On the power of ordering in linear arithmetic theories

Dmitry Chistikov
Centre for Discrete Mathematics and its Applications (DIMAP) &
Department of Computer Science, University of Warwick, Coventry, UK
d.chistikov@warwick.ac.uk

Christoph Haase
Department of Computer Science, University College London, London, UK
c.haase@ucl.ac.uk

Abstract

We study the problems of deciding whether a relation definable by a first-order formula in linear rational or linear integer arithmetic with an order relation is definable in absence of the order relation. Over the integers, this problem was shown decidable by Choffrut and Frigeri [*Discret. Math. Theor. C.*, 12(1), pp. 21–38, 2010], albeit with non-elementary time complexity. Our contribution is to establish a full geometric characterisation of those sets definable without order which in turn enables us to prove coNP-completeness of this problem over the rationals and to establish an elementary upper bound over the integers. We also provide a complementary Π_2^{P} lower bound for the integer case that holds even in a fixed dimension. This lower bound is obtained by showing that universality for ultimately periodic sets, i.e., semilinear sets in dimension one, is Π_2^{P}-hard, which resolves an open problem of Huynh [*Elektron. Inf.verarb. Kybern.*, 18(6), pp. 291–338, 1982].

2012 ACM Subject Classification Theory of computation → Automated reasoning

Keywords and phrases logical definability, linear arithmetic theories, semi linear sets, ultimately periodic sets, numerical semigroups

Related Version An extended version of the paper including full proofs is available from the authors' webpages.

Funding *Christoph Haase*: This work is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 852769, ARiAT).

1 Introduction

A central topic in mathematical and computational logic is to investigate the expressive power of first-order formulas in a given structure. Notable results in this context include Robinson's seminal work showing that the integers are first-order definable in the structure $\langle \mathbb{Q}, +, \cdot \rangle$, thus rendering its first-order theory undecidable [13]. Another example is the celebrated theorem of Muchnik [10] showing decidability of the problem of determining whether a relation first-order definable in the structure $\langle \mathbb{Z}, +, V_k, < \rangle$ (whose first-order theory is known as *Büchi arithmetic*) is definable in the weaker substructure $\langle \mathbb{Z}, +, < \rangle$ (known as *linear integer arithmetic* or *Presburger arithmetic*). Here, V_k is the function mapping an integer to the highest power of k dividing it without remainder. Muchnik's approach yields a quadruply exponential time algorithm for this problem when the relation is given as a deterministic finite-state automaton; a polynomial-time algorithm was later claimed by Leroux [9]. It has recently been shown that this problem can be solved in quasi-linear time for unary relations [1].

In this paper, we investigate the computational complexity of the *weak definability problem*, which is the problem of deciding whether a relation first-order definable in the structures

$\langle \mathbb{Q}, +, < \rangle$ or $\langle \mathbb{Z}, +, < \rangle$ is definable in its weak counterpart, which is obtained from replacing the order relation with the equality relation. It follows from elementary model-theoretic arguments that such *weak linear arithmetic theories* are strictly less expressive compared to their counterparts that include the order relation, since $h(x) := -x$ is an automorphism in structures without order but fails to be one in the presence of the order relation. For Presburger arithmetic, the weak definability problem was shown decidable by Choffrut and Frigeri, albeit with non-elementary time complexity [5]. To the best of the authors' knowledge, no decidability results on weak definability for the structure $\langle \mathbb{Q}, +, < \rangle$ are known. Following [5], *weak Presburger arithmetic* (alternatively *weak linear integer arithmetic)* refers to the first-order theory of $\langle \mathbb{Z}, +, = \rangle$ whereas *weak linear rational arithmetic* refers to the theory of $\langle \mathbb{Q}, +, = \rangle$.

The main contribution of this paper is to significantly improve existing algorithmic upper bounds for the weak definability problem by establishing elementary upper bounds over both $\mathbb{Q}$ and $\mathbb{Z}$: we show that weak definability is in coNP for rational linear arithmetic, and decidable in elementary time for Presburger arithmetic. To this end, we develop geometric criteria that precisely characterize when a relation is definable without order. We also establish complementary lower bounds. While a matching coNP lower bound in the rational case is easy to obtain, we furthermore establish a Π_2^{P} lower bound for the weak definability problem for Presburger arithmetic that holds even in a fixed dimension. This lower bound is obtained by establishing Π_2^{P}-completeness of the universality problem for ultimately periodic sets, i.e., semi-linear sets in dimension one. This positively answers one of the longest-standing open problems in the theory of semi-linear sets posed by Huynh in 1982 [8], who asked whether the universality problem for semi-linear sets in dimension one is Π_2^{P}-complete. Our lower bound moreover strengthens previously known Π_2^{P} lower bounds for the inclusion problems for linear sets which have only been obtained in recent years, see [4, Thm. 12] and [16].

2 The weak definability problem

Everywhere in this paper, we denote by $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{N}$ and $\mathbb{N}_+$ the rational numbers, integers, natural numbers including zero and the positive integers, respectively. Given $a, b \in \mathbb{Z}$, we write $[a, b]$ for $\{a, a+1, \ldots, b\}$ and $[a]$ as a shortcut for $[1, a]$. Throughout this paper, numbers are assumed to be encoded in binary.

Linear arithmetic theories. Quantifier-free formulas ψ of the linear arithmetic theories we consider in this paper are obtained as Boolean combinations of linear constraints of the form $a_1 \cdot x_1 + \cdots + a_n \cdot x_n \sim b$, where all a_i and b are integer constants, the x_i are first-order variables, and $\sim$ is a relation symbol $<$ or $=$. If ψ only contains equality symbols it is a formula of a weak linear arithmetic theory, in which case we call ψ *weak*. We can always without loss of generality assume that ψ is in negation normal form, and if ψ is not weak, we can furthermore assume that ψ is negation-free. Formulas Φ of linear arithmetic theories additionally allow for quantification over first-order variables and are of the form $Q_1 x_1 \cdots Q_n x_n \, \psi$, where each Q_i is $\exists$ or $\forall$, and ψ is a quantifier-free formula. We write $\|\Phi\|$ to denote the maximum of two and the maximum of the absolute values of all constants occurring in Φ, and $|\Phi|$ to denote the length of Φ which is the number of symbols required to write down Φ.

The semantics of such formulas is given with respect to the structures of linear rational arithmetic $\langle \mathbb{Q}, +, < \rangle$ and linear integer arithmetic $\langle \mathbb{Z}, +, < \rangle$. We call Φ weak whenever ψ

is weak. Let $x_1, \dots, x_n$ be the free variables of Φ, and let $\mathbb{D} \subseteq \mathbb{Q}$; often $\mathbb{D}$ will be $\mathbb{Q}$ or $\mathbb{Z}$. We write $[\![\Phi]\!]_{\mathbb{D}} \subseteq \mathbb{D}^n$ for the set of all variable assignments making Φ a true sentence in the structure $\langle \mathbb{D}, +, < \rangle$. We may drop the subscript $\mathbb{D}$ when $\mathbb{D}$ is clear from the context.

Definable relations and weak definability. Fix $\mathbb{D}$ to be either $\mathbb{Q}$ or $\mathbb{Z}$. A relation $R \subseteq \mathbb{D}^n$ is called $\mathbb{D}$-definable whenever there is a linear arithmetic formula $\Phi(x_1, \dots, x_n)$ such that $(m_1, \dots, m_n) \in R$ if and only if $(m_1, \dots, m_n) \in [\![\Phi]\!]_{\mathbb{D}}$. In particular, we call R *weakly* $\mathbb{D}$*-definable* if the witnessing formula Φ is weak. The weak definability problem, the main decision problem considered in this paper, asks whether a relation definable by an arbitrary linear arithmetic formula is definable without order: *Given a linear arithmetic formula* Φ*, is the set* $[\![\Phi]\!]_{\mathbb{D}}$ *weakly* $\mathbb{D}$*-definable?*

Obviously, for a relation to be definable without order, it has to also be definable with order. In other words, weak definability, somewhat ironically, is a stronger property than definability.

▶ **Example 1.** Let $\Phi(x) := x < 0 \vee x \geq 1$. We observe that $[\![\Phi]\!]_{\mathbb{Q}}$ is not weakly $\mathbb{Q}$-definable. However, $[\![\Phi]\!]_{\mathbb{Z}}$ is weakly $\mathbb{Z}$-definable by $\Phi'(x) := \neg(x = 0)$, showing that weak $\mathbb{Z}$-definability does not imply weak $\mathbb{Q}$-definability. Conversely, weak $\mathbb{Q}$-definability does not imply weak $\mathbb{Z}$-definability either. Let $\Phi(x) := \exists y\, (x > 0 \wedge x = 2 \cdot y) \vee (x < 0 \wedge x = 3 \cdot y)$. Then $[\![\Phi]\!]_{\mathbb{Z}}$ is not weakly $\mathbb{Z}$-definable, but $[\![\Phi]\!]_{\mathbb{Q}}$ is weakly $\mathbb{Q}$-definable via the same $\Phi'(x)$ as above.

When studying the complexity of weak definability problems, it makes sense to restrict the input formulas we consider. On the one hand, due to quantifier elimination (Fourier-Motzkin and Presburger's quantifier elimination procedures, respectively), quantifier-free formulas of linear rational arithmetic can define all sets definable in linear rational arithmetic, and existential formulas those in linear integer arithmetic. On the other hand, satisfiability can be reduced to deciding weak definability, meaning that if we allow arbitrary formulas as input, deciding weak definability is at least as hard as deciding linear rational and integer arithmetic, respectively. This does, however, blur the inherent complexity of deciding weak definability. For these reasons, we restrict formulas in instances of the weak definability problem to the most restricted fragments that are still expressively complete:

Weak $\mathbb{Q}$-definability: Given a *quantifier-free* formula Φ of linear arithmetic, decide whether $[\![\Phi]\!]_{\mathbb{Q}}$ is weakly $\mathbb{Q}$-definable.

Weak $\mathbb{Z}$-definability: Given an *existential* quantifier-free formula Φ of linear arithmetic, decide whether $[\![\Phi]\!]_{\mathbb{Z}}$ is weakly $\mathbb{Z}$-definable.[1]

By establishing an analogue of semilinear sets characterizing sets definable in weak integer arithmetic, Choffrut and Frigeri [5] have shown that weak $\mathbb{Z}$-definability is decidable. As the main result of the present paper, we show that deciding weak $\mathbb{Q}$-definability is coNP-complete, and deciding weak $\mathbb{Z}$-definability is elementary and Π_2^{P}-hard.

3 Preliminaries

Linear algebra. We denote by $\boldsymbol{e}_i$ the i-th unit vector in any dimension. For $\boldsymbol{v} = (v_1, \dots, v_d) \in \mathbb{Z}^d$, we denote $\|\boldsymbol{v}\| := \max\{2, \max_{1 \leq i \leq d} |v_i|\}$, and for $V \subseteq \mathbb{Q}^d$ we set $\|V\| := \max_{\boldsymbol{v} \in V} \|\boldsymbol{v}\|$. Likewise, for a matrix $\boldsymbol{A}$, $\|\boldsymbol{A}\|$ is defined as the maximum over $\|\boldsymbol{v}\|$ for every column vector

[1] One could alternatively take quantifier-free formulas with additional divisibility predicates $c \mid \cdot$ for every $c \in \mathbb{N}_+$.

$\boldsymbol{v}$ of $\boldsymbol{A}$. We sometimes treat finite sets $V \subseteq \mathbb{Q}^m$ with n elements as $m \times n$ matrices (e.g., by ordering the vectors in V lexicographically), which we denote by $\boldsymbol{V}$. Given $A, B \subseteq \mathbb{Q}^d$, the Minkowski sum of A and B is defined as $A + B := \{a + b : a \in A,\ b \in B\}$. If A or B is a singletons, we omit set brackets and write, e.g., $a + B$ instead of $\{a\} + B$. Analogously, we define $A \cdot B := \{a \cdot b : a \in A,\ b \in B\}$. We write $A \triangle B$ for the symmetric difference of A and B, i.e., $A \triangle B := (A \setminus B) \cup (B \setminus A)$.

A set $A \subseteq \mathbb{Q}^d$ is an *affine subspace* if $A = \boldsymbol{a} + V$ for some $\boldsymbol{a} \in \mathbb{Q}^d$ and a linear subspace $V \subseteq \mathbb{Q}^d$. Affine subspaces are sometimes called flats. The affine hull of any set $V \subseteq \mathbb{Q}^d$ is the smallest affine subspace containing V and is equal to

$$\operatorname{aff} V := \left\{ \sum_{i=1}^{n} \lambda_i \cdot \boldsymbol{v}_i : n \in \mathbb{N},\ \boldsymbol{v}_i \in V,\ \lambda_i \in \mathbb{Q},\ i \in [n],\ \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$

The *dimension* of an affine subspace A, denoted $\dim A$, is defined as the dimension of the associated linear space A_0 such that $A = \boldsymbol{v} + A_0$ for some $\boldsymbol{v} \in \mathbb{Q}^d$. It is standard that this is well-defined. If affine subspaces A_1, A_2 are such that $A_1 \subseteq A_2$, then $\dim A_1 \leq \dim A_2$ (and moreover $A_1 = A_2$ iff $\dim A_1 = \dim A_2$).

A set S in $\mathbb{Q}^d$ (resp. $\mathbb{Z}^d$) *is vanishing (is zero, has measure zero) with respect to an affine subspace* $A \subseteq \mathbb{Q}^d$ if the set S is contained in a finite union of affine subpaces of A of dimension strictly less than A. Note that $S \subseteq A$ for any such S. For example, all finite subsets of $\mathbb{Q}^d$ have measure zero with respect to $\mathbb{Q}^d$ unless $d = 0$. If we do not specify the affine subspace A explicitly, $A = \mathbb{Q}^d$ is implicitly assumed.

The geometry of linear arithmetic. We recall some definitions and results from the literature on the geometry of solutions to systems of linear constraints.

It is well known that the set of solutions of a system of linear equations $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ over the rationals is an affine subspace. Conversely, every affine subspace of $\mathbb{Q}^d$ is the set of solutions to a system of (linearly independent) equations with integer coefficients. For a representation $A = \{\boldsymbol{x} \in \mathbb{Q}^d : \boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}\}$, we write $\|A\|$ to denote $\max\{\|\boldsymbol{B}\|, \|\boldsymbol{c}\|\}$, when $\boldsymbol{B}$ and $\boldsymbol{c}$ are determined by the context.

The Minkowski–Weyl theorem states that the set of points in a rational polyhedron $\boldsymbol{A} \cdot \boldsymbol{x} \geq \boldsymbol{b}$ can be represented as the sum of a bounded polytope and a convex cone, and vice versa. Given a finite set of vectors $V = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} \subseteq \mathbb{Q}^d$, we define

$$\operatorname{conv} V := \left\{ \sum_{i=1}^{n} \lambda_i \cdot \boldsymbol{v}_i : \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_i \in \mathbb{Q}_{\geq 0},\ i \in [n] \right\} \quad \text{and}$$

$$\operatorname{cone} V := \left\{ \sum_{i=1}^{n} \lambda_i \cdot \boldsymbol{v}_i : \lambda_i \in \mathbb{Q}_{\geq 0},\ i \in [n] \right\},$$

the *convex hull* of V and the *convex cone* generated by V, respectively. It follows that sets in $\mathbb{Q}^d$ definable in linear rational arithmetic can be obtained as finite unions of sets of the form $(\operatorname{conv} B + \operatorname{cone} P) \setminus A$, where A is a finite union of affine subspaces of $\mathbb{Q}^d$.

An effective version of the Minkowski–Weyl theorem over $\mathbb{Q}$ states that the two representations (as the intersection of half-spaces and as $\operatorname{conv} F + \operatorname{cone} G$ for finite sets F, G of generators) can be translated from one to another with a single exponential blowup (see, e.g., [15, Chapter 10]).

In the discrete setting, semilinear sets [11] characterize the sets of integer vectors definable in linear integer arithmetic [6]. For technical purposes, we give a slightly more generic definition. Let $\boldsymbol{b} \in \mathbb{Z}^d$ and $P = \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n\} \subseteq \mathbb{Z}^d$, we call $(\boldsymbol{b}, P)$ a *generator tuple* with

base $\boldsymbol{b}$ and *periods* P. Fix $\mathbb{D}$ to be some subset of $\mathbb{Q}^d$, then the $\mathbb{D}$-*linear set* $L_{\mathbb{D}}(\boldsymbol{b}, P)$ generated by the tuple $(\boldsymbol{b}, P)$ is defined as

$$L_{\mathbb{D}}(\boldsymbol{b}, P) = \boldsymbol{b} + \mathbb{D} \cdot \boldsymbol{p}_1 + \cdots + \mathbb{D} \cdot \boldsymbol{p}_n \,.$$

Thus, $\mathbb{N}$-linear sets recover standard linear sets [11] as defined in the literature. Given a set $M \subseteq \mathbb{Z}^d$, we say that M is $\mathbb{D}$-*linear* if there is a generator tuple $(\boldsymbol{b}, P)$ such that $M = L_{\mathbb{D}}(\boldsymbol{b}, P)$. Clearly, $\mathbb{N}$-linear sets contain all $\mathbb{Z}$-linear sets, since $L_{\mathbb{Z}}(\boldsymbol{b}, P) = L_{\mathbb{N}}(\boldsymbol{b}, P \cup -P)$, but the converse does not hold: it is easy to prove that $\mathbb{N}$ is $\mathbb{N}$-linear but not $\mathbb{Z}$-linear.

We say that M is a *hybrid* $\mathbb{D}$-*linear set* if it is a union of the form

$$L_{\mathbb{D}}(B, P) = \bigcup_{\boldsymbol{b} \in B} L_{\mathbb{D}}(\boldsymbol{b}, P)$$

where both sets $B, P \subseteq \mathbb{Z}^d$ are finite. A $\mathbb{D}$-linear or hybrid $\mathbb{D}$-linear set, as defined above, is k-*dimensional* if the linear span of the set P has dimension k; it is *full-dimensional* if $k = d$. One may think of k-dimensional $\mathbb{Z}$-linear sets as of shifted lattices (or cosets of lattices) inside k-dimensional affine subspaces.

We say that M is $\mathbb{D}$-*semilinear* if it is a finite union of $\mathbb{D}$-linear sets. We call a semilinear set *proper* if all sets of period vectors of those linear sets are linearly independent.

Hybrid $\mathbb{N}$-linear sets (often just *hybrid linear sets* for short) are discrete analogues of convex polyhedra; they are exactly sets of integer solutions to systems of linear inequalities $\boldsymbol{A} \cdot \boldsymbol{x} \geq \boldsymbol{c}$ [17, 2], which implies that sets definable in linear integer arithmetic are $\mathbb{N}$-semilinear [6]. Similarly, integer solutions of systems of linear equations are $\mathbb{Z}$-linear:

▶ **Proposition 2.** *Suppose $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{c}$ has a solution in $\mathbb{Z}^d$. Then its set of solutions in $\mathbb{Z}^d$ is a proper $\mathbb{Z}$-linear set $L_{\mathbb{Z}}(\boldsymbol{b}, P)$, with $\|\boldsymbol{b}\| \leq 2^{O(d^4 \log d)} \cdot \|\boldsymbol{A}\|^{O(d^4)} \cdot \|\boldsymbol{c}\|$ and $\|P\| \leq 2^{O(d^4 \log d)} \cdot \|\boldsymbol{A}\|^{O(d^4)}$. Moreover, $|P| = d - \operatorname{rank} \boldsymbol{A}$ and the vectors in P are fully determined by $\boldsymbol{A}$ (i.e., independent of $\boldsymbol{c}$).*

It should be noted, however, that $\mathbb{Z}$-semilinear sets do not fully characterize sets definable in weak linear integer arithmetic. Choffrut and Frigeri [5] have shown that $M \subseteq \mathbb{Z}^d$ is definable in weak linear arithmetic if and only if $M = \bigcup_{i \in I} S_i \setminus T_i$, where the S_i are proper $\mathbb{Z}$-linear sets, T_i proper $\mathbb{Z}$-semilinear sets, and $T_i \subseteq S_i$ for all $i \in I$.

Descriptional complexity. The complexity upper bounds we obtain in this paper rely on bounds on the constants in the generator representation of sets definable in linear arithmetic theories. Given a $\mathbb{D}$-semilinear set $S \subseteq \mathbb{Z}^d$ in generator representation, $S = \bigcup_{i \in I} L(B_i, P_i)$, we define $\|S\| := \max_{i \in I} \max \|B_i \cup P_i\|$.

▶ **Proposition 3.** *Let Φ be an existential formula of linear integer arithmetic and $S = [\![\Phi]\!]_{\mathbb{Z}}$. Then $S = \bigcup_{i \in I} L(B_i, P_i)$ such that $\log\|S\|, \log|I| \leq \operatorname{poly}(|\Phi|)$.*

Proof. For a system of linear inequalities $\boldsymbol{A} \cdot \boldsymbol{x} \geq \boldsymbol{c}$, the statement follows, e.g., from [17]. Moreover, the disjunctive normal form of Φ consists of at most $2^{|\Phi|}$ conjunctions, each of which is a system of linear inequalities, from which the statement follows. ◀

4 A characterisation of weak $\mathbb{Q}$- and $\mathbb{Z}$-definability

We now establish properties that fully characterize when a subset of $\mathbb{Q}^d$ is weakly $\mathbb{Q}$- and weakly $\mathbb{Z}$-definable, respectively.

▶ **Definition 4.** *Suppose $X \subseteq \mathbb{Q}^d$. We say that X:*

- *has* 0–1 property with respect to an affine subspace A *if either $X \cap A$ or $A \setminus X$ is contained in a finite union of affine subspaces of dimension* $\dim A - 1$,
- *has* hierarchical 0–1 property with respect to an affine subspace A *if it has* 0–1 *property with respect to A, where the subspaces of lower dimension are some $H_1, \ldots, H_m$, and, if* $\dim A > 1$, *it has hierarchical* 0–1 *property with respect to each H_i, $1 \leq i \leq m$,*
- *has* (hierarchical) 0–1 property *if it has (hierarchical)* 0–1 *property with respect to $\mathbb{Q}^d$,*
- *has* global 0–1 property *if, for every affine subspace $A \subseteq \mathbb{Q}^d$, it has* 0–1 *property with respect to A,*
- *has ℓ-*bounded 0–1 property *if, for every affine subspace $A = \{\boldsymbol{x} \in \mathbb{Q}^d : \boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}\}$ with $\|\boldsymbol{B}\|, \|\boldsymbol{c}\| \leq \ell$, it has* 0–1 *property with respect to A.*

The term "0–1 property" refers to the intuition that a (weakly $\mathbb{Q}$-definable) set must either vanish ("zero") or fill almost all the space ("one"). The hierarchical version of the property basically says, "and the same holds recursively for these subspaces of lower dimension."

The following theorem, whose proof is deferred to the full version of the paper, shows how these properties relate to weak $\mathbb{Q}$-definability.

▶ **Theorem 5.** *For all sets $X \subseteq \mathbb{Q}^d$ the following statements are equivalent:*

- *X is weakly $\mathbb{Q}$-definable,*
- *X has hierarchical* 0–1 *property,*
- *X has global* 0–1 *property, and*
- *X has $\|\Phi\|$-bounded* 0–1 *property, where Φ is a quantifier-free formula of linear rational arithmetic such that $X = [\![\Phi]\!]$.*

▶ **Example 6.** To illustrate the global version of the property, we demonstrate how from Theorem 5 we can derive that $\Phi(x, y) := x \geq 0 \wedge y = 0$ is not weakly $\mathbb{Q}$-definable. Observe that although $[\![\Phi]\!]$ satisfies the 0–1 property, it fails the global 0–1 property with $A = \{(x, 0) : x \in \mathbb{Q}\}$. Indeed, both $[\![\Phi]\!] \cap A$ and $A \setminus [\![\Phi]\!]$ are infinite sets, whereas every affine subspace of A of dimension lower than A, i.e., of dimension zero, is a single point. As a consequence, $[\![\Phi]\!] \cap A$ is not contained in a finite union of such subspaces, and neither is $A \setminus [\![\Phi]\!]$.

In comparison, $\Psi(x, y) := y = 0$ satisfies the global 0–1 property. We observe, for example, that for $A = \{(x, 0) : x \in \mathbb{Q}\}$ the set $A \setminus [\![\Psi]\!]$ is empty and thus contained in a finite union of subspaces of A of dimension $0 < \dim A$.

Note that the ℓ-bounded version of the property will later give a decision procedure for weak $\mathbb{Q}$-definability.

Weakly $\mathbb{Z}$-definable sets do not have to satisfy an immediate analogue of the 0–1 property: $2 \cdot \mathbb{Z} \subseteq \mathbb{Q}$ is an example of a weakly $\mathbb{Z}$-definable set failing the property. The following definition bridges this gap and requires instead that almost all the space is "tiled" by the set X in a periodic manner. Empty tiling (almost all space is empty, i.e., X vanishes) and full tiling ($\mathbb{Z}^d \setminus X$ vanishes in $\mathbb{Q}^d$) are special cases of this.

▶ **Definition 7.** *Suppose $X \subseteq \mathbb{Z}^d$. We say that X:*

- *has* mosaic property with respect to an affine subspace A *if there exists a hybrid $\mathbb{Z}$-linear set $F \subseteq A$ of dimension* $\dim A$ *such that $(X \cap A) \triangle F$ is contained in a finite union of affine subspaces inside A of dimension* $\dim A - 1$,
- *has* hierarchical mosaic property with respect to an affine subspace A *if it has mosaic property with respect to A, where the subspaces of lower dimension are some $H_1, \ldots, H_m$, and, if* $\dim A > 1$, *it has hierarchical mosaic property with respect to each H_i, $1 \leq i \leq m$,*

- *has* (hierarchical) mosaic property *if it has (hierarchical) mosaic property with respect to* $\mathbb{Q}^d$*,*
- *has* global mosaic property *if, for every affine subspace* $A \subseteq \mathbb{Q}^d$*, it has mosaic property with respect to* A*,*
- *has* ℓ-bounded mosaic property *if, for every affine subspace* $A = \{\boldsymbol{x} \in \mathbb{Q}^d : \boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}\}$ *with* $\|\boldsymbol{B}\|, \|\boldsymbol{c}\| \leq \ell$*, it has mosaic property with respect to* A*.*

The intuition behind the variants of the mosaic property is the same as for the 0–1 property.

▶ **Example 8.** We show how the sets defined by the following logical formulae over the integers satisfy the hierarchical mosaic property.

1. $\Phi_1(x, y) := \exists u \exists v.(x = 3u \vee x = 3u + 1) \wedge y = 2v$. The mosaic property w.r.t. $\mathbb{Q}^2$ holds with $F = L_{\mathbb{Z}}(\{(0,0),(1,0)\},\{(3,0),(0,2)\})$, because $[\![\Phi_1]\!]_{\mathbb{Z}} = F$. So the hierarchical mosaic property (w.r.t. $\mathbb{Q}^2$) holds too, with $m = 0$.
2. $\Phi_2(x, y) := (x = 0) \vee (y > 0) \vee (y < 0)$. The mosaic property w.r.t. $\mathbb{Q}^2$ holds with $F = \mathbb{Z}^2 = L_{\mathbb{Z}}(\boldsymbol{0}, \{\boldsymbol{e}_1, \boldsymbol{e}_2\})$, because we can pick an "exceptional" subspace $H_1 = \{(x, y) \in \mathbb{Q}^2 : y = 0\}$. Going inside H_1, we notice that $[\![\Phi_2]\!]_{\mathbb{Z}}$ also has the mosaic property w.r.t. H_1, because its intersection with H_1 is the singleton $\{\boldsymbol{0}\}$, so we can pick $F' = \emptyset$ and zero-dimensional subspace $H_{1,1} = \{\boldsymbol{0}\}$. Therefore, $[\![\Phi_2]\!]_{\mathbb{Z}}$ satisfies the hierarchical mosaic property (w.r.t. $\mathbb{Q}^2$).
3. $\Phi_3(x, y, z) := (y = 0 \wedge z = 0) \vee (x = 0 \wedge (\exists t.y = 2t) \wedge z = 0) \vee (x = 0 \wedge y = 0 \wedge (\exists t.z = 3t+1))$. For the mosaic property w.r.t. $\mathbb{Q}^3$, we can pick $F = \emptyset$ with $H_1 = [\![x = 0]\!]_{\mathbb{Q}}$ and $H_2 = [\![y = 0]\!]_{\mathbb{Q}}$. The mosaic property w.r.t. H_1 holds with $F_1 = \emptyset$ and $H_{1,1} = [\![x = 0 \wedge z = 0]\!]_{\mathbb{Q}}$, $H_{1,2} = [\![x = 0 \wedge y = 0]\!]_{\mathbb{Q}}$. Similarly, the mosaic property w.r.t. H_2 holds with $F_2 = \emptyset$ and $H_{2,1} = [\![y = 0 \wedge z = 0]\!]_{\mathbb{Q}}$, $H_{2,2} = H_{1,2}$. We leave it as an exercise to the reader to check that $[\![\Phi_3]\!]_{\mathbb{Z}}$ has the mosaic property with respect to each of $H_{1,1}$, $H_{1,2}$, and $H_{2,1}$. In conclusion, $[\![\Phi_3]\!]_{\mathbb{Z}}$ has the hierarchical mosaic property (w.r.t. $\mathbb{Q}^3$).

The proof of the following theorem is deferred to the full version of this paper:

▶ **Theorem 9.** *There exists a function* $f(s, d) \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$*,* $f(n, d) = s^{d^{O(d)}}$*, such that for all sets* $X \subseteq \mathbb{Z}^d$ *the following statements are equivalent:*

- X *is weakly* $\mathbb{Z}$*-definable,*
- X *has hierarchical mosaic property,*
- X *has global mosaic property, and*
- X *has* $f(\|\Phi\|, d)$*-bounded mosaic property, where* Φ *is an existential formula of linear integer arithmetic such that* $X = [\![\Phi]\!]$*.*

The function f gives an upper bound on the magnitude of coefficients in systems of linear equations that specify relevant affine subspaces. We can see from the theorem that, for fixed dimension d, the function f is at most polynomial; and in arbitrary dimension, the bit size of required coefficients is at most $d^{O(d)} \cdot \log\|\Phi\|$, a single exponential in d.

4.1 More on mosaic property

We introduce the following problem.

Piece of Mosaic

INPUT: Semilinear set $S \subseteq \mathbb{Z}^d$ in generator representation, system of equations $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ with integer coefficients defining an affine subspace $A \subseteq \mathbb{Q}^d$.

OUTPUT: Hybrid $\mathbb{Z}$-linear set F represented as $F = \bigcup_{j\in J} L_{\mathbb{Z}}(C_j, Q_j)$ and such that, if S has mosaic property w.r.t. A, then $(S \triangle F) \cap A$ vanishes w.r.t. A.

Note that this is a promise problem: if S does *not* satisfy the mosaic property w.r.t. A, then no restriction is imposed on F.

The intuition for the name of this problem is that an algorithm for it "finds a piece of mosaic", F, in the set S: if S as a whole satisfies the mosaic property (w.r.t. A), then $S \cap A$ should really look like F everywhere in A (except for maybe lower-dimensional exceptions). But it is also possible that some parts of S are unlike F, and then S does not satisfy the mosaic property (in which case the algorithm may output any hybrid $\mathbb{Z}$-linear set).

Also note that, in the case that S satisfies the mosaic property, although the set F must be hybrid $\mathbb{Z}$-linear, the algorithm is permitted to output a $\mathbb{Z}$-semilinear representation of it. This is because the bit size and the norm can grow significantly if a hybrid $\mathbb{Z}$-linear representation is required. (For example, in dimension one, the set $\bigcup_{i=1}^{n} L_{\mathbb{Z}}(0, i)$ is hybrid $\mathbb{Z}$-linear, but all its representations as $L_{\mathbb{Z}}(C, Q)$ have bit size superpolynomial in n.)

▶ **Lemma 10.** *Let $t := (\|S\| + \|A\|)^{\mathrm{poly}(d)}$. The following statements hold:*

1. *If a semilinear set S satisfies the mosaic property with respect to an affine subspace A, then*
   - **(a)** *the set F in the definition of the property is determined uniquely and*
   - **(b)** *$(S \cap A) \triangle F$ is contained in a finite collection of affine subspaces of A of dimension $\dim A - 1$ each, defined by linear equations with integer coefficients of absolute value at most t.*
2. *There is a t-time algorithm that solves the* Piece of Mosaic *problem, which always produces an $F = \bigcup_{j\in J} L_{\mathbb{Z}}(C_j, Q_j)$ such that $\|C_j\|, \|Q_j\| \le t$.*

The proof is given in a separate subsubsection below.

▶ **Lemma 11.** *There is an algorithm with running time $2^{(\|A\|+\|S\|)^{\mathrm{poly}(d)}}$ that, given a semilinear set $S = \bigcup_{i\in I} L(B_i, P_i)$ and a system of equations $\boldsymbol{B}\cdot\boldsymbol{x} = \boldsymbol{c}$ with integer coefficients defining an affine subspace $A \subseteq \mathbb{Q}^d$, decides whether S has the mosaic property with respect to A.*

Proof. Run the algorithm from Lemma 10(2) to obtain a hybrid $\mathbb{Z}$-linear set F such that $\|F\| \le t$ with t defined as in the lemma. To check whether S has mosaic property, we determine whether $(S \cap A) \triangle F$ is vanishing. By Proposition 2, $A \cap \mathbb{Z}^d = L(C, Q)$ such that $\|C\|, \|Q\| \le \|A\|^{\mathrm{poly}(d)}$. Then from [2, Thm. 6], we get that $S \cap A = \bigcup_{k\in K} L(D_k, E_k)$ such that $\|D_k\|, \|E_k\| \le (\|S\| + \|A\|)^{\mathrm{poly}(d)}$. We compute the semilinear representation M of $((S \cap A) \setminus F) \cup (F \setminus (S \cap A))$. It follows the bounds on the description size of (each) set difference [2, Cor. 22] that $\log\|M\| \le (\|A\| + \|S\|)^{\mathrm{poly}(d)}$. In order to check whether $(S \cap A) \triangle F$ vanishes w.r.t. to A, it remains to iterate over all hybrid linear sets defining M and to check whether the dimension of the affine hull of the period vectors in each set is strictly less than $\dim A$. We remark that the upper bound on the running time holds, because $\|A\|$ and $\|S\|$ are at most exponential in the bit size of the encoding. ◀

▶ **Lemma 12.** *There is a doubly exponential algorithm that, given $\Phi \subseteq \mathbb{Z}^d$, decides whether $[\![\Phi]\!]$ has the $f(\|\Phi\|, d)$-bounded mosaic property, where f is defined as in Theorem 9.*

Proof. It follows from Proposition 3 that we can compute in time $2^{\mathrm{poly}(|\Phi|)}$ the semilinear representation of $[\![\Phi]\!]$. The lemma then follows by an application of Lemma 11; note that the running time of the algorithm of that lemma has only a single exponential dependency on $\|S\| + \|A\|$. ◀

5 Computational complexity of weak $\mathbb{Q}$- and $\mathbb{Z}$-definability

Building upon the results in the previous sections, we now prove the main theorem of this paper.

▶ **Theorem 13.** *The weak $\mathbb{Q}$-definability problem is coNP-complete, and the weak $\mathbb{Z}$-definability problem is Π_2^{P}-hard and can be decided by an algorithm with elementary running time.*

We first outline the upper bounds of this theorem and then the lower bounds. The lower bound for weak $\mathbb{Z}$-definability already holds in a fixed dimension and is obtained from showing that the universality problem for one-dimensional semi-linear sets is Π_2^{P}-hard. We give the proof of Π_2^{P}-hardness of this universality problem in a separate Section 6.

5.1 Upper bounds for deciding weak $\mathbb{Q}$- and $\mathbb{Z}$-definability

By Theorem 5, a set $X \subseteq \mathbb{Q}^d$ given by a quantifier-free formula Φ of linear rational arithmetic is weakly $\mathbb{Q}$-definable if and only if for every affine subspace A, given by a system of linear equations $\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{c}$ such that $\|\boldsymbol{A}\|, \|\boldsymbol{c}\| \leq \|\Phi\|$, either $X \cap A$ or $A \setminus X = A \cap \overline{X}$ is contained in a finite union of affine subspaces of dimension $\dim A - 1$. To prove that Φ is not weakly $\mathbb{Q}$-definable, we attempt to find an affine subspace A such that neither $X \cap A$ nor $A \setminus X$ is contained in a finite union of affine subspaces of dimension $\dim A - 1$. Note that due to Φ being quantifier-free, $\neg\Phi$ is also a quantifier-free formula of linear rational arithmetic. Hence it will suffice to only discuss the case $X \cap A$; and, in this case, we can also assume without loss of generality that Φ is negation-free. Let Ψ be the disjunctive normal form of Φ. We clearly have $[\![\Psi]\!] = X$, and $[\![\Psi]\!] \cap A$ is not contained in a finite union of hyperplanes of dimension $\dim A - 1$ if and only if for some polyhedron P, defined by a conjunction of Ψ, the polyhedron $P \cap A$ has dimension $\dim A$. We now outline how to decide in polynomial time whether $\dim(P \cap A)$ equals $\dim A$.

Both P and A are given as systems of linear constraints, and hence we immediately obtain $P \cap A$ as a system of linear constraints. The dimension $\dim(P \cap A)$ is, by definition, equal to $\dim \operatorname{aff}(P \cap A)$. When $P \cap A$ is given by a system of non-strict linear inequalities, one can obtain in polynomial time a representation of $\operatorname{aff}(P \cap A)$ as the intersection of at most d implicit equalities of $P \cap A$, each obtained from a row of the system defining $P \cap A$ [15, p. 100]. One can then compute $\dim \operatorname{aff}(P \cap A)$, by (a variant of) Gaussian elimination [15, Section 3.3]. In general, the system of constraints defining P may contain strict inequalities though. However, this does not cause any problems, since for a polyhedron P, $\dim(P) = \dim(\operatorname{cl} P)$, where $\operatorname{cl} P$ is the closure of P. If $P \neq \emptyset$ and given as a system of linear constraints, $\operatorname{cl} P$ can be obtained by making all strict inequalities in the defining system of P non-strict.

From this line of reasoning, we obtain a coNP upper bound for deciding weak $\mathbb{Q}$-definability as follows. We guess $\boldsymbol{A}$ and $\boldsymbol{c}$ above in non-deterministic polynomial time. While the disjunctive normal forms of Φ and $\neg\Phi$ (both assumed negation-free, with no loss of generality) can be exponentially long, we can — given the formulas Φ and $\neg\Phi$ — guess a single conjunction of their disjunctive normal forms in non-deterministic polynomial time, by inspecting the structure of Φ and $\neg\Phi$. These two conjunctions induce polyhedra P and P'. We then decide in polynomial time whether the dimension of the polyhedra $P \cap A$ and $P' \cap A$ equals $\dim A$. We reject if and only if this is the case for both conjunctions.

We now turn towards a sketch of the upper bound for weak $\mathbb{Z}$-definability. To this end, let $X \subseteq \mathbb{Z}^d$ be defined by an existential formula $\Phi(\boldsymbol{x})$ of linear integer arithmetic. For an

elementary upper bound, following Theorem 9 we iterate over all affine subspaces A given by systems of equations $\boldsymbol{B} \cdot \boldsymbol{x} = \boldsymbol{c}$ such that $\|\boldsymbol{B}\|, \|\boldsymbol{c}\|$ are at most $f(\|\Phi\|, d)$ and check whether Φ has mosaic property with respect to A. By Lemma 12, there is a doubly exponential algorithm that achieves this.

5.2 Lower bounds for deciding weak $\mathbb{Q}$-definability and $\mathbb{Z}$-definability

Lower bound for weak $\mathbb{Q}$-definability. We show a matching coNP-lower bound for deciding weak $\mathbb{Q}$-definability by a reduction from the problem of deciding whether a Boolean 3-DNF formula is a tautology. Let $\psi = \psi_1 \vee \cdots \vee \psi_k$ be in 3-DNF over Boolean variables $X_1, \ldots, X_d$. Let ϕ_1 be the formula of rational arithmetic obtained from ψ by applying the function h to every literal, with h defined as $h(X_i) := x_i = 1$ and $h(\neg X_i) := x_i = 0$. In addition, define

$$\phi_2 := \bigvee_{1 \leq i \leq d} x_i > 1 \vee x_i < 0 \vee (0 < x_i < 1)$$

Observe that $[\![\phi_1 \vee \phi_2]\!]$ is the whole of $\mathbb{Q}^d$ except possibly a finite set of points corresponding to those truth assignments that do not make ψ evaluate to true, i.e., $[\![\phi_1 \vee \phi_2]\!] = \mathbb{Q}^d$ if and only if ψ is a tautology. Now define $\Phi := \phi_1 \vee \phi_2 \vee u > 0$, where u is a fresh variable. Note that $[\![\Phi]\!]$ is the whole of $\mathbb{Q}^{d+1}$ except possibly several half-lines that correspond to assignments falsifying ψ. If a half-line is missing then $[\![\Phi]\!]$ does not satisfy the global 0–1 property and is therefore not weakly $\mathbb{Q}$-definable, by Theorem 5. Otherwise no half-line is missing, ψ is a tautology, $[\![\Phi]\!]$ is equal to $\mathbb{Q}^d$ and is weakly $\mathbb{Q}$-definable.

Lower bound for weak $\mathbb{Z}$-definability. We show a Π_2^{P}-lower bound for weak $\mathbb{Z}$-definability via a reduction from the universality problem for semilinear sets. Given a semilinear set $M \subseteq \mathbb{N}^d$ in the generator representation, the *universality problem* is to decide whether $M = \mathbb{N}^d$. It was asked by Huynh [8] whether this problem is Π_2^{P}-hard when he established a Π_2^{P}-upper bound for this problem. We show in the next section that this is the case, even in dimension one, and assume hardness for now to show our lower bound for weak $\mathbb{Z}$-definability.

To this end, let $M = \bigcup_{i \in I} L(\boldsymbol{b}_i, P_i) \subseteq \mathbb{N}^d$. One easily constructs an existential formula $\psi(x_1, \ldots, x_d)$ of linear integer arithmetic such that $M = [\![\psi]\!]$. Now consider $\Phi(u, x_1, \ldots, x_d) := \psi \vee u > 0 \vee \bigvee_{i \in [d]} x_i < 0$, where u is a fresh variable. Analogously to what we showed above, $[\![\Phi]\!]$ is the whole of $\mathbb{Z}^{d+1}$ except possibly several "discrete half-lines" corresponding to elements $\boldsymbol{v} \in \mathbb{N}^d \setminus M$. By the global mosaic property (Theorem 9), $[\![\Phi]\!]$ is weakly $\mathbb{Z}$-definable if and only if no half-line is missing.

6 A lower bound for deciding universality of semilinear sets

We exclusively deal with $\mathbb{N}$-semilinear sets in this section and so, when presenting such semilinear sets in generator presentation, for readability, instead of writing, e.g., $L_{\mathbb{N}}(\boldsymbol{b}, P)$ we subsequently drop the subscript $\mathbb{N}$ and simply write $L(\boldsymbol{b}, P)$. The main result of this section is a Π_2^{P} lower bound for the universality problem for ultimately periodic sets, which are semilinear sets in dimension one:

Ultimately Periodic Set Universality

INPUT: Finite set I and, for each $i \in I$, a number $b_i \in \mathbb{N}$ and a finite set $P_i \subseteq \mathbb{N}$.
QUESTION: Is $\bigcup_{i \in I} L(b_i, P_i) = \mathbb{N}$?

Since a set $M \subseteq \mathbb{N}^d$ is universal if and only if $L(\boldsymbol{0}, \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}) \subseteq M$, this universality problem is a special case of *inclusion* for semilinear sets, which asks to decide if $N \subseteq M$ for

Simultaneous Subset Sum

INPUT: Finite set $W \subseteq \mathbb{N}$, and $t, h, 2^m \in \mathbb{N}$ such that $t < h$.
QUESTION: For every $i \in [0, 2^m - 1]$, does there exist a $W' \subseteq W$ such that $\sum W' = t + i \cdot h$?

Restricted Simultaneous Subset Sum

INPUT: Finite set $W \subseteq \mathbb{N}$ and $t, 2^k, 2^m \in \mathbb{N}$ such that $t < 2^k$.
QUESTION: For every $i \in [0, 2^m - 1]$, does there exist a $W' \subseteq W$ such that $\sum W' = t + i \cdot 2^k$?

Figure 1 The simultaneous subset sum problem introduced in [4] and its restricted version.

Ray Cover

INPUT: Finite set $P \subseteq \mathbb{N}^d$ and $a, 2^k \in \mathbb{N}$ such that $a < 2^k$.
QUESTION: Does $L(\mathbf{0}, P) \supseteq L(\boldsymbol{a}, \boldsymbol{b})$, where $\boldsymbol{a} = (a, \mathbf{1})$ and $\boldsymbol{b} = (2^k, \mathbf{0})$?

Bounded Ray Cover

INPUT: Finite set $P \subseteq \mathbb{N}^d$ and $a, 2^k, 2^m \in \mathbb{N}$ such that $a < 2^k$.
QUESTION: Does $L(\mathbf{0}, P) \supseteq L(\boldsymbol{a}, \boldsymbol{b}) \cap [0, B]^d$, where $\boldsymbol{a} = (a, \mathbf{1})$, $\boldsymbol{b} = (2^k, \mathbf{0})$, and $B = a + (2^m - 1) \cdot 2^k$?

Figure 2 Ray cover problems in arbitrary dimensions.

semilinear sets M, N. While inclusion for semilinear sets has been known to be Π_2^{P}-complete since the 1980s [8], the lower bound has only been strengthened to hold for inclusion of *linear* sets as recently as 2018 [4, Thm. 12], and then also to linear sets in dimension one [16]. It has also recently been shown that universality for linear sets is decidable in polynomial time, even for hybrid linear sets of the form $L(B, P) := \bigcup_{\boldsymbol{b} \in B} L(\boldsymbol{b}, P)$ [3].

Our new Π_2^{P} lower bound is based on a chain of reductions between intermediate problems that we now introduce. The overall reduction chain is displayed in Figure 4. Our starting point is the Simultaneous Subset Sum problem introduced in [4] from which we derive a slightly restricted version. Both problems are defined in Figure 1 and are variants of the classical subset sum problem. They ask whether all elements in a finite arithmetic progression can be obtained as sums of subsets of a given set $W \subseteq \mathbb{N}$.

Via a reduction from Restricted Simultaneous Subset Sum, we prove Π_2^{P}-hardness of two special cases of the semilinear set inclusion problem, Bounded Ray Cover and Ray Cover, which are defined in Figure 2. The problem Ray Cover asks whether a discrete ray in $\mathbb{N}^d$ (basically an arithmetic progression) is contained in an integer cone (linear set) — this is a restricted variant of linear set inclusion. The problem Bounded Ray Cover is the same but only concerns a finite segment of the ray. The Π_2^{P}-hardness of Ray Cover follows from that of Bounded Ray Cover.

Next, by a reduction from Bounded Ray Cover, we show Π_2^{P}-hardness of the one-dimensional versions of the ray cover problems, formally defined in Figure 3. A reduction from Bounded 1D Ray Cover will then give the desired Π_2^{P}-lower bound of Ultimately Periodic Set Universality.

▶ **Theorem 14.** *All six problems in Figure 4 are Π_2^{P}-complete.*

All upper bounds in Theorem 14 are easily obtained from the observation that the respective problems either reduce to semi-linear set inclusion, which is in Π_2^{P} [8], or involve sets of numbers of polynomial bit size.

1D Ray Cover

INPUT: Finite set $P \subseteq \mathbb{N}$ and $a, 2^k \in \mathbb{N}$ such that $a < 2^k$.
QUESTION: Does $L(0, P) \supseteq L(a, b)$, where $b = 2^k$?

Bounded 1D Ray Cover

INPUT: Finite set $P \subseteq \mathbb{N}$ and $a, 2^k, 2^m \in \mathbb{N}$ such that $a < 2^k$.
QUESTION: Does $L(0, P) \supseteq L(a, b) \cap [0, B]$, where $b = 2^k$ and $B = a + (2^m - 1) \cdot b$?

Figure 3 Ray cover problems in dimension one.

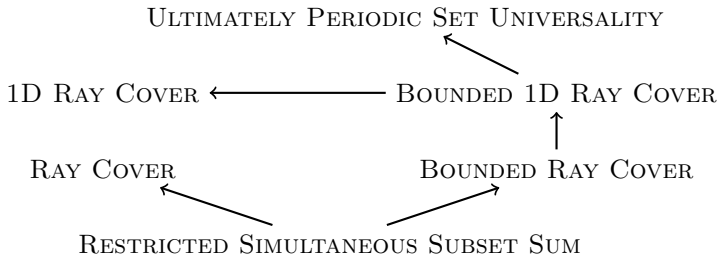

Figure 4 Reductions in the proof of Theorem 14. $X \to Y$ denotes a logarithmic-space reduction from X to Y.

To prove the lower bounds, once the intermediate problems have been identified, one of the key insights is in the reduction from Bounded Ray Cover to Bounded 1D Ray Cover, which maps a problem from $\mathbb{N}^d$ into $\mathbb{N}$. In the remainder of this section, we focus on the techniques that enable us to overcome this challenge.

6.1 Aggregation of several dimensions into one

Aggregation is an important technique in the theory of integer programming (see, e.g., Schrijver's book [15, Sections 16.6 and 18.2]). It has been used by Huynh [8] and Simon [16] for showing lower bounds for the inclusion problem for (semi)linear sets in dimension one. Aggregation can be achieved in the setting of linear Diophantine equations over nonnegative (integer) variables, following a classic result of Glover and Woolsey [7], which we extend.

As we already mentioned, we apply aggregation in order to reduce Bounded Ray Cover to Bounded 1D Ray Cover. In the former problem, think of vectors $\boldsymbol{v} \in L(\boldsymbol{a}, \boldsymbol{b}) \cap [0, B]^d$ as targets to hit. Each of them is hit if and only if $L(\boldsymbol{0}, P)$ contains it, i.e., if the system of equations $\boldsymbol{P} \cdot \boldsymbol{x} = \boldsymbol{v}$ has a solution in $\mathbb{N}^d$. The instance is a yes-instance if and only if all 2^m targets are hit. While Glover and Woolsey's result would allow us to aggregate *one* system of equations like this (for a single target) into a single equation, the key technical challenge in our setting is that — in contrast to [7, Thm. 3] — we need to aggregate *several* such systems, for 2^m different targets (each into its own one equation). That is, these systems have identical coefficients, but different constant terms.

For the following lemma, we will explain the rationale behind the constraints "α is a power of two" and "$\alpha/\beta > \max B$" subsequently.

▶ **Lemma 15.** *Let A, B be finite subsets of $\mathbb{N}_+$. Then there exist $\alpha, \beta \in \mathbb{N}$ such that α is a*

power of two, $\alpha/\beta > \max B$*, and, for each pair* $(a,b) \in A \times B$*, every system of equations*

$$\begin{cases} \boldsymbol{c} \cdot \boldsymbol{x} = a \\ \boldsymbol{d} \cdot \boldsymbol{x} = b \end{cases}$$

with $\boldsymbol{c}, \boldsymbol{d} \in \mathbb{N}^d$ *and* $(a,b) \in A \times B$ *has the same set of solutions in* $\mathbb{N}^d$ *as the single equation*

$$(\alpha \boldsymbol{c} + \beta \boldsymbol{d}) \cdot \boldsymbol{x} = \alpha a + \beta b\,.$$

Moreover, α *and* β *are polynomial-time computable from* $\max A$ *and* $\max B$ *and are independent of* $\boldsymbol{c}$ *and* $\boldsymbol{d}$*.*

Proof. It follows from [7, Thm. 3] of Glover and Woolsey that when A and B are singletons $A = \{a\}$ and $B = \{b\}$ the statement holds if

$$\alpha > b, \qquad \beta > a, \quad \text{and} \quad \gcd(a,b) = 1\,. \tag{1}$$

We show how to find a *single pair* of coefficients $\alpha, \beta \in \mathbb{N}$ that make Glover and Woolsey's result applicable to the system in question for *all* pairs $(a,b) \in A \times B$. Indeed, choose r as the smallest power of two such that $r-1$ strictly exceeds both $\max A$ and $\max B$. Pick $\beta = r - 1 > \max A$ and $\alpha = r^2$. We now check that α and β satisfy conditions (1). Observe that

$$\begin{aligned} &\alpha = r^2 > r - 1 > b, && \text{for all } b \in B, \\ &\beta = r - 1 > a, && \text{for all } a \in A, \end{aligned}$$

and $\gcd(\alpha,\beta) = \gcd\left((r-1)(r+1)+1, r-1\right) = 1$. It remains to note that $\alpha/\beta \geq r > \max B$. ◀

Note that Lemma 15 aggregates several systems using *the same* α and β. It guarantees that the numbers α and β stay small (have polynomial size) and satisfy additional constraints, to make the subsequent reduction from Bounded 1D Ray Cover to Ultimately Periodic Set Universality possible. We now discuss how these constraints arise.

6.2 Additional constraints and the final reduction

Let us take a step forward in the chain of reductions in Figure 4. We illustrate the significance of the additional requirements on the input of Bounded 1D Ray Cover using the following simple observation on the representation of the complement of a segment of a linear progression.

▶ **Lemma 16.** *Let* $a, 2^k, 2^m \in \mathbb{N}$*. Denote* $b = 2^k$ *and* $B = a + (2^m - 1) \cdot b$*, as in the input of* Bounded 1D Ray Cover*. Then the set* $\mathbb{N} \setminus (L(a,b) \cap [0,B])$ *is semilinear with a generator representation of size* $O(\lfloor a/b \rfloor \log a + k^2 + m)$*.*

Proof. Observe that $\mathbb{N} \setminus (L(a,b) \cap [0,B]) = \{n \in \mathbb{N} : n < a,\ n \equiv a \mod b\} \cup \{n \in \mathbb{N} : n \not\equiv a \mod b\} \cup \{n \in \mathbb{N} : n > B\}$. The third set on the right-hand side is $L(B+1,1)$, and the first set is a finite set with $\lfloor a/b \rfloor$ elements, i.e., a union of $\lfloor a/b \rfloor$ linear sets of the form $L(x,0)$, $x < a$. For the second set, we will rely on the assumption that $b = 2^k$. Suppose that in binary we have $a \bmod b = a_{k-1} \cdot 2^{k-1} + \cdots + a_1 \cdot 2^1 + a_0 \cdot 2^0$, $a_i \in \{0,1\}$. Notice that an arbitrary $n \in \mathbb{N}$ is not congruent to a modulo 2^k if and only if, for some $j \in \{0,\ldots,k-1\}$, it is congruent to $c_j := \overline{a}_j \cdot 2^j + a_{j-1} \cdot 2^{j-1} \cdots a_1 \cdot 2^1 + a_0 \cdot 2^0$ modulo 2^{j+1}, where $\overline{a}_j = 1 - a_j$. Therefore, the second set in the equation above is equal to $\bigcup_{j=0}^{k-1} L(c_j, 2^{j+1})$. ◀

Given Lemma 16, the final reduction from Bounded 1D Ray Cover to Ultimately Periodic Set Universality is simple. Indeed, $L(0,P) \supseteq L(a,b) \cap [0,B]$ if and only if $L(0,P) \cup (\mathbb{N} \setminus (L(a,b) \cap [0,B])) = \mathbb{N}$, where the second set in the union is described by a semilinear set in generator representation of polynomial size. The bound on the size holds, because k and m cannot exceed the bit size of the input instance and $\lfloor a/b \rfloor = \lfloor a/2^k \rfloor = 0$ by the promise that $a < 2^k$ in the definition of Bounded 1D Ray Cover.

Note that the proof of Lemma 16 would work just as well if we had b equal to a power of 3, or a power of any other fixed number (other forms are also possible). For the input of Bounded 1D Ray Cover, any of these constraints is not difficult to satisfy on its own. In comparison, the dependence on $\lfloor a/b \rfloor$ in Lemma 16 is more important and more difficult to handle. We ensure in our chain of reductions from Restricted Simultaneous Subset Sum that $a/b < 1$; in particular, in our aggregation procedure (Lemma 15) we require that the coefficients α and β satisfy $\alpha/\beta > b$ The condition that α is a power of two comes from our preference to use an arithmetic progression with a difference of the form 2^v, $v \in \mathbb{N}$, in Lemma 16: $L(a,b)$, $b = 2^k$.

Detailed Π_2^{P}-hardness proofs for the lower bounds in Theorem 14 can be found in the full version of this paper.

7 Conclusion

Choffrut and Frigeri left open the question whether there is an algorithm with an elementary running time deciding weak $\mathbb{Z}$-definability [5]. We have shown in this article that this is the case. There still remains a significant gap between our Π_2^{P} lower bound and the rather crude upper bound that we obtained. Our algorithm is based on a geometric characterisation of sets definable in weak Presburger arithmetic that complements the generator characterisation obtained by Choffrut and Frigeri [5].

While weak definability is an interesting problem in its own right, another motivation for our work stems from the fact it is an open problem whether deciding weak Presburger arithmetic is computationally easier than deciding full Presburger arithmetic. In fact, in his original article [12], Presburger only showed decidability of weak Presburger arithmetic and remarked that his proof could be adapted to also work for Presburger arithmetic. We are unable to give an answer to this open problem at the present stage, but we believe that the geometric insights that enable us to show the elementary upper bounds of the weak definability problems may eventually help settling the computational complexity of weak Presburger arithmetic. Note that deciding weak linear rational arithmetic has essentially the same complexity as linear rational arithmetic [14].

References

1 B. Boigelot, I. Mainz, V. Marsault, and M. Rigo. An efficient algorithm to decide periodicity of b-recognisable sets using MSDF convention. In *International Colloquium on Automata, Languages, and Programming, ICALP*, volume 80 of *LIPIcs*, pages 118:1–118:14, 2017. doi:10.4230/LIPIcs.ICALP.2017.118.

2 D. Chistikov and C. Haase. The taming of the semi-linear set. In *International Colloquium on Automata, Languages, and Programming, ICALP*, volume 55 of *LIPIcs*, pages 128:1–128:13, 2016. doi:10.4230/LIPIcs.ICALP.2016.128.

3 D. Chistikov and C. Haase. On the complexity of quantified integer programming. In *International Colloquium on Automata, Languages, and Programming, ICALP*, volume 80 of *LIPIcs*, pages 94:1–94:13, 2017. doi:10.4230/LIPIcs.ICALP.2017.94.

4 D. Chistikov, C. Haase, and S. Halfon. Context-free commutative grammars with integer counters and resets. *Theor. Comput. Sci.*, 735:147–161, 2018. doi:10.1016/j.tcs.2016.06.017.

5 C. Choffrut and A. Frigeri. Deciding whether the ordering is necessary in a Presburger formula. *Discret. Math. Theor. C.*, 12(1):21–38, 2010. URL: http://dmtcs.episciences.org/510.

6 S. Ginsburg and E.H. Spanier. Bounded ALGOL-like languages. *T. Am. Math. Soc.*, pages 333–368, 1964. doi:10.2307/1994067.

7 F.W. Glover and R.E.D. Woolsey. Aggregating diophantine equations. *Zeitschr. für OR*, 16(1):1–10, 1972. doi:10.1007/BF01917186.

8 T.-D. Huynh. The complexity of semilinear sets. *Elektron. Inf.verarb. Kybern.*, 18(6):291–338, 1982.

9 J. Leroux. A polynomial time Presburger criterion and synthesis for number decision diagrams. In *Symposium on Logic in Computer Science, LICS*, pages 147–156, 2005. doi:10.1109/LICS.2005.2.

10 A.A. Muchnik. The definable criterion for definability in Presburger arithmetic and its applications. *Theor. Comput. Sci.*, 290(3):1433–1444, 2003. doi:10.1016/S0304-3975(02)00047-6.

11 R. Parikh. On context-free languages. *J. ACM*, 13(4):570–581, 1966. doi:10.1145/321356.321364.

12 M. Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Comptes Rendus du I congres de Mathematiciens des Pays Slaves*, pages 92–101. 1929.

13 J. Robinson. Definability and decision problems in arithmetic. *J. Symb. Log.*, 14(2):98–114, 1949. doi:10.2307/2266510.

14 A. Ronquist. A lower bound on the complexity of real addition without order. Master's thesis, University of Oxford, United Kingdom, 2019.

15 A. Schrijver. *Theory of linear and integer programming.* John Wiley & Sons, 1986.

16 H.U. Simon. On the containment problem for linear sets. In *Symposium on Theoretical Aspects of Computer Science, STACS*, volume 96 of *LIPIcs*, pages 55:1–55:12, 2018. doi:10.4230/LIPIcs.STACS.2018.55.

17 J. von zur Gathen and M. Sieveking. A bound on solutions of linear integer equalities and inequalities. *P. Am. Math. Soc.*, 72(1):155–158, 1978. doi:10.1090/S0002-9939-1978-0500555-0.