Discrete Mathematics and Theoretical Computer Science DMTCS vol. **12:1**, 2010, 20–38

Deciding whether the ordering is necessary in a Presburger formula†

Christian Choffrut[1] and Achille Frigeri[2]

[1] *L.I.A.F.A., Université Paris 7, 2 Pl. Jussieu – 75 251 Paris Cedex – France*
[2]*Dipartimento di Matematica, Politecnico di Milano & LIAFA, Université Paris 7 via Bonardi, 9 – 20133 Milano, Italia*

received March 30 2009, *revised August 26 2009*, *accepted January 14 2010*.

We characterize the relations which are first-order definable in the model of the group of integers with the constant 1. This allows us to show that given a relation defined by a first-order formula in this model enriched with the usual ordering, it is recursively decidable whether or not it is first-order definable without the ordering.

Keywords: Presburger arithmetics

Introduction

Presburger arithmetic is the fragment of arithmetic concerning the integers with addition and order. Presburger's supervisor considered the decidability of this fragment too modest a result to deserve a Ph.D. degree and he accepted it only as a Master's Thesis in 1928. Looking at the number of citations, we may say that history revised this depreciative judgment long ago. There still remains, at least as far as we can see, some confusion concerning the domain of the structure: $\mathbb{Z}$ or $\mathbb{N}$? with or without the order relation? (the main popular mathematical web sites disagree on that respect). The original paper deals with the additive group of positive and negative integers with no binary relation, but in a final remark of the original communication, the author asserts that the same result, namely quantifier elimination, holds when the structure is enriched with the binary relation "$<$". In (1), which is the main reference on the subject, Presburger arithmetic is defined as the elementary theory of integers with equality, addition and having 0 and 1 as constant symbols and "$<$" as binary predicate, see also (20). On the other hand, the majority of the "modern" papers referring to Presburger arithmetic is concerned with the natural numbers where the order relation is unnecessary as it is first-order expressible in $\langle\mathbb{N}; +\rangle$.

The origin of the present work is the simple remark that concerning the set of integers $\mathbb{Z}$, the binary relation matters. Here we study the decidability of the definability in the structure $\langle\mathbb{Z}; +, 0, 1\rangle$ which we call the weak Presburger arithmetic, for a given relation defined in $\langle\mathbb{Z}; +, <, 0, 1\rangle$. We show that it is

†A preliminary version of these results was published in the proceedings of the Conference ICTCS 2007, held in Rom, Italy, 2007, (5)

indeed recursively decidable and we prove this result by revisiting the notion of linear subsets introduced by Ginsburg and Spanier (9; 10; 11; 12) in the sixties for n-ary relations on $\mathbb{N}$. Other problems of definability in substructures of Presburger arithmetic introduced in (15) have been solved positively in (4): given a formula in the Presburger arithmetic, it is recursively decidable whether or not it can be expressed in the structure with domain $\mathbb{Z}$ (resp. $\mathbb{N}$) and with the unique predicates of the form $x - y > a$ and $x > a$ for all $a \in \mathbb{Z}$ (resp. $a \in \mathbb{N}$).

Despite its simplicity, Presburger arithmetic is central in many areas of theoretical and applied computer science. From a theoretical point of view, it has many remarkable properties: 1) it admits quantifier elimination (1; 19; 20) and therefore it is decidable, 2) given a formula on the expansion of the structure obtained by adding the function which to each integer assigns the maximal power of 2 which divides it, it is decidable whether or not it is definable by a Presburger formula over $\mathbb{N}$, (18); the claim of a polynomial time algorithm can be found in (17). Moreover, there is a strong and old connection between language theory, Presburger definable sets and rational relations on $\mathbb{Z}$ and $\mathbb{N}$ dating back to the sixties (2; 7; 9). The concept is also widely used in many application areas, such as program analysis and model-checking and more specifically timed automata since it models infinite systems, see for example (3).

The paper is organized as follows. Section 1 recalls known results which are necessary for the rest of the paper, essentially around the notion of linear sets in $\mathbb{Z}^n$. Section 2 investigates the closure properties of the linear sets. This allows us to give in Section 3 a characterization of the subsets of $\mathbb{Z}^n$ which are definable in the weak Presburger arithmetic, along with a decision procedure.

1 Preliminaries

1.1 Variants of Presburger arithmetic

In this paper we deal with first-order languages with equality, i.e., the signature of a structure implicitly contains the symbol "$=$ ". As observed in the introduction, a source of confusion is the lack of agreement in the definition of Presburger arithmetic itself. We make the convention of calling *weak Presburger arithmetic* the structure $\mathcal{Z}^W = \langle \mathbb{Z}; +; 0, 1\rangle$ originally studied in (19) (see also (21) for an English translation and commentary), while with $\mathcal{Z}$ we mean the *(standard) Presburger arithmetic* $\langle \mathbb{Z}; +, <, 0, 1\rangle$. Observe that the predicate $<$, as restriction of the order on $\mathbb{Z}$ to $\mathbb{N}$, is already definable in $\langle \mathbb{N}; +\rangle$. These structures are decidable in the sense that given a closed formula, it is recursively decidable whether or not it holds in that structure. In particular $\mathcal{Z}^W$ admits quantifier elimination in the augmented language with the additional unary functional symbol "$-$ " and the (recursive) set of binary relations $\{\equiv_m\}_{m\in\mathbb{N}\setminus\{0,1\}}$, having the usual meaning of opposite and modulo. As for $\mathcal{Z}$, quantifier elimination is obtained in the same augmented language enriched with the binary relation $<$.

1.2 Logical definability

Here we are concerned with the *definability* issue. Consider a structure $\mathcal{D}$ with domain D and a first-order formula on this structure, say $\phi(x_1, \ldots, x_n)$, where $x_1, \ldots, x_n$ are its free variables. Then the n-ary relation R *defined* by ϕ is the set of n-tuples $(a_1, \ldots, a_n)$ such that ϕ holds true when the variable x_i is assigned the value a_i, i.e., $R = \{(a_1, \ldots, a_n) \in D^n \mid \mathcal{D} \models \phi(a_1, \ldots, a_n)\}$. A relation is $\mathcal{D}$-*definable* or simply *definable* when the structure is understood, if it can be defined by a first-order formula on $\mathcal{D}$.

The following is the main result on the integer arithmetic without multiplication. It proves that it admits quantifier elimination and it is due to Presburger, cf. (19).

Theorem 1.1 (Presburger) *A subset X of $\mathbb{Z}^n$ is $\mathcal{Z}$-definable if and only if it is a Boolean combination of relations defined by predicates of the form*

$$t < t' \quad \textit{and} \quad t \equiv t' \pmod{b} \tag{1}$$

where t and t' are linear expressions on the integer variables x_i of the form $a_0 + \sum_{i=1}^{n} a_i x_i$ with $a_i \in \mathbb{Z}$, and $b \in \mathbb{N}$, $b > 1$.
A subset X of $\mathbb{Z}^n$ is $\mathcal{Z}^W$-definable if and only if it is a Boolean combination of relations defined as in (1) where the binary relation $<$ is replaced by the equality.

1.3 Linear sets

The following definitions could be given for arbitrary finitely generated commutative monoids but we are mainly interested in the free commutative group $\mathbb{Z}^n$. We consider it as a subset of the $\mathbb{Q}$-vector space $\mathbb{Q}^n$ and view its elements as row vectors. The operation of addition is extended from elements to subsets: if $X, Y \subseteq \mathbb{Z}^n$, then the *sum* $X + Y \subseteq \mathbb{Z}^n$ is the set of all sums $x + y$ where $x \in X$ and $y \in Y$. When X is a singleton $\{x\}$ we simply write $x+Y$. Given x in $\mathbb{Z}^n$, the expression $\mathbb{N}x$ represents the subset of all vectors nx where n range over $\mathbb{N}$ and similarly for $\mathbb{Z}x$. For example, $\mathbb{Z}x + \mathbb{Z}y$ represents the subgroup generated by the vectors x and y. We extend also the matricial notation to subsets, e.g., if A is an $(n \times p)$-matrix with integer or rational entries, then $\mathbb{Q}^n A$ stands for the set of all p-row vectors of the form xA for some n-row vector $x \in \mathbb{Q}^n$.

We now briefly recall the classical theory of the linear and semilinear subsets of $\mathbb{Z}^n$ as exposed in (7).

Definition 1.2 *A subset of $\mathbb{Z}^n$ is* $\mathbb{N}$-linear *if it is of the form*

$$a + \sum_{i=1}^{k} \mathbb{N} b_i, \qquad a, b_i \in \mathbb{Z}^n \;\; i = 1, \ldots, k. \tag{2}$$

It is $\mathbb{N}$-simple *if the vectors b_i are linearly independent as vectors of $\mathbb{Q}^n$. It is* $\mathbb{N}$-semilinear *if it is a finite union of linear sets.*

Ginsburg and Spanier proved (11) the following equivalent statements for $\mathbb{N}^n$, but it can readily be seen to hold for $\mathbb{Z}^n$. Together with Theorem 1.1, it can be interpreted as saying that the first-order definable sets in $\langle \mathbb{Z}; +, 0, 1, < \rangle$ are exactly the rational subsets of $\mathbb{Z}^n$. Actually Eilenberg and Schützenberger in (7), and independently Ito in (14), strengthened the third condition by proving that the simple sets may be assumed disjoint.

Theorem 1.3 *Given a subset X of $\mathbb{Z}^n$, the following assertions are equivalent:*

1. *X is first-order definable in $\mathcal{Z}$;*
2. *X is $\mathbb{N}$-semilinear;*
3. *X is a finite union of $\mathbb{N}$-simple sets.*

Furthermore, each of the specifications can be effectively transformed into another.

When dealing with a linear subset, this result allows us to assume that it is given as a finite union of simple subsets.

The following notion of dimension will be useful in the sequel. It is defined on linear sets but extended to arbitrary sets in a natural way.

Definition 1.4 *The* dimension *of the simple set $a + \sum_{i=1}^{m} \mathbb{N}b_i$ is the integer m. More generally, the dimension of an arbitrary nonempty subset $X \subseteq \mathbb{Z}^n$ is the minimum integer m, denoted $\dim(X)$, such that X is included in a finite union of simple sets of dimension at most m. The dimension of the empty set is equal to -1.*

Observe that the dimension of a finite union of simple sets does not depend on the specific expression which defines it. This is seen by observing that an expression of a simple set of dimension m as a finite union of simple sets contains a simple set of dimension m and no simple set of dimension greater than m and that equality $\text{dim}(A \cup B) = \max\{\text{dim}(A), \text{dim}(B)\}$ holds. Indeed, since dimension is a nondecreasing operator, we have $\text{dim}(A \cup B) \geq \text{dim}(A)$ and $\text{dim}(A \cup B) \geq \text{dim}(B)$ and therefore $\text{dim}(A \cup B) \geq \max\{\text{dim}(A), \text{dim}(B)\}$. The other direction is trivial. As a result we have the following proposition.

Proposition 1.5 *Given a subset $X \subseteq \mathbb{Z}^n$ specified by some first-order formula in $\mathcal{Z}$ and some integer $0 \leq d \leq n$, it is recursively decidable whether or not it has dimension at most d.*

Proof: Indeed, using Ginsburg and Spanier's construction in (11, Theorem 1.3.), an expression as a finite union of simple sets for X can be effectively computed from the first-order formula specifying it. A simple inspection suffices to determine whether or not the linear sets composing it have dimension less than or equal to d. □

1.4 $\mathbb{Z}$-linear sets

We now extend the notion of $\mathbb{N}$-linear subsets defined in Section 1.3 to $\mathbb{Z}$-linear subsets by allowing the coefficients of expression (2) to be in $\mathbb{Z}$.

Definition 1.6 *A subset of $\mathbb{Z}^n$ is* $\mathbb{Z}$-linear *if it is of the form*

$$a + \sum_{i=1}^{k} \mathbb{Z}b_i, \qquad a, b_i \in \mathbb{Z}^n,\ i = 1, \dots, n. \tag{3}$$

The notions of $\mathbb{Z}$-simple *and* $\mathbb{Z}$-semilinear *sets are defined with the obvious modifications of Definition 1.2.*

There exists a notion which is halfway between that of simple and that of arbitrary disjoint unions of simple sets. We use the standard geometric notion of translation, restricted to $\mathbb{Z}^n$, which is a mapping of the form $x \mapsto a + x$ for some given vector $a \in \mathbb{Z}^n$. A *translate* of a subset is its image in a translation.

Definition 1.7 *A subset is* quasi-simple *if it is a finite union of translates of a given simple set S. This is equivalent to saying that it is a set of the form $A + S$, where $A \subseteq \mathbb{Z}^n$ is finite and S is a subgroup of $\mathbb{Z}^n$. The set S is called* template.

Clearly, a finite union of $\mathbb{Z}$-linear subsets is also a finite union of $\mathbb{N}$-linear subsets.

Proposition 1.8 *Every $\mathbb{Z}$-linear subset of $\mathbb{Z}^n$ is a finite union $\mathbb{N}$-linear subsets.*

Proof: Indeed, a $\mathbb{Z}$-linear subset as in expression (3) is equal to the union of all $a + \sum_{i=1}^{k} \mathbb{N}(\epsilon_i b_i)$ where $(\epsilon_1, \ldots, \epsilon_k)$ ranges over the set $\{\pm 1\}^k$. $\Box$

The converse does not hold, e.g., the subset $\mathbb{N}$ is not expressible as a finite union of $\mathbb{Z}$-linear subsets (this is a direct consequence of Theorem 1.1 and Theorem 1.3 but can be worked out directly too).

2 Closure properties

2.1 Properties of $\mathbb{Z}$-linear sets

Our decidability result is based on the equivalence between definable subsets in the structure $\mathcal{Z}^W$ and the family of finite unions of subset differences of $\mathbb{Z}$-semilinear subsets, see subsection 2.2. This characterization is obtained by a careful study of the closure properties of the family of $\mathbb{Z}$-linear sets carried out here.

We will show that the class of $\mathbb{Z}$-linear sets enjoys many properties such as closure under finite sum as defined in subsection 1.3, projection, direct product and even more interestingly intersection, making it a more robust family than the family of $\mathbb{N}$-linear sets since $\mathbb{N}$-linear sets are not closed under intersection. On the opposite, the family of $\mathbb{N}$-semilinear sets is more robust than that of $\mathbb{Z}$-semilinear sets, in fact while the first is closed under all Boolean operations, the second is not. The sets belonging to the Boolean closure of the $\mathbb{Z}$-linear sets have a more complex representation (see Theorem 3.1).

We start with a property which shows how different the $\mathbb{N}$- and the $\mathbb{Z}$-linear subsets are. In the case of the nonnegative integers, every linear set is a finite disjoint union of simple sets. For $\mathbb{Z}$, we have a stronger property which is a direct consequence of a classical algebraic result.

Proposition 2.1 *Every $\mathbb{Z}$-linear set is $\mathbb{Z}$-simple.*

Proof: Consider the set $X = a + \sum_{i=1}^{m} \mathbb{Z}b_i$, with $b_i \in \mathbb{Z}^n$ for $1 \leq i \leq m$. It suffices to consider the case where a is null. Then X is a subgroup of $\mathbb{Z}^n$ and therefore it is free, (16, Theorem 4, (I,§9))), i.e., there exists $h \leq n$ linearly independent vectors $b'_i \in \mathbb{Z}^n$ such that $X = \sum_{i=1}^{h} \mathbb{Z}b'_i$. $\Box$

Using standard manipulations we get the following proposition.

Proposition 2.2 *The family of $\mathbb{Z}$-linear sets is closed under projection, direct product and finite sum.*

Observe that $\mathbb{N}$-linear sets are closed under direct product and finite sum but they are not closed under projection. Indeed, consider the $\mathbb{N}$-linear (actually $\mathbb{N}$-simple) set $\mathbb{N}(2,1) + \mathbb{N}(3,1)$. Its projection onto the first component is the set $\{0\} \cup \{n \mid n \geq 2\}$.

The following is a kind of closure property under composition.

Proposition 2.3 *Let $X = a + \sum_{i=1}^{p} \mathbb{Z}b_i \subseteq \mathbb{Z}^n$ and $Y = c + \sum_{j=1}^{m} \mathbb{Z}d_j \subseteq \mathbb{Z}^p$ be two $\mathbb{Z}$-linear subsets. Then the set*

$$\{a + \sum_{i=1}^{p} \lambda_i b_i \mid (\lambda_1, \ldots, \lambda_p) \in Y\} \tag{4}$$

is $\mathbb{Z}$-linear.

Proof: Indeed, denote by B the $(p \times n)$-matrix consisting of the p row-vectors b_i and by D the $(m \times p)$-matrix consisting of the m row-vectors d_j. Then the set defined in (4) is equal to

$$a + (c + \mathbb{Z}^m D)B = (a + cB) + \mathbb{Z}^m(DB)$$

□

The intersection of two $\mathbb{N}$-linear subsets is a finite union of $\mathbb{N}$-linear subsets, not necessarily an $\mathbb{N}$-linear subset. Here we have a stronger property: the intersection of two $\mathbb{Z}$-linear subsets is a $\mathbb{Z}$-linear subset. We prove this result by resorting to a different approach from that in (7) which is purely combinatorial and proceeds by induction on the dimension of the space while we make use of a classical theorem of linear algebra. We recall that $\mathrm{GL}_n(\mathbb{Z})$ represents the unimodular group which consists of all the invertible $n \times n$-matrices with entries in $\mathbb{Z}$. We denote by rank(A) the rank of the matrix A.

Theorem 2.4 (Smith normal form (6, §2.4.4)) *Let $A \in \mathbb{Z}^{n\times m}$ be an integer matrix of rank s. Then there exist two unique invertible matrices $U \in GL_n(\mathbb{Z})$ and $V \in GL_m(\mathbb{Z})$, such that*

$$A' = UAV = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

where $D = (d_{ij}) \in \mathbb{Z}^{s\times s}$ is an integer square diagonal matrix such that d_{jj} divides d_{ii}, for $1 \le i \le j \le s$.

The matrix A' is called the *Smith normal form* of A. The following is a preparatory result.

Proposition 2.5 *Given $A \in \mathbb{Z}^{n\times m}$ and $b \in \mathbb{Z}^n$, the set of solutions in $\mathbb{Z}^m$ of the linear system $Ax = b$ is a $\mathbb{Z}$-simple set of dimension $m -$ rank(A).*

Proof: Let rank$(A) = s$ and $A' = UAV$ be the Smith normal form of A. The given system is equivalent to $UAVV^{-1}x = Ub$. Setting $V^{-1}x = y$ and $Ub = b'$, the set of solutions of $A'y = b'$ is the $\mathbb{Z}$-simple set $S' = c + \sum_{i=s+1}^{m} \mathbb{Z}e_i$, where $c_j = b'_j/a_{jj}$ for $1 \le j \le s$ and $c_j = 0$ for $s < j \le m$ and the e_i are the vectors of the canonical basis (observe that the system admits solutions in $\mathbb{Z}^m$ if and only if a_{jj} divides b'_j, for $1 \le j \le s$). Then the set of solutions of the equation $Ax = b$ is equal to $S = VS'$, and so if $\hat{c} = Vc$ and $\hat{e}_i = Ve_i$, we have $S = \hat{c} + \sum_{i=s+1}^{m} \mathbb{Z}\hat{e}_i$. Because V is unimodular, the vectors $\hat{e}_i$ are again linearly independent, so the set has dimension equal to $m -$ rank(A). □

Consequently, we have the following theorem.

Theorem 2.6 *If P and Q are $\mathbb{Z}$-simple sets of $\mathbb{Z}^n$ of dimension p and q respectively, then $P \cap Q$ is a $\mathbb{Z}$-simple set of dimension less than or equal to $\min\{p, q\}$. Furthermore, an expression for this intersection can be computed effectively.*

Proof: Let $P = a + \sum_{i=1}^{p} \mathbb{Z}b_i$ and $Q = c + \sum_{j=1}^{q} \mathbb{Z}d_j$ be two $\mathbb{Z}$-simple sets in $\mathbb{Z}^n$ with $p \le q$ and suppose that $P \cap Q \neq \emptyset$. Let us consider the linear system $\sum_{i=1}^{p} b_i x_i + \sum_{i=p+1}^{p+q} d_{i-q} x_i = c - a$ and let S be its set of solutions in $\mathbb{Z}^{p+q}$. Then S is a $\mathbb{Z}$-simple set in $\mathbb{Z}^{p+q}$ and so, its projection T on the first p components is a $\mathbb{Z}$-linear set in $\mathbb{Z}^p$, say $T = e + \sum_{i=1}^{\ell} \mathbb{Z}f_i$, for some $e, f_i \in \mathbb{Z}^p$ and $\ell \le p$. Considering the first p components of S, we have the following expression for $P \cap Q$:

$$P \cap Q = \left\{ a + \sum_{i=1}^{p} x_i b_i \mid (x_1, \ldots x_p) \in T \right\}.$$

It then suffices to apply Proposition 2.3. Concerning the complexity of the computation, it directly follows from the fact that Smith normal form of A can be computed in polynomial time, see (22). □

We now turn to the union operation. Its is clear that the union of two $\mathbb{Z}$-simple sets in $\mathbb{Z}^n$ is not necessarily simple, e.g., in $\mathbb{Z}$, the singletons $\{0\}$ and $\{1\}$ are simple but their union is not. However, when the simple sets have maximal dimension, i.e., when their dimension equals n, the union is almost simple. This is expressed rigorously in the next result.

Proposition 2.7 *Every finite union of $\mathbb{Z}$-simple sets of maximal dimension is a $\mathbb{Z}$-quasi-simple set.*

Proof: Let $S \subseteq \mathbb{Z}^n$ be a $\mathbb{Z}$-semilinear set of the form

$$S = \bigcup_{i=1}^{m} X_i,$$

where for all $1 \leq i \leq m$, $X_i = a^{(i)} + \sum_{j=1}^{n} \mathbb{Z} b_j^{(i)}$ is a $\mathbb{Z}$-simple set. Since for each fixed $i = 1, \ldots, m$, the vectors $b_j^{(i)}$, $1 \leq j \leq n$, are linearly independent, these vectors generate a subgroup of $\mathbb{Z}^n$ of finite index. In particular $X_i - a^{(i)}$ is the inverse image of the unit of a finite commutative group. In other words there exists a surjective morphism φ_i of $\mathbb{Z}^n$ into a finite commutative group G_i such that $X_i = \varphi_i^{-1}(\varphi_i(a^{(i)}))$. Consider the morphism φ of $\mathbb{Z}^n$ into the direct product $G_1 \times \cdots \times G_m$ defined by $\varphi(u) = (\varphi_1(u), \ldots, \varphi_m(u))$ and set $K = \{x \in G_1 \times \cdots \times G_m \mid x_i = \varphi_i(a^{(i)}) \text{ for some } 1 \leq i \leq m\}$. Then $S = \varphi^{-1}(K)$, which means that S is a union of cosets of the subgroup $\varphi^{-1}(e)$ where e is the unit of the direct product $G_1 \times \cdots \times G_m$. □

The following proposition expresses also an interesting property.

Proposition 2.8 *Let S, T be $\mathbb{Z}$-simple sets in $\mathbb{Z}^n$. If dim(S) = dim(T) and $T \subseteq S$, then $X = S \setminus T$ is a $\mathbb{Z}$-quasi-simple set with template T.*

Proof: Let $\dim(S) = \dim(T) = m \leq n$, then $S = a + \sum_{i=1}^{m} \mathbb{Z} b_i$ and $T = c + \sum_{i=1}^{m} \mathbb{Z} d_i$, with $a, c, b_i, d_i \in \mathbb{Z}^n$ for $1 \leq i \leq m$. From $T \subseteq S$ we have $c = a + \sum_{i=1}^{m} \gamma_i b_i$, so

$$\begin{aligned} S &= c + \textstyle\sum_{i=1}^{m} (-\gamma_i) b_i + \sum_{i=1}^{m} \mathbb{Z} b_i \\ &= c + \textstyle\sum_{i=1}^{m} (\mathbb{Z} - \gamma_i) b_i = c + \sum_{i=1}^{m} \mathbb{Z} b_i. \end{aligned} \tag{5}$$

Observing that for $A, B \subseteq \mathbb{Z}^n$, $v \in \mathbb{Z}^n$, we have $(A + v) \cap (B + v) = (A \cap B) + v$, we can suppose without loss of generality that $c = 0$, so that $S = \sum_{i=1}^{m} \mathbb{Z} b_i$ and $T = \sum_{i=1}^{m} \mathbb{Z} d_i$. Then S and T are additive groups of finite index and moreover T is a subgroup of S, so $S \setminus T$ can be written as a finite union of cosets of T, which completes the proof. □

Corollary 2.9 *Let S and T be $\mathbb{Z}$-simple sets of maximal dimension n in $\mathbb{Z}^n$. Then $X = S \setminus T$ is a $\mathbb{Z}$-quasi-simple set of dimension n with template T.*

Proof: Observe that $S \setminus T = S \setminus (S \cap T)$. From Theorem 2.6, $S \cap T$ is again a $\mathbb{Z}$-simple set of dimension n. Then $(S \cap T) \subseteq S$ holds so that we can apply Proposition 2.8. □

The following proposition provides us with a remarkable family of $\mathbb{Z}$-simple sets which are the traces of the affine $\mathbb{Q}$-spaces in $\mathbb{Z}^n$.

Proposition 2.10 *Let $H = a + \sum_{i=1}^{p} \mathbb{Q}b_i \subseteq \mathbb{Q}^n$ with $0 \leq p \leq n$ and $a, b_i \in \mathbb{Z}^n$, $i = 1, \ldots, p$. Then $H' = H \cap \mathbb{Z}^n$ is a $\mathbb{Z}$-simple set of $\mathbb{Z}^n$ of the same dimension as that of H. Moreover a $\mathbb{Z}$-simple expression for H' is effectively computable.*

Proof: Without loss of generality we can suppose that $0 \in H$ and that $\dim(H) = p < n$ holds so that $H = \mathbb{Q}^p A$, where A is the $(p \times n)$-matrix whose rows are the p vectors b_i. From Theorem 2.4 there exist two integer invertible matrices $U \in GL_p(\mathbb{Z})$ and $V \in GL_n(\mathbb{Z})$ such that

$$UAV = A' = \begin{pmatrix} d_1 & \cdots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & 0 & \vdots \\ 0 & \ldots & d_p & 0 & \ldots & 0 \end{pmatrix}$$

where all the coefficients d_i are integers. Since U and V are invertible we have

$$\mathbb{Q}^p U^{-1} A' V^{-1} \cap \mathbb{Z}^n V V^{-1} = (\mathbb{Q}^p U^{-1} A' \cap \mathbb{Z}^n V) V^{-1},$$

and, from the unimodularity of U and V, $\mathbb{Z}^n V = \mathbb{Z}^n$ and $\mathbb{Q}^p U^{-1} = \mathbb{Q}^p$. Furthermore, we clearly have

$$\mathbb{Q}^p A' = \mathbb{Q}^p \times \overbrace{\{0\} \times \cdots \times \{0\}}^{n-p \text{ times}},$$

so we we obtain

$$\begin{aligned} \mathbb{Q}^p A \cap \mathbb{Z}^n &= \mathbb{Q}^p U^{-1} A' V^{-1} \cap \mathbb{Z}^n V V^{-1} = (\mathbb{Q}^p U^{-1} A' \cap \mathbb{Z}^n V) V^{-1} \\ &= (\mathbb{Q}^p A' \cap \mathbb{Z}^n) V^{-1} = ((\mathbb{Q}^p \times \overbrace{\{0\} \times \cdots \times \{0\}}^{n-p \text{ times}}) \cap \mathbb{Z}^n) V^{-1} \\ &= (\mathbb{Z}^p \times \overbrace{\{0\} \times \cdots \times \{0\}}^{n-p \text{ times}}) V^{-1}. \end{aligned}$$

Denote by V' the matrix obtain from V^{-1} by substituting 0 for all entries in the last $n-p$ rows, we finally get:

$$\mathbb{Q}^p A \cap \mathbb{Z}^n = \mathbb{Z}^p V',$$

which completes the proof. □

As a rephrasing of the previous proposition we obtain the following result which involves a notion of saturation.

Corollary 2.11 *For all $\mathbb{Z}$-simple sets $X = a + \sum_{i=1}^{p} \mathbb{Z}b_i$, the set*

$$\big(a + \sum_{i=1}^{p} \mathbb{Q}b_i\big) \cap \mathbb{Z}^n$$

is $\mathbb{Z}$-simple and has the same dimension as X.

2.2 Boolean closure

The previous results show that the $\mathbb{Z}$-linear subsets enjoy stronger closure properties than the $\mathbb{N}$-linear subsets but that they still fail to form a Boolean algebra. As the family of finite unions of $\mathbb{N}$-linear sets is closed under the Boolean operations, we may wonder whether or not so is the family of finite unions of $\mathbb{Z}$-linear subsets. This is not so and here is how it can be seen. Consider the singleton $\{0\}$ in $\mathbb{Z}$. Its complement cannot be expressed as a finite union of $\mathbb{Z}$-linear subsets. Indeed, such a finite union would consist of a finite subset of $\mathbb{Z}$ and, by Proposition 2.7, a union of cosets of a subgroup of the form $p\mathbb{Z}$, i.e., it would be cofinite if and only if it were equal to $\mathbb{Z}$, contradiction. So, we are led to consider the family of finite unions of differences of $\mathbb{Z}$-semilinear sets.

Definition 2.12 *We denote by $\mathcal{F}$ the family of finite unions of subsets of the form $A \setminus B$ where A and B are $\mathbb{Z}$-semilinear sets.*

We shall see that $\mathcal{F}$ is the Boolean closure of the $\mathbb{Z}$-linear subsets. In the meantime, we show that these subsets may be written in different relatively simple ways.

Proposition 2.13 *Let $X \subseteq \mathbb{Z}^n$. Then the following conditions are equivalent:*

(i) *the set X belongs to $\mathcal{F}$;*

(ii) *the set X is a finite union of subsets of the form*

$$\bigcap_{i=1}^{m} (S_i \setminus T_i), \tag{6}$$

where $T_1, \ldots, T_m$ and $S_1, \ldots, S_m$ are $\mathbb{Z}$-simple sets and $T_i \subseteq S_i$ for $1 \le i \le m$;

(iii) *the set X is a finite union of subsets of the form*

$$S \setminus \Big(\bigcup_{i=1}^{m} T_i\Big), \tag{7}$$

where $T_1, \ldots, T_m$ and S are $\mathbb{Z}$-simple sets and $T_i \subseteq S$ for $1 \le i \le m$;

(iv) *the set X is a finite union of subsets of the form*

$$S \setminus \Big(\bigcup_{i=1}^{m} T_i\Big), \tag{8}$$

where $T_1, \ldots, T_m$ and S are $\mathbb{Z}$-simple sets and $\dim(T_i) < \dim(S)$, for $1 \le i \le m$.

Proof: We prove that condition **(i)** implies **(ii)**. It suffices to start with a subset of the form

$$\Big(\bigcup_{i=1}^{l} S_i\Big) \setminus \Big(\bigcup_{j=1}^{m} T_j\Big) \tag{9}$$

where the sets $S_1,\dots,S_l,T_1,\dots,T_m$ are $\mathbb{Z}$-simple. Applying the three rules of set differences $\big(X\cup Y\big)\setminus Z=\big(X\setminus Z\big)\cup\big(Y\setminus Z\big)$, $X\setminus\big(Y\cup Z\big)=\big(X\setminus Y\big)\cap\big(X\setminus Z\big)=\big(X\setminus Y\big)\setminus Z$ and $X\setminus\big(Y\cap X\big)=\big(X\setminus Y\big)$ the subset defined in (9) becomes

$$\bigcup_{i=1}^{l}\Big(S_i\setminus(\bigcup_{j=1}^{m}T_j)\Big)=\bigcup_{i=1}^{l}\Big(\bigcap_{j=1}^{m}(S_i\setminus T_j)\Big)=\bigcup_{i=1}^{l}\bigg(\bigcap_{j=1}^{m}\Big(S_i\setminus(T_j\cap S_i)\Big)\bigg),$$

so the first implication is proved.
Now let us prove that condition **(ii)** implies condition **(iii)**. It suffices to start with a subset of the form

$$\bigcap_{i=1}^{m}(S_i\setminus T_i) \tag{10}$$

where the sets $S_1,\dots,S_m,T_1,\dots,T_m$ are $\mathbb{Z}$-simple and $T_i\subseteq S_i$ for all $1\le i\le m$. Then the set in (10) can be written as

$$\Big(\bigcap_{i=1}^{m}S_i\Big)\cap\Big(\bigcap_{i=1}^{m}\overline{T_i}\Big)=\Big(\bigcap_{i=1}^{m}S_i\Big)\setminus\Big(\bigcup_{i=1}^{m}T_i\Big)=\Big(\bigcap_{i=1}^{m}S_i\Big)\setminus\Big(\bigcup_{i=1}^{m}T_i'\Big),$$

where $T_i'=T_i\cap\big(\bigcap_{i=1}^{m}S_i\big)$. Observe that the set $\big(\bigcap_{i=1}^{m}S_i\big)$ is $\mathbb{Z}$-simple by Theorem 2.6 and that so is each of the sets T_i' for the same reason.
To prove the implication from condition **(iii)** to condition **(iv)**, it suffices to show that each $\mathbb{Z}$-linear subsets T_i may be furthermore assumed of dimension less than s where $s=\dim(S)$. Indeed, after possible renumbering of the sets T_i, we may assume that the p first of them are of dimension s and that the remaining are of lesser dimension. Write

$$S\setminus\Big(\bigcup_{i=1}^{m}T_i\Big)=((\dots(S\setminus T_1)\setminus\dots)\setminus T_p)\setminus\Big(\bigcup_{i=p+1}^{m}T_i\Big). \tag{11}$$

By Proposition 2.8 the set i $\big(\big(\dots\big(S\setminus T_1\big)\setminus\dots\big)\setminus T_p\big)$ is a finite union of $\mathbb{Z}$-simple set of dimension s, say $S_1,\dots,S_r$. Since the set difference distributes over the union the above expression (11) is the union over i $j=1,\dots,r$ of the subsets

$$S_j\setminus\Big(\bigcup_{i=p+1}^{m}(T_i\cap S_j)\Big)$$

which completes the verification of this implication.
The implication from condition **(iv)** to condition **(i)** is trivial. □

We are now in a position to prove the main result of this section.

Theorem 2.14 *The family $\mathcal{F}$ is an effective Boolean algebra, i.e., given two subsets $X,Y\subseteq\mathbb{Z}^n$ specified as in Definition 2.12, there exists a procedure that computes the specification of the subsets $X\cup Y$, $X\cap Y$ and $\mathbb{Z}^n\setminus X$.*

Proof: Since $\mathcal{F}$ is defined as a collection of finite unions of sets, in order to prove the closure under complement we first have to prove the closure under intersection. Remembering that intersection distributes over union, it suffices to consider intersections of the form $P \cap Q$ with

$$P = S \setminus \Big(\bigcup_{i=1}^{l} T_i\Big) \quad \text{and} \quad Q = S' \setminus \Big(\bigcup_{j=1}^{m} T'_j\Big),$$

where all the sets $S, S', T_1, \ldots, T_l, T'_1, \ldots, T'_m$ are $\mathbb{Z}$-simple. Applying the general set theoretical equivalence $(X \setminus Y) \cap (Z \setminus T) = (X \cap Z) \setminus (Y \cup T)$, we obtain

$$P \cap Q = (S \cap S') \setminus \Big(\Big(\bigcup_{i=1}^{l} T_i\Big) \cup \Big(\bigcup_{j=1}^{m} T'_j\Big)\Big),$$

and then applying Theorem 2.6 to $S \cap S'$, we get the required form for the intersection $P \cap Q$. Concerning the complement, and due to the fact that $\mathcal{F}$ is trivially closed under union and, as we have proved, also intersection, is suffices to consider the complement of the set

$$P = S \setminus \Big(\bigcup_{i=1}^{l} T_i\Big),$$

defined as above. Then its complement $\overline{P}$ can be equivalently written as

$$\overline{P} = \Big(\overline{S} \cup \bigcup_{i=1}^{l} T_i\Big) = \bigcup_{i=1}^{l} (T_i \setminus S),$$

namely in the required form.

Effectiveness follows from Theorem 2.6 since intersection between $\mathbb{Z}$-simple sets is the only real operation we have to perform. □

We can interpret the previous result as a bound for nested Boolean operations, in the sense that no matter what Boolean combination of $\mathbb{Z}$-simple sets, it can be always written in such way that its derivation tree has no more than three levels.

For our decision procedure we will need more detailed information about Boolean operations between sets belonging to the family $\mathcal{F}$. In particular we can be more precise when simple subsets of maximal dimension are involved.

Proposition 2.15 *Let X a subset of $\mathbb{Z}^n$ such that*

$$X = \bigcup_{i=1}^{m} \Big(S_i \setminus T_i\Big), \tag{12}$$

where the $S_1, \ldots, S_m$ are $\mathbb{Z}$-simple of maximal dimension n, each $T_1, \ldots, T_m$ belongs to the family $\mathcal{F}$ and is included into a finite union of $\mathbb{Z}$-simple of dimension strictly less than n. Moreover if we have that $T_i \subseteq S_i$ for all $1 \leq i \leq n$, then X can be written in the form

$$\Big(\bigcup_{i=1}^{m} S_i\Big) \setminus T,$$

where T belongs to the family $\mathcal{F}$ and is included in a finite union of $\mathbb{Z}$-simple sets of dimension strictly less than n.

Observe that we may not assume that T is a finite union of simple sets of dimension less than n, even if this is true for all sets T_i. Indeed, consider for example the set X defined as

$$X = (S_1 \setminus T_1) \cup (S_2 \setminus T_2),$$

where

$$\begin{array}{llll} S_1 & = \mathbb{Z}^2, & T_1 & = \mathbb{Z} \times \{0\}, \\ S_2 & = 2\mathbb{Z} \times \mathbb{Z}, & T_2 & = \{0\} \times \mathbb{Z}. \end{array}$$

Then

$$X = \mathbb{Z}^2 \setminus \big((\mathbb{N} \setminus \{0\}) \times \{0\} \cup (-\mathbb{N} \setminus \{0\}) \times \{0\}\big) = \mathbb{Z}^2 \setminus \big((\mathbb{Z} \setminus \{0\}) \times \{0\}\big),$$

and we already know that the set $\big((\mathbb{Z} \setminus \{0\}) \times \{0\}\big)$ is not $\mathbb{Z}$-semilinear.

Proof: Without loss of generality we can suppose $n = 2$. Then from set theoretical rules, we have the equivalence

$$\big(S_1 \setminus T_1\big) \cup \big(S_2 \setminus T_2\big) = \big(S_1 \cup S_2\big) \setminus \Big(\big(\overline{S_2} \cap T_1\big) \cup \big(\overline{S_1} \cap T_2\big) \cup \big(T_1 \cap T_2\big)\Big).$$

From Proposition 2.7 it follows that $S_1 \cup S_2$ is a $\mathbb{Z}$-quasi-simple set of maximal dimension n. From Proposition 2.8, the subsets $\overline{S_1}$ and $\overline{S_2}$ are again $\mathbb{Z}$-quasi-simple. Finally from Theorem 2.6 it follows that the union of the last three terms of the above expression is included into a finite union of $\mathbb{Z}$-simple sets of dimension less than n and moreover belongs to $\mathcal{F}$ as an immediate consequence of Theorem 2.14. □

3 The weak Presburger arithmetic

3.1 Definable sets in the weak Presburger arithmetic

We are now in a position to prove the main result of this section which gives an algebraic characterization of the $\mathcal{Z}^W$-definable sets.

Theorem 3.1 *The family $\mathcal{F}$ coincides with the family of $\mathcal{Z}^W$-definable sets.*

Proof: Since $\mathcal{F}$ is the Boolean closure of the family of $\mathbb{Z}$-simple sets, in order to prove that each element of $\mathcal{F}$ is $\mathcal{Z}^W$-definable, it suffices to prove that an arbitrary $\mathbb{Z}$-simple set

$$X = a + \sum_{i=1}^{m} \mathbb{Z} b_i \subseteq \mathbb{Z}^n$$

is $\mathcal{Z}^W$-definable. Denote by a_j and b_{ij} respectively the j-th component of the vectors a and b_i. For simplicity, identify assignments of variables with variables themselves. Then $(x_1, \ldots, x_n)$ belongs to X if and only if the variables $x_1, \ldots, x_n$ satisfy the formula φ with

$$\varphi(x_1, \ldots, x_n) \equiv \exists y_1 \ldots \exists y_m \bigwedge_{j=1}^{n} x_j = a_j + y_1 b_{1j} + \cdots + y_m b_{mj}.$$

We now prove the converse. To that order, we use Presburger's elimination of quantifiers which asserts that every formula $\varphi(x_1, \ldots, x_n)$ with n free variables is equivalent to a Boolean combination of formulas of the form $\sum_{i=1}^{n} a_i x_i + b = 0$ and of the form $\sum_{i=1}^{n} a_i x_i \equiv_m b$, with $a_i, b \in \mathbb{Z}$. Concerning the first type of predicate, Proposition 2.5 where the matrix A is reduced to a unique row, guarantees that the set of solutions is linear. As for the second type, we prove by induction on n that it defines a $\mathbb{Z}$-semilinear set. For $n = 1$, we get $ax \equiv_m b$. If the equation $ax = b$ has no solution in the finite cyclic group $\mathbb{Z}/m\mathbb{Z}$, then the relation defined by φ is empty. Otherwise, let $0 \leq c_1 < \cdots < c_s < m$ be the set of solutions in $\mathbb{Z}/m\mathbb{Z}$ (we identify the nonnegative integers less than m with their natural image in the group $\mathbb{Z}/m\mathbb{Z}$). Then the formula φ defines the subset of integers x such that $x = c + m\mathbb{Z}$, for some $c \in \{c_1, \ldots, c_s\}$, i.e., the $\mathbb{Z}$-semilinear set

$$\bigcup_{i=1}^{s} (c_i + \mathbb{Z}m)\,.$$

Now let $n > 1$. The condition $\sum_{i=1}^{n} a_i x_i \equiv_m b$ is equivalent to the disjunction

$$\bigvee_{j=0}^{m-1} \Big((\sum_{i=1}^{n-1} a_i x_i \equiv_m j) \wedge a_n x_n \equiv_m (b - j) \Big).$$

By induction hypothesis, the subformula $\sum_{i=1}^{n-1} a_i x_i \equiv_m j$ defines a $\mathbb{Z}$-semilinear subset, say $T_j \subseteq \mathbb{Z}^{n-1}$, and $a_n x_n \equiv_m (b-j)$, as proved above, another $\mathbb{Z}$-semilinear subset, say $R_j \subseteq \mathbb{Z}$. Because of Proposition 2.2, the family of finite union of $\mathbb{Z}$-simple sets is closed under direct product, so the relation is equal to

$$\bigcup_{j=0}^{m-1} \big(T_j \times R_j\big).$$

$\square$

The last result asserts, loosely speaking, that a Presburger (resp. weak Presburger) definable subset of $\mathbb{Z}^n$ included in a simple set, is a Presburger (resp. weak Presburger) definable subset of that subspace.

Proposition 3.2 *Let $X = a + \sum_{i=1}^{p} \mathbb{Z}b_i \subseteq \mathbb{Z}^n$ be a simple set and let $\tau : X \to \mathbb{Z}^p$ be the mapping of X onto $\mathbb{Z}^p$ which assigns to each $a + \sum_{i=1}^{p} \lambda_i b_i$ the element $(\lambda_1, \ldots, \lambda_p) \in \mathbb{Z}^p$.*
Then a subset Y of X is $\mathcal{Z}$-definable (resp. $\mathcal{Z}^W$-definable) if and only if the set $\tau(Y)$ is $\mathcal{Z}$-definable (resp. $\mathcal{Z}^W$-definable).

Proof: Observe that the function τ is well-defined since the vectors b_i are linearly independent. Now, if $\varphi(x_1, \ldots, x_n)$ is a $\mathcal{Z}$- (resp. $\mathcal{Z}^W$-) formula defining a subset $Y \subseteq X$, then the subset $\tau(Y)$ is defined by the $\mathcal{Z}$- (resp. $\mathcal{Z}^W$-) formula

$$\exists x_1 \ldots \exists x_n \left(\varphi(x_1, \ldots, x_n) \wedge \bigwedge_{j=1}^{n} \big(x_j = a_j + \sum_{i=1}^{p} b_{ij} y_j\big) \right),$$

where a_j and b_{ij} represent the j-th component of the vectors a and b_i.
Conversely, if for some subset Y of X the set $\tau(Y)$ is definable by some $\mathcal{Z}$- (resp. $\mathcal{Z}^W$-) formula

$\psi(y_1, \ldots, y_p)$, then Y is definable by the $\mathcal{Z}$- (resp. $\mathcal{Z}^W$-) formula

$$\exists y_1 \ldots \exists y_p \left(\psi(y_1, \ldots, y_p) \wedge \bigwedge_{j=1}^{n} \left(x_j = a_j + \sum_{i=1}^{p} b_{ij} y_j\right) \right).$$

□

3.2 Decidability of weak Presburger logic in Presburger logic

We recall that the problem of deciding whether or not a relation definable in Büchi arithmetic can be actually defined in Presburger arithmetic has been positively answered in (18). Here instead of considering Presburger arithmetic with a superstructure, we consider it with a substructure: we prove that it is decidable whether or not a given formula of the structure $\langle \mathbb{Z}; +, <, 0, 1\rangle$ is equivalent to some formula in the structure $\langle \mathbb{Z}; +, 0, 1\rangle$.

The procedure for deciding weak Presburger definability proceeds by induction on the dimension of the subsets. To that purpose, we introduce the following notation. Let X and Y be two subsets of $\mathbb{Z}^n$. If $\dim(X) = \dim(Y) = k > 0$, we write $X \sim Y$ if the symmetric difference $X \Delta Y = X \setminus Y \cup Y \setminus X$ has dimension less than k. Otherwise, i.e., if $\dim(X), \dim(Y) \leq 0$, we put $X \sim Y$ if and only if $X = Y$. The relation is clearly an equivalence relation and plays an important role in our decision procedure.

3.2.1 A special case

Our decision procedure relies on the following structural characterization of the $\mathcal{Z}$-definable subsets of maximal dimension. The general case, i.e., when no assumption is made on the dimension, is, as we shall see, a reduction to this special case. Therefore, the following theorem can be considered as the crux of the algorithm. Given an $\mathbb{N}$-simple set $S = a + \sum_{i=1}^{m} \mathbb{N} b_i$ we denote by $S^{\mathbb{Z}}$ the $\mathbb{Z}$-simple set $a + \sum_{i=1}^{m} \mathbb{Z} b_i$ and we extend this notation to finite unions of $\mathbb{N}$-simple sets. Also we denote by $S^{\mathbb{Q}}$ the $\mathbb{Z}$-simple sets $\left(a + \sum_{i=1}^{m} \mathbb{Q} b_i\right) \cap \mathbb{Z}^n$, cf. Corollary 2.11.

Theorem 3.3 *Let $X \subseteq \mathbb{Z}^n$ be a $\mathcal{Z}$-definable set,*

$$X = T \cup \bigcup_{i=1}^{m} Y_i,$$

where the sets Y_i are $\mathbb{N}$-simple sets of dimension n and where T is a finite union of $\mathbb{N}$-simple sets of dimension less than n.
Set $P = \bigcup_{i=1}^{m} Y_i^{\mathbb{Z}}$. Then X is $\mathcal{Z}^W$-definable if, and only if, the sets $X \setminus P$ and $P \setminus X$ are $\mathcal{Z}^W$-definable subsets of dimension less than n.

Proof: The subset P is clearly $\mathcal{Z}^W$-definable. Now we have

$$X = \big(X \setminus P\big) \cup \big(P \setminus (P \setminus X)\big) \tag{13}$$

which shows that the condition is sufficient. Let us prove that it is necessary.
If X is $\mathcal{Z}^W$-definable so are $X \setminus P$ and $P \setminus X$. Since (13) always holds, it suffices to prove that $X \setminus P$ and $P \setminus X$ are of dimension less than n and since $X \setminus P \subseteq T$ holds, we prove that $P \setminus X$ has dimension

less that n. Because of propositions 2.13 and 2.15, we can isolate all the simple sets of dimension n and write

$$X = Q \cup \big((A + \sum_{j=1}^{n} \mathbb{Z}d_j) \setminus R\big)$$

where $A \subseteq \mathbb{Z}^n$ is finite, the vectors d_j are linearly independent and Q and R are $\mathcal{Z}^W$-definable sets of dimension less than n. For all $i = 1, \dots, m$, set $Y_i = a^{(i)} + \sum_{j=1}^{n} \mathbb{N}b_j^{(i)}$ where the $b_j^{(i)}$ are linearly independent. Now we prove

$$\bigcup_{i=1}^{m} Y_i^{\mathbb{Z}} = \bigcup_{i=1}^{m} \big(a^{(i)} + \sum_{j=1}^{n} \mathbb{Z}b_j^{(i)}\big) \sim \big(A + \sum_{j=1}^{n} \mathbb{Z}d_j\big). \tag{14}$$

Clearly, if we take a vector in $A + \sum_{j=1}^{n} \mathbb{Z}d_j$ not belonging to $R \cup T$ (which is contained in $R \cup T^{\mathbb{Z}}$, that is, a finite union of $\mathbb{Z}$-simple sets of dimension less than n), it must belong to one of the Y_i, and thus to $\bigcup_{i=1}^{m} \big(a^{(i)} + \sum_{j=1}^{n} \mathbb{Z}b_j^{(i)}\big)$.
Conversely, consider one of the sets Y_i, namely $Y = a + \sum_{j=1}^{n} \mathbb{N}b_j$ (we drop the upper indices to simplify the notation). We observe that Q is in particular $\mathcal{Z}$-definable, so $Q = \bigcup_{1 \le i \le \ell} Q_i$, where every Q_i is a $\mathbb{N}$-simple set of dimension less than n. Set

$$U = \bigcup_{1 \le i \le \ell} Q_i^{\mathbb{Z}}$$

(recall the definition of $Q^{\mathbb{Z}}$ before the theorem) which, by Corollary 2.11, is a finite union of simple sets of dimension less than n.

Now we show by induction on j that the set

$$W_j = (a + \mathbb{Z}b_1 + \cdots + \mathbb{Z}b_{j-1} + \mathbb{N}b_j + \cdots + \mathbb{N}b_n) \setminus U$$

is included in $A + \sum_{j=1}^{n} \mathbb{Z}d_j$. The case $j = 1$ is obvious, so we suppose $1 < j \le n$. By induction hypothesis, for every $h < j$ we have $W_h \subseteq \big(A + \sum_{j=1}^{n} \mathbb{Z}d_j\big)$. Consider a vector $v \in W_j$ and observe that the line $v + \mathbb{N}b_j$ intersects a subset $Q_i^{\mathbb{Z}}$ in at most one point so that its intersection with U is a finite set. Indeed, set $Q_i = c + \sum_k \mathbb{N}e_k$ and assume that for two integers $\alpha, \beta \in \mathbb{Z}$ we have that $v + \alpha b_j$ and $v + \beta b_j$ belong to $c + \sum_k \mathbb{Q}e_k$. This implies that b_j belongs to the $\mathbb{Q}$-vector space generated by the vectors e_k, so that $v \in \big(c + \sum_k \mathbb{Q}e_k\big) \cap \mathbb{Z}^n = Q_i^{\mathbb{Z}}$, contradiction.
Consequently, for every sufficiently large integer s, we have $v + sb_j = a_s + \sum_{i=1}^{n} \lambda_s^{(i)} d_i$, where $a_s \in A$ and $\lambda_s^{(i)} \in \mathbb{Z}$. Since A is finite, so there exists $s_1 \in \mathbb{N}$ and $0 \le r \le |A|$ such that $a_{s_1} = a_{s_1+r}$. By computing $v + (s_1 + r)b_j - (v + s_1 b_j)$, we obtain $rb_j = \sum_{i=1}^{n} \mu_r^{(i)} d_i$, where $\mu_r = \lambda_{s_1+r} - \lambda_{s_1}$. Now for all integers s let $m \in \mathbb{Z}$ and $0 \le r' \le r$ be such that $s = s_1 + mr + r'$, then we have

$$\begin{aligned} v + sb_j &= v + ((s_1 + r') + mr)b_j \\ &= a_{s_1+r'} + \textstyle\sum_{i=1}^{n} (\lambda_{s_1+r'}^{(i)} d_i) + m \sum_{i=1}^{n} (\mu_r^{(i)} d_i) \in \big(A + \sum_{j=1}^{n} \mathbb{Z}d_j\big). \end{aligned}$$

Now we have $X \sim (A + \sum_{j=1}^{n} \mathbb{Z}d_j) \sim P$ which completes the proof. □

3.2.2 The procedure

The procedure solves the following decision problem:

Input: a $\mathcal{Z}$-definable set X given as a finite union of $\mathbb{N}$-simple sets

Question: decide whether or not the set X is $\mathcal{Z}^W$-definable and, in the affirmative case, give a representation as a Boolean combination of $\mathbb{Z}$-simple sets.

We cannot directly use Theorem 3.3, because it requires that the finite union has an $\mathbb{N}$-simple subset of maximal dimension. We thus proceed as follows. The set $X \subseteq \mathbb{Z}^n$ is given as a union of sets of the form

$$X_i = a^{(i)} + \sum_{j=1}^{r_i} \mathbb{N} b_j^{(i)}, \qquad 1 \le i \le m. \tag{15}$$

We use an induction on $\dim(X) = \max\{r_i \mid 1 \le i \le m\}$. When this dimension is at most 0 the answer is "yes". Assume thus that $d = \dim(X) > 0$. Set $H_i = X_i^{\mathbb{Q}}$ for $1 \le i \le m$ and observe that it is $\mathbb{Z}$-simple by Corollary 2.11, thus $\mathcal{Z}^W$-definable by Theorem 3.1. Suppose, after possibly changing some indices, that $H_1, \ldots, H_p$, $p \le m$, are the maximal elements for the subset inclusion, of the collection of sets $\{H_i \mid 1 \le i \le m\}$. We claim that X is $\mathcal{Z}^W$-definable if and only if every intersection $X \cap H_i$, $1 \le i \le p$, is $\mathcal{Z}^W$-definable. Indeed, this is clearly necessary since H_i is $\mathcal{Z}^W$-definable. Conversely, if every $X \cap H_i$ is $\mathcal{Z}^W$-definable, because each X_j is a subset of H_i for some $1 \le i \le p$, i.e, because of the inclusion $X \subseteq \bigcup_{i=1}^p H_i$ holds, we have

$$X = X \cap \Big(\bigcup_{i=1}^{p} H_i\Big) = \bigcup_{i=1}^{p} \big(X \cap H_i\big).$$

For each H_i of dimension less than d, we call the procedure recursively. If one of these calls returns "no", then the procedure returns "no" and stops . Now, for each H_i of dimension d we call the procedure recursively as described below and we return "yes" if all these calls return "yes". We now explain how we treat each H_i, denoted H for simplification. Write $H = a + \sum_{j=1}^d \mathbb{Z} c_j$ where the c_js are a set of vectors generating H freely. Let $\tau : H \to \mathbb{Z}^d$ be the mapping which assigns the d-tuple $(\lambda_1, \ldots, \lambda_d) \in \mathbb{Z}^d$ to the element $a + \sum_{j=1}^d \lambda_j c_j$. By Proposition 3.2, for all $Y \subseteq H$ we have: Y is $\mathcal{Z}^W$-definable if, and only if, $\tau(Y)$ is $\mathcal{Z}^W$-definable. Let I be the set of indices $1 \le j \le m$ such that $X_j \subseteq H$ has dimension d and let T be the union of all other X_js included in H. Then we have

$$\tau(X \cap H) = \tau(T) \cup \bigcup_{j \in I} \tau(X_j),$$

so that every $\tau(X_j)$ for $j \in I$, has maximal dimension in $\tau(H) = \mathbb{Z}^d$. Then we are in the conditions of Theorem 3.3 and it suffices to check the equivalence

$$\bigcup_{j \in I} \tau(X_j) \sim \bigcup_{j \in I} \tau(X_j)^{\mathbb{Z}}$$

which is decidable by Proposition 1.5 since it involves Boolean operations on $\mathbb{N}$-semilinear subsets. If the equivalence does not hold, then the procedure returns "no" and stops. Otherwise we call the procedure recursively with $\tau(X \cap H) \setminus P$ and $P \setminus (P \setminus \tau(X \cap H))$ where $P = \bigcup_{i \in I} \tau(X_i)^{\mathbb{Z}}$.

Considering Ginsburg and Spanier's construction, the previous procedure has the following consequence.

Theorem 3.4 *Given a first-order formula over* $\langle \mathbb{Z}; +, <, 0, 1\rangle$ *it is recursively decidable whether or not it is expressible in* $\langle \mathbb{Z}; +, 0, 1\rangle$.

4 Further works

We have presented an algorithm based on an algebraic characterization of the sets of integers that can be expressed in first order logic with the binary function of sum and the constants 0 and 1. We did not tackle the problem of evaluating the complexity of the algorithm (which is probably nonelementary as written) nor that of the problem. Concerning the former, we think that even more elementary issues should be addressed first relative to the constructions of Ginsburg and Spanier to which we have alluded at several instances. Little has been undertaken since their publications. For example, Huynh proved in (13) that inequality of two finite unions of $\mathbb{N}$-linear sets is log-complete in Σ_2^p, but other questions remain unsettled: the intersection of two $\mathbb{N}$-linear sets is a finite union of $\mathbb{N}$-linear sets, but we ignore the size of the output as a function of the size of the two $\mathbb{N}$-linear sets. This leaves room for further research which might simplify or improve the present results.

References

[1] Paul Bernays and David Hilbert. *Grundlagen der Mathematik I*, chapter 7, pages 368–377. Springer-Verlag, Berlin, 2nd edition, 1970.

[2] Alexis Bès. A survey of arithmetical definability. *Bullettin of the Belgian Mathematical Society. Simon Stevin*, suppl.:1–54, 2001.

[3] Véronique Bruyère, Emmanuel Dall'Olio, and Jean-François Raskin. Durations, parametric model-checking in timed automata with presburger arithmetic. In *STACS*, pages 687–698, 2003.

[4] Christian Choffrut. Deciding whether a relation defined in Presburger logic can be defined in weaker logics. *Theoretical Informatics and Applications*, 42:121–135, 2008.

[5] Christian Choffrut and Achille Frigeri. Definable sets in weak Presburger arithmetic. In G. F. Italiano et al., editor, *Theoretical Computer Science,* 10^{th} *Italian Conference, ICTCS 2007*, pages 175–186, 2007.

[6] Henri Cohen. *A Course in Computational Algebraic Number Theory*, volume 138 of *Graduate Texts in Mathematics*. Springer-Verlag, 1993.

[7] Samuel Eilenberg and Marcel-Paul Schützenberger. Rational sets in commutative monoids. *Journal of Algebra*, 13:173–191, 1969.

[8] Herbert B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, New York, 1972.

[9] Seymour Ginsburg and Edwin H. Spanier. Bounded ALGOL-like languages. *Transactions of the American Mathematical Society*, 113:333–368, 1964.

[10] Seymour Ginsburg and Edwin H. Spanier. Bounded regular sets. *Proceedings of the American Mathematical Society*, 17:1043–1049, 1966.

[11] Seymour Ginsburg and Edwin H. Spanier. Semigroups, Presburger formulas, and languages. *Pacific Journal of Mathematics*, 16:285–296, 1966.

[12] Seymour Ginsburg and Edwin H. Spanier. AFL with the semilinear property. *Journal of Computer and System Science*, 5:365–396, 1971.

[13] T.-D. Huynh. The complexity of semilinear sets. In *ICALP*, pages 324–337, 1980.

[14] Ryuichi Ito. Every semilinear set is a finite union of disjoint linear sets. *Journal of Computer and System Sciences*, 3:221–231, 1969.

[15] Manolis Koubarakis. Complexity results for first-order theories of temporal constraints. In Jon Doyle et al., editor, *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*, pages 379–390, 1994.

[16] Serge Lang. *Algebra*. Addinson Wesley, 1965.

[17] Jérôme Leroux. A polynomial time Presburger criterion and synthesis for number decision diagrams. In *Proceedings of the* 20^{th} *IEEE Symposium on Logic in Computer Science*, pages 147–156, 2005.

[18] Andrei Al'bertovich Muchnik. The definable criterion for definability in Presburger arithmetic and its applications. *Theoretical Computer Science*, 290:1433–1444, 2003.

[19] Mojżesz Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Sprawozdanie z I Kongresu matematyków krajów słowiańskich, Warszawa 1929 (Comptes-rendus du I Congrés des Mathématiciens des Pays Slaves, Varsovie 1929)*, pages 92–101, and an unnumbered addendum after page 394, Warsaw, 1930.

[20] Craig Smoryński. *Logical Number Theory I*, chapter III, pages 307–329. Universitext. Springer-Verlag, Berlin, 1991.

[21] Ryan Stansifer. Presburger's article on integer arithmetic: Remarks and translation. Technical Report TR84–639, Cornell University, Computer Science Department, `http://ecommons.library.cornell.edu/ handle/1813/6478`, September 1984.

[22] Arne Storjohann. Near optimal algorithms for computing Smith normal forms of integer matrices. In Y. N. Lakshman, editor, *Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation, ISSAC '96*, pages 1–8. ACM Press, 1996.