Computability 1 (2012) 1–32
DOI
IOS Press

Undecidability of Satisfiability of Expansions of FO[<] over Words with a FO[+]-definable set.

Arthur Milchior
IRIF, Université Paris 7 - Denis Diderot, France
CNRS UMR 8243, Université Paris Diderot - Paris 7, Case 7014
75205 Paris Cedex 13
Arthur.Milchior@liafa.univ-paris-diderot.fr
http://www.liafa.univ-paris-diderot.fr/web9/equiprech/fichepers_fr.php?id=353

Université Paris-Est,
LACL (EA 4219), UPEC,
F-94010 Créteil, France

Abstract. Two new characterizations of $\mathsf{FO}[<, \mathrm{mod}]$-definable sets, i.e. sets of integers definable in first-order logic with the order relation and modular relations, are provided. Those characterizations are used to prove that satisfiability of first-order logic over words with an order relation and a $\mathsf{FO}[+]$-definable set that is not $\mathsf{FO}[<, \mathrm{mod}]$-definable is undecidable.

Keywords: Finite model theory, first order logic, arithmetic, undecidability

1. Introduction

A classical result of descriptive complexity theory (see [Imm99]) states that a language of finite words over an alphabet α belongs to the circuit complexity class alternative log-time uniform AC^0 if and only if it is $\mathsf{FO}\left[+, \times, (P_a)_{a\in\alpha}\right]$-definable (where $P_a(x)$ holds if and only if the x-th letter carries the letter a). One can consider variants where $\{+, \times\}$ is replaced by weaker arithmetical relations. Two important examples are the fragments $\mathsf{FO}\left[<, (P_a)_{a\in\alpha}\right]$ and $\mathsf{FO}\left[<, \mathrm{mod}, (P_a)_{a\in\alpha}\right]$, where mod denotes the set of modular relations. The fragment $\mathsf{FO}\left[<, (P_a)_{a\in\alpha}\right]$ captures the class of star-free regular languages [MP71]. The fragment $\mathsf{FO}\left[<, \mathrm{mod}, (P_a)_{a\in\alpha}\right]$ also captures a subclass of the regular languages, which enjoys an (effective) algebraic characterization (see [BCST92, Corollary 10]); moreover $\mathsf{FO}\left[<, \mathrm{mod}, (P_a)_{a\in\alpha}\right]$ is maximal with respect to regular languages, in the sense that every non-trivial extension of the signature $\{<, \mathrm{mod}\}$ with numerical relations allows to define non-regular languages [Pél92]. For more information on this logic, the reader is referred to the book [Str94].

To our knowledge, there exist only few results on the expressivity of logics which lay between $\mathsf{FO}\left[<, \mathrm{mod}, (P_a)_{a\in\alpha}\right]$ and $\mathsf{FO}\left[+, \times, (P_a)_{a\in\alpha}\right]$. The situation contrasts with the extensive literature on definability of fragments of arithmetic *over non-negative integers*, but is not surprising since many classical (un)definability techniques and results cannot be transferred to finite models. On the one hand, the first-order theory of addition over non-negative integers, $\mathsf{FO}[+]$, (i.e. Presburger Arithmetic), is decidable [Pre27], and there exist several characterizations of sets definable in this logic [GS66, Muc03, MV96]. On the other hand, the expressive power of $\mathsf{FO}\left[+, (P_a)_{a\in\alpha}\right]$ over finite words is not completely understood. Recently Choffrut & al. [CMMP10] proved several closure properties of $\mathsf{FO}\left[+, (P_a)_{a\in\alpha}\right]$-definable languages, and provided a partial characterization.

A natural approach to evaluate the expressive power of these logics is to study the complexity of the satisfiability problem. On the one hand this problem is undecidable for $\mathsf{FO}[+, \times]$, as a direct corollary of Trakhtenbrot's Theorem [Tra50]. Lange [Lan04, Lemma 6.3] proved that undecidability occurs even for $\mathsf{FO}\left[+, (P_a)_{a\in\alpha}\right]$ over words. On the other hand satisfiability is decidable for $\mathsf{FO}\left[<, \mathrm{mod}, (P_a)_{a\in\alpha}\right]$, since this fragment is contained in $\mathsf{MSO}[<]$ for which satisfiability is decidable [Bü60, Elg61, Tra61]. In this paper the decidability frontier for logics between $\mathsf{FO}\left[<, \mathrm{mod}, (P_a)_{a\in\alpha}\right]$ and $\mathsf{FO}\left[+, (P_a)_{a\in\alpha}\right]$ is specified, by proving that for every integer $d \geq 1$ and every relation $R \subseteq \mathbb{N}^d$, if R is $\mathsf{FO}[+]$-definable but not $\mathsf{FO}[<, \mathrm{mod}]$-definable, then satisfiability is undecidable for $\mathsf{FO}\left[<, R, (P_a)_{a\in\alpha}\right]$.

One should note that the previous result does not hold anymore when the condition that R is $\mathsf{FO}[+]$-definable is removed. It can be shown for instance that satisfiability is decidable for $\mathsf{FO}\left[<, \mathrm{mod}, R, (P_a)_{a\in\alpha}\right]$ when R denotes the set of factorials or the set of powers of two (this is a direct consequence of Elgot-Rabin' result that satisfiability of $\mathsf{MSO}[<, R]$ is decidable in this case [ER66]).

In order to obtain our undecidability result, general results about definability in fragments of Presburger Arithmetic are proven. It is known that sets definable in Presburger Arithmetic coincide with semilinear sets [GS66]. Two characterizations of semilinear sets are given in [Muc03, MV96] in terms of sets of smaller arity and local properties. Similar results are proven for the logic $\mathsf{FO}[<, \mathrm{mod}]$, and also for logics $\mathsf{FO}[<, \mathrm{mod}\, m]$ where only modulo m relation is used for some fixed non-negative integer m. Note that the $\mathsf{FO}[<, \mathrm{mod}]$-definable sets are sometimes called *regular sets* because they correspond to languages accepted by synchronous multi-tape automata which read integers in base 1 (e.g. [Str94]).

The first result, Theorem 4.4, states that for each $d \in \mathbb{N}$, a set $R \subseteq \mathbb{N}^d$ is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable if and only if:

- Every section of R (i.e. every subset of $\mathbb{N}^{d-1}$ obtained from R by fixing a component) is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable and
- apart from a finite number of sections, R is equal to $\{(x_0 + m, \ldots, x_{d-1} + m) \mid (x_0, \ldots, x_{d-1}) \in R\}$, that is, R translated by m in all directions.

The logics $\mathsf{FO}[<, \mathrm{mod}\, m]$ and $\mathsf{FO}[<, \mathrm{mod}]$ admit another characterization: Theorem 4.18 states that a set R is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable if and only if there exists a $\mathsf{FO}[<, R]$-definable set of integers which is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Similarly Theorem 4.15 states that a $\mathsf{FO}[+]$-definable set R is not $\mathsf{FO}[<, \mathrm{mod}]$-definable if and only if there exists a unary $\mathsf{FO}[<, R]$-definable function which is not $\mathsf{FO}[<, \mathrm{mod}]$-definable.

It should be noted that $\mathsf{FO}[<]$ has the same expressive power as $\mathsf{FO}[<, \mathrm{mod}\, 1]$, hence every result about $\mathsf{FO}[<, \mathrm{mod}\, 1]$ holds for $\mathsf{FO}[<]$.

Then in Section 5, Theorem 5.1 states that satisfiability of $\mathsf{FO}\left[<, R, (P_a)_{a\in\alpha}\right]$ is undecidable over words when R is a $\mathsf{FO}[+]$-definable set which is not $\mathsf{FO}[<, \mathrm{mod}]$-definable.

2. Definitions and notations

In this section, useful definitions are recalled and some notations are fixed. To avoid ambiguity, the "=" symbol is used for mathematical equality and definitions, but in logical formulas, the equality relation is denoted by "$\doteq$". Let $\mathbb{N}$ denote the set of non-negative integers, let $\mathbb{Z}$ denote the set of integers and let $\mathbb{Q}$ denote the set of rational numbers. For $S \subseteq \mathbb{Q}$ and $c \in \mathbb{Q}$, let $S^{>c}$ (respectively, $S^{\geq c}$) be the set of elements of S which are greater than (respectively, greater or equal to) c.

For a set S, let $\#S$ denote its cardinality. For $d \in \mathbb{N}^{>0}$, let S^d denote the set of d-tuples of elements of S. Let $a, b \in \mathbb{N}$, then $[b] = \{n \in \mathbb{N} \mid 0 \leq n \leq b\}$ and $[a, b] = \{n \in \mathbb{N} \mid a \leq n \leq b\}$. For $m \in \mathbb{N}^{>0}$ and $k \in [m-1]$, let $m\mathbb{N} + k$ be the set of non-negative integers congruent to k modulo m. For $d \in \mathbb{N}$, bold letters are used to denote d-tuples of variables, such as $\boldsymbol{x} \in \mathbb{N}^d$, which is an abbreviation for $(x_0, \cdots, x_{d-1})$. For $i \in [d-1]$, the variable x_i is called the i-th component of $\boldsymbol{x}$. Let $(\overline{m, \ldots, m}^d)$ denote the d-tuple $(m, \cdots, m)$. For $\boldsymbol{x} \in \mathbb{N}^d$, let $\min(\boldsymbol{x})$ denote $\min\{x_i \mid i \in [d-1]\}$, and let $\|\boldsymbol{x}\|$ denote $\sum_{i=0}^{d-1} x_i$. Finally, let $\boldsymbol{x} + \boldsymbol{y}$ denote $(x_0 + y_0, \cdots, x_{d-1} + y_{d-1})$ and let $\boldsymbol{x} - \boldsymbol{y}$ denote $(x_0 - y_0, \cdots, x_{d-1} - y_{d-1})$.

2.1. First-order logic over a vocabulary $\mathcal{V}$ ($\mathsf{FO}[\mathcal{V}]$)

In this section, the definitions concerning the logical formalisms of this paper are introduced.

Definition 2.1 (Universe). A *universe* $\mathcal{U}$ is a set. In this paper, $\mathcal{U}$ is always equal, either to $\mathbb{N}$ or to $[n]$ for $n \in \mathbb{N}$.

Definition 2.2 (Vocabulary). A *vocabulary* is a set of the form

$$\mathcal{V} = \{(R_i/d_i)_{i<n}, (c_i)_{i<q}\},$$

where n and q are either integers or ω (the cardinality of the set of integers).

For $i < n$, the R_i's are the *relation* symbols and their arity is d_i. For $i < q$, the c_i's are the *constant* symbols.

Definition 2.3 (Structure). Let $\mathcal{V}$ be a vocabulary. A $\mathcal{V}$-*structure* $\mathcal{S}$ over the universe $\mathcal{U}$ is a tuple

$$(\mathcal{U}, (R_i^{\mathcal{S}})_{i<n}, (c_i^{\mathcal{S}})_{i<q})$$

where $R_i^{\mathcal{S}} \subseteq \mathcal{U}^{d_i}$ for $i < n$, and $c_i^{\mathcal{S}} \in \mathcal{U}$ for $i < q$.

For ς a relational symbol with arity d and $\iota \subseteq \mathcal{U}^{[d-1]}$, let $\mathcal{S}[\varsigma/\iota]$ denote the $\mathcal{V} \cup \{\varsigma\}$-structure of universe $\mathcal{U}$ where $\varsigma^{\mathcal{S}[\varsigma/\iota]} = \iota$ and $\tau^{\mathcal{S}[\varsigma/\iota]} = \tau^{\mathcal{S}}$ for every symbol $\tau \in \mathcal{V} \setminus \{\varsigma\}$. The same definition is used for a constant symbol ς and $i \in \mathcal{U}$.

Definition 2.4 (Finite structure $\mathcal{S}_{|n}$). Let $n \in \mathbb{N}$, let $\mathcal{V}$ be a vocabulary and let $\mathcal{S}$ be a $\mathcal{V}$-structure of cardinality at least n and such that for each constant symbol c_i, $c_i^{\mathcal{S}} < n$ holds. Then $\mathcal{S}_{|n}$ is the $\mathcal{V}$-structure over the universe $[n-1]$ such that $c_i^{\mathcal{S}_{|n}} = c_i^{\mathcal{S}}$ and $R_i^{\mathcal{S}_{|n}} = R_i^{\mathcal{S}} \cap [n-1]^{d_i}$.

In this paper, if $\mathcal{V}$ contains a symbol with a standard interpretation over $\mathbb{N}$, such as "+" or "$<$" then the only $\mathcal{V}$-structures $\mathcal{S}$, of universe $\mathcal{U}$, which are considered, are the ones such that each symbol has its standard interpretation. For example, it is always assumed that $+^{\mathcal{S}} = \{(i,j,k) \in \mathcal{U}^3 \mid i+j=k\}$. In particular, for a vocabulary $\mathcal{V}$ which only contains symbols with a standard interpretation over $\mathbb{N}$, let $\mathcal{N}$ denote the structure which associates to each symbol its standard interpretation.

The *fragments* of first- and second-order logic used in this paper are now defined.

Definition 2.5 ($\mathcal{V}$-Formulas). Let $\mathcal{V}$ be a vocabulary. The set of quantifier-free $\mathcal{V}$-formulas, denoted by $\Sigma_0[\mathcal{V}]$ and $\Pi_0[\mathcal{V}]$, is defined by the grammar:

$$\Sigma_0[\mathcal{V}] ::= \neg\phi_0 \mid \phi_0 \wedge \phi_1 \mid \phi_0 \vee \phi_1 \mid R_i(c_0, \dots, c_{d_i-1}) \mid c_0 \doteq c_1$$

where the c_i's are constant symbols of $\mathcal{V}$ and the ϕ_i's are $\Sigma_0[\mathcal{V}]$-formulas.

The sets $\Sigma_i[\mathcal{V}]$, $\Pi_i[\mathcal{V}]$ and $D_i[\mathcal{V}]$ are defined by mutual recursion on i. For $i \in \mathbb{N}$ let $D_i[\mathcal{V}]$ be defined by the grammar:

$$D_i[\mathcal{V}] ::= \neg\phi_0 \mid \phi_0 \wedge \phi_1 \mid \phi_0 \vee \phi_1 \mid \psi$$

where ϕ_i belongs to $D_i[\mathcal{V}]$ and ψ belongs to $\Sigma_i[\mathcal{V}]$ or to $\Pi_i[\mathcal{V}]$.

For $i \in \mathbb{N}^{\geq 0}$, let $\Sigma_i[\mathcal{V}]$ be defined by the grammar :

$$\Sigma_i[\mathcal{V}] ::= \exists x.\psi \mid \phi_0 \wedge \phi_1 \mid \phi_0 \vee \phi_1 \mid \neg\chi \mid \xi$$

where $\psi \in \Sigma_i[\mathcal{V}, x]$, the ϕ_i's are $\Sigma_i[\mathcal{V}]$-formulas, χ is a $\Pi_i[\mathcal{V}]$-formula and ξ is a $D_{i-1}[\mathcal{V}]$-formula.

Similarly, let $\Pi_i[\mathcal{V}]$ be defined by the grammar:

$$\Pi_i[\mathcal{V}] ::= \forall x.\psi \mid \phi_0 \wedge \phi_1 \mid \phi_0 \vee \phi_1 \mid \neg\chi \mid \xi$$

where $\psi \in \Pi_i[\mathcal{V}, x]$, the ϕ_i's are $\Pi_i[\mathcal{V}]$-formulas, χ is a $\Sigma_i[\mathcal{V}]$-formula and ξ is a $D_{i-1}[\mathcal{V}]$-formula.

Let $\mathsf{FO}[\mathcal{V}]$ be the union of the fragments $\Pi_i[\mathcal{V}]$ for $i \in \mathbb{N}$.

For a logical fragment $\mathcal{L}[\mathcal{V}]$, let $\exists\mathsf{MSO}\mathcal{L}[\mathcal{V}]$ denote the set of formulas of the form $\exists R_0, \dots, R_n.\psi$ where ψ is a $\mathcal{L}[\mathcal{V}, R_0, \dots, R_n]$-formula and where the arity of the R_i's is 1.

Let $\phi_0 \implies \phi_1$ be an abbreviation for $\neg\phi_0 \vee \phi_1$ and let $\phi_0 \iff \phi_1$ be an abbreviation for $\phi_0 \implies \phi_1 \wedge \phi_1 \implies \phi_0$.

The arity and the curly brackets are omitted in logics' notations. For instance, $\mathsf{FO}[+]$ is written instead of $\mathsf{FO}[\{+/3\}]$.

For a vocabulary $\mathcal{V}$, and a logical fragment $\mathcal{L}$, the $\mathcal{L}[\mathcal{V} \cup \{x_0, \cdots, x_{d-1}\}]$-formulas are said to be $\mathcal{L}[\mathcal{V}]$-formulas with arity d. The x_i's for $i \in [d-1]$, are called the free variables and do not belong to $\mathcal{V}$. Given some $\mathcal{V}$-structure $\mathcal{S}$, the semantic of a formula is defined recursively as usual.

Definition 2.6 (Definability). Let $d \in \mathbb{N}$ and $\phi(x_0, \cdots, x_{d-1})$ be a formula with d free variables in a logic $\mathcal{L}[\mathcal{V}]$. For a $\mathcal{V}$-structure $\mathcal{S}$, the formula ϕ defines in $\mathcal{S}$ the d-ary set $\phi(\boldsymbol{x})^{\mathcal{S}} = \{\boldsymbol{x} \in \mathcal{U}^d \mid \mathcal{S} \models \phi(\boldsymbol{x})\}$.

A set $R \subseteq \mathcal{U}^d$ is said to be $\mathcal{L}[\mathcal{V}]$-definable in $\mathcal{S}$ if there exists $\phi(x_0, \cdots, x_{d-1}) \in \mathcal{L}[\mathcal{V}]$ such that $R = \phi(x_0, \cdots, x_{d-1})^{\mathcal{S}}$. Furthermore, if f is a function from $\mathcal{U}^d$ to $\mathcal{U}$, then f is $\mathcal{L}[\mathcal{V}]$-definable if its graph is $\mathcal{L}[\mathcal{V}]$-definable.

In some proofs, it is needed to define a set within another set.

Definition 2.7 (Defining in a set). Let $F \subseteq \mathbb{N}^d$, the formula $\phi(\boldsymbol{x})$ defines R in F if, for all $\boldsymbol{n} \in F$, $\mathcal{S}[\boldsymbol{x}/\boldsymbol{n}] \models \phi(\boldsymbol{x})$ if and only if $\boldsymbol{n} \in R$.

In this paper many definability results are given, and they are obtained by explicitly constructing a formula that defines the definable set. Hence for the sake of readability, some notations are introduced. When a formula which defines a set S is constructed, it is sometimes needed to state "If there exists some $(i \in F, \boldsymbol{x} \in \mathbb{N}^d)$ such that $\phi_i(\boldsymbol{x})$ holds, then we define S by $\chi_i(x)$, otherwise we define it by ψ". In this paper, there is always at most one i and one $\boldsymbol{x}$ for which the property $\phi_i(\boldsymbol{x})$ holds. Hence the following notation is introduced.

Definition 2.8. For $i \in F$, $j \in \mathbb{N}$, let $\phi_i(\boldsymbol{x}), \chi_i(\boldsymbol{x}) \in \Sigma_j[\mathcal{V}]$ and $\psi \in \Pi_j[\mathcal{V}]$ be formulas. Let $\langle \bigvee_{i \in F} \exists \boldsymbol{x}.\phi_i(\boldsymbol{x}) \mid \chi_i(\boldsymbol{x}) \mid \psi \rangle$ denote the $D_j[\mathcal{V}]$-formula:

$$[\bigvee_{i \in F} \exists \boldsymbol{x}.\phi_i(\boldsymbol{x}) \wedge \chi_i(\boldsymbol{x})] \vee [\bigwedge_{i \in F} \forall \boldsymbol{x}.\neg\phi_i(\boldsymbol{x}) \wedge \psi].$$

Finally, a notation must also be defined to consider functions.

Definition 2.9. Let $\phi \in \mathsf{FO}[<, R]$ be a formula with $\boldsymbol{x}, \boldsymbol{y}$ as free variables. Let $\phi(\boldsymbol{x}; \boldsymbol{y})$ denote the fact that for all $\boldsymbol{x}$, there exists exactly one $\boldsymbol{y}(\boldsymbol{x})$ such that $\phi(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{x}))$ holds. Then ϕ can be seen as a function mapping $\boldsymbol{x}$ to $\boldsymbol{y}(\boldsymbol{x})$ and $\phi(\boldsymbol{x})$ is used to denote $\boldsymbol{y}(\boldsymbol{x})$.

Functions Let us say a word about functions in this paper. Usually in finite model theory, the vocabulary does not contain function symbols, hence the only terms are the constants. Instead, formally, unary functions $f(x)$ are replaced by a binary relation $f(x, y)$, such that for every $n \in \mathcal{U}$ there is at most one $r \in \mathcal{U}$ such that $f(n, r)$ holds. For example, the addition of 1 is denoted as the binary relation $+_1(x, y)$ interpreted in $[n]$ by $\{(x, x+1) \mid x < n\}$. The distinction is important because over the universe $[n]$, the value of $n+1$ is undefined.

On the other hand, for the sake of clarity, "$+_1(x, y)$" is written "$x + 1 \doteq y$", and more generally if $f(x)$ is a function, then "$f(x, y)$" is written "$f(x) \doteq y$".

In this paper, the number of alternations of quantifiers of formulas is specified. For the sake of simplicity, we introduce abbreviations such as $f(g(x))$. The following lemmas explain how to encode formally all of those abbreviations in fragments of the logic.

Lemma 2.10. *Let f be a unary function symbol and x be a constant symbol. Let $\mathcal{V}$ be a vocabulary which contains x and f, and let $\mathcal{S}$ be a $\mathcal{V}$-structure. Let $\psi(x)$ be a $\Sigma_i[\mathcal{V}]$-formula with $i > 0$, then $\psi(f(x))$ is equivalent to a $\Sigma_i[\mathcal{V}]$-formula in $\mathcal{S}$.*

Proof. The formula $\psi(f(x))$ is equivalent to the $\Sigma_i[\mathcal{V}]$-formula:

$$\exists y.f(x) \doteq y \wedge \psi(y).$$

Clearly, if $f^{\mathcal{S}}(x^{\mathcal{S}})$ is not defined then this formula does not hold in $\mathcal{S}$, which is the intended behaviour as $f(x)$ is not defined either. □

The preceding lemma extends easily to the compositions of several function symbols.

2.2. Words

An alphabet α is a finite non-empty set. The set α^+ stands for the set of finite non-empty sequences of elements of α, which are called words.

It is now explained how to associate a $\mathcal{V}$-structure $\mathcal{S}$ to a word w. Unary predicates $(P_a)_{a\in\alpha}$ are introduced in the following definition and are added to $\mathcal{V}$ in order to encode w in $\mathcal{S}$. The predicates P_a are interpreted by the set of positions of the a's in the word w.

Definition 2.11. Let $\mathcal{V}$ be a vocabulary and let $\mathcal{S}$ be a $\mathcal{V}$-structure of universe $\mathbb{N}$. Let $w = w[0]\cdots w[n-1]$ with $w[i] \in \alpha$ be a word of length n over the alphabet α. Then let $\mathcal{S}(w)$ be the $\mathcal{V} \cup \left\{(P_a)_{a\in\alpha}\right\}$-structure $\mathcal{S}[(P_a/\{i \mid w[i] = a\})_{a\in\alpha}]$ of universe $[n-1]$.

2.3. Relations

The vocabularies $\mathcal{V}$ considered in this paper contain relation symbols whose interpretation is fixed. In this section, those symbols are listed, and their interpretation over the universe $\mathbb{N}$ is given. For the sake of readability, the infix notation is used when it is more standard than the prefix one.

Let x, y be variables. Let $c \in \mathbb{N}$, $r \in \mathbb{Q}^{>0}$, $m \in \mathbb{N}^{>0}$, $N \subseteq \mathbb{N}$ and $a \in [m-1]$ be constants.

- For $c \in \mathbb{N}$, the symbol c is also used as a constant symbol of the vocabulary. Then N represents the set of constant symbols belonging to N.
- Let $+_c$ be the relation $\{(n, n+c) \mid n \in \mathbb{N}\}$. Clearly $+_0$ is the equality relation and $+_1$ is the successor relation. The notation "$x + c \doteq y$" is used instead of "$+_c(x, y)$". Let $+_N$ denote the set of relations $\{+_c \mid c \in N\}$.
- Let "$\equiv a \bmod m$" denote the set $\{x \in \mathbb{N} \mid x \equiv a \mod m\}$, and let "$\bmod m$" be the set containing the m relations $\equiv a \mod m$. For $M \subseteq \mathbb{N}^{>0}$, let "$\bmod M$" denote the union of the relations $\bmod m$ for $m \in M$, and "$\bmod$" denote $\bmod \mathbb{N}^{>0}$. Then the notation "$x \equiv a \mod m$" is used instead of "$\equiv a \mod m(x)$". Let "$x + c \equiv y \mod m$" be an abbreviation for $\bigvee_{a=0}^{m-1} x \equiv a \mod m \wedge y \equiv (a+c) \mod m$.

3. First-order logic

All results concerning the manipulation of formulas and of sets of tuples of integers used in this paper are introduced in this section.

Section 3.1 shows how to adapt formulas about $\mathbb{N}$ to formulas about finite structures. Then Section 3.2 introduces the notions of sections, diagonals and subspaces of a set. Then Section 3.3 considers sets of integers and unary functions which are $\mathsf{FO}[<, \bmod]$-definable, $\mathsf{FO}[<, \bmod m]$-definable and $\mathsf{FO}[+]$-definable. Finally Section 3.4 shows that the set of $\mathsf{FO}[<, \bmod]$-formulas admits elimination of quantifiers.

3.1. Logic over $\mathbb{N}$ and over finite models.

In Section 5, finite models are considered, using results given in this section. Hence, a notion of *convergence* is needed to state that properties of formulas over $\mathbb{N}$ can also be used over finite models.

Definition 3.1 (Set convergence). Let $\mathcal{V}$ be a vocabulary. Let $\mathcal{S}$ be a $\mathcal{V}$-structure over $\mathbb{N}$. Let $\phi(\boldsymbol{x})$ be a formula with arity d such that its interpretation in $\mathcal{S}_{|n}$ (respectively, in $\mathcal{S}$) is a set E_n (respectively, E). Then $\phi(\boldsymbol{x})$ is said to be converging to E in $\mathcal{S}$, if for all $\boldsymbol{c} \in \mathbb{N}^d$, there exists $N \in \mathbb{N}$ such that for all $i \geq N$, $\boldsymbol{c} \in E_i$ if and only if $\boldsymbol{c} \in E$.

For formulas that define functions, this notion is equivalent to pointwise convergence.

Example 3.1. Let f_n be defined on $[n-1]$ by $f_n(c) = c + \lfloor \frac{2c}{n} \rfloor$. Then for all $c \in \mathbb{N}$, $(f_n(c))_{n>c}$ is a sequence of integers converging to c, so $(f_n)_{n\in\mathbb{N}}$ converges to the identity function.

In fact, in our proofs, the values of f_n are only studied on $[c_n - 1]$ for some $c_n \leq n$ and c_n increasing to infinity, i.e. the values of $f_n(b)$ for $b \geq c_n$ are not considered.

The following lemma is straightforward from the definition.

Lemma 3.2. *Let $f, g : \mathbb{N} \to \mathbb{N}$ be two converging functions, then $f \circ g$ is a converging function.*

3.2. Sections, diagonals and subsets

In this section some notations are introduced to transform sets into sets which are, intuitively, smaller.

Definition 3.3 (Section, diagonal and straight subspace). Let $d \geq 1$, $R \subseteq \mathbb{N}^d$, i, j be distinct elements of $[d-1]$, and $c \in \mathbb{N}$. Then we define the section $\sec(R; x_i = c) \subseteq \mathbb{N}^{d-1}$ as the set obtained from R by fixing the i-th component to c:

$$\sec(R; x_i = c) = \left\{(x_0, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{d-1}) \in \mathbb{N}^{d-1} \mid (x_0, \cdots, x_{i-1}, c, x_{i+1}, \cdots, x_{d-1}) \in R\right\}.$$

Similarly, the *diagonal* $\mathrm{diag}(R; x_i = x_j + c) \subseteq \mathbb{N}^{d-1}$ is defined as follows:

$$\mathrm{diag}(R; x_i = x_j + c) = \left\{(x_0, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{d-1}) \in \mathbb{N}^{d-1} \mid (x_0, \cdots, x_{i-1}, x_j + c, x_{i+1}, \cdots x_{d-1}) \in R\right\}.$$

The set of *straight subspaces* of R of dimension $d' \in [d-1]$ is defined by induction on $d - d'$ as $\{R\}$ if $d' = d$, and as the set of sections and diagonals of straight subspaces of dimension $d' + 1$ otherwise.

Example 3.2. The sections and diagonals of the addition relation $x_0 + x_1 = x_2$ are now studied. Let $c \in \mathbb{N}$. One has:

$$\begin{array}{llll}
\sec(+; x_0 = c) & = \sec(+; x_1 = c) & = \{(n, n + c) \mid n \in \mathbb{N}\}, & \sec(+; x_2 = c) = \{(n, c - n) \mid n \leq c\}, \\
\mathrm{diag}(+; x_0 = x_1 + c) & = \mathrm{diag}(+; x_1 = x_0 + c) & = \{(n, 2n + c) \mid n \in \mathbb{N}\}, & \mathrm{diag}(+; x_2 = x_1 + c) = \{(c, n) \mid n \in \mathbb{N}\}, \\
\mathrm{diag}(+; x_0 = x_2 + c + 1) & = \mathrm{diag}(+; x_1 = x_2 + c + 1) = \emptyset & & \mathrm{diag}(+; x_2 = x_0 + c) = \{(n, c) \mid n \in \mathbb{N}\}.
\end{array}$$

A notation for straight subspaces is now introduced. A pair of d-tuples $(\boldsymbol{t}, \boldsymbol{c})$ is associated with each straight subspace. The d-tuple $\boldsymbol{t}$ belongs to $(\{\mathrm{var}, \mathrm{const}\} \cup \{\mathrm{add}(i) \mid i \in [d-1]\})^d$. Intuitively $t_i = \mathrm{const}$ means that the i-th dimension is fixed, $t_i = \mathrm{add}(j)$ means that the distance between the j-th and the i-th dimension is fixed, and $t_i = \mathrm{var}$ means that the i-th dimension is free.

The formal definition is now introduced.

Definition 3.4. Let $d \in \mathbb{N}^{>0}$, let $\boldsymbol{t} \in (\{\mathrm{var}, \mathrm{const}\} \cup \{\mathrm{add}(i) \mid i \in [d-1]\})^d$ and let $\boldsymbol{c} \in \mathbb{N}^d$. Let $\dim(t)$ be the number of indices i such that $t_i = \mathrm{var}$. Then let $\mathrm{sub}\,(R; \boldsymbol{t}, \boldsymbol{c})$ denote the set

$$\left\{(x_i)_{t_i=\mathrm{var}} \in \mathbb{N}^{\dim(t)} \mid (x_0, \ldots, x_{d-1}) \in R, x_i = c_i \text{ if } t_i = \mathrm{const}, x_i = x_j + c_i \text{ if } t_i = \mathrm{add}(j)\right\}.$$

The following lemma is an easy consequence of the definition.

Lemma 3.5. *Let $d \in \mathbb{N}^{>0}$, $R \subseteq \mathbb{N}^d$, $d' < d$ and $S \subseteq \mathbb{N}^{d'}$. The two following statements are equivalent:*

1. *the set S is a straight subspace of $R \subseteq \mathbb{N}^d$,*
2. *there exists $(\boldsymbol{t}, \boldsymbol{c})$ such that $S = \mathrm{sub}\,(R; (\boldsymbol{t}, \boldsymbol{c}))$ and there are d' integers i such that $t_i = \mathrm{var}$.*

A formula which defines $\mathrm{sub}\,(R; \boldsymbol{t}, \boldsymbol{c})$ is now introduced.

Definition 3.6. Let $\mathcal{L}$ be a fragment of logic and let $\mathcal{V}$ be a vocabulary. Let $\chi(\boldsymbol{x})$ be an $\mathcal{L}[\mathcal{V}]$-formula. Let $\boldsymbol{t}, \boldsymbol{c}$ be as in Notation 3.4.

Then let $\mathrm{sub}(\chi, \boldsymbol{t}, \boldsymbol{c})(\boldsymbol{x})$ denote the $\mathcal{L}[\mathcal{V}, \mathbb{N}, +_{\mathbb{N}}]$-formula $\chi(\boldsymbol{y})$ defined as the conjunction of the following formulas, for each $i \in [d-1]$:

- $y_i = x_h$, if $t_i = \mathrm{var}$, and h is the i-th value such that $t_h = \mathrm{var}$,
- $y_i = c_i$ if $t_i = \mathrm{const}$ and
- $y_i = x_h + c_i$ if $t_i = \mathrm{add}(j)$, and h is the j-th value such that $t_h = \mathrm{var}$.

In particular, $\mathrm{sub}(R, \boldsymbol{t}, \boldsymbol{c})$ is a $\Sigma_0[\mathbb{N}, +_{\mathbb{N}}, R]$-formula.

The following lemma is a straightforward application of the definitions.

Lemma 3.7. *Let $d \in \mathbb{N}$, let $\mathcal{V}$ be a vocabulary and let $\mathcal{S}$ be a $\mathcal{V}$-structure. Let $\mathcal{L}$ be a logical fragment.*
Let $\chi(\boldsymbol{x})$ be an $\mathcal{L}[\mathcal{V}]$-formula, then $\mathrm{sub}(\chi(\boldsymbol{x}), \boldsymbol{t}, \boldsymbol{c})(\boldsymbol{x})^{\mathcal{S}}$ equals $\mathrm{sub}\left(\chi(\boldsymbol{x})^{\mathcal{S}}; \boldsymbol{t}, \boldsymbol{c}\right)$.

Definition 3.8 (Fixed order). Let $\mathbb{P}_d$ be the set of permutations of $[d-1]$. For every permutation σ of $\mathbb{P}_d$, let $\mathbb{N}^d_\sigma$ be the subset of $\mathbb{N}^d$ where for each $i \in [d-2]$, the relation $x_{\sigma(i)} \leq x_{\sigma(i+1)}$ holds. It is defined by the $\Sigma_0[<]$-formula $\bigwedge_{i=0}^{d-2} x_{\sigma(i)} \leq x_{\sigma(i+1)}$.

It should be noted that $\mathbb{N}^d$ is the union of the $\mathbb{N}^d_\sigma$'s. Therefore, for every vocabulary $\mathcal{V}$ which contains $<$, for every logical fragment $\mathcal{L}$, a set $R \subseteq \mathbb{N}^d$ is $\mathcal{L}[\mathcal{V}]$-definable if and only if for each $\sigma \in \mathbb{P}_d$, the set R is $\mathcal{L}[\mathcal{V}]$-definable in $\mathbb{N}^d_\sigma$.

3.3. FO[+]-, FO[<, mod]- and FO[<, mod m]-definable sets

Some well-known facts about FO[+]-, FO[<, mod]- and FO[<, mod m]-definable sets are now stated. The first lemma concerns unary sets which are FO[+] or FO[<, mod]-definable. Before giving this lemma, the following definition is needed.

Definition 3.9 (Ultimately (m-)periodic). A set $R \subseteq \mathbb{N}$, is *ultimately m-periodic* if there exists an integer $\tau \in \mathbb{N}$ (which is called the threshold) such that for all $n \geq \tau$, $n \in R$ if and only if $n + m \in R$. In particular the ultimately 1-periodic unary sets are the finite or co-finite sets. The set R is *ultimately periodic* if it is ultimately m-periodic for some $m \in \mathbb{N}^{>0}$.

For R an ultimately periodic set of integers, let $\mathrm{thres}_m(R) \in \mathbb{N}$ denote the minimal integer such that for all $n \geq \mathrm{thres}_m(R)$, $n \in R$ is equivalent to $n + m \in R$.

Let $\mathrm{mod}_m(R) \subseteq [m-1]$ denote the subset of $[m-1]$ such that, for all $n \geq \mathrm{thres}_m(R)$, $n \in R$ if and only if $n \equiv p \mod m$ for some $p \in \mathrm{mod}_m(R)$.

Intuitively, $\mathrm{mod}_m(R)$ is the set of congruence classes modulo m which ultimately belong to R, and $[m-1] \setminus \mathrm{mod}_m(R)$ is the set of congruence classes modulo m which are ultimately disjoint from R.

Lemma 3.10 ([Pre27]). *Let R be a subset of $\mathbb{N}$.*

- *The set R is FO[<, mod m]-definable if and only if it is ultimately m-periodic.*
- *The set R is FO[+]-definable if and only if it is ultimately periodic.*
- *The set R is FO[<, mod]-definable if and only if it is ultimately periodic.*

The second lemma concerns unary functions over integers which are FO[+]-definable or FO[<, mod m]-definable.

Lemma 3.11 ([FL08]). *A unary function $f : \mathbb{N} \to \mathbb{N}$ is FO[+]-definable if and only if there exist $m, \tau \in \mathbb{N}$, $r_0, \cdots, r_{m-1} \in \mathbb{Q}^{\geq 0}$ and $s_0, \cdots, s_{m-1} \in \mathbb{Q}$ such that $f(n) = r_j n + s_j$ for all $n \geq \tau$ such that $n \equiv j \mod m$.*
Furthermore, f is FO[<, mod]-definable if and only if $r_k \in \{0, 1\}$ for every $k \in [m-1]$.

It should be noted that if f is a unary $\mathsf{FO}[+]$-definable function which is not $\mathsf{FO}[<, \mathrm{mod}]$-definable, then one of the r_i is different from 0 and 1.

In particular, several proofs of this paper involve $\mathsf{FO}[+]$-definable functions which are increasing. They are now characterized.

Lemma 3.12. *A unary increasing function $f : \mathbb{N} \to \mathbb{N}$ is $\mathsf{FO}[+]$-definable if and only if there exist $m \in \mathbb{N}^{>0}$, $\tau \in \mathbb{N}$, $r \in \mathbb{Q}^{\geq 0}$ and $s_0, \cdots, s_{m-1} \in \mathbb{Q}$ such that $f(n) = rn + s_j$ for all $n \geq \tau$ such that $n \cong j \mod m$.*

The value of r is called the *slope* of f.

Proof. By Lemma 3.11, there exist $m \in \mathbb{N}^{>0}$, $\tau \in \mathbb{N}$, $r_0, \cdots, r_{m-1} \in \mathbb{Q}^{\geq 0}$ and $s_0, \cdots, s_{m-1} \in \mathbb{Q}$ such that $f(n) = r_j n + s_j$ for all $n \geq \tau$ such that $n \cong j \mod m$. It suffices to prove that $r_i = r_j$ for all $i, j \in [m-1]$.

For the sake of contradiction, let us assume that there exist $i, j \in [m-1]$ such that $r_i > r_j$. Without loss of generality, let us assume that $j > i$. Since $j > i$ and f is increasing, $f(n+j) \geq f(n+i)$.

Let $n \in \mathbb{N}$ be such that $n \equiv 0 \mod m$ and $n > \max\left(\frac{r_j(j-i)+s_j-s_i}{r_i-r_j}, \tau\right)$. Then:

$$\begin{aligned} n > \tfrac{r_j j - r_i i + s_j - s_i}{r_i - r_j} &\iff (r_i - r_j)n &> r_j j - r_i i + s_j - s_i, \\ &\iff r_i(n+i) + s_i &> r_j(n+j-i) + s_j, \\ &\iff f(n+i) &> f(n+j). \end{aligned}$$

which is a contradiction. □

Some proofs of this paper consider increasing $\mathsf{FO}[+]$-definable functions whose slope is strictly greater than 1. It is now proven that such a function allows to define a function whose slope is as great as needed.

Lemma 3.13. *Let f be a unary function symbol and let $\mathcal{S}$ be a $\{f, <\}$-structure such that $f^{\mathcal{S}}$ is an increasing $\mathsf{FO}[+]$-definable function with slope $c > 1$. Let $c', e \in \mathbb{N}^{>0}$. There exists a converging $\Sigma_1[\mathbb{N}, <, +_{\mathbb{N}}, f]$-formula $\mu_{f\mathcal{S}}(x; y)$ which defines a function $g : \mathbb{N} \to \mathbb{N}$ in $\mathcal{S}$ such that for all $n \in \mathbb{N}$, $g(n) \geq c'n + e$.*

Proof. The proof consists in two parts: defining a function $h : \mathbb{N} \to \mathbb{N}$ with slope at least c', and then defining the function g.

By an easy induction on i, the i-th iterate f^i of f, is increasing, $\mathsf{FO}[+]$-definable, and its slope is c^i. Let $k = \max(1, \lceil \log_c(c') \rceil)$. It suffices to set h equal to f^k.

Let us assume that h is increasing, $\mathsf{FO}[+]$-definable, with slope at least c. Then there exists τ, m and $s_0, \ldots, s_{m-1}$ as in Lemma 3.11. Then it suffices to take $g(n)$ to be:

- $c'n + e$ for $n < \tau$, and
- $h(n) + \lceil e - \min(\{s_i \mid i \in [d-1]\} \cup \{e\}) \rceil$ for $n \geq \tau$

It should be noted that $\min(\{s_i \mid i \in [d-1]\} \cup \{e\}) \leq e$, hence $e - \min(\{s_i \mid i \in [d-1]\} \cup \{e\}) \geq 0$.

The formula $\mu_{f\mathcal{S}}(x; y)$ is then:

$$y \doteq f^k(x) + e - \min(\{s_i \mid i \in [d-1]\} \cup \{e\}).$$

By Lemma 3.2, the set of converging functions is closed by composition, hence f^k is converging. □

3.4. Quantifier elimination for $\mathsf{FO}[<, \mathrm{mod}]$ and $\mathsf{FO}[<, \mathrm{mod}\, m]$

Cooper's algorithm [Coo72] transforms a $\mathsf{FO}[+]$-formula into an equivalent $\Sigma_0[\mathbb{N}, +, <, \mathrm{mod}\, m]$-formula for some $m \in \mathbb{N}$.

Let $M \subseteq \mathbb{N}$ be closed under the least common multiple operation. In this paper, M is always equal, either to a singleton or to $\mathbb{N}$. Cooper's algorithm is now given for the vocabulary $\{\mathbb{N}, <, +_{\mathbb{N}}, \mathrm{mod}\, M\}$. The algorithm Cooper:

- takes as input a $\mathsf{FO}[\mathbb{N}, <, +_{\mathbb{N}}, \mathrm{mod}\, M]$-formula $\phi(\boldsymbol{x})$ and
- returns a boolean combination of atomic formulas of the form $x_i + c \sim x_j$, $x_i \sim c$, and $x_i \equiv k \mod m$ with $\sim$ belonging to $\{<, \doteq, >\}$, $c \in \mathbb{N}$, $m \in M$ and $k \in [m-1]$.

Definition 3.14 (Cooper). Let $\phi(\boldsymbol{x})$ be a $\mathsf{FO}[\mathbb{N}, <, +_{\mathbb{N}}, \mathrm{mod}\, M]$-formula. Then, let $\mathrm{Cooper}(\phi)(\boldsymbol{x})$ be the $\Sigma_0[\mathbb{N}, <, +_{\mathbb{N}}, \mathrm{mod}\, M]$-formula defined by induction on ϕ as follows:

- If $\phi(\boldsymbol{x})$ is an atomic formula then $\mathrm{Cooper}(\phi)(\boldsymbol{x}) = \phi(\boldsymbol{x})$.
- If $\phi(\boldsymbol{x})$ is of the form $\neg\psi(\boldsymbol{x})$, then $\mathrm{Cooper}(\phi)(\boldsymbol{x}) = \neg\, \mathrm{Cooper}(\phi)(\boldsymbol{x})$.
- If $\phi(\boldsymbol{x})$ is of the form $\psi_0(\boldsymbol{x}) \vee \psi_1(\boldsymbol{x})$, then $\mathrm{Cooper}(\phi)(\boldsymbol{x}) = \mathrm{Cooper}(\psi_0)(\boldsymbol{x}) \vee \mathrm{Cooper}(\psi_1)(\boldsymbol{x})$.
- If $\phi(\boldsymbol{x})$ is of the form $\psi_0(\boldsymbol{x}) \wedge \psi_1(\boldsymbol{x})$, then $\mathrm{Cooper}(\phi)(\boldsymbol{x}) = \mathrm{Cooper}(\psi_0)(\boldsymbol{x}) \wedge \mathrm{Cooper}(\psi_1)(\boldsymbol{x})$.
- If ϕ is of the form $\exists y.\psi(y, \boldsymbol{x})$: let $\psi'(y, \boldsymbol{x}) = \mathrm{Cooper}(\psi)(y, \boldsymbol{x})$. By induction, $\psi'(y, \boldsymbol{x})$ is a boolean combination of atomic formulas of the form:
  - $y < t_i$ or $y = t_i$ or $y > t_i$, where t_i is either a non-negative integer or a term of the form $x_i + c$ with $c \in \mathbb{Z}$, or
  - $y \equiv c_i \mod m_i$, where m_i belongs to M and $c_i \in [m_i - 1]$.

  Let m be the least common multiple of the m_i, it belongs to M. Let $T(\psi'(y, \boldsymbol{x}))$ be the set containing the integer 0 and the terms t_i. Then let ψ'' be

$$\bigvee_{t_i \in T(\psi'(y,\boldsymbol{x}))} \bigvee_{j \in [m]} \psi'(t_i + j, \boldsymbol{x}).$$

  It should be noted that $t_i + j$ is of the form $x_i + c + j$ with $c \in \mathbb{Z}$, hence $c + j$ may be negative. In this case $\psi''(\boldsymbol{x})$ does not belong to $\Sigma_0[\mathbb{N}, <, +_{\mathbb{N}}, \mathrm{mod}\, M]$. It remains to modify the formula to restrict additions to positive integers.

  Let $\mathrm{Cooper}(\phi)(\boldsymbol{x})$ be the formula $\psi''(\boldsymbol{x})$ where:
  - each atomic formula of the form $x_i + c \sim x_i + c'$, with $\sim$ being $<$, $\doteq$ or $>$, is replaced by true if $c \sim c'$, and is replaced by false otherwise,
  - each atomic formula of the form $x_i + c \sim x_j + c'$ with $i \neq j$, $c \leq c'$ and $\sim$ belonging to $\{<, \doteq, >\}$, is replaced by $x_i \sim x_j + (c' - c)$,
  - each atomic formula of the form $x_i + c \sim c'$ with $c \leq c'$ and $\sim$ belonging to $\{<, \doteq, >\}$, is replaced by $x_i \sim (c' - c)$,
  - each atomic formula of the form $x_i + c \sim c'$ with $c > c'$ and $\sim$ being $<$, $\doteq$ or $>$, is replaced by the formula false, false and true respectively, and,
  - each atomic formula of the form $x_i + c \equiv c' \mod m$ is replaced by $x_i \equiv c' - c \mod m$.
- Finally, if $\phi(\boldsymbol{x})$ is of the form $\forall y.\psi(y, \boldsymbol{x})$, then $\mathrm{Cooper}(\phi)(\boldsymbol{x}) = \neg\, \mathrm{Cooper}(\exists y.\neg\psi)(y, \boldsymbol{x})$.

Example 3.3. Let ϕ be the formula $\exists x.(x \doteq y \wedge x \doteq 0) \vee (x + 4 \leq y \wedge x \equiv 0 \mod 2)$. Then $T(\phi) = \{y, y-4, 0\}$ and $m = 2$. Then ψ' is equal to :

$$\begin{array}{l}
\bigvee_{i=0}^{2} \{(\quad i \doteq y \wedge \quad i \doteq 0) \vee (\quad i + 4 \leq y \wedge \quad i \equiv 0 \mod 2)\} \vee \\
\bigvee_{i=-4}^{-2} \{(y + i \doteq y \wedge y + i \doteq 0) \vee (y + i + 4 \leq y \wedge y + i \equiv 0 \mod 2)\} \vee \\
\bigvee_{i=0}^{2} \{(y + i \doteq y \wedge y + i \doteq 0) \vee (y + i + 4 \leq y \wedge y + i \equiv 0 \mod 2)\},
\end{array}$$

hence $\mathrm{Cooper}(\phi)$ is equal to

$$\begin{array}{ll}
\bigvee_{i=0}^{2} & \{(i \doteq y \ \wedge \ i \doteq 0 \ ) \vee (i + 4 \leq y \ \wedge \ i \equiv 0 \mod 2)\} \vee \\
\bigvee_{i=-3}^{-2} & \{(\ \mathrm{false} \wedge y \doteq -i) \vee (\quad \mathrm{false} \wedge y \equiv i \mod 2)\} \vee \\
 & \{(\ \mathrm{false} \wedge y \doteq 4 \ ) \vee (\quad \mathrm{true} \wedge y \equiv 0 \mod 2)\} \vee \\
\bigvee_{i=1}^{2} & \{(\ \mathrm{false} \wedge \ \mathrm{false} \ ) \vee (\quad \mathrm{false} \wedge y \equiv i \mod 2)\} \vee \\
 & \{(\ \mathrm{true} \ \wedge y \doteq 0 \ ) \vee (\quad \mathrm{false} \wedge y \equiv 0 \mod 2)\}.
\end{array}$$

Let us prove that the algorithm Cooper behaves as expected.

Proposition 3.15. *Let $\phi(\boldsymbol{x}) \in \mathsf{FO}[\mathbb{N}, <, +_{\mathbb{N}}, \operatorname{mod} M]$, then $\operatorname{Cooper}(\phi)(\boldsymbol{x})$ belongs to $\Sigma_0[\mathbb{N}, <, +_{\mathbb{N}}, \operatorname{mod} M]$.*
For all $\boldsymbol{n} \in \mathbb{N}^d$, let $\mathcal{N}' = \mathcal{N}[\boldsymbol{x}/\boldsymbol{n}]$. Then $\mathcal{N}' \models \phi(\boldsymbol{x})$ is equivalent to $\mathcal{N}' \models \operatorname{Cooper}(\phi)(\boldsymbol{x})$.

The proof of this proposition follows closely the proof of [Coo72].

Proof. It is trivial to check by induction on $\phi(\boldsymbol{x})$ that $\operatorname{Cooper}(\phi)(\boldsymbol{x})$ belongs to $\Sigma_0[\mathbb{N}, <, +_{\mathbb{N}}, \operatorname{mod} M, \boldsymbol{x}]$. Let $\boldsymbol{n} \in \mathbb{N}^d$ and let $\mathcal{N}' = \mathcal{N}[\boldsymbol{x}/\boldsymbol{n}]$. Let us prove that $\mathcal{N}' \models \phi(\boldsymbol{x})$ is equivalent to $\mathcal{N}' \models \operatorname{Cooper}(\phi)(\boldsymbol{x})$. The proof goes by induction on ϕ. The only non-trivial case is when $\phi(\boldsymbol{x})$ is of the form $\exists x.\psi(y, \boldsymbol{x})$.

Let $\psi'(y, \boldsymbol{x})$, $T(\psi')$ and m be as in Definition 3.14. By the induction hypothesis $\psi(y, \boldsymbol{x})$ is equivalent to $\psi'(y, \boldsymbol{x})$. Furthermore, $\mathcal{N}' \models \psi''(\boldsymbol{x})$ is easily equivalent to $\mathcal{N}' \models \operatorname{Cooper}(\phi)(\boldsymbol{x})$. Hence, it remains to prove that $\mathcal{N}' \models \exists y.\psi'(y, \boldsymbol{x})$ is equivalent to $\mathcal{N}' \models \psi''(\boldsymbol{x})$.

Let us first assume that $\mathcal{N}' \models \psi''(\boldsymbol{x})$, and let us prove that $\mathcal{N}' \models \exists y.\psi'(y, \boldsymbol{x})$. Since $\mathcal{N}' \models \psi''(\boldsymbol{x})$, there exists $t \in T(\psi')$ and $j \in [m]$ such that $\mathcal{N}' \models \psi'(t_i + j, \boldsymbol{x})$, hence $\mathcal{N}'[y/t_i'^{\mathcal{N}'} + j] \models \psi'(y, \boldsymbol{x})$ and then $\mathcal{N}' \models \exists y.\psi'(y, \boldsymbol{x})$.

Let us now assume that $\mathcal{N}' \models \exists y.\psi'(y, \boldsymbol{x})$, and let us prove that $\mathcal{N}' \models \psi''(\boldsymbol{x})$. Let $n \in \mathbb{N}$ be the least integer such that $\mathcal{N}'[y/n] \models \psi'(y, \boldsymbol{x})$. Let us prove that n is of the form $t + j$ with $t \in T(\phi')$ and $j \in [m]$, which implies that $\mathcal{N}' \models \psi''(\boldsymbol{x})$. For the sake of contradiction, let us assume that n is not of the form $t + j$ with $t \in T(\phi')$. Hence $n > m$. Then, by an easy induction on the subterms ξ of ϕ', $\mathcal{N}'[y/n] \models \xi$ is equivalent to $\mathcal{N}'[y/n - m] \models \xi$, which contradicts the minimal hypothesis about n. □

4. Characterization of FO $[<, \operatorname{mod}]$- and of FO $[<, \operatorname{mod} m]$-definable sets

In this section, two characterizations of FO$[<, \operatorname{mod} m]$-definable sets and one characterization of FO$[<, \operatorname{mod}]$-definable sets are given.

In Section 4.1, a first characterization of FO$[<, \operatorname{mod} m]$-definable sets is given, in term of local properties and recursive properties. In Section 4.2, it is proven that, from a FO $[+]$-definable set which is not FO$[<, \operatorname{mod}]$-definable, a unary functions which is not FO$[<, \operatorname{mod}]$-definable can be defined. Finally, in Section 4.3, it is proven that, from a set of tuples of integers which is not FO$[<, \operatorname{mod} m]$-definable, a set of integers which is not FO$[<, \operatorname{mod} m]$-definable can be defined.

4.1. Local and recursive characterization of FO$[<, \operatorname{mod} m]$-definable sets

In this section, a characterization of FO$[<, \operatorname{mod} m]$-definable-sets is given. Then some easy consequences of this characterization are given. This characterization is similar to Muchnik's characterization of FO $[\mathbb{N}, +]$, which is first recalled. For this, two notions are first introduced.

The notion of cube is now introduced.

Definition 4.1. Let $d \in \mathbb{N}$, $R \subseteq \mathbb{N}^d$, $\boldsymbol{x} \in \mathbb{N}^d$ and let $k \in \mathbb{N}$. Then let $C(\boldsymbol{x}, k)$ denote the cube with size k and whose minimal element is $\boldsymbol{x}$. Formally, it is defined as:

$$C(\boldsymbol{x}, k) = \left\{\boldsymbol{x} + \boldsymbol{c} \mid \boldsymbol{c} \in [k-1]^d\right\}$$

The notion of P-periodicity for a set P of tuples is now introduced.

Definition 4.2 ($\boldsymbol{p}$-periodicity in $\boldsymbol{F}$). Let $S, F \subseteq \mathbb{N}^d$, and $\boldsymbol{p} \in \mathbb{N}^d$. Then S is $\boldsymbol{p}$-periodic in F if for all $\boldsymbol{x} \in F$ such that $\boldsymbol{x} + \boldsymbol{p} \in F$, $\boldsymbol{x} \in S$ is equivalent to $\boldsymbol{x} + \boldsymbol{p} \in S$.

For $m \in \mathbb{N}$, S is said to be m-periodic in F if it is $(\overline{m, \dots, m}^d)$-periodic in F.

For $P \subseteq \mathbb{Z}^d$, a set of possible periodicities, the set S is said to be P-periodic in F if S is p-periodic for some $p \in P$.

Muchnik's theorem [Muc03, BHMV94] which characterizes FO $[+]$ can now be stated.

Theorem 4.3 ([Muc03, Theorem 1]). *Let $d \in \mathbb{N}^{>0}$ and $R \subseteq \mathbb{N}^d$. The following properties are equivalent;*

1. *The set R is* $\mathsf{FO}[+]$*-definable.*
2. *(a) All sections of R are* $\mathsf{FO}[+]$*-definable and*
   (b) there exists a finite set $P \subseteq \mathbb{Z}^d \setminus \{(\overline{0,\dots,0}^d)\}$ such that for every $k \in \mathbb{N}$, there exists $t \in \mathbb{N}$ such that for all $\boldsymbol{x} \in \mathbb{N}^d$ with $\|\boldsymbol{x}\| > t$, R is P-periodic in $C(\boldsymbol{x}, k)$.

Moreover the latter property only has to be verified for $k = \sum_{p\in P} \|p\|$.

Intuitively the set P is the set of possible periods, k is the size of the cube and t is the "threshold" for the periodicity.

A similar theorem for $\mathsf{FO}[<, \operatorname{mod} m]$-definable sets is now given.

Theorem 4.4. *Let $d \in \mathbb{N}^{>0}$, $R \subseteq \mathbb{N}^d$, and $m \in \mathbb{N}$. The following three statements are equivalent:*

1. *The set R is* $\mathsf{FO}[<, \operatorname{mod} m]$*-definable.*
2. *(a) There exists $t \in \mathbb{N}$ such that R is m-periodic in $\left(\mathbb{N}^{\geq t}\right)^d$ and*
   (b) if $d > 1$, then all sections and diagonals of R are $\mathsf{FO}[<, \operatorname{mod} m]$*-definable.*
3. *For every straight subspace S of dimension $d' \geq 1$ of R, there exists $t \in \mathbb{N}$ such that S is $(\overline{m,\dots,m}^{d'})$-periodic in $\left(\mathbb{N}^{\geq t}\right)^{d'}$.*

The notation $\operatorname{thres}_m(R)$ is introduced in Definition 3.9 for $R \subseteq \mathbb{N}$. It is now generalized to any dimension $d \in \mathbb{N}^{>0}$.

Definition 4.5. Let $d \in \mathbb{N}^{>0}$, $R \subseteq \mathbb{N}^d$, and $m \in \mathbb{N}$. If R satisfies Condition 2b of Theorem 4.4, let $\operatorname{thres}_m(R)$ denote the least integer t which satisfies this condition.

An example of application of Theorem 4.4 is given first.

Example 4.1. As seen in Example 3.2, all strict straight subspaces of + are $\mathsf{FO}[<, \operatorname{mod}]$-definable sets. So + satisfies the criterion (2b) of Theorem 4.4. For each $l \in \mathbb{N}$, one has $(l, l, 2l) \in +$. For every $m \in \mathbb{N}$, the 3-tuple $(l+m, l+m, 2l+m)$ does not belong to the relation +, so + does not satisfy the criterion (2a).

Let us consider another example. Let $R = \left\{(0, n^2) \mid n \in \mathbb{N}\right\}$. The set R is 1-periodic in $(\mathbb{N}^{\geq 1})^2$, so it satisfies the criterion (2a). But clearly $R^{x_0=0} = \left\{n^2 \mid n \in \mathbb{N}\right\}$ is not $\mathsf{FO}[<, \operatorname{mod}]$-definable, so it does not satisfy the criterion (2b).

Theorem 4.4 is now proven.

Proof of Theorem 4.4. Let us show by induction on $d \in \mathbb{N}$ that properties (1), (2) and (3) are equivalent for every $R \subseteq \mathbb{N}^d$.

Let us first assume that $d = 1$. Lemma 3.10 proves the equivalence of (1) and of (2). Since R is the only straight subspace of R of dimension at least 1, (3) and (2a) are equivalent. Moreover, (2b) holds since $d = 1$. Hence (2) and (3) are equivalent.

Let us now assume that $d > 1$ and that the property holds for $d - 1$.

Proof of (1) $\implies$ (3) Assume that R is $\mathsf{FO}[<, \operatorname{mod} m]$-definable and let us prove that for every integer $d' \leq d$, and every straight subspace S of R of dimension d', there exists $\operatorname{thres}_m(S)$ such that S is $(\overline{m,\dots,m}^{d'})$-periodic in $(\mathbb{N}^{\geq \operatorname{thres}_m(S)})^{d'}$.

By Property 3.15, there exists a quantifier-free $\Sigma_0[\mathbb{N}, <, +_{\mathbb{N}}, \operatorname{mod} M]$-formula $\chi(\boldsymbol{x})$ which defines R. Let S be a straight subspace of R, then by Lemma 3.7, there exists a $\Sigma_0[\mathbb{N}, <, +_{\mathbb{N}}, \operatorname{mod} M]$-formula $\chi_S(\boldsymbol{x})$ which defines S. Let $\operatorname{thres}_m(S) = \max\{c \in \mathbb{N} \mid x_i \sim c \text{ appears in } \chi_S(\boldsymbol{x}), \sim \text{ being } <,= \text{ or } >\} + 1$.

Let us prove by induction on the subformulas $\xi(\boldsymbol{x})$ of $\chi_S(\boldsymbol{x})$ that $\xi(\boldsymbol{x})^{\mathcal{N}}$ is $(\overline{m,\dots,m^d})$-periodic in $(\mathbb{N}^{\geq \mathrm{thres}_m(S)})^{d'}$. Let $\boldsymbol{n} \in \mathbb{N}^{d'}$ with $\min(\boldsymbol{n}) > \mathrm{thres}_m(S)$, it must be proven that $\xi(\boldsymbol{n})$ is equivalent to $\xi(\boldsymbol{n} + (\overline{m,\dots,m^d}))$. The induction case is trivial. Let us consider the base case:

- if $\xi(\boldsymbol{x})$ is of the form $x_i \equiv a \mod m$, then $n_i \equiv a \mod m$ is equivalent to $n_i + m \equiv a \mod m$,
- if $\xi(\boldsymbol{x})$ is of the form $x_i - x_j \sim c$ for $\sim$ being $<$ or $\doteq$, then $n_i - n_j \sim c$ is equivalent to $(n_i + m) - (n_j + m) \sim c$,
- if $\xi(\boldsymbol{x})$ is of the form $x_i < c$ or $x_i \doteq c$, then by hypothesis $c < \mathrm{thres}_m(S)$, hence neither $\xi(\boldsymbol{n})$ nor $\xi(\boldsymbol{n} + (\overline{m,\dots,m^d}))$ hold.
- if $\xi(\boldsymbol{x})$ is of the form $x_i > c$ then by hypothesis $c < \mathrm{thres}_m(S)$, hence $\xi(\boldsymbol{n})$ and $\xi(\boldsymbol{n} + (\overline{m,\dots,m^d}))$ hold.

Proof of (3) $\implies$ (2) The set R is a straight subspace of R hence by hypothesis, there exists $\mathrm{thres}_m(R) \in \mathbb{N}$ such that R is $(\overline{m,\dots,m^d})$-periodic in $(\mathbb{N}^{\geq \mathrm{thres}_m(R)})^d$.

Proof of (2) $\implies$ (1) Let R be such that all sections $\mathrm{sec}(R; x_i = c)$ and diagonals $\mathrm{diag}(R; x_i = x_j + c)$ are $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Then, for all distinct elements $i, j \in [d-1]$ and for all $c \in \mathbb{N}$, let $\psi_{\mathrm{sec}(R;x_i=j)}(x_0, \cdots, x_{d-2})$ and $\psi_{\mathrm{diag}(R;x_i=x_j+c)}(x_0, \cdots, x_{d-2})$ be $\mathsf{FO}[\mathbb{N}, <, +_{\mathbb{N}}, \mathrm{mod}\, M]$-formulas which define $\mathrm{sec}(R; x_i = c)$, and $\mathrm{diag}(R; x_i = x_j + c)$, respectively. Those formulas exist by hypothesis.

Figure 1 serves as example for the proof. It represents the set R defined by the formula $\phi(x_0, x_1)$ equal to

$$\{(x_0 + 2 \geq x_1 \wedge x_0 + 4 \not\doteq x_1) \vee x_1 \doteq 1\} \wedge x_0 \equiv 1 \mod 2. \tag{1}$$

In this example, $m = 2$ and $\mathrm{thres}_m(R) = 2$.

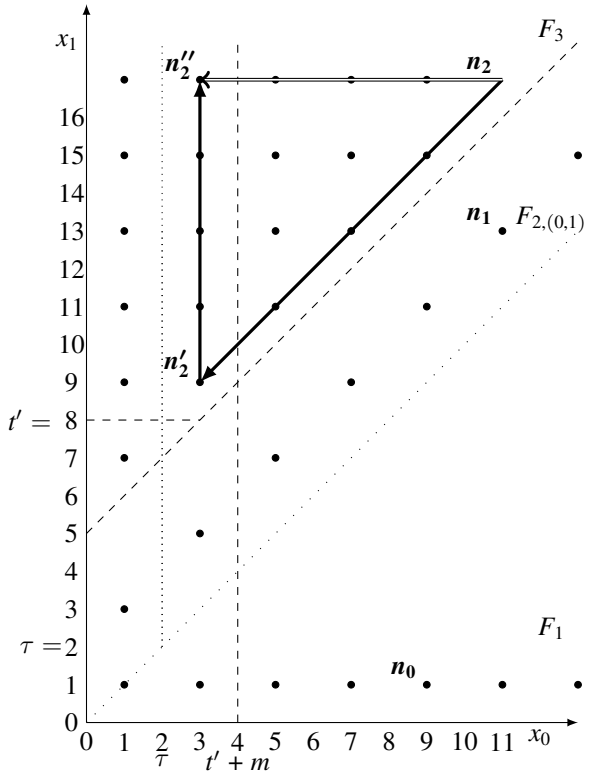

Figure 1: $\{(x_0 + 2 \geq x_1 \wedge x_0 + 4 \not\doteq x_1) \vee x_1 \doteq 1\} \wedge x_0 \equiv 1 \mod 2$

A $\subseteq$-decreasing sequence $F_0, F_1, F_{2,\sigma}, F_3, F_4$ of subsets of $\mathbb{N}^d$, with $F_0 = \mathbb{N}^d$, is introduced. The proof consists in reducing the problem of defining R in F_i to the problem of defining R in F_{i+1}. A formula $\phi_i(\boldsymbol{x})$ is then given, such that $\phi_i(\boldsymbol{x})^{\mathcal{N}}$ defines R in F_i. The formula ϕ_i uses a formula which defines R in F_{i+1}, a formula which defines R in $F_i \setminus F_{i+1}$ and the formula ϕ_{i+1}.

It is proven that $\phi_i(\boldsymbol{n})$ defines R in F_i.

The set F_0 Let $\tau = \operatorname{thres}_m(R)$ and let $F_1 = (\mathbb{N}^{\geq\tau})^d$. The set F_1 is considered, because, by definition of τ, the set R is m-periodic in F_1.

The formula ϕ_0 is constructed using a formula which defines R in $F_0 \setminus F_1$, a formula which defines R in F_1 and a formula which defines F_1. It should be noted that $F_0 \setminus F_1$ is included in a finite union of sections of R, and hypothesis (2a) can be used on each of those sections. Let us assume that $\phi_1(\boldsymbol{x})$ defines R in F_1. Then let ϕ_0 be:

$$\phi_0(\boldsymbol{x}) = \langle \bigvee_{i=0}^{d-1} \bigvee_{j=0}^{l-1} x_i \doteq j \mid \psi_{\sec(R;x_i=j)}(x_0, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{d-2}) \mid \phi_1(\boldsymbol{x}) \rangle.$$

In the example of Figure 1, let $\boldsymbol{n}_0 = (9,1)$, $\boldsymbol{n}_1 = (11,11)$ and $\boldsymbol{n}_2 = (11,17)$, then $\boldsymbol{n}_0 \notin F_1$ and $\boldsymbol{n}_1, \boldsymbol{n}_2 \in F_1$. The vertical and horizontal dotted line represent the boundary of F_1.

Let us prove that $\boldsymbol{n} \in R$ if and only if $\mathcal{N}' \models \phi_0(\boldsymbol{x})$ holds, where $\mathcal{N}' = \mathcal{N}[\boldsymbol{x}/\boldsymbol{n}]$. Let us first assume that $\boldsymbol{n} \in R$ and let us prove that $\mathcal{N}' \models \phi_0(\boldsymbol{x})$. Two cases must be considered, depending on whether $\min(\boldsymbol{n}) < \tau$ or not.

- Let us assume that $\min(\boldsymbol{n}) < \tau$, then there exist $i \in [d-1]$ and $j \in [\tau - 1]$ such that $n_i = j$. Hence $\mathcal{N}' \models x_i \doteq j$. Let $\boldsymbol{n'}$ be $\boldsymbol{n}$ without its i-th component. By definition of section, $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n'} \in \sec(R; x_i = j)$, and by definition of $\psi_{\sec(R;x_i=j)}(\boldsymbol{x})$, it implies that $\mathcal{N}' \models \psi_{\sec(R;x_i=j)}(\boldsymbol{n'})$. Hence $\mathcal{N}' \models \phi_0(\boldsymbol{x})$.
- Let us now assume that $\min(\boldsymbol{n}) \geq \tau$, then, $\mathcal{N}' \models \neg \left(\bigvee_{i=0}^{d-1} \bigvee_{j=0}^{l-1} x_i \doteq j \right)$ and by hypothesis about $\phi_1(\boldsymbol{x})$, $\boldsymbol{n} \in R$ is equivalent to $\mathcal{N}' \models \phi_1(\boldsymbol{x})$. Hence $\mathcal{N}' \models \phi_0(\boldsymbol{x})$.

Finally, in both cases, $\mathcal{N}' \models \phi_0(\boldsymbol{x})$.

Let us now assume that $\mathcal{N}' \models \phi_0(\boldsymbol{x})$ and let us prove that $\boldsymbol{n} \in R$. Again, two cases must be considered. Since ϕ_0 is a disjunction, one of the two subformulas holds over $\mathcal{N}'$. Either $\mathcal{N}' \models \bigvee_{i=0}^{d-1} \bigvee_{j=0}^{l-1} x_i \doteq j \wedge \psi_{\sec(R;x_i=j)}(x_0, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{d-2})$ or $\mathcal{N}' \models \neg \left(\bigvee_{i=0}^{d-1} \bigvee_{j=0}^{l-1} x_i \doteq j \right) \wedge \phi_1(\boldsymbol{x})$. In the first case, there exists $i \in [d-1]$ such that $n_i < \tau$ and for $\boldsymbol{n'}$ defined as above, $\boldsymbol{n'}$ belongs to $\sec(R; x_i = n_i)$, hence $\boldsymbol{n} \in R$. Otherwise, by hypothesis, $\boldsymbol{n} \in R$.

The set F_1 Let us now consider the set F_1.

For all $\sigma \in \mathbb{P}_d$, let $F_{2,\sigma} = F_1 \cap \mathbb{N}^d_\sigma$. The sets $F_{2,\sigma}$ are considered because, for each $\boldsymbol{n} \in F_{2,\sigma}$, $n_{\sigma(0)}$ is the smallest of the n_i's. The set $R \cap F_1$ is equal to the union of the sets $R \cap F_{2,\sigma}$ for $\sigma \in \mathbb{P}_d$. Using this idea, the formula ϕ_1 is constructed as the disjunction of the $d!$ formulas which state that $\boldsymbol{x} \in \mathbb{N}^d_\sigma$ and that $\phi_{2,\sigma}(\boldsymbol{x})$ holds (for each $\sigma \in \mathbb{P}_d$).

Let us assume that $\phi_{2,\sigma}(\boldsymbol{x})$ defines the set $R \cap \mathbb{N}^d_\sigma$ in $F_{2,\sigma}$, then let $\phi_1(\boldsymbol{x})$ be:

$$\bigvee_{\sigma \in \mathbb{P}_d} \left(\bigwedge_{i=0}^{d-2} x_{\sigma(i)} \leq x_{\sigma(i+1)} \right) \wedge \phi_{2,\sigma}(\boldsymbol{x}).$$

In the example of Figure 1, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ belongs to $F_{2,(1,0)}$ while $\boldsymbol{n}_0$ belongs to $F_{2,(0,1)}$. The dashed diagonal line represents the boundary between those two sets.

Let us prove that $\boldsymbol{n} \in R$ if and only if $\mathcal{N}' \models \phi_0(\boldsymbol{x})$ holds. Let us assume that $\boldsymbol{n} \in R$ and let us prove that $\mathcal{N}' \models \phi_0(\boldsymbol{x})$. Let σ be a permutation such that $\boldsymbol{n} \in F_{2,\sigma}$, then by hypothesis about $\phi_{2,\sigma}(\boldsymbol{x})$, $\mathcal{N}' \models \phi_{2,\sigma}(\boldsymbol{x})$, and furthermore $\mathcal{N}' \models \bigwedge_{i=0}^{d-2} x_{\sigma(i)} \leq x_{\sigma(i+1)}$, hence $\mathcal{N}' \models \phi_0$.

Let us assume that $\mathcal{N}' \models \phi_0(\boldsymbol{x})$ and let us prove that $\boldsymbol{n} \in F_{2,\sigma}$. Let $\sigma \in \mathbb{P}_d$ be such that $\mathcal{N}' \models (\bigwedge_{i=0}^{d-2} x_{\sigma(i)} \leq x_{\sigma(i+1)} \wedge \phi_{2,\sigma}(\boldsymbol{x}))$. Then $\boldsymbol{n} \in \mathbb{N}^d_\sigma$, and by hypothesis about $\phi_{2,\sigma}(\boldsymbol{x})$, it implies that $\boldsymbol{n} \in R$.

The sets $F_{2,\sigma}$ For the sake of simplicity, let us consider the permutation $\sigma = (0, \dots, d-1)$ and let $F_2 = F_{2,(0,\dots,d-1)}$. For all $\boldsymbol{n} \in F_2$, for all $i \in [d-2]$, $n_i \leq n_{i+1}$.

It follows from hypothesis (2a) that, for all $p \in \mathbb{N}$, the section $\mathrm{sec}(R; x_0 = p)$ is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. By the induction hypothesis there exists $t'_p = \mathrm{thres}_m(\mathrm{sec}(R; x_0 = p)) \in \mathbb{N}$ such that $\mathrm{sec}(R; x_0 = p)$ is m-periodic in $(\mathbb{N}^{\geq t'_p})^{d-1}$. Let $t' = \max\{t'_p \mid p \in [\tau, \tau + m - 1]\}$.

In the example of Figure 1, $\mathrm{thres}_m(\mathrm{sec}(R; x_0 = 2)) = 0$ and $\mathrm{thres}_m(\mathrm{sec}(R; x_0 = 3)) = 8$. Hence $t' = 8$. The line $x_1 = 8$ is represented by the horizontal dashed line. It should be noted that the sections $\mathrm{sec}(R; x_0 = 2)$ and $\mathrm{sec}(R; x_0 = 3)$ are 2-periodic on the right side of this line.

Let $F_3 = \{\boldsymbol{x} \in F_2 \mid x_0 + t' \leq x_1\}$. The set F_3 is considered for two reasons:

- for $p \in [\tau, \tau + m - 1]$, $\mathrm{sec}(R; x_0 = p)$ is m-periodic in $\mathrm{sec}(F_3; x_0 = p)$, and
- it is proven below that, for $p \geq \tau + m$, the study of sections $\mathrm{sec}(R; x_0 = p)$ can be reduced to the study of the m sections $\mathrm{sec}(R; x_0 = p')$ for $p' \in [\tau, \tau + m - 1]$.

As for F_1, the formula $\phi_2(\boldsymbol{x})$ which defines R in F_2 is constructed, using the formula $\phi_3(\boldsymbol{x})$ which defines R in F_3, a formula which defines R in $F_2 \setminus F_3$, and a formula which defines F_3 in F_2. By the induction hypothesis, for each $i \in \mathbb{N}$, there exists a formula $\psi_{\mathrm{diag}(R; x_1 = x_0 + i)}(x_0 \cdots, x_{d-2})$ which defines the diagonal $\mathrm{diag}(R; x_1 = x_0 + i)$ in $\mathbb{N}^d$, hence in F_2. Let us assume that $\phi_3(\boldsymbol{x})$ is a formula which defines R in F_3. Then R is defined in F_2 by the formula $\phi_2(\boldsymbol{x})$:

$$\langle \bigvee_{i=0}^{t'-1} x_0 + i \doteq x_1 \mid \psi_{\mathrm{diag}(R; x_1 = x_0 + i)}(x_0, x_2 \cdots, x_{d-1}) \mid \phi_3(\boldsymbol{x}) \rangle.$$

In the example of Figure 1, $\boldsymbol{n}_1$ belongs to F_3 and $\boldsymbol{n}_2$ does not belong to F_3. The diagonal dashed line represents the boundary of F_3.

The proof that $\boldsymbol{n} \in R$ is equivalent to $\mathcal{N}' \models \phi_3$ is similar to the proof for F_1, apart that diagonals are considered instead of sections.

The set F_3 Let us now consider the set F_3. Note that $n_0 + t' \leq n_1$ for all $\boldsymbol{n} \in F_3$.

For all $\boldsymbol{n} \in F_3$, let n''_0 the be least integer greater than τ equivalent to n_0 modulo m. Equivalently, n''_0 is the only element of $[\tau, \tau + m - 1]$ equivalent to n_0 modulo m.

In the example of Figure 1, the vertical dashed line represents $x_0 = 4 = \tau + m$.

Then, let $\boldsymbol{n''} \in \mathbb{N}^d$ be the d-tuple such that $n''_i = n_i$ for each $i > 0$.

For Figure 1 when $\boldsymbol{n} = \boldsymbol{n_2}$, then $\boldsymbol{n''}$ is equal to $(3, 17)$, which is denoted as $\boldsymbol{n''_2}$ in the figure.

Let us claim that for all $\boldsymbol{n} \in F_3$, $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n''} \in R$. In this case, R is defined in F_3 by the formula $\phi_3(\boldsymbol{x})$:

$$\bigvee_{k=\tau}^{\tau+m-1} x_0 \equiv k \mod m \wedge \psi_{\mathrm{sec}(R; x_0 = k)}(x_1, \cdots, x_{d-1}).$$

It remains to prove that $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n''} \in R$. Let $k = \frac{n_0 - n''_0}{m}$ and let $\boldsymbol{n'} \in \mathbb{N}^d = \boldsymbol{n} - k(\overline{m, \dots, m^d})$. It suffices to prove that $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n'} \in R$ and that $\boldsymbol{n'} \in R$ is equivalent to $\boldsymbol{n''} \in R$. For Figure 1 with $\boldsymbol{n} = \boldsymbol{n}_2$, the value $\boldsymbol{n'}$ is $\boldsymbol{n'_2}$ in the figure.

Let us first prove that $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n'} \in R$. By induction on $i \in \mathbb{N}$, $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n} - i(\overline{m, \dots, m^d}) \in R$ when $n_0 - im \geq \tau$. And since $\boldsymbol{n'} = \boldsymbol{n} - k(\overline{m, \dots, m^d})$ and $n'_0 = n''_0 \geq \tau$ by construction, $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n'} \in R$.

Let us now prove that $\boldsymbol{n'} \in R$ is equivalent to $\boldsymbol{n''} \in R$. Let $\boldsymbol{N'}$ and $\boldsymbol{N''}$ be $\boldsymbol{n'}$ and $\boldsymbol{n''}$ without their first component, respectively. Since $n'_0 = n''_0$, it suffices to prove that $\boldsymbol{N'} \in \sec(R; x_0 = n'_0)$ is equivalent to $\boldsymbol{N''} \in \sec(R; x_0 = n'_0)$.

It should be noted that $n_0 \geq 0$ and $n_1 - n_0 \geq t'$, hence $n_1 \geq t'$. And by definition of t', $\sec(R; x_0 = n'_0)$ is m-periodic in $(\mathbb{N}^{\geq t'})^{d-1}$. Hence, by induction on $i \in \mathbb{N}$, $\boldsymbol{N'} \in \sec(R; x_0 = n'_0)$ is equivalent to $\boldsymbol{N'} + i(\overline{m, \dots, m}^{d-1}) \in \sec(R; x_0 = n'_0)$. In particular, $\boldsymbol{N''} = \boldsymbol{N'} + k(\overline{m, \dots, m}^{d-1})$, hence $\boldsymbol{N'} \in \sec(R; x_0 = n'_0)$ is equivalent to $\boldsymbol{N''} \in \sec(R; x_0 = n'_0)$. □

The above-mentioned Theorem 4.3 of [Muc03] admits the following corollary.

Theorem 4.6 ([Muc03, Theorem 2]). *There exists a* $\mathsf{FO}[+, R]$*-formula* μ_d *such that* $\mathcal{S} \models \mu_d$ *if and only if* $R^{\mathcal{S}}$ *is* $\mathsf{FO}[+]$*-definable.*

Since the characterization given in Theorem 4.4 is $\mathsf{FO}[<]$-expressible, Theorem 4.4 admits the following similar corollary.

Corollary 4.7. *Let* $d, m \in \mathbb{N}^{>0}$*, and* R *be a relation symbol of arity* d*. There exists a* $\Pi_3[<, +, R]$*-formula* $\mu_{d,m}$ *which holds over* $\{R\}$*-structures* $\mathcal{S}$ *such that* $R^{\mathcal{S}}$ *is* $\mathsf{FO}[<, \bmod m]$*-definable.*

Proof. The proof mostly uses the formula $\mathrm{sub}(R, \boldsymbol{t}, \boldsymbol{c})$ introduced in Notation 3.6.

$$\mu_{d,m} = \bigwedge_{\boldsymbol{t} \in (\{\mathrm{var}, \mathrm{const}\} \cup \{\mathrm{add}(i) | i \in [d-1]\})^d} \forall \boldsymbol{c}. \exists N. \forall (n_i > N)_{t_i = \mathrm{var}} \mathrm{sub}(R, \boldsymbol{t}, \boldsymbol{c})(\boldsymbol{n}) \iff \mathrm{sub}(R, \boldsymbol{t}, \boldsymbol{c})(\boldsymbol{n} + (\overline{m, \dots, m}^{\dim(t)}).$$

□

The formula which defines the value of $\mathrm{thres}_m(R)$ is now given.

Lemma 4.8. *Let* $d \in \mathbb{N}^{>0}$*, let* $m \in \mathbb{N}$*. Let* R *be a relation symbol with arity* d*.*
There exists a $\Pi_1[+_m, <, R]$*-formula* $\theta_{d,m}(x)$ *such that if* $R^{\mathcal{S}}$ *is* $\mathsf{FO}[<, \bmod m]$*-definable then:*

- $\theta_{d,m}(x)^{\mathcal{S}} = \{\mathrm{thres}_m(R^{\mathcal{S}})\}$ *and*
- $\theta_{d,m}(x)^{\mathcal{S}_{|n}} = \{\mathrm{thres}_m(R^{\mathcal{S}})\}$ *for all* $n > \mathrm{thres}_m(R^{\mathcal{S}}) + m$*.*

It should be noted that the interpretation of $\theta_{d,m}(x)^{\mathcal{S}_{|n}}$ for $n < \mathrm{thres}_m(R^{\mathcal{S}}) + m$ is undefined.

Proof. The $\Pi_1[<, +_{\mathbb{N}}, R]$-formula $\theta_{d,m}(x)$ is the conjunction of two subformulas. The first one, $\phi_{\geq}(x)$ states that the threshold is at most x. It is:

$$\forall n_0, \cdots, n_{d-1} \left(\bigwedge_{i=0}^{d-1} n_i \geq x \right) \implies [R(n_0, \cdots, n_{d-1}) \iff R(n_0 + m, \cdots, n_{d-1} + m)].$$

The second formula, $\phi_{>-1}(x)$, states that the threshold is strictly greater than $x - 1$. That is, either $x = 0$, or there exists $\boldsymbol{n} \in \mathbb{N}^d$ with $\min(\boldsymbol{n}) = x - 1$ and $\boldsymbol{n} \in R$ is not equivalent to $\boldsymbol{n} + (\overline{m, \dots, m}^d) \in R$. It is:

$$x \doteq 0 \vee \exists \boldsymbol{y}. \left(\bigvee_{i=0}^{d-1} y_i + 1 \doteq x \wedge \bigwedge_{i=0}^{d-1} y_i \geq x - 1 \wedge \left(R(y) \iff \neg R(y + (\overline{m, \dots, m}^d)) \right) \right).$$

Then, let $\theta_{d,m}(x)$ be $\phi_{\geq}(x) \vee \phi_{>-1}(x)$.

For any universe of cardinality greater than $\mathrm{thres}_{d,m}\left(R^{\mathcal{S}}\right) + m$, the integer $\mathrm{thres}_{d,m}\left(R^{\mathcal{S}}\right)$ is the only value which satisfies the formula. Hence $\theta_{d,m}(x)$ converges. □

In the two following sections, proofs consider the evolution of thresholds of sections.

Definition 4.9. Let $d \in \mathbb{N}^{>1}$, and let $R \subseteq \mathbb{N}^d$. Then let $\tau_{R,m}(n)$ be the function which maps $n \in \mathbb{N}$ to the greatest threshold of section $\mathrm{sec}(R; x_0 = i)$ for $i \leq n$. Formally:

$$\tau_{R,m}(n) = \max\left\{\mathrm{thres}_m\left(\mathrm{sec}(R; x_0 = r)\right) \mid r \leq n\right\}.$$

Lemma 4.10. *Let $d \in \mathbb{N}^{>1}$, let R be a relation symbol with arity d and let $\mathcal{S}$ be a $\{<, +_{\mathbb{N}}, R\}$-structure. There exists a $\Pi_2\left[<, +_{\mathbb{N}}, R\right]$-formula $\tau_{d,m}(N; x)$ which defines the function $\tau_{R^{\mathcal{S}},m}$.*

The naive method would be to universally quantify the value $n \leq N$ and use the formula $\mathrm{thres}_m[R](x_1, \ldots, x_{d-1})$ on the sections $\mathrm{sec}(R; x_0 = n)$. But such a formula does not belong to $\Pi_2\left[<, +_{\mathbb{N}}, R\right]$.

Proof. The formula $\tau_{R^{\mathcal{S}},m}$ is now constructed. It is similar to the formula of Lemma 4.8.

The formula $\tau_{d,m}(N; x)$ is the conjunction of two formulas. The first formula $\phi_{\geq}(N, x)$, states that the threshold of every sections $\mathrm{sec}(R; x_0 = n)$ is less than x for all $n \leq N$. Let $\phi_{\geq}(N, x)$ be the $\Pi_2\left[<, +_{\mathbb{N}}, R\right]$-formula:

$$\forall n < N.\, (\forall z_1 \geq x, \cdots, z_{d-1} \geq x.) \implies (R(n, z_1, \cdots, z_{d-1}) \iff R(z, z_1 + m, \cdots, z_{d-1} + m)).$$

The second formula, $\phi_{>-1}(N, x)$, states that there exists some $n \leq N$ such that the threshold of the sections $\mathrm{sec}(R; x_0 = n)$ is strictly greater than $x - 1$. It is possible either if $x = 0$, or if there exists $\boldsymbol{n} \in \mathbb{N}^d$ with $n_0 \leq N$ and $\min\{n_1, \ldots n_{d-1}\} = x - 1$ such that $\boldsymbol{n} \in R$ is not equivalent to $\boldsymbol{n} + (0, \overline{m, \ldots, m}^{d-1}) \in R$. Let $\phi_{>-1}(N, x)$ be the $\Pi_2\left[<, +_{\mathbb{N}}, R\right]$-formula:

$$x \doteq 0 \vee \exists \boldsymbol{n}. \left[n_0 \leq N \wedge \bigvee_{i=1}^{d-1} n_i \doteq x - 1 \wedge \left(R(\boldsymbol{n}) \iff \neg R(\boldsymbol{n} + (\overline{m, \ldots, m}^{d}))\right) \right].$$

Then let $\tau_{d,m}(N; x)$ be $\phi_{\geq}(N, x) \wedge \phi_{>-1}(N, x)$. □

Theorem 4.6 of [Muc03] admits the following easy corollary:

Theorem 4.11 ([Muc03, Theorem 3]). *Let M be a finite automaton whose input is a d-tuple of natural numbers written in positional system (all numbers have the same base and are aligned). One can decide whether the set recognized by M is definable in Presburger Arithmetic.*

Similarly, Corollary 4.7 admits the following corollary:

Corollary 4.12. *Let M be a finite automaton as in Theorem 4.11. It is decidable whether the set recognized by M is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable (respectively, $\mathsf{FO}[<, \mathrm{mod}]$-definable).*

The proof follows closely the proof of [Muc03, Theorem 3].

Proof. The addition and order relation are recognizable by finite automata. The family of recognizable sets is closed under logical operations. Therefore, it is possible to construct an automaton which encodes the formula $\mu_{d,m}$ (respectively, $\exists m.\mu_{d,m}$). It recognizes the empty set if $\mu_{d,m}$ (respectively, $\exists m.\mu_{d,m}$) does not hold, hence if the set recognized by M is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable (respectively, not $\mathsf{FO}[<, \mathrm{mod}]$-definable). □

Corollary 4.7 admits another corollary, similar to Corollary 4.12, considering formulas as inputs instead of automata.

Corollary 4.13. *It is decidable whether a* $\mathsf{FO}[+]$*-formula* $\phi(x)$ *defines a* $\mathsf{FO}[<, \mathrm{mod}]$*-definable set (respectively, a* $\mathsf{FO}[<, \mathrm{mod}\, m]$*-definable set).*

Another algorithm to decide whether $\phi(\boldsymbol{x})$ defines a $\mathsf{FO}[<, \mathrm{mod}]$-definable set is given in [Cho08, Theorem 10].

Proof. Since $\mathsf{FO}[+]$ is decidable, by [Pre27], it suffices to test whether the formula $\exists m.\mu_{d,m}$ (respectively, $\mu_{d,m}$), where $R(\boldsymbol{x})$ is replaced by $\phi(\boldsymbol{x})$, holds. □

4.2. Extracting a unary function from a non $\mathsf{FO}[<, \mathrm{mod}]$-definable set

In this section, a characterization of $\mathsf{FO}[<, \mathrm{mod}]$-definable sets is given. It is similar to the characterization [MV96, Theorem 5.1] of $\mathsf{FO}[+]$. This characterization is first recalled.

Theorem 4.14 ([MV96]). *Let* $d \in \mathbb{N}^{>0}$*, R be a relation symbol with arity d and* $\mathcal{S}$ *be a* $\{+, R\}$*-structure such that* $R^{\mathcal{S}}$ *is not* $\mathsf{FO}[+]$*-definable, then there is a* $\mathsf{FO}[+, R]$*-definable set of integers which is not* $\mathsf{FO}[+]$*-definable.*

Theorem 4.14 does not directly extend to $\mathsf{FO}[<, \mathrm{mod}]$. Indeed, the addition relation is not $\mathsf{FO}[<, \mathrm{mod}]$-definable but Lemma 3.10 states that every $\mathsf{FO}[<, +]$-definable subset of $\mathbb{N}$ is ultimately periodic, hence $\mathsf{FO}[<, \mathrm{mod}]$-definable.

Hence, instead of considering sets of integers, the following theorem considers unary function.

Theorem 4.15. *Let* $d \in \mathbb{N}^{>0}$*, R be a relation symbol with arity d. Let* $\mathcal{S}$ *be a* $\{<, R\}$*-structure with universe* $\mathbb{N}$ *such that* $R^{\mathcal{S}}$ *is* $\mathsf{FO}[+]$*-definable and not* $\mathsf{FO}[<, \mathrm{mod}]$*-definable.*

There exists a converging formula $\nu_R(x; y)$ *in* $\Pi_2[<, +_{\mathbb{N}}, R]$ *such that* $\nu_R(x; y)^{\mathcal{S}}$ *is the graph of an increasing function with slope greater than 1.*

This result can be restated as: the function f is of the form $n :\mapsto rn + g(n)$ with $r > 1$ and $g(n)$ bounded.

By Lemma 3.11, a function is $\mathsf{FO}[<, \mathrm{mod}]$-definable if and only if its slope is 0 or 1. Hence the function of Theorem 4.15 is not $\mathsf{FO}[<, \mathrm{mod}]$-definable. Hence, Theorem 4.15 implies that a $\mathsf{FO}[+]$-definable set R is $\mathsf{FO}[<, \mathrm{mod}]$-definable if and only if every unary $\mathsf{FO}[<, R]$-definable function is $\mathsf{FO}[<, \mathrm{mod}]$-definable.

Two lemmas must first be proved. The following lemma states that, when $\mathsf{FO}[+]$-definable sets are considered, one can consider only one periodicity for all straight subspaces of dimension 1.

Lemma 4.16. *Let* $d \in \mathbb{N}^{>0}$ *and* $R \subseteq \mathbb{N}^d$ *be a* $\mathsf{FO}[+]$*-definable set. Then there exists* $m \in \mathbb{N}^{>0}$ *such that every straight subspace of R of dimension 1 is* $\mathsf{FO}[<, \mathrm{mod}\, m]$*-definable.*

Proof. By Proposition 3.15, it can be assumed that R is defined by a quantifier-free formula $\phi \in \mathsf{FO}[+, <, \mathrm{mod}]$. Without loss of generality, it can be assumed that all modular relations are of the form $x_i \equiv a \mod m$ with only one value m.

Let T be a straight subspace of R of dimension 1. By Lemma 3.5, there exists $\boldsymbol{t}$ such that there is exactly one $i \in [d-1]$ such that $t_i = \mathrm{var}$, and there exists $\boldsymbol{c} \in \mathbb{N}^d$ such that $T = \mathrm{sub}(R; \boldsymbol{t}, \boldsymbol{c})$. Then T is defined by the formula $\mathrm{sub}(R, \boldsymbol{t}, \boldsymbol{c})(x)$ given in Notation 3.6.

Since x is the only variable, each occurrence of x in atomic formulas occurring in ψ is of the form $x \equiv k \mod m$ or $q \times x > p$ or $q \times x \doteq p$ or $q \times x < p$ with $k \in [m-1]$, $q \in \mathbb{N}^{>0}$, $p \in \mathbb{Z}$. Let $\phi'(x)$ be the formula obtained from $\phi(x)$ by replacing the atomic formulas $q \times x > p$ (respectively, $q \times x < p$, $q \times x \doteq p$) are rewritten as $x > \lfloor \frac{p}{q} \rfloor$ (respectively, $x < \lceil \frac{p}{q} \rceil$, $x \doteq \frac{p}{q}$ if q divides p and false otherwise). The formula $\phi'(x)$ is equivalent to $\phi(x)$ and belongs to $\mathsf{FO}[<, \mathrm{mod}\, m]$. □

The following technical lemma states that, given some specific conditions about the evolution of thresholds of sections, it can be claimed that a set is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. This lemma is used both for Theorem 4.15 and for Theorem 4.18. Recall that the notation $\mathbb{N}^d_{0,\cdots,d-1}$, has been introduced in Definition 3.8.

Lemma 4.17. *Let $d \geq 2$, $m \in \mathbb{N}^{>0}$ and $R \subseteq \mathbb{N}^d_{0,\cdots,d-1}$, be such that:*

- *all of its sections and diagonals are $\mathsf{FO}[<, \bmod m]$-definable, and*
- *the function $\tau_{R,m}$ is $\mathsf{FO}[+]$-definable and its slope r is at most 1.*

Then R is $\mathsf{FO}[<, \bmod m]$-definable.

It should be noted that if $\tau_{R,m}$ is $\mathsf{FO}[<, \bmod]$-definable then its slope is 0 or 1. Hence this lemma implies that if $\tau_{R,m}$ is $\mathsf{FO}[<, \bmod]$-definable then R is $\mathsf{FO}[<, \bmod m]$-definable.

Moreover the function $\tau_{R,m}$ is increasing, hence the notion of slope is defined for $\tau_{R,m}$.

Proof. Since $\tau_{R,m}$ is $\mathsf{FO}[<, \bmod]$-definable, it is $\mathsf{FO}[<, \bmod m']$-definable for some $m' \in \mathbb{N}$. The Figure 2 illustrates the proof, with the function $\tau_{R,m}$ represented by unfilled circles.

In the example, $\tau_{R,m}(0) = 0$ and $\tau_{R,m}(n) = 3\lceil \frac{n}{3} \rceil$ for all positive n, thus $\tau_{R,m}$ is $\mathsf{FO}[<, \bmod 3]$ definable.

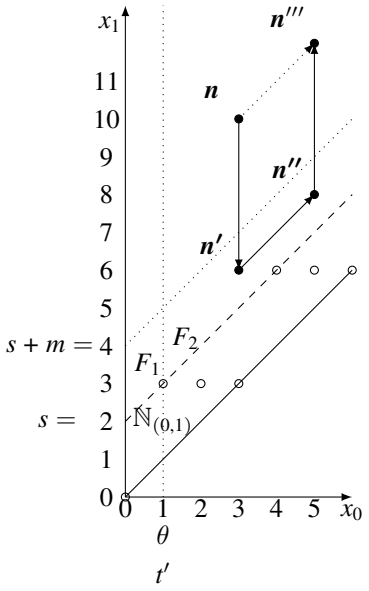

Figure 2: Example for Lemma 4.17

The diagonal line of Figure 2 represents the boundary of $\mathbb{N}^2_{0,1}$. By hypothesis, $R \subseteq \mathbb{N}^2_{0,1}$, hence R belongs to the upper-left half of the figure.

By Lemma 3.11 there exists $\theta \in \mathbb{N}$, a rational $r < 1$ and $s_0, \ldots, s_{m-1} \in \mathbb{Q}$ such that for all $i \geq \theta$, $\tau_{R,m}(I) = ri + s_{(i \bmod m')}$. Let $s = \max\{s_i \mid i \in [m' - 1]\}$.

For the example of Figure 2, $r = 1$, $s_0 = 0$, $s_1 = 2$ and $s_2 = 1$ and then $s = 2$.

The $\mathsf{FO}[<, \bmod m]$-formula ϕ_R which defines R is now constructed. The formula which defines $\tau_{R,m}$ is not used in ϕ_R.

The structure of the proof is similar to the part "(2) $\implies$ (1)" of the proof of Lemma 4.4. A $\subseteq$-decreasing sequence F_0, F_1, F_2 of subsets of $\mathbb{N}^d$, with $F_0 = \mathbb{N}^d$ is introduced. The proof consists in proving that R is definable in F_i by proving that it is definable in F_{i+1} and in $F_i \setminus F_{i+1}$.

The set F_0 Let $F_1 = \{\boldsymbol{x} \in F_0 \mid x_0 + s \leq x_1\}$. The set F_1 is considered, because, by definition of θ, for all $i \geq \theta$ each section $\mathrm{sec}(R; x_0 = i)$ is m-periodic in $\mathrm{sec}(F_1; x_0 = i)$. On Figure 2, the dashed diagonal represents the boundary of F_1.

The set $R \setminus F_1$ is equal to

$$\bigcup_{i=0}^{s-1} \{(x_0, x_0 + i, x_1, \dots, x_{d-2}) \mid (x_0, \dots, x_{d-2}) \in \mathrm{diag}(R; x_1 = x_0 + i)\},$$

and by hypothesis, each set $\mathrm{diag}(R; x_1 = x_0 + i)$ is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable, hence $R \setminus F_1$ is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Since F_1 is also $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable it remains to prove that R is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable in F_1.

The set F_1 Let us now consider F_1. Let t' be the maximal threshold of the diagonals $\mathrm{diag}(R^{\mathcal{S}}; x_1 = x_0 + i)$ for $i \in [s, s + m - 1]$. Let $t'' = \max\{t', \theta\}$. Then let $F_2 = \{\boldsymbol{x} \in F_1 \mid x_0 \geq t''\}$. The set F_2 is considered for two reasons:

- all diagonals $\mathrm{diag}(R; x_1 = x_0 + i)$ are m-periodic for $i \in [s, s + m - 1]$,
- it is proven below that, for $i \geq s$, the study of diagonals $\mathrm{diag}(R; x_0 = x_1 + s)$ can be reduced to the study of the m diagonals $\mathrm{diag}(R; x_0 = x_1 + i')$ for $i' \in [s, s + m - 1]$.

The dotted vertical line of Figure 2 represents the boundary of F_2.

The set R in $F_1 \setminus F_2$ is the union of the sets $\mathrm{sec}(R; x_0 = i)$ for $i \in [t'' - 1]$. By hypothesis, those sections are $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable hence R is also $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable in $F_1 \setminus F_2$. Furthermore F_2 is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable, hence it remains to prove that R is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable in F_2.

The set F_2 Let us now consider the set F_2. Let us prove that R is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable in F_2. By Theorem 4.4 it suffices to prove Property 2 of Theorem 4.4. Property 2b holds by hypothesis. Let us prove Property 2a with the threshold being t''. That is, let $\boldsymbol{n} \in F_2$ and let $\boldsymbol{n'''} = \boldsymbol{n} + (\overline{m, \dots, m}^d)$. It remains to prove that $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n'''} \in R$.

In the example of Figure 2, let $m = 2$, $\boldsymbol{n} = (3, 11)$ and $\boldsymbol{n'''} = (5, 12)$.

The equivalence between $\boldsymbol{n} \in R$ and $\boldsymbol{n'''} \in R$ follows from the following equivalences:

- Let $k = \lfloor \frac{n_1 - (n_0 + s)}{m} \rfloor$ and let $\boldsymbol{n'} = \boldsymbol{n} - k(0, \overline{m, \dots, m}^{d-1})$, let us prove that $\boldsymbol{n} \in R$ is equivalent to $\boldsymbol{n'} \in R$.
  For the example of Figure 2, $k = 2$ and $\boldsymbol{n'} = (5, 8)$.
  Let $\boldsymbol{N}$ and $\boldsymbol{N'}$ be $\boldsymbol{n}$ and $\boldsymbol{n'}$ without their first component. It suffices to prove that $\boldsymbol{N} \in \mathrm{sec}(R; x_0 = n_0)$ is equivalent to $\boldsymbol{N'} \in \mathrm{sec}(R; x_0 = n_0)$. By definition of threshold, it suffices for this to prove that $N'_i \geq \mathrm{thres}_m(\mathrm{sec}(R; x_0 = n_0))$ for all $i \in [d - 2]$. It is a consequence of the following inequalities:
  - $\mathrm{thres}_m(\mathrm{sec}(R; x_0 = n_0)) \leq \tau_{R,m}(n_0)$ by definition of $\tau_{R,m}$,
  - since $\boldsymbol{n} \in F_2$ by definition of F_2 and of t'', $n_0 \geq t'' \geq \theta$. By definition of θ, $\tau_{R,m}(n_0) = rn_0 + s_{(n_0 \mod m)}$,
  - since $r \leq 1$, then $rn_0 + s_{(n_0 \mod m)} = n_0 + s_{(n_0 \mod m)}$,
  - since $s \geq s_k$ for all $k \in [m - 1]$, then $s_{n_0 \mod m} \leq n_0 + s$,
  - by definition of k, $n_0 + s = n_1 - (n_1 - (n_0 + s)) \leq n_1 - \lfloor \frac{n_1 - (n_0 + s)}{m} \rfloor m = n_1 - km = n'_1 = N'_0$ and
  - for all $i \in [d - 3]$, let us prove that $N'_i \leq N'_{i+1}$:
    * $N'_i = n'_{i+1}$ by definition of $\boldsymbol{N'}$,
    * $n'_{i+1} = n_{i+1} - km$ by definition of $\boldsymbol{n'}$
    * $n_{i+1} - km \leq n_{i+2} - km$ since $\boldsymbol{n} \in \mathbb{N}^d_{(0,\dots,d-1)}$
    * $n_{i+2} - km = n'_{i+2}$ by definition of $\boldsymbol{n'}$,
    * $n'_{i+2} = N'_{i+1}$ by definition of N'.

    Thus $N'_i = n'_{i+1} = n_i - km \leq n_{i+2} - km = n'_{i+2} = N'_{i+1}$.

  Thus, for all $i \in [d - 2]$, one has:

$$\mathrm{thres}_m(\mathrm{sec}(R; x_0 = n_0)) \leq \tau_{R,m}(n_0) = rn_0 + s_{n_0 \mod m} = n_0 + s_{(n_0 \mod m)} \leq n_0 + s \leq N'_0 \leq \dots \leq N'_i.$$

- Let $\boldsymbol{n''} = \boldsymbol{n'} + (\overline{m, \dots, m}^d)$ and let $k' = n'_1 - n'_0$.

For the example of Figure 2, $\boldsymbol{n''} = (7, 10)$ and $k' = 3$. Let us prove that $\boldsymbol{n'} \in R$ if and only if $\boldsymbol{n''} \in R$. Let $\boldsymbol{N''}$ be $\boldsymbol{n''}$ without its first component. It suffices to prove that $\boldsymbol{N'} \in \mathrm{diag}(R; x_1 = x_0 + k')$ is equivalent to $\boldsymbol{N''} \in \mathrm{diag}(R; x_1 = x_0 + k')$. By definition of threshold, it suffices to prove that for all $i \in [d-2]$, $N'_i \geq \mathrm{thres}_m(\mathrm{diag}(R; x_1 = x_0 + k'))$. It follows from the following inequalities:
- Let us claim that $k' \in [s, s+m-1]$, then $\mathrm{thres}_m(\mathrm{diag}(R; x_1 = x_0 + k')) \leq \theta$ by definition of t',
- by definition of t'', $t' \leq t''$,
- $t'' \leq n_0$ since $\boldsymbol{n} \in F_2$, by definition of F_2,
- $n_0 = n'_0$ by definition of $\boldsymbol{n'}$,
- $n'_0 < n'_1 = N'_0$ as proven in the previous case,
- for all $i \in [d-2]$, $N'_i \leq N'_{i+1}$ as proven in the previous case.

Thus, for all $i \in [d-2]$, one has:

$$\mathrm{thres}_m(\mathrm{diag}(R; x_1 = x_0 + k')) \leq \theta \leq t'' \leq n_0 = n'_0 < n'_1 = N'_0 \leq \cdots \leq N'_{i+1}.$$

It remains to prove that $k' \in [s, s+m-1]$. It follows from the following statements:
- $\frac{n_1-(n_0+s)}{m} - 1 < \lfloor \frac{n_1-(n_0+s)}{m} \rfloor \leq \frac{n_1-(n_0+s)}{m}$,
- $\frac{n_1-(n_0+s)}{m} - 1 < k \leq \frac{n_1-(n_0+s)}{m}$, by definition of k,
- $n_1 - (n_0 + s) - m < km \leq n_1 - (n_0 + s)$,
- $-n_1 + n_0 + s \leq -km < -n_1 + n_0 + s + m$, by multiplying by -1,
- $n_0 + s \leq n_1 - mk < n_0 + s + m$, by adding n_1,
- $n'_0 + s \leq n'_1 < n'_0 + s + m$, by definition of $\boldsymbol{n'}$,
- $s \leq n'_1 - n'_0 < s + m$.

- Finally, the proof that $\boldsymbol{n''} \in R$ if and only if $\boldsymbol{n'''} \in R$ is the same as the proof of equivalence of $\boldsymbol{n} \in R$ and $\boldsymbol{n'} \in R$.

□

Theorem 4.15 can now be proven.

Proof. The proof goes by induction on d. In the case of $d = 1$, there is no $\mathsf{FO}[+]$-definable set which is not $\mathsf{FO}[<, \mathrm{mod}]$-definable, hence the theorem trivially holds. Let us now assume that $d > 1$ and that it holds for $d-1$.

By Lemma 4.16, there exists $m \in \mathbb{N}^{>0}$ such that every straight subspace of dimension 1 of R is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable.

Two cases must be considered, depending on whether all sections of R are $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable or not. Let us first assume that there is a section $S = \mathrm{sec}(R; x_i = c)$ for $i \in [d-1]$ and $c \in \mathbb{N}$ such that S is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Then by the induction hypothesis, there exists a $\Pi_2[<, +_{\mathbb{N}}, S]$-formula $\nu_S(x; y)$ which defines a non $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable unary function in $\mathcal{N}[S/\mathrm{sec}(R^{\mathcal{S}}; x_i = c)]$. Then let $\nu_{m,R}(x)$ be the $\Pi_2[<, +_{\mathbb{N}}, R]$-formula $\nu_S(x; y)$ where $S(x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{d-1})$ is replaced by $R(x_0, \ldots, x_{i-1}, c, x_{i+1}, \ldots, x_{d-1})$.

Let us now assume that all sections are $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Similarly to the last case, two cases must be considered, depending on whether all diagonals of R are $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable or not. If some diagonal is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable, the proof is similar to the case where a section is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable.

Let us now assume that all sections and diagonals of $R^{\mathcal{S}}$ are $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Then there exists a permutation $\sigma \in \mathbb{P}_d$ such that $R^{\mathcal{S}} \cap \mathbb{N}^d_\sigma$ is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. For the sake of simplicity, let us assume that $\sigma = (0, \ldots, d-1)$ and let $R' = R^{\mathcal{S}} \cap \mathbb{N}^d_\sigma$.

By Lemma 4.8, the function $\tau_{R',m}$ is defined by a $\Pi_2[R', <, +_{\mathbb{N}}]$-formula $\tau_{d,m}(N; x)$ which converges. To define $\nu_R(x; y)$, it suffices to replace in $\tau_{d,m}(N; x)$ the predicate $R'(\boldsymbol{x})$ by $R(\boldsymbol{x}) \wedge \bigwedge_{i=0}^{d-2} x_i < x_{i+1}$.

Since R is $\mathsf{FO}[+]$-definable $\tau_{R',m}$ is $\mathsf{FO}[+]$-definable. Furthermore $\tau_{R',m}$ is increasing. Let r be its slope. The set R' is not $\mathsf{FO}[<, \mathrm{mod}]$-definable, thus $r > 1$ by Lemma 4.17. Hence $\nu_R(x; y)$ satisfies the conditions of the theorem. □

In the above theorem, R must be $\mathsf{FO}[+]$-definable because the proof of the theorem requires that there exists a positive integer m such that all straight subspaces of dimension 1 are $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. To the best of our

knowledge, it is an open question to know whether a function f which is not $\mathsf{FO}[<, \mathrm{mod}]$-definable can be created for any R which is not $\mathsf{FO}[+]$-definable.

4.3. Extracting a unary set from a non $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable set

In this section, a theorem similar to Theorem 4.14 is proved, for the logic $\mathsf{FO}[<, \mathrm{mod}\, m]$.

Theorem 4.18. *Let $d, m \in \mathbb{N}^{>0}$, and R be a relation symbol with arity d. Let $\mathcal{S}$ be a $\{<, R\}$-structure with universe $\mathbb{N}$ such that $R^{\mathcal{S}}$ is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable.*

There exists a $\mathsf{FO}[<, R]$-formula $\nu_{m,R}(x)$ such that $\nu_{m,R}(x)^{\mathcal{S}}$ is not ultimately m-periodic.

In other words, a set R is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable if and only if every unary $\mathsf{FO}[<, R]$-definable set is ultimately m-periodic.

The following proposition, which considers the simpler case of an increasing function, must be proved first.

Proposition 4.19. *Let $m \in \mathbb{N}^{>0}$. Let f be a unary function symbol. Let $\mathcal{S}$ be a $\{<, f\}$-structure with universe $\mathbb{N}$ such that $f^{\mathcal{S}}$ is increasing, unbounded and not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Then there exists a $\mathsf{FO}[<, f]$-formula $\nu_{m,f}(x)$ such that $\nu_{m,f}(x)^{\mathcal{S}}$ is not ultimately m-periodic.*

To prove this proposition, the two following lemmas must first be proved. The first lemma is similar to Proposition 4.19. The only difference is that $\mathrm{mod}\, m$ also belongs to the vocabulary. The second lemma lets us encode the modular predicate using only the increasing function f and the order relation.

Lemma 4.20. *Let $m \in \mathbb{N}^{>0}$. Let f be a unary function symbol. Let $\mathcal{S}$ be a $\{<, f\}$-structure with universe $\mathbb{N}$ such that $f^{\mathcal{S}}$ is increasing, unbounded and not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable. Then there exists a $\mathsf{FO}[<, \mathrm{mod}\, m, f]$-formula $\nu'_{m,f}(x)$ such that $\nu'_{m,f}(x)^{\mathcal{S}}$ is not ultimately m-periodic.*

The proof of Lemma 4.20 consists in a disjunction between three cases. Let us first give an example of three applications of this lemma, considering each of the three cases.

Example 4.2. Let $m = 2$. Some $\{<, f\}$-structures $\mathcal{S}$ are given, such that $f^{\mathcal{S}}$ is not ultimately m-periodic. Then, for each structure, a $\mathsf{FO}[<, f]$-formula ϕ is given such that $\phi^{\mathcal{S}}$ is not ultimately 2-periodic.

- Let $f^{\mathcal{S}}(n) = 4n$. In this case, it suffices to consider the image of f. let $\phi(x)$ be $\exists y. f(y) \doteq x$, then $\phi(x)^{\mathcal{S}} = 4\mathbb{N}$, which is not ultimately 2-periodic.
- Let $f^{\mathcal{S}}(n) = \lfloor \frac{n}{2} \rfloor$. Its image is $\mathbb{N}$ which is ultimately 2-periodic. In this case, it suffices to consider the pre-image of an equivalence class. Let $\phi(x)$ be $f(x) \equiv 0 \mod 2$, then $\phi(x)^{\mathcal{S}} = 4\mathbb{N} \cup (4\mathbb{N} + 1)$, which is not ultimately 2-periodic.
- Let $f^{\mathcal{S}}(n) = 2\lfloor \frac{n}{4} \rfloor$. Its image is $2\mathbb{N}$, the pre-image of $2\mathbb{N}$ and of $2\mathbb{N} + 1$ are $\mathbb{N}$ and $\emptyset$, which are ultimately 2-periodic. In this case, it suffices to consider the part of $\mathbb{N}$ where f does not increase. Let $\phi(x)$ be $f(x) \doteq f(x + 2)$, then $\phi(x)^{\mathcal{S}} = 4\mathbb{N} \cup (4\mathbb{N} + 1)$, which is not ultimately 2-periodic.

Lemma 4.20 is now proved.

Proof of Theorem 4.20. Since $f^{\mathcal{S}}$ is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable, according to Theorem 3.11 there exists $k \in [m - 1]$ such that there are no $N, r \in \mathbb{N}$ such that $f(n) = n + r$ for every $n > N$ and $n \equiv k \mod m$. From now on, let us consider $f_{|m\mathbb{N}+k} : m\mathbb{N} + k \to \mathbb{N}$, the restriction of $f^{\mathcal{S}}$ on the set $m\mathbb{N} + k$, and its image $Im(f_{|m\mathbb{N}+k})$.

Two cases must be considered, depending on whether $Im(f_{|m\mathbb{N}+k})$ is ultimately m-periodic or not. Let us first assume that $Im(f_{|m\mathbb{N}+k})$ is not ultimately m-periodic. Then, let $\nu'_{m,f}(x)$ be

$$\exists y \equiv k \mod m. f(y) \doteq x.$$

Let us now assume that the image $Im(f_{|m\mathbb{N}+k})$ of $f_{|m\mathbb{N}+k}$ is ultimately m-periodic. Since f is unbounded and increasing, then $Im(f_{|m\mathbb{N}+k})$ is infinite. Hence, there exists $\tau = \text{thres}_m\left(Im(f_{|m\mathbb{N}+k})\right) \in \mathbb{N}$ and $P = \text{mod}_m(Im(f_{|m\mathbb{N}+k})) \subseteq [m-1]$ as in Theorem 3.9. That is, for all $n \geq \tau$, $n \in Im(f_{|m\mathbb{N}+k})$ if and only if $n \equiv p \mod m$ for some $p \in \mathbb{N}$. Since $Im(f_{|m\mathbb{N}+k})$ is infinite, then P is non-empty.

Two cases must be considered depending on whether $\#P = 1$ or $\#P \geq 2$. Let us first assume that P contains two distinct elements q and q'. Then the inverse of $m\mathbb{N}+q$ (respectively of $m\mathbb{N}+q'$) by $f_{|m\mathbb{N}+k}$ contains an infinite number of elements. It implies that the inverse of $m\mathbb{N}+q'$ by $f_{|m\mathbb{N}+k}$ is infinite, and its complement in $m\mathbb{N}+k$ is also infinite. Hence the inverse of $m\mathbb{N}+q'$ by $f_{|m\mathbb{N}+k}$ is not ultimately m-periodic. Then, let $\nu'_{m,f}(x)$ be:

$$f(x) \equiv q \mod m.$$

Let us now assume that P is the singleton $\{p\}$. Let us claim that the set

$$E = \left\{n \equiv k \mod m \mid f_{|m\mathbb{N}+k}(n) = f_{|m\mathbb{N}+k}(n+m)\right\}$$

is not ultimately m-periodic. In this case, let $\nu'_{m,f}(x)$ be:

$$f(x) \doteq f(x+m).$$

Let us now prove that E is not ultimately m-periodic. It suffices to prove that E is infinite and $(m\mathbb{N}+k) \setminus E$ is infinite.

Let us first prove that $(m\mathbb{N}+k) \setminus E$ is infinite. Let $N \equiv k \mod m$ and let us prove that there exists $n > N$ such that $n \notin E$, that is, such that $f(n) \neq f(n+m)$. Since $f_{|m\mathbb{N}+k}$ is unbounded, there exists $n' \equiv k \mod m$ such that $f(n') > f(N+m)$, and since f is increasing, $n' > N+m$. Let us assume that n' is chosen minimal. Let us prove that $n = n' - m$ is such that $f(n) \neq f(n+m)$.

By the minimality of $n', f(n'-m) \leq f(N+m)$. Finally, $f(n) = f(n'-m) \leq f(N+m) < f(n') = f(n+m)$.

Let us now prove that E is infinite. Let $N > \max(\tau, 2m)$ be such that $N \equiv k \mod m$ and let us prove that there exists $n \geq N$ belonging to E. Let $r = f_{|m\mathbb{N}+k}(N-m) - (N-m)$. Let n be the minimal integer such that $n \geq N+m, n \equiv k \mod m$ and $f_{|m\mathbb{N}+k}(n) \neq n+r$, by hypothesis about k it exists. Let us prove that $n-m \in E$, note that $n-m \geq N$. By definition of E, it suffices to prove that $f_{|m\mathbb{N}+k}(n) = f_{|m\mathbb{N}+k}(n-m)$. Since n is minimal, either $n = N+m$ or $f_{|m\mathbb{N}+k}(n-m) = n-m+r$. Note that $f(N-m) = N-m+r$ hence $n \neq N-m$. The two preceding statements implies that $f_{|m\mathbb{N}+k}(n-m) = n-m+r$.

For the sake of contradiction, let us now assume that $f_{|m\mathbb{N}+k}(n) \neq f_{|m\mathbb{N}+k}(n-m)$. Let us claim that $f_{|m\mathbb{N}+k}(n) \geq n+r+m$. Since $n+r \equiv p \mod m$ and $n+r \geq \tau$, by definition of τ, there exists $s \in m\mathbb{N}+k$ such that $f_{|m\mathbb{N}+k}(s) = n+r$. Since $n+r-m < n+r < n+r+m \leq f_{|m\mathbb{N}+k}(n), n+r-m = f_{|m\mathbb{N}+k}(n-m), n+r = f_{|m\mathbb{N}+k}(s)$ and $f_{|m\mathbb{N}+k}$ is increasing, $n-m < s < n$. Since $s \equiv p$ and $p \equiv n \mod m$ then $s \equiv n \mod m$. Having both $s \equiv n \mod m$ and $n-m < s < n$ is a contradiction.

Let us now prove that $f_{|m\mathbb{N}+k}(n) \geq n+r+m$. It follows from the following (in)equalities:

- Since $f_{|m\mathbb{N}+k}$ is increasing, $f_{|m\mathbb{N}+k}(n) \geq f_{|m\mathbb{N}+k}(n-m)$.
- Since $f_{|m\mathbb{N}+k}(n) \neq f_{|m\mathbb{N}+k}(n-m), f_{|m\mathbb{N}+k}(n) > f_{|m\mathbb{N}+k}(n-m)$.
- By minimality of $n, f_{|m\mathbb{N}+k}(n-m) = n-m+r$, hence $f_{|m\mathbb{N}+k}(n) > f_{|m\mathbb{N}+k}(n-m) = n-m+r$.
- Since $n \geq \tau, f_{|m\mathbb{N}+k}(n) \equiv p \mod m$. Let us claim that $n+r \equiv p \mod m$ then $f_{|m\mathbb{N}+k}(n) \geq n+r$. Let us now prove that $n+r \equiv p \mod m$. Since $n-m > \tau, n+r \equiv n+r-m = f_{|m\mathbb{N}+k}(n-m) \equiv p \mod m$.
- By $f_{|m\mathbb{N}+k}(n) \geq n+r$ and by definition of $n, f_{|m\mathbb{N}+k}(n) \neq n+r$, hence $f_{|m\mathbb{N}+k}(n) > n+r$.
- Having $f_{|m\mathbb{N}+k}(n) \equiv p \mod m, n+r \equiv p \mod m$ and $f_{|m\mathbb{N}+k}(n) > n+r$ imply $f_{|m\mathbb{N}+k}(n) \geq n+r+m$.

□

The following lemma states how to define a modular relation using an ultimately periodic set of integers.

Lemma 4.21. *Let R be a monadic relation symbol. Let $\mathcal{S}$ be an $\{<, R\}$-structure such that $R^{\mathcal{S}}$ is ultimately periodic. Let $m \in \mathbb{N}^{>0}$ be the least period of $R^{\mathcal{S}}$. Then $m\mathbb{N}$ is $\mathsf{FO}[<, R]$-definable in $\mathcal{S}$.*

Proof. Since $R^{\mathcal{S}}$ is ultimately m-periodic, there exist $\tau \in \mathbb{N}$ and $\mathrm{mod}_m(R^{\mathcal{S}}) \subseteq [m-1]$ as in Definition 3.9. That is, for all $n \geq \tau$, $n \in Im(f_{|m\mathbb{N}+k})$ if and only if $n \equiv p \mod m$ for some $p \in \mathrm{mod}_m(R^{\mathcal{S}})$.

Let us assume that $\phi_{\geq\tau}(x)$ defines $m\mathbb{N}$ in $\mathbb{N}^{\geq\tau}$. Then $m\mathbb{N}$ is defined by

$$\phi(x) = \langle i < \tau \mid \bigvee_{i=0}^{\frac{\tau}{m}} x = im \mid \phi_{\geq\tau}(x) \rangle.$$

Let us now define $\phi_{\geq\tau}(x)$ which defines $m\mathbb{N}$ in $\mathbb{N}^{\geq\tau}$. Let us assume that for all $n \equiv 0 \mod m$ for all $n \geq \tau$ if and only if $n + p \in R^{\mathcal{S}}$ for all $p \in \mathrm{mod}_m(R^{\mathcal{S}})$, and $n + p \notin R^{\mathcal{S}}$ for all $p \in [m-1] \setminus \mathrm{mod}_m(R^{\mathcal{S}})$. Then let $\phi_{\geq\tau}(x)$ be:

$$\bigwedge_{p \in \mathrm{mod}_m(R^{\mathcal{S}})} R(x+p) \wedge \bigwedge_{p \in [m-1] \setminus \mathrm{mod}_m(R^{\mathcal{S}})} \neg R(x+p).$$

Let us prove that for all $n \geq \tau$, the two following statements are equivalent:

1. $n \equiv 0 \mod m$ and
2. for all $p \in \mathrm{mod}_m(R^{\mathcal{S}})$, $n + p \in R^{\mathcal{S}}$ and for all $p \in [m-1] \setminus \mathrm{mod}_m(R^{\mathcal{S}})$, $n + p \notin R^{\mathcal{S}}$.

The fact that (1) implies (2) follows directly from the definition of τ and $\mathrm{mod}_m(R^{\mathcal{S}})$. Let us prove that (2) implies (1).

For the sake of contradiction, let us assume that there exists $k \in [m-1] \setminus \{0\}$ such that $n \equiv k \mod m$. It implies that for all $p \in [m-1]$, $p \in \mathrm{mod}_m(R^{\mathcal{S}})$ is equivalent to $p + k \in \mathrm{mod}_m(R^{\mathcal{S}})$ if $p + k < m$, and to $p + k - m \in \mathrm{mod}_m(R^{\mathcal{S}})$ otherwise.

Let us prove that $R^{\mathcal{S}}$ is ultimately k-periodic, which contradicts the minimal hypothesis about m. It suffices to prove that for all $n \geq \tau$, $n \in R^{\mathcal{S}}$ if and only if $n + k \in R^{\mathcal{S}}$.

Let $n \geq \tau$. The following statements are equivalent:

- $n \in R^{\mathcal{S}}$,
- by definition of $\mathrm{mod}_m(R^{\mathcal{S}})$: $n \mod m \in \mathrm{mod}_m(R^{\mathcal{S}})$,
- by hypothesis about k: $n + k \mod m \in \mathrm{mod}_m(R^{\mathcal{S}})$,
- by definition of $\mathrm{mod}_m(R^{\mathcal{S}})$: $n + k \in R^{\mathcal{S}}$.

□

Proposition 4.19 is now proved.

Proof of Theorem 4.19. The proof consists in exhibiting a $\mathsf{FO}[<,f]$-formula which defines a set which is not ultimately

$$m$$

-periodic.

Let $m_f \in \mathbb{N}^{>0}$ be the greatest integer which divides m and is such that $m_f\mathbb{N}$ is $\mathsf{FO}[<,f]$-definable in $\mathcal{S}$. Note that 1 divides m and $1\mathbb{N}$ is $\mathsf{FO}[<,f]$-definable in $\mathcal{S}$, hence m_f is correctly defined. Let ϕ_{m_f} be the $\mathsf{FO}[<,f]$-formula which defines $m_f\mathbb{N}$ in $\mathcal{S}$.

Since f is not $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable and m_f divides m, f is not $\mathsf{FO}[<, \mathrm{mod}\, m_f]$-definable. By Theorem 4.20 there exists a $\mathsf{FO}[<, f, \mathrm{mod}\, m_f]$-formula ξ'_E which defines a set $E \subseteq \mathbb{N}$ which is not ultimately

$$m_f$$

-periodic. Let ξ_E be the $\mathsf{FO}[<,f]$-formula ξ'_E where each occurrence of $x \equiv a \bmod m_f$ is replaced by $\phi_{m_f}(x - a + m_f)$, it also defines E.

Let m_E be the least integer such that E is ultimately

$$m_E$$

-periodic. By Theorem 4.21, there exists a $\mathsf{FO}[<,E]$-formula ψ_{m_E} which defines $m_E\mathbb{N}$ in $\mathcal{S}$. Let ϕ_{m_E} be the $\mathsf{FO}[<,f]$-formula ψ_{m_E} where $E(x)$ is replaced by ξ_E; it defines $m_E\mathbb{N}$ in $\mathcal{S}$. Note that E is ultimately

$$n$$

-periodic if and only if m_E divides n, for all $n \in \mathbb{N}^{>0}$. Two cases must be considered, depending on whether m_E divides m or not. If m_E does not divide m, then E is not ultimately

$$m$$

-periodic, hence the result is proved. Let us now assume that m_E divides m. Since E is not ultimately

$$m_f$$

-periodic, m_E does not divide m_f.

Let $m_{\mathrm{lcm}} = \mathrm{lcm}(m_f, m_E)$. Since m_E does not divide m_f, $m_{\mathrm{lcm}} > m_f$. Since m_f and m_E divides m, m_{lcm} divides m. The set $m_{\mathrm{lcm}}\mathbb{N}$ is $\mathsf{FO}[<,f]$-defined by $\phi_{m_f} \wedge \phi_{m_E}$. Since $m_{\mathrm{lcm}}\mathbb{N}$ is $\mathsf{FO}[<,f]$-definable and m_{lcm} divides m, by maximality hypothesis about m_f, $m_{\mathrm{lcm}} \leq m_f$. Having $m_{\mathrm{lcm}} > m_f \geq m_{\mathrm{lcm}}$ is a contradiction. □

Let us now prove Theorem 4.18.

Proof. The proof is by induction on d. If $d = 1$ then let $\nu_{m,R^{\mathcal{S}}}(x)$ be the formula $R(x)$. It is now assumed that $d > 1$ and that the theorem holds for $d - 1$.

Two cases must be considered, depending on whether there exists a section E of $R^{\mathcal{S}}$ which is $\mathsf{FO}[<, \bmod m]$-definable or not. Let us first assume that there is a section $E = \mathrm{sec}(R^{\mathcal{S}}; x_i = c)$ for $i \in [d-1]$ and $c \in \mathbb{N}$ such that E is not $\mathsf{FO}[<, \bmod m]$-definable. Then by the induction hypothesis, there exists a formula $\nu_{m,S}(x)$ such that $\nu_{m,S}(x)$ defines E in $\mathcal{S}$. Then let $\nu_{m,R^{\mathcal{S}}}(x)$ be the formula $\nu_{m,S}(x)$ where $E(x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{d-1})$ is replaced by $R(x_0, \ldots x_{i-1}, c, x_{i+1}, \ldots, x_{d-1})$.

Let us now assume that all sections of $R^{\mathcal{S}}$ are $\mathsf{FO}[<, \bmod m]$-definable. Similarly, two cases must be considered, depending on whether there exists a diagonal E of $R^{\mathcal{S}}$ which is not $\mathsf{FO}[<, \bmod m]$-definable. If there exists such a diagonal, the proof is similar to the preceding case. Let us now assume that all diagonals of $R^{\mathcal{S}}$ are $\mathsf{FO}[<, \bmod m]$-definable.

The problem is reduced to the study of a set included in $\mathbb{N}^d_\sigma$ which is not $\mathsf{FO}[<, \bmod m]$-definable for some permutation σ.

The set $R^{\mathcal{S}}$ is the union of the $d!$ sets $R^{\mathcal{S}} \cap \mathbb{N}^d_\sigma$ for $\sigma \in \mathbb{P}_{d-1}$. Since $R^{\mathcal{S}}$ is not $\mathsf{FO}[<, \bmod m]$-definable, there exists $\sigma \in \mathbb{P}_{d-1}$ such that $R^{\mathcal{S}} \cap \mathbb{N}^d_\sigma$ is not $\mathsf{FO}[<, \bmod m]$-definable. For the sake of simplicity, let us assume that σ is the identity permutation, that it $(0, \ldots, d-1)$.

Let $R' = R^{\mathcal{S}} \cap \mathbb{N}^d_{(0,\ldots,d-1)}$ and let $\mathcal{S}'$ be $\mathcal{S}[R/R']$. Let us assume that there exists a $\mathsf{FO}[<, R']$-formula $\phi_{R'}$ which defines in $\mathcal{S}'$ a set $E \subseteq \mathbb{N}$ which is not $\mathsf{FO}[<, \bmod m]$-definable. Then it suffices to set $\nu_{m,R}(x)$ as the formula obtained from $\phi_{R'}$ by replacing every atomic formula $R'(\boldsymbol{x})$ by $R(\boldsymbol{x}) \wedge \bigwedge_{i=0}^{d-2} x_i \leq x_{i+1}$.

By Lemma 4.17, the function $\tau_{m,R'}$ is not $\mathsf{FO}[<, \mathrm{mod}]$-definable. Then, by Proposition 4.19, there exists a set $E \subseteq \mathbb{N}$ which is $\mathsf{FO}[<, R]$-definable and which is not ultimately m-periodic. Since $\tau_{R^{\mathcal{S}},m}$ is $\mathsf{FO}[<, R]$-definable in $\mathcal{S}$, E is $\mathsf{FO}[<, R]$-definable. □

Let us give an example of application of Theorem 4.18. Some $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable sets R are considered. From them, some subsets of $\mathbb{N}$ which can be defined from R are considered.

Example 4.3. Let $m = 1$ and $R = \{\boldsymbol{x} \mid x_1 + x_2 = x_3 x_0\}$. Fixing x_0 to 1 gives the addition relation, which is not $\mathsf{FO}[<]$-definable. So a $\mathsf{FO}[R, <]$-definable set which is not $\mathsf{FO}[<]$-definable can be obtained by exhibiting a $\mathsf{FO}[+, <]$-definable set which is not $\mathsf{FO}[<]$-definable.

Every section of + is $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable, so induction can not be used anymore. Let $\sigma = (0, 1, 2)$, then $S = + \cap \mathbb{N}^3_{(0,1,2)} = \{\boldsymbol{x} \mid x_0 \leq x_1, x_1 + x_0 = x_2\}$. Let $T_n = \mathrm{sec}(S; x_0 = n) = \{(x, x + n) \mid x \leq n\}$. Then $(n, 2n) \in T_n$ and $(n + 1, 2n + 1) \notin T_n$, so $\mathrm{thres}_1(n) > 2n$. Having $y, z > n$, $(y, z) \notin T_n$ and $(y + 1, z + 1) \notin T_n$ imply $\mathrm{thres}_1(n) = 2n + 1$, and so $\tau_{R,m}(n) = 2n + 1$.

The image of $\tau_{R^S,m}$ is the set of odd integers, which is not $\mathsf{FO}[<]$-definable. Note that this set is $\mathsf{FO}[<, \mathrm{mod}\, 2]$-definable.

5. About (finite) satisfiability of some class of existential-monadic formulas

This section considers the (finite) satisfiability problem for a logic with a $\mathsf{FO}[+]$-definable predicate which is not $\mathsf{FO}[<, \mathrm{mod}]$-definable.

Let α be an alphabet and ϕ be a $\mathsf{FO}\left[\mathcal{V}, (P_a)_{a\in\alpha}\right]$-formula without free variable. The formula Φ is said to be *(finitely) satisfiable* over a $\mathcal{V}$-structure $\mathcal{S}$ if there exists a (finite) word w such that $\mathcal{S}_w \models \phi$.

The (finite) satisfiability problem of $\mathsf{FO}\left[\mathcal{V}, (P_a)_{a\in\alpha}\right]$ is said to be decidable if there exists an algorithm which takes as input a formula, and accepts if and only if ϕ is (finitely) satisfiable.

The main theorem of this section is now stated.

Theorem 5.1. *Let $d \in \mathbb{N}^{>0}$ and let R be a relation symbol with arity d. Let $\mathcal{S}$ be a $\{\mathbb{N}, <, +_{\mathbb{N}}, R\}$-structure such that $R^{\mathcal{S}}$ is $\mathsf{FO}[+]$-definable and not $\mathsf{FO}[<, \mathrm{mod}]$-definable. The finite satisfiability of $\Pi_3[\mathbb{N}, <, +_{\mathbb{N}}, R, P_a, P_b]$, as well as the satisfiability of $\Pi_3[\mathbb{N}, <, +_{\mathbb{N}}, R, P_a, P_b]$ over $\mathbb{N}$, are undecidable in $\mathcal{S}$.*

Note that our theorem considers a binary alphabet. It easily extends to any alphabet of cardinality at least 2.

5.1. 2-counter automata

Our undecidability result is obtained by a reduction from the halting problem for 2-counter automata which is undecidable [Min60]. Let us briefly recall the definition of a 2-*counter automaton*.

Definition 5.2 (2-counter automaton). A 2-counter automaton A consists of a list of instructions. Let $\#A$ denote the number of instruction of A. The instructions are "incr(h)", "decr(h)", "jmp(j)", "jz(h,j)" with $h \in \{a, b\}$, $j \in [\#A - 1]$ and "Halt". The j-th instruction is written A_j. Without loss of generality, it is assumed that only one Halt instruction appears in the list and that it appears as the last instruction.

Then it is explained how those automata compute.

Definition 5.3 (Configuration and Simulation). Let A be a 2-counter automaton. A *configuration* of A is a 3-tuple of natural numbers (q, n_0, n_1) where q is the next instruction of the automaton and n_j is the value of the j-th counter.

The *simulation* of A is a 3-tuple of sequences $(\kappa[i], c_0[i], c_1[i])_{i\in I}$ (where I is an initial segment of $\mathbb{N}$) that satisfies the following properties:

- The first configuration, which is denoted by $(\kappa[1], c_0[1], c_1[1])$, equals $(0, 0, 0)$.
- For every $l \in \mathbb{N}$ such that the l-th configuration is $(\kappa[l], c_0[l], c_1[l])$ and $A_{\kappa[l-1]} \neq$ Halt (that is $\kappa[l - 1] < \#A - 1$), then the l-th configuration, denoted by $(\kappa[l + 1], c_0[l + 1], c_1[l + 1])$ is defined as follows:

$$
\begin{array}{llll}
\text{if } A_{\kappa[l]} = incr(i) & \text{then} & c_i[l+1] = c_i[l] + 1, c_{1-i}[l+1] = c_{1-i}[l], \text{ and} & \kappa[l+1] = \kappa[l] + 1, \\
\text{if } A_{\kappa[l]} = decr(i) & \text{then} & c_i[l+1] = c_i[l] - 1, c_{1-i}[l+1] = c_{1-i}[l], \text{ and} & \kappa[l+1] = \kappa[l] + 1, \\
\text{if } A_{\kappa[l]} = jmp(m) & \text{then} & \forall i. c_i[l+1] = c_i[l], \text{ and} & \kappa[l+1] = m, \\
\text{if } A_{\kappa[l]} = jz(i,m) & \text{then} & \forall i. c_i[l+1] = c_i[l], \text{ if } c_i[l] = 0 \text{ then } \kappa[l+1] = m \text{ and otherwise} & \kappa[l+1] = \kappa[l] + 1.
\end{array}
$$

- The last configuration, if it exists, satisfies $\kappa[\text{last}] = \#A - 1$.

It should be noted that those automata do not have any input. Hence an automaton admits at most one computation.

It is now explained how to transform a sequence of configuration $(\kappa[i], c_0[i], c_1[i])_{i\in I}$ into a language $L(\kappa, c_0, c_1)$. This transformation is such that two distinct sequences correspond to disjoint languages.

A language $L(q, n_0, n_1)$ is first associated to each configuration (q, n_0, n_1).

Definition 5.4. The definition of $L(q, n_0, n_1)$ uses another language $L'(n_0, n_1)$, which is defined by induction on n_0 and n_1 as follows:

$$L'(n_0, n_1) = \begin{cases} \epsilon, & \text{if } n_0 = n_1 = 0, \\ bab^*L'(n_0 - 1, n_1) & \text{if } n_0 > 0 \\ baab^*L'(n_0, n_1 - 1) & \text{if } n_1 > 0. \end{cases}$$

Let $q \in \mathbb{N}$. The language $L(q, n_0, n_1)$ is defined by the rational expression:

$$ba^{q+3}b^*L'(n_0, n_1)$$

Finally, for a finite or infinite sequence $(\kappa[i], c_0[i], c_1[i])_{i\in I}$, let $L(\kappa, c_0, c_1)$ be the concatenation of the languages $L(\kappa[i], c_0[i], c_1[i])$ for $i \in [n - 1]$.

Intuitively, in $L(q, n_0, n_1)$, the letter b serves as separator and the number of successive a contains all the information.

Let $(\kappa[i], c_0[i], c_1[i])_{i\in I}$ be a sequence of configurations. Let w be a word of $L(\kappa, c_0, c_1)$. Let $(s_i)_{i\in[n-1]}$ be the successive positions in w such that $w[s_i]w[s_i + 1]w[s_i + 2]w[s_i3] = baaa$. Then let $[s_i, s_{i+1} - 1]$ be called the i-th segment, denoted by S_i.

The values of $\kappa[i]$, $c_0[i]$ and of $c_1[i]$ can be read in S_i since:

- $\kappa[i]$ is the length of the first sequence of a, minus 2,
- $c_0[i]$ is the number of factors bab in S_i, and
- $c_1[i]$ is the number of factors $baab$ in S_i.

An example of word encoding the two first configurations of some automaton is now given.

Example 5.1. Let A be a 2-counter automaton with 2 instructions: $incr(0)$ and $jmp(0)$. Let (κ, c_0, c_1) be a simulation of this automaton. Equation (2) represents a word of $L(\kappa[0], c_0[0], c_1[0])L(\kappa[1], c_0[1], c_1[1])$.

The $||$'s represent the beginning of each sequence, and the $|$'s represent the beginning of each part which corresponds to a counter. The $\ldots$'s corresponds to sequences of b's.

$$\left\|\begin{matrix}0&1&2&3&4&5&6\\ b&a&a&a&b&b&b\end{matrix}\right\|\left|\begin{matrix}7&8&9&10&11\\ b&a&a&a&a\end{matrix}\right|\left|\begin{matrix}12&13&14&15&\ldots\\ b&a&b&b&\ldots\end{matrix}\right\|\left|\begin{matrix}21&22&23&24&25&\ldots\\ b&a&a&a&b&\ldots\end{matrix}\right|\left|\begin{matrix}31&32\\ b&a\end{matrix}\right. \tag{2}$$

5.2. Undecidability

Theorem 5.1 is now proved.

Proof. The proof goes by reduction from the halting problem for 2 counter automaton. Let A be a 2 counter automaton with $\#A$ states.

The first part of the proof consists in defining in $\Pi_2\,[\mathbb{N}, <, +_{\mathbb{N}}, R, P_a, P_b]$ an increasing function g in $\mathcal{S}$ such that, for all $n \in \mathbb{N}$, $g(n) \geq 2n + \#A + 5$. The second part consists in creating an $\Pi_3\,[\mathbb{N}, <, +_{\mathbb{N}}, R, P_a, P_b]$-formula which holds in $\mathcal{S}$ if and only if A halts. The formula of the second part uses the function g.

Defining g By Theorem 4.15, there exists a converging formula $\nu_R(x;y)$ in $\Pi_2[\mathbb{N},<,+_\mathbb{N},R]$ such that $\nu_R(x;y)^\mathcal{S}$ is the graph of an increasing function f with slope l greater than 1.

By Lemma 3.13, there exists a $\Sigma_1[\mathbb{N},<,+_\mathbb{N},f]$-definable function $g:\mathbb{N}\to\mathbb{N}$ such that for all n, $g(n)\geq 4n+\#A+1$. Hence g is $\Pi_3[\mathbb{N},<,+_\mathbb{N},R]$-definable in $\mathcal{S}$.

Reduction from a 2-counter automata The simulation of a 2-counter automaton A is encoded by a formula ϕ_A in $\Pi_3[\mathbb{N},<,+_\mathbb{N},R,P_a,P_b]$ such that there exists a word $w\in\{a,b\}^+$ such that $\mathcal{S}(w)\models\phi_A$ if and only if A halts.

More precisely, the formula ϕ_A states that the word w belongs to $L(\kappa,c_0,c_1)$ where (κ,c_0,c_1) is a simulation of A.

Some notations are now introduced:

- for $q\in[\#A-1]$, let Q_q be the set of positions p such that the factor of length $q+5$ of w beginning at position p is $ba^{q+3}b$,
- for $i\in\{0,1\}$, let C_i be the set of positions p such that the factor of length $i+3$ of w beginning at position p, is $ba^{i+1}b$.
- let S be the union of the sets Q_q for $q\in[\#A-1]$,

The set Q_q (respectively, C_i) is defined by the $\Sigma_1[+_\mathbb{N},P_a,P_b]$-formula $\varsigma_{q+3}(x)$ (respectively, $\varsigma_{i+1}(x)$), where $\varsigma_n(x)$ states that the factor of w of length $n+2$ beginning at position x is ba^nb. The formula $\varsigma_n(x)$ is

$$P_b(x)\wedge\bigwedge_{j=1}^{n}P_a(x+j)\wedge P_b(x+n+1).$$

The set S is defined by $\bigvee_{q=0}^{\#A-1}Q_q(x)$.

Defining S Let us assume that the simulation halts after n steps. Then the finite structure is partitioned into n consecutive segments of the form $S_i=[s_i,s_{i+1}-1]$ for $i\in[n]$, with $s_0=0$ and $s_{i+1}=g^\mathcal{S}(s_i)$ for all $i\in[n]$. The i-th segment, denoted by S_i, encodes the i-th step of the simulation.

In Example 5.1 corresponds to the case $g(n)=2n+7$ (with $\#A=2$).

By an easy induction on i, S_i has length at least $4^i(\#A+5)$.

For the sake of readability, the notation $g^{-1}(n)$ is used for any n belonging to the image of g. For any formula $\psi(x)$, let $\psi(g^{-1}(n))$ be an abbreviation for the $\Sigma_3[n,S,g]$ formula:

$$\exists n'.g(n')\doteq n\wedge S(n')\wedge\psi(n').$$

A $\Pi_4[\mathbb{N},<,+_\mathbb{N},R,S]$-formula ϕ_S which asserts that S is the image of 0 by a sequence of iterations of g is now given. An integer n belongs to S if and only if $n=s_0=0$, or there exists $n'<n$ such that $g^\mathcal{S}(n')=n$ and $S(n')$ holds. Let ϕ_S be:

$$\forall n.\left\{S(n)\iff\left[n\doteq 0\vee S(g^{-1}(n))\right]\right\}.$$

Simulation Let us now consider a run of A. Let us first consider the states. For $i\in[n]$, for $q\in[\#A-1]$, the formula ϕ_q requires that $Q_q(s_i)$ holds if and only if $\kappa(i)=q$.

Let us now consider the counters. For $i\in[n]$, the formula ϕ_q requires that $\#(C_j\cap S_i)=c_j[i]$.

The main issue with this encoding is that in order to ensure that two successive segments encode two successive configurations, it is needed to compare the cardinality of two sets, which does not seem possible in our logic. To overcome this, the properties of the function $g^\mathcal{S}$ are used. For example if the i-th step is a *jump*, then $c_j[i]=c_j[i+1]$, and in this case C_j can be chosen in S_{i+1} to be the image by $g^\mathcal{S}$ of C_j in S_i. If the i-th step is $incr(j)$ then it is enough to add a single position to C_j in S_{i+1}. Note that the properties of the function $g^\mathcal{S}$ ensure that such a position exists.

Formally, the formula ϕ_q states the following requirements:

- The initial state is 0, which is defined by:

$$Q_0(0),$$

- The last state is $\#A$, which is defined by:

$$\exists x.\neg\exists y.g(x) \doteq y \wedge Q_{\#A-1}(x),$$

In this formula, x represents the beginning of the last segment of the universe.
Let s be an integer of the form s_n, then let $s' = s_{n+1}$ and $s'' = s_{n+2}$. Let q be the i-th step of the computation. Let us now give formulas ϕ_q which assert that two successive segments $[s, s'-1]$ and $[s', s''-1]$ encode a step of the computation

- If the i-th step is $jmp(q')$, that is $Q_q(s_i)$ with $A_{\kappa(q)} = jmp(q')$, then:
  - $Q_{q'}(s_{i+1})$,
  - for $p \in [s_{i+1}, s_{i+2}-1]$, for $j \in \{0,1\}$, $C_j(p)$ holds if and only if there exists p' with $g^{\mathcal{S}}(p') = p$ and $C_j(p')$ holds. This can be expressed by the $\Pi_4\ [\mathbb{N}, <, +_{\mathbb{N}}, R, P_a, P_b]$-formula $\phi_{\text{copy } j}$:

$$\forall p \in [s', s''-1].(C_j(p) \iff C_j(g^{-1}(p'))).$$

This can be expressed by the $\Pi_4\ [\mathbb{N}, <, +_{\mathbb{N}}, R, P_a, P_b]$-formula ϕ_q equal to

$$Q_{q'}(s_{i+1}) \wedge \phi_{\text{copy } 0} \wedge \phi_{\text{copy } 1}$$

- If the i-th step is $jz(z, q')$, that is $Q_q(s_i)$ with $A_{\kappa(q)} = jz(z, q')$, then:
  - For $p \in [s_{i+1}, s_{i+2}-1]$, for $j' \in \{0,1\}$, $C_{j'}(p)$ holds if and only if there exists p' with $g^{\mathcal{S}}(p') = p$ and $C_{j'}(p')$ holds.
  - If there exists a position $p \in [s_i, s_{i+1}-1]$ such that $C_z(p)$ then $Q_{q+1}(s_{i+1})$, and $Q_{q'}(s_{i+1})$ otherwise. This can be expressed by the $\Sigma_1\ [\mathbb{N}, <, +_{\mathbb{N}}, R, s, s', s'', P_a, P_b]$-formula ϕ_{jz}:

$$\langle \exists p \in [s, s'-1].C_j(p) \mid Q_{q'}(s') \mid Q_{q+1}(s') \rangle$$

Then ϕ_q is the $D_1\ [\mathbb{N}, <, +_{\mathbb{N}}, R, s, s', s'', P_a, P_b]$-formula :

$$\phi_{jz} \wedge \phi_{\text{copy } 0} \wedge \phi_{\text{copy } 1}$$

- If the i-th step is $decr(j)$, that is $Q_q(s_i)$ with $\kappa(q) = decr(j)$, then
  - $Q_{q+1}(s_{i+1})$,
  - there exists exactly one integer $p \in [s_i, s_{i+1}-1]$ such that $C_j(p)$ holds and $C_j(g(p))$ does not hold .
    Formally, it is expressed as the conjunction of two formulas, a $\Sigma_3\ [<, +_1, s, s', C_j, Q_{q+1}]$-formula $\phi_{q,\exists}$ which states the existence of such a p:

$$\exists p \in [s, s'-1].C_j(p) \wedge \neg C_j(g(p)),$$

and a $\Pi_4\ [<, +_1, s', s'', C_j, g]$-formula $\phi_{q,!}$ which states that there are no two such p:

$$(\forall p_0, p_1 \in [s, s'-1].C_j(p_0) \wedge \neg C_j(g(p_0)) \wedge C_j(p_1) \wedge \neg C_j(g(p_1)))\, p_0 \doteq p_1.$$

  - for all $q \in [s_{i+1}, s_{i+2}-1]$, if $(g^{\mathcal{S}})^{-1}(q)$ is not defined or $C_j((g^{\mathcal{S}})^{-1}(q))$ does not hold, then $C_j(q)$ does not hold. This can be expressed by the $\Pi_4\ [\mathbb{N}, <, +_{\mathbb{N}}, R, s', s'', P_a, P_b]$-formula ϕ_j:

$$\forall p \in [s', s''-1]\left(\neg C_j(g^{-1}(p)) \implies \neg C_j(p)\right).$$

  - for all $q \in [s_{i+1}, s_{i+2}-1]$, $C_{1-j}(q)$ holds if and only if there exists q' with $g^{\mathcal{S}}(q') = q$ and $C_{1-j}(q')$ holds.

This can be expressed by the $\Pi_4\ [<, +_1, s', s'', C_0, C_2, g]$-formula ϕ_q:

$$Q_{q+1}(s') \wedge \phi_{q,\exists} \wedge \phi_{q,!} \wedge \phi_j \wedge \phi_{\text{copy } 1-j}$$

- If the i-th step is $incr(j)$, that is $Q_q(s_i)$ with $\kappa(q) = incr(j)$, then
  - $Q_{q+1}(s_{i+1})$,
  - there exists exactly one $p \in [s_{i+1}, s_{i+2} - 1]$ such that $C_j(p)$ holds and there is no p' such that $g(p') = p$ and $C_j(p')$ hold.
    Formally, it is expressed by two formulas, a $\Sigma_3\ [<, +_1, s, s', C_j, Q_{q+1}]$-formula which states the existence of such a p, $\phi_{q,\exists}$:

$$\exists p \in [s', s'' - 1].C_j(p) \wedge \neg C_j(g^{-1}(p)),$$

    and a $\Pi_4\ [<, +_1, s', s'', C_j, g]$-formula $\phi_{q,!}$ which states that there are no two such p,

$$\left(\forall p_0, p_1 \in [s', s'' - 1].C_j(p_0) \wedge \neg C_j(g^{-1}(p_0)) \wedge C_j(p_1) \wedge \neg C_j(g^{-1}(p_1))\right) p_0 \doteq p_1.$$

  - for all $q \in [s_{i+1}, s_{i+2} - 1]$, if $C_j((g^{\mathcal{S}})^{-1}(q))$ holds, then $C_j(q)$ holds. This can be expressed by the $\Pi_4\ [\mathbb{N}, <, +_{\mathbb{N}}, R, s', s'', P_a, P_b]$-formula ϕ_j:

$$\forall p \in [s', s'' - 1]\ \left(C_j(g^{-1}(p)) \implies C_j(p)\right).$$

  - for all $q \in [s_{i+1}, s_{i+2} - 1]$, $C_{1-j}(q)$ holds if and only if there exists q' with $g^{\mathcal{S}}(q') = q$ and $C_{1-j}(q')$ holds.
    This can be expressed by the $\Pi_4\ [<, +_1, s', s'', C_0, C_2, g]$-formula ϕ_q:

$$Q_{q+1}(s') \wedge \phi_{q,\exists} \wedge \phi_{q,!} \wedge \phi_j \wedge \phi_{\text{copy } 1-j}$$

- finally if the i-th step is *Halt*, that is $Q_q(s_i)$ with $A_{\kappa(q)} = Halt$, then ϕ_q is *true*.

The above-mentioned formula ϕ_A which accepts encoding of simulations of A is now given. Let ϕ_A be the $\Pi_4\ [\mathbb{N}, <, +_{\mathbb{N}}, R, P_a, P_b]$-formula:

$$\begin{aligned}&\phi_S \wedge Q_0(0) \wedge \exists x.Q_{\#A-1}(x) \wedge \\ (\forall s, s', s''.S(s) \wedge g(s) = s' \wedge g(s') = s'') \implies &\bigwedge_{q=0}^{\#A-1} (Q_q(s) \wedge \phi_q).\end{aligned}$$

Universe $\mathbb{N}$ Let us now consider the satisfiability problem over $\mathbb{N}$. The proof is the same as finite case excepts that it must be asserted that the word ends by an infinite sequence of b's and only the letters up to the last a are considered.

Formally, one quantifies existentially over a variable m which represents the last occurrence of a. Then all quantifications of ϕ_A are restricted to integers less than or equal to m.

□

6. Conclusion

It is proven in this paper that a set R is $\mathsf{FO}[<, \mathrm{mod}]$-definable if and only if every unary $\mathsf{FO}\ [<, R]$-definable set of integers is ultimately periodic. It is also proven that the class of $\mathsf{FO}[<, \mathrm{mod}\, m]$-definable sets also admits a Muchnik-like property in terms of local and recursive properties.

It has also been proved that finite satisfiability of $\mathsf{FO}\left[<, R, (P_a)_{a\in\alpha}\right]$ is undecidable for every set R which is $\mathsf{FO}\ [+]$-definable and not $\mathsf{FO}[<, \mathrm{mod}]$-definable.

We see many directions for further research. We may want to minimize the number of alternations of quantifiers needed to get undecidability, and conversely to find an algorithm to decide those logics with one or two alternations.

Let us note that a construction similar to the proof of Theorem 5.1 can be used to prove undecidability of some other related logics:

- $\mathsf{FO}\left[R, +1, (P_a)_{a\in\alpha}\right]$ when R is $\mathsf{FO}\left[+\right]$-definable and not $\mathsf{FO}[<, \mathrm{mod}]$-definable,
- $\mathsf{FO}\left[+1, g\right]$ as soon as g is not interpreted, and
- $\mathsf{FO}\left[<, g, (P_a)_{a\in\alpha}\right]$ when g is ultimately greater than any function $+c$ for $c \in \mathbb{N}$.

Another direction for further research is to extend Corollary 4.12 in order to obtain a polynomial-time algorithm.

Acknowledgments We thank the anonymous referees of the first version of this paper for their remarks and suggestion to improve the paper. We thank Alexis Bès for his help during the preparation of this paper. We thank the organizers of Computatibility in Europe 2013 conference, where this result was first publicly presented.

References

[BCST92] David A. Mix Barrington, Kevin Compton, Howard Straubing, and Denis Thérien. Regular languages in nc1. *Journal of Computer and System Sciences*, 44(3):478 – 499, 1992.

[BHMV94] Véronique Bruyère, Georges Hansel, Christian Michaux, and Roger Villemaire. Logic and p-recognizable sets of integers. *Bull. Belg. Math. Soc*, 1:191–238, 1994.

[Bü60] J. Richard. Büchi. Weak second-order arithmetic and finite automata. *Zeitschrift für Mathematische Logik und Grudlagen der Mathematik*, 6:66–92, 1960.

[Cho08] Christian Choffrut. Deciding whether a relation defined in Presburger logic can be defined in weaker logics. *RAIRO ITA*, 42(1):121–135, 2008.

[CMMP10] Christian Choffrut, Andreas Malcher, Carlo Mereghetti, and Beatrice Palano. On the expressive power of FO[+]. In Adrian Horia Dediu, Henning Fernau, and Carlos Martín-Vide, editors, *LATA*, volume 6031 of *Lecture Notes in Computer Science*, pages 190–201. Springer, 2010.

[Coo72] D. C. Cooper. Theorem proving in arithmetic without multiplication. *Machine Intelligence 7*, pages 91–99, 1972.

[Elg61] Calvin C. Elgot. Decision problems of finite automata design and related arithmetics. *Trans. Amer. Math. Soc.*, 98:21–52, 1961.

[ER66] Calvin C. Elgot and Michael O. Rabin. Decidability and undecidability of extensions of second (first) order theory of (generalized) successor. *J. Symb. Log.*, 31(2):169–181, 1966.

[FL08] Alain Finkel and Jérôme Leroux. Presburger functions are piecewise linear. Research Report LSV-08-08, Laboratoire Spécification et Vérification, ENS Cachan, France, March 2008. 9 pages.

[GS66] Seymour Ginsburg and Edwin Henry Spanier. Semigroups, Presburger formulas, and languages. *Pacific Journal of Mathematics*, 16:285–296, 1966.

[Imm99] Neil Immerman. *Descriptive Complexity*. Springer, 1999.

[Lan04] Klaus-Jörn Lange. Some results on majority quantifiers over words. In *IEEE Conference on Computational Complexity*, pages 123–129. IEEE Computer Society, 2004.

[Min60] Marvin L. Minsky. *Recursive Unsolvability of Post's Problem of "tag": And Other Topics in Theory of Turing Machines*. Group report. Massachusetts Institute of Technology, Lincoln Laboratory, 1960.

[MP71] Robert McNaughton and Seymour Papert. *Counter-free automata*. M.I.T. Press Cambridge, Mass., 1971.

[Muc03] Andrei A. Muchnik. The definable criterion for definability in Presburger arithmetic and its applications. *Theor. Comput. Sci.*, 290(3):1433–1444, 2003.

[MV96] Christian Michaux and Roger Villemaire. Presburger arithmetic and recognizability of sets of natural numbers by automata: New proofs of cobham's and semenov's theorems. *Ann. Pure Appl. Logic*, 77(3):251–277, 1996.

[Pél92] Pierre Péladeau. Formulas, regular languages and boolean circuits. *Theor. Comput. Sci.*, 101(1):133–141, 1992.

[Pre27] Mojżesz Presburger. Über die vollständigkeit eines gewissen systems der arithmetik ganzer zahlen, in welchen, die addition als einzige operation hervortritt. In *Comptes Rendus du Premier Congrès des Mathématicienes des Pays Slaves*, pages 92–101, 395, Warsaw, 1927.

[Str94] Howard Straubing. *Finite Automata, Formal Logic, and Circuit Complexity*. Birkhäuser, 1994.

[Tra50] B. A. Trakhtenbrot. Impossibility of an algorithm for the decision problem in finite classes. *Doklady Akademii Nauk SSSR*, 70:569–572, 1950. (in Russian).

[Tra61] B. A. Trakhtenbrot. Finite automata and logic of monadic predicates (in Russian). *Dokl. Akad. Nauk SSSR*, 140:326–329, 1961.