The main difference is that, in order to represent correctly subformulas of the second type, we add a third dimension representing a multiple of p^2. Moreover, the representation of the intervals in formula of type 2 creates complications as the representation we build now includes false point (though for high value of the coefficients).

Preliminary assumptions

Let ϕ be a parametric Presburger formula describing a $(1, 1)$-LpSl set with parameter p that has the simple form described in Lemma 2. 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 or second type as described in Lemma 2. For every formula ϕ_i, we will build a three dimensional semi-linear set L_{ϕ_i} and set $L_\phi = \cup_{i \in I} L_{\phi_i}$.

Considering one sub-formula ϕ_i which is thus of the form $\Lambda((b_0, c_0) \mid (b_1), \dots (b_n))$. We denote for all $j \leq n$, $b_j = f_j p + g_j$, and $c_0 = f'_0 p + g'_0$.

We consider three possibilities:

• If ϕ_i is a formula of the first type and for all $1 \leq j \leq n$, $f_j = 0$. We order the periods so that g_n is the largest integer. Then we guess the set

$$\begin{aligned} G_i = &\{(r,s,t) \in \mathbb{N}^3 \mid r \leq n g_n \wedge s \leq n g_n \wedge t \leq g_n \\ &\wedge \exists k_1, \dots, k_n \in \mathbb{N}, f_0 p + g_0 + \sum_{i=1}^{n} k_i g_i = r p^2 + s p + t\}. \end{aligned}$$

In particular $(0, f_0, g_0) \in G_i$.

Thanks to the conditions that $r \leq n g_n$, $s \leq n g_n$ and $t \leq g_n$, there is finitely many different guesses that can be made for the set G_i. The construction (and analysis) which follows apply for every parameter p for which the guess is correct. Hence, our guess produces a partition of $\mathbb{N}$ and we study the equality of sets separately for each element of this partition.

• If ϕ_i is a formula of the first type and for all $1 \leq j \leq n$, $f_j \neq 0$. We order the periods in growing order of f_j, meaning that $f_n = \max_j f_j$. We set M_i to be such that $f_j M_i \geq g_j$ for all $j = 0, \dots, n$ and we guess the set

$$\begin{aligned} G_i = &\{(r,s,t) \in \mathbb{N}^3 \mid r \leq n f_n \wedge s \leq n M_i f_n \wedge t \leq M_i f_n \\ &\wedge \exists k_1, \dots, k_n \in \mathbb{N}, f_0 p + g_0 + \sum_{i=1}^{n} k_i (f_i p + g_i) = r p^2 + s p + t\}. \end{aligned}$$

In particular $(0, f_0, g_0) \in G_i$.

• If ϕ_i is a formula of the second type, we fix $h = \max_j g_j$ with $h = 0$ if $n = 0$. We guess the sets

$$\begin{aligned} G_i = &\{(r,s,t) \in \mathbb{N}^3 \mid r \leq n f \wedge s \leq n h \wedge t \leq h \\ &\wedge \exists k_1, \dots, k_n \in \mathbb{N}, f_0 p + g_0 + \sum_{i=1}^{n} k_i (f_i p + g_i) = r p^2 + s p + t\}, \\ D_i = &\{(r,s,t) \in \mathbb{N}^3 \mid r \leq n f \wedge s \leq n h \wedge t \leq h \\ &\wedge \exists k_1, \dots, k_n \in \mathbb{N}, f'_0 p + g'_0 + \sum_{i=1}^{n} k_i (f_i p + g_i) = r p^2 + s p + t\}, \\ H_i = &\{(r,s,0) \in \mathbb{N}^3 \mid \exists d, t \in \mathbb{N}, r \leq n f \wedge s \leq n h \wedge t \leq h \wedge f_0 < d < f'_0 \\ &\wedge \exists k_1, \dots, k_n \in \mathbb{N}, dp + \sum_{i=1}^{n} k_i (f_i p + g_i) = r p^2 + s p + t\} \end{aligned}$$

In particular $(0, f_0, g_0) \in G_i$, $(0, f'_0, g'_0) \in D_i$ and for all $f_0 < d < f'_0$, $(0, d, 0) \in H_i$.

Constructing a semi-linear set representing ϕ_i

Intuitively, we will build a three dimensional semi-linear set L_{ϕ_i} such that for "low values" of x, y and z, $(x, y, z) \in L_{\phi_i}$ if $xp^2 + yp + z$ is accepted by ϕ_i, and such that the reverse implication partially holds. We separate the construction following the same three cases of our preliminary assumptions.
• If ϕ_i is a formula of the first type and every period of ϕ_i is not an integer period, we build the semi-linear set:

$$L_{\phi_i} = \bigcup_{(x_0,y_0,z_0)\in G_i} \{(x, y, z) \in \mathbb{N}^3 \mid \exists k_1, \ldots, k_n, t_1, \ldots, t_n \in \mathbb{N},$$
$$(x, y, z) = (x_0, y_0, z_0) + \sum_{j=1}^{n} k_j(0, f_j, g_j) + \sum_{j=1}^{n} t_j(f_j, g_j, 0)\}.$$

Lemma 6. *Let $x, y, z \in \mathbb{N}$. If $(x, y, z) \in L_{\phi_i}$, $\phi_i(p)(xp^2 + yp + z)$ is true for all p satisfying the guess G_i.*

Conversely, for p satisfying G_i, if $\phi_i(p)(k)$ is true for some $k \in \mathbb{N}$, then there exists $x, y, z \in \mathbb{N}$ such that $k = xp^2 + yp + z$, $z \leq M_i(y + f_n)$ and $(x, y, z) \in L_{\phi_i}$.

We omit the proof of this lemma as it simply follows the structure of Lemma 3.
• If ϕ_i is a formula of the first type and every period of ϕ_i is an integer period, we build the semi-linear set:

$$L_{\phi_i} = \bigcup_{(x_0,y_0,z_0)\in G_i} \{(x, y, z) \in \mathbb{N}^3 \mid \exists k_1, \ldots, k_n, t_1, \ldots, t_n, r_1, \ldots, r_n \in \mathbb{N},$$
$$(x, y, z) = (x_0, y_0, z_0) + \sum_{j=1}^{n} k_j(0, 0, g_j) + \sum_{j=1}^{n} t_j(0, g_j, 0) + \sum_{j=1}^{n} r_j(g_j, 0, 0)\}.$$

Lemma 7. *Let $x, y, z \in \mathbb{N}$. If $(x, y, z) \in L_{\phi_i}$, $\phi_i(p)(xp^2 + yp + z)$ is true for all p satisfying the guess G_i.*

Conversely, for p satisfying G_i, if $\phi_i(p)(k)$ is true for some $k \in \mathbb{N}$, then for all $x, y, z \in \mathbb{N}$ such that $k = xp^2 + yp + z$ and either $y \geq f_0$ or $x > 1$, we have $(x, y, z) \in L_{\phi_i}$.

We omit the proof of this lemma as it simply follows the structure of Lemma 4.
• If ϕ_i is a formula of the second type, we build the semi-linear set: L_{ϕ_i} as the union of three sets $L^G_{\phi_i}, L^D_{\phi_i}$ and $L^H_{\phi_i}$ defined by

$$L^G_{\phi_i} = \bigcup_{(x_0,y_0,z_0)\in G_i} \{(x, y, z) \in \mathbb{N}^3 \mid \exists k_1, k_2, t_1, t_2, r \in \mathbb{N},$$
$$(x, y, z) = (x_0, y_0, z_0) + (\sum_{j=1,2} k_j(0, f, g_j)) + (\sum_{j=1}^{n} t_j(f^2, g_j, 0)) + (0, 0, r)\},$$

$$L^D_{\phi_i} = \bigcup_{(x_0,y_0,z_0)\in D_i} \{(x,y,r) \in \mathbb{N}^3 \mid \exists k_1, k_2, t_1, t_2, r \in \mathbb{N},$$
$$r \leq z \wedge (x,y,z) = (x_0,y_0,z_0) + (\sum_{j=1,2} k_j(0,f,g_j)) + (\sum_{j=1}^{n} t_j(f^2,g_j,0))\},$$

and

$$L^H_{\phi_i} = \bigcup_{(x_0,y_0,z_0)\in H_i} \{(x,y,z) \in \mathbb{N}^3 \mid \exists k_1, k_2, t_1, t_2 \in \mathbb{N},$$
$$\wedge (x,y,0) = (x_0,y_0,0) + (\sum_{j=1,2} k_j(0,f,0)) + (\sum_{j=1}^{n} t_j(f^2,g_j,0))\},$$

As mentioned earlier, we set $L_\phi = \cup_{i\in I} L_{\phi_i}$.

An interval interpretation of L_ϕ

The set L_ϕ we built, due to how we handled the interval produced by formulas of the second type, contains spurious points: the interval $(0,p)$ is translated into $(0,1,0)$ and all the points $(0,0,z)$ with $z \in \mathbb{N}$, hence, for all value of p, our set would contain $(0,0,p+1)$ and $p+1$ is not in the interval.

Instead of directly considering the sets, we will thus see L_ϕ as generating a set of intervals instead of points. More precisely, an interval of L_ϕ is defined by a pair of triple (x,y,z) and (x',y',z') where

- $(x,y,z) \in L_\phi$ "begins" a set of accepted points meaning that one of the following holds:
  - $z \neq 0$ and $(x,y,z-1) \notin L_\phi$,
  - $z = 0$, $y \neq 0$ and for all $N \in \mathbb{N}$, there exists $k \geq n$ such that $(x,y-1,k) \notin L_\phi$,
  - if $y = z = 0$, $x \neq 0$ and for all $N \in \mathbb{N}$, there exists $k, k' \geq n$ such that $(x-1,k,k') \notin L_\phi$,
  - $x = y = z = 0$.
- $(x',y',z') \in L_\phi$ "ends" an interval, meaning that $(x',y',z'+1) \notin L_\phi$.
- (x',y',z') is the closest triple ending an interval: if a triple (x'',y'',z'') ends an interval, then (1) $x'' > x'$, (2) $x'' = x'$ and $y'' > y'$ or (3) $x'' = x'$, $y'' = y'$ and $x'' > x'$[4].

We also consider the interval of the form $((x,y,z),\infty)$ if (x,y,z) begins an interval and there is no triple greater than (x,y,z) ending it.

Let I_ϕ be the set of all the intervals defined by L_ϕ. Given an interval $\Delta \in I_\phi$, we denote by $[\![\Delta]\!]_p$ the set of integers contained in this interval: if $\Delta = ((x,y,z),(x',y',z'))$, $[\![\Delta]\!]_p = \{k \in \mathbb{N} \mid xp^2 + yp + z \leq k \leq x'p^2 + y'p + z'\}$ and if $\Delta = ((x,y,z),\infty)$, $[\![\Delta]\!]_p = \{k \in \mathbb{N} \mid xp^2 + yp + z \leq k\}$

[4] In other words, we ordered the triples in a lexicographical order and (x',y',z') is thus the smallest triple ending an interval and greater than (x,y,z).

Lemma 8. *Let p be a parameter value satisfying all the guesses related to the subformulas*

Given $k \in \mathbb{N}$ such that $\phi(p)(k)$ holds iff there exists an interval $\Delta \in I_\phi$ such that $k \in [\![\Delta]\!]_p$.

Proof. Let p be a parameter value satisfying all the guesses related to the subformulas

• Let $k \in \mathbb{N}$ such that $\phi(p)(k)$ holds, let us find an interval of I_ϕ containing k. Let i such that $\phi_i(p)(k)$ holds. By construction, given $\Delta \in I_{\phi_i}$, there exists $\Delta' \in I_\phi$ such that $[\![\Delta]\!]_p \subseteq [\![\Delta']\!]_p$. Hence, finding an interval of I_{ϕ_i} containing k is sufficient.

If ϕ_i is a formula of the first type, then by Lemma 6 and Lemma 7, there exists x, y, z such that $(x, y, z) \in L_{\phi_i}$ and $k = xp^2 + yp + z$. By definition of the intervals, there exists an interval $(x', y', z') \in I_{\phi_i}$ such that $x'p^2 + y'p + z' \leq k$. Select the one closest to (x, y, z). As (x', y', z') is the closest triple starting an interval below (x, y, z), every triple between (x', y', z') and (x, y, z) belong to L_{ϕ_i}. In particular, this means that the end of this interval is either ∞ or a triple greater than (x, y, z). Thus, denoting Δ this interval, $k = xp^2 + yp + z \in [\![\Delta']\!]_p$

If ϕ_i is a formula of the second type, $\Lambda((f_0p+g_0, f'_0p+g'_0) \mid (fp+g_1)(fp+g_2))$, then there exists k_1 and k_2 such that $left = f_0p + g_0 + \sum_{j=1,2} k_j(fp + g_j) \leq k \leq f'_0p + g'_0 + \sum_{j=1,2} k_j(fp + g_j) = right$. By construction of L_{phi_i}, there exists (x, y, z) and (x, y', z') in L_{phi_i} such that $left = xp^2 + yp + z$, $right = xp^2 + y'p + z'$ and every triple in between (x, y, z) and (x, y', z') also belongs to L_{phi_i}. Therefore, either $((x, y, z), (x, y', z')) \in I_{\phi_i}$, or there exists $\Delta \in I_{\phi_i}$ that extends $((x, y, z), (x, y', z'))$ and thus $k \in [\![((x, y, z), (x, y', z'))]\!]_p \subseteq [\![\Delta]\!]_p$.

• Conversely, let $\Delta \in I_\phi$ and $k \in [\![\Delta]\!]_p$. The intervals of I_ϕ can be obtained by regrouping the intervals of the I_{ϕ_i} thus there exists i such that $k \in I_{\phi_i}$.

Assuming for simplicity that $\Delta = ((x, y, z), (x', y', z'))$ (hence ignoring the possibility of the infinite end), there exists (x'', y'', z'') such that $k = x''p^2 + y'' + z''$ and (x'', y'', z'') is between (x, y, z) and (x', y', z') in the lexicographic order. Hence, by definition of the interval, $(x'', y'', z'') \in L_{\phi_i}$.

If ϕ_i is a formula of the first type, by Lemma 6 and Lemma 7, this implies that $\phi_i(p)(k)$ holds, hence $\phi(p)(k)$ holds as well.

If ϕ_i is a formula of the second type, by construction of L_{ϕ_i}, (x, y, z) was attained in $L^G_{\phi_i}$ and (x', y', z') belong to the corresponding element (same choice of constants) in $L^D_{\phi_i}$. Thus there exists k_1 and k_2 such that $xp^2 + yp + z = f_0p + g_0 + \sum_{j=1,2} k_j(fp + g_j)$ and $x'p^2 + y'p + z' = f'_0p + g'_0 + \sum_{j=1,2} k_j(fp + g_j)$. As $xp^2 + yp + z \leq k \leq x'p^2 + y'p + z'$, we have that $\phi_i(p)(k)$ holds.

Conclusion on Theorem 3

Let ϕ and ψ be two parametric Presburger formulas describing two $(1, 1)$-LpSl sets S_ϕ and S_ψ and let $N_\phi, N_\psi \in \mathbb{N}$, and $\bigvee_{i \in I_\phi} \phi_i$ and $\bigvee_{i \in I_\psi} \phi_i$ be given by applying Lemma 2 on ϕ and ψ. We have the following relation (miroring Lemma 5):

Lemma 9. *There exists N such that if $p > N$ and p satisfies all the guesses G_i for $i \in I_\phi \cup I_\psi$, then $L_\phi = L_\psi$ iff $S_\phi(p) = S_\psi(p)$.*

Proof. Let $N_1 = \max(N_\phi, N_\psi)$. Let H be the constant computed in Corollary 23 of [4] on L_ϕ and L_ψ. We fix $N = \max(N_1, H)$ and assume that $p > N$ and that p satisfies the guesses G_i, D_i and H_i for $i \in I_\phi \cup I_\psi$.

Assume first that $L_\phi = L_\psi$. Then $I_\phi = I_\psi$ and by Lemma 8, $S_\phi = S_\psi$.

Conversely, assume that $L_\phi \neq L_\psi$. By Corollary 23 of [4], there exists a triple (x, y, z) that distinguishes the two sets (say, $(x, y, z) \in L_\phi \setminus L_\psi$) such that the modulus of x, y and z are smaller than H. Assume there exists $(x', y', z') \in L_\psi$ different from (x, y, z) and such that $xp^2+yp+z = x'p^2+y'p+z'$. Then $x' < x$ as y and z are smaller than H and thus smaller than p. Moreover, either thanks to the guesses or application of p periods in one step, if one can reach (x', y', z'), one can reach any other writing of the same number with a greater first component. Hence, $(x, y, z) \in L_\psi$, raising a contradiction. □

Proof (Proof of Theorem 3). Let S_ϕ and S_ψ be two $(1, 1)$-LpSl sets described by two formulas ϕ and ψ.

The integers can be effectively finitely partitioned $\mathbb{N} = \bigcup_{i \in I} K_i$ based on which set of guesses G_i, D_i and H_i they satisfy. For any set K_i, applying Lemma 9, one obtains a constant N_i such that for any $p \in K_i$, if $p > N_i$, then the equality of $S_\phi(p)$ and $S_\psi(p)$ 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 built for 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 $S_\phi(p)$ and $S_\psi(p)$. The subset K'_0 of parameter values from K_0 achieving equality of the two $(1, 1)$-LpSl sets can be computed as there are finitely many integers within K_0 and for a fixed value of p $S_\phi(p)$ and $S_\psi(p)$ are semi-linear sets. We thus have that the set of parameter values of p such that $S_\phi(p) = S_\psi(p)$ is $K'_0 \bigcup_{i \in J} K_i$. □