DEFINABILITY IN ARITHMETICS AND COMPUTABILITY

Déjà parus

CAHIER 1 — épuisé
Intuitionnisme et théorie de la démonstration.

CAHIER 2 — épuisé
Textes de Jean Pieters.
– *Origines de la découverte par Leibniz du calcul infinitésimal.*
– *Frege et le projet des Grundlagen.*
– *La théorie des types de Russell et Whitehead.*

CAHIER 3 — épuisé
J.L. Moens. *Forcing et sémantique de Kripke–Joyal.*

CAHIER 4 — épuisé
La théorie des ensembles de Quine.

CAHIER 5
T.E. Forster. *Quine's New Foundations.*

CAHIER 6
Logique et informatique.

CAHIER 7
L'antifondation en logique et en théorie des ensembles.

CAHIER 8
Ph. de Groote (ed.). *The Curry–Howard Isomorphism.*

CAHIER 9
A. Pétry (éd.). *Méthodes et analyse non standard.*

CAHIER 10
M.R. Holmes. *Elementary Set Theory with a Universal Set.*

CAHIERS DU CENTRE DE LOGIQUE

11

DEFINABILITY IN ARITHMETICS AND COMPUTABILITY

C. Michaux

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
DÉPARTEMENT DE PHILOSOPHIE

BRUYLANT–ACADEMIA – LOUVAIN-LA-NEUVE – 2000

CAHIERS DU CENTRE DE LOGIQUE

Volume 11 sous la direction de :
Chr. Michaux,
Université de Mons-Hainaut.

Composition :
D. Dzierzgowski.

Centre de logique
Département de philosophie
Place Mercier 14 – B-1348 Louvain-la-Neuve (Belgique)

D/2000/4910/09 ISBN 2–87209–577–2

Imprimé en Belgique.

Préface

Le onzième volume des Cahiers du Centre de logique *est une collection d'articles concernant la définissabilité en arithmétique et la calculabilité, et, plus particulièrement, la décidabilité de certaines extensions de l'arithmétique de Presburger.*

L'idée de ce volume est née pendant l'année académique 1995–1996, année où je donnai à Bruxelles un séminaire sur la décidabilité en arithmétique. La première partie de ces exposés était consacrée à des travaux récents (publiés en 1993) par Bateman, Jockush et Woods sur la décidabilité avec un prédicat pour les nombres premiers. Ces travaux contenaient aussi la première preuve publiée d'un important critère de décidabilité dû à Semenov. La seconde partie de ce séminaire fut consacrée à un survol de plusieurs articles importants publiés par Semenov dans les années septante et quatre-vingt, concernant la décidabilité de certaines extensions de l'arithmétique de Presburger. Ces exposés furent suivis de nombreuses discussions.

Pendant cette même période, plusieurs participants du séminaire travaillaient sur les interactions entre la définissabilité et la décidabilité en arithmétique. Le présent volume contient une partie de ces travaux.

Ce volume s'ouvre sur une contribution de A. Maes. Il y fait une relecture personnelle des travaux de A. L. Semenov sur certaines extensions de l'arithmétique de Presburger; il met particulièrement en lumière la filiation des méthodes utilisées avec celles de M. Presburger dans sa célèbre preuve de la décidabilité de la théorie des naturels avec l'addition.

L'article suivant est une courte contribution par Th. Lavendhomme et A. Maes. Les auteurs y donnent une nouvelle preuve d'un résultat de M. Boffa sur l'indécidabilité de la théorie du premier ordre des naturels avec l'addition et un prédicat pour les nombres premiers d'une progression arithmétique.

Dans le troisième article, M. Margenstern et L. Pavlotskaïa développent la notion de fonctions calculables par une machine de Turing sur un ensemble donné de mots et montrent que cette notion est très dépendante de la notion de calcul choisie, en particulier pour les machines de Turing universelles.

Finalement le volume se ferme sur une contribution de Fr. Point. Elle étudie des extensions de l'arithmétique de Presburger liées à certains systèmes de numérations (dits de Bertrand). Par des méthodes modèle-théoriques, elle obtient plusieurs résultats d'élimination (relative) des quantificateurs et de décidabilité pour ces extensions de l'arithmétique de Presburger.

Ce volume a été réalisé avec le soutien du Centre national de recherches de logique. *Je remercie les auteurs pour leurs contributions ainsi que D. Dzierzgowski qui a pris en charge la réalisation matérielle de ce volume.*

CHR. MICHAUX,
Université de Mons-Hainaut,
Novembre 1999.

Preface

The eleventh volume of the Cahiers du Centre de logique *is a collection of contributions to the study of definability in arithmetics and computability; special emphasis is put on the decidability of some extensions of Presburger arithmetic.*

During years 1995–1996 I gave a seminar on decidability in arithmetics. The first part of these lectures was devoted to recent work (published in 1993) by Bateman, Jockush and Woods on decidability with a predicate for the prime natural numbers. This paper also contains the first written proof of an important decidability criterium (due to Semenov). The second part of my lectures surveys several papers published by Semenov during the 70's and 80's. They dealt with certain extensions of Presburger arithmetic. These lectures were followed by numerous discussions and bore the idea to have this volume.

At that time several participants to the seminar worked on the interactions between decidability and arithmetics. This volume contains part of these works.

In the first of these articles, A. Maes revisits A.L. Semenov's work on some extensions of Presburger arithmetic. He sheds a particular light on the filiation between Semenov's methods and the celebrated proof by Presburger that the theory of natural numbers with addition is decidable.

The next contribution, due to T. Lavendhomme and A. Maes, provides *a new proof of a recent result by M. Boffa on the undecidability of Presburger arithmetic enriched with a predicate for the prime numbers of an arithmetical progression.*

In the third paper M. Margenstern and L. Pavlotskaïa introduce the notion of a function computable by a Turing machine on a fixed set of words. They show that this notion is very dependent on the notion of computation which has been chosen, in particular for universal Turing machines.

F. Point, in the last contribution to this volume studies extension of Presburger arithmetic closely related to Bertrand numeration systems. By model-theoretic methods she proves several (relative) quantifier elimination and decidability results.

This volume has been realized with the support of the Centre national de recherches de logique. I want to thank the authors for their contributions together with D. Dzierzgowski who formatted this volume.

C. MICHAUX,
Université de Mons-Hainaut,
November 1999.

Contents

Cahiers du Centre de logique
Volume 11

Revisiting Semenov's Results about Decidability of Extensions of Presburger Arithmetic

by

A. Maes [1]
Université de Mons-Hainaut

1. Introduction

The aim of this paper is to give a different sight on some of Semenov's results [10]. These extend and precise classical results about the decidability of arithmetic theories.

Langford [5] shows that the theory of the structure $\langle\omega; S, <, 0\rangle$ is decidable (the theory of the natural numbers with the unary function 'successor' [2], the order relation and the constant 0). This is proved by showing that this theory admits quantifiers elimination (QE) (see [4, §3.2]).

Presburger [8] shows that the theory of the structure $\langle\omega; S, +, <, 0\rangle$ is decidable as well. Although this theory does not admit QE, it is interdefinable with the theory of the structure $\langle\omega; S, +, <, \equiv_2, \equiv_3, \ldots, 0\rangle$, where $\equiv_n$ is the binary relation of congruence modulo $n \in \mathbb{N}$, and this last theory admits QE.

Semenov [10] considers several possible extensions of these results. First,

1. Aspirant du Fonds national belge de la recherche scientifique.
2. The successor function maps x onto $x + 1$.

given a set $\mathcal{P}$ of unary predicates, he gives a necessary and sufficient condition such that the theory of the structure $\langle \omega; S, <, 0, \mathcal{P} \rangle$ admits QE. Moreover, he precises those sets $\mathcal{P}$ such that this theory is decidable. His characterization uses the notion of *almost-periodicity* defined in Section 2. The proof of his result, given in Section 3, is some kind of generalization of Presburger's proof: in order to remove an existential quantifier acting on an open formula, it suffices to check the compatibility of 'bounds' as well as the satisfaction of the formula on a segment of fixed length. An example of predicate is given in Appendix A such that the theory of the corresponding structure is undecidable but admits QE [3].

Semenov also considers a non-almost-periodic infinite predicate R and wonders about the decidability of the theories of structures like $\langle \omega; S, <, R, 0 \rangle$ and $\langle \omega; S, +, <, R, 0 \rangle$. Given such a predicate, and in order to answer these questions, we first expand our language: we define a larger structure built on two domains (both being ω), one of them containing the elements of R, the other being used for 'indexing' these elements. Let $T_R^<$ (resp. T_R^+) represent the theory on this new structure. We show in Section 4 (resp. 5) that whenever R satisfies certain conditions, we may link the decidability of $T_R^<$ (resp. T_R^+) to that of the 'index-theory' (a single-domain 'subtheory' of $T_R^<$ (resp. T_R^+)). Using the previous characterization, we obtain the decidability of this 'index-theory', implying that of $T_R^<$ (resp. T_R^+). We finally show in Section 6 how to link these results back to the initial problem about $\langle \omega; S, <, R, 0 \rangle$ and $\langle \omega; S, +, <, R, 0 \rangle$.

Although everything lies in [10], this paper tries to give a different approach by formalizing these two-domain structures and gives complete proofs of Semenov's results.

As further reading, we may cite another paper by Semenov [11] where similar results are drawn for extension of Presburger arithmetic with unary functions. More recently, Bateman, Jockusch and Woods [1] shown that the linear case of Schinzel's Hypothesis implies the undecidability of the first order theory of $\langle \omega; +, P \rangle$ (see also [6] in this Cahier). Finally, Cobham's and Semenov's theorems ([3] and [9]) give other examples of methods of definability and of recognizability. See [7] to have a recent proof these results.

3. This will emphasize the notion of *effectiveness*.

2. Almost Periodicity

Let us introduce the notion of almost-periodicity for words and predicates. Basically, we say that an infinite word W on a given alphabet [4] L (we write $W \in L^\omega$) is almost-periodic if it satisfies the following property: for any finite word w on L (we write $w \in L^*$), if w appears infinitely often as subword of W, then the distance between two consecutive occurrences of w in W is bounded.

Notation. Given an infinite word on an alphabet L, say $W = c_1 c_2 \ldots \in L^\omega$, and given two natural numbers $0 \leqslant x \leqslant y$, we denote by $W[x,y]$ the finite subword $c_x c_{x+1} \ldots c_{y-1} c_y$. We also define $W[x, \infty[$ in the obvious way.

We denote the length of a finite word w by $|w|$. ⌟

Definition 1. *Let $W \in L^\omega$ be an infinite word. We say that W is* almost-periodic (a-p) *if and only if, for any finite word $w \in L^*$, there exists a natural number Δ_w such that*

- *either w does not appear in $W[\Delta_w, \infty[$,*
- *or for any $x \in \mathbb{N}$, w appears in $W[x, x + \Delta_w]$.*

If W is an almost-periodic word, then we say that W is effectively almost-periodic (e-a-p) *if there exists an algorithm providing Δ_w for any given word $w \in L^*$.*

We say that Δ_w is an *almost-period* of w in W. Notice that any natural number greater than Δ_w is an almost-period as well.

Definition 2. *We extend this notion to unary predicates on $\mathbb{N}$: we say that a predicate $P \subseteq \mathbb{N}$ is* (effectively) almost-periodic *if and only if so is its characteristic word $W_P \in \{0,1\}^\omega$ defined by*

$$\forall x \in \mathbb{N} \quad W_P[x,x] = \text{`1'} \Leftrightarrow P(x)$$

Example 3. Any (ultimately) periodic word W is almost-periodic.

4. Practically, our alphabet will be $L = \{0,1\}$ and words will be characteristic words of predicates.

Indeed, suppose W is periodic and let π be its period (the case where W is ultimately periodic is very similar). A word w appearing in W in position $x \in \mathbb{N}$ will also appear in position $x+\pi, x+2\pi, \ldots$ so that any subword of W of length $|w|+\pi$ will contain a copy of w. Thus we may take $\Delta_w = |w| + \pi$. ⌟

Example 4. Thue-Morse word is obtained as 'limit' of the iteration

- $t_0 =$ '0'
- t_i is made out of t_{i-1} by replacing each symbol '0' with '01' and each '1' with '10'.

So Thue-Morse word starts with 0110100110010110...

We show in Appendix A that this word is almost-periodic. ⌟

See Appendix A for other examples of almost-periodic words, as well as for an example of non-effectively almost-periodic word.

We extend the notion of almost-periodicity to sets of words by requiring the existence of an almost-period when searching *simultaneously* on finitely many of these words:

Definition 5. *We say that a set of infinite words $\mathcal{W} = \{W_i \mid i \in I\}$ on an alphabet L forms an* almost-periodic system *if, for any finite sequence $(W_{i_1}, \ldots, W_{i_n})$ of words of $\mathcal{W}$, and any finite sequence $(w_1, \ldots, w_n)$ of finite words of L^*, there exists a natural number Δ such that*

- *either there is no $x > \Delta$ such that $w_j = W_{i_j}[x, x+|w_j|-1]$ for all j,*
- *or for any $x \in \mathbb{N}$, there exists a y such that $x \leqslant y \leqslant x + \Delta$ and*

$$w_j = W_{i_j}[y, y+|w_j|-1] \text{ for all } j.$$

Denote by $\genfrac{}{}{0pt}{}{W_{i_1}}{\overset{\cdots}{W_{i_n}}}$ the infinite word (seen as a word on L^n) obtained by 'stacking' the words $W_{i_1}, \ldots, W_{i_n}$. The previous condition just consists in requiring the word $\genfrac{}{}{0pt}{}{W_{i_1}}{\overset{\cdots}{W_{i_n}}}$ to be almost-periodic.

The system is effectively almost-periodic if there exists an algorithm providing Δ for any choice of $n, i_1, \ldots, i_n$ and $w_1, \ldots, w_n$.

Definition 6. *Again, we extend this notion to sets of unary predicates in the obvious way.*

Example 7. The set $\mathcal{E} = \{\cdot \equiv_m c \mid c, m \in \mathbb{N}, c < m \neq 0, 1\}$ of congruences modulo m forms an *a-p* system.

This comes from the fact that a word obtained by stacking elements of $\mathcal{E}$ is periodic (a period is the lowest common multiple of the m's of the words forming it). ⌟

Remark 8. Basic properties of almost-periodic systems of predicates are:

- Any subset of an *a-p* system still is an *a-p* system.
- They may be closed by boolean operations, that is, for any predicates P_1, P_2 of an *a-p* system $\mathcal{P}$,

  $$\mathcal{P} \cup \{P_1 \vee P_2\},\ \mathcal{P} \cup \{P_1 \wedge P_2\} \text{ and } \mathcal{P} \cup \{\neg P_1\}$$

  still form an *a-p* system [5].

  This is clear for '$\wedge$' from the definition. '$\neg$' comes from the fact that for any word $w \in \{0,1\}^* w$ appears somewhere in $W_{\neg P_1} = \overline{W_{P_1}}$ if and only if $\overline{w}$ appears at the same place in W_{P_1}. (We denote by $\overline{W}$ the word W where every '0' has been changed into a '1' and conversely). This implies the result for '$\vee$'. □
- They may be closed by translations, that is, for any predicate P of an *a-p* system $\mathcal{P}$ and for any integer c,

  $$\mathcal{P} \cup \{P^{+c}\}$$

  still forms an *a-p* system [6], where P^{+c} is the unary predicate defined by

  $$P^{+c}(x) \Leftrightarrow x \geqslant -c \wedge P(x+c).$$

5. As the condition for *a-p* systems only requires to be satisfied on finite sets of predicates, we see that the boolean closure of $\mathcal{P}$ is indeed an *a-p* system.
6. The same remark applies: $\{P^{+c} \mid P \in \mathcal{P}, c \in \mathbb{Z}\}$ is also an *a-p* system.

Indeed, the characteristic word of P^{+c} is the characteristic word of P whose the c first symbols have been removed (when $c \geqslant 0$) or in front of which $|c|$ symbols '0' have been added (when $c < 0$).

$$\text{Example:}\quad \begin{array}{rcr} \text{if } P & = & 10111110111110\ldots \\ P^{+2} & = & 111110111110\ldots \\ P^{-2} & = & 0010111110111110\ldots \end{array}$$

Using this, it is easy to see that 'extended systems' are *a-p*: in order to find an almost-period for a word

$$u = \begin{pmatrix} u_{11} & \ldots & u_{1m} \\ u_{n1} & \ldots & u_{nm} \end{pmatrix} \in (\{0,1\}^n)^*$$

in the characteristic word of $\begin{smallmatrix} P_1^{+c_1} \\ \cdots \\ P_n^{+c_n} \end{smallmatrix}$, it suffices to look in $\begin{smallmatrix} P_1 \\ \cdots \\ P_n \end{smallmatrix}$ for an almost-period of the 'incomplete' word v obtained from u by shifting the i^{th} line by c_i units (to the right when $c_i > 0$, to the left when $c_i < 0$, $1 \leqslant i \leqslant n$) [7].

Let M be the set of all words of $(\{0,1\}^n)^*$ of length $|v|$ that coincide with v where v is defined. Remark that there are finitely many such words. For each word $\tilde{v} \in M$, there exists a $\tilde{v}$-almost-period $\Delta_{\tilde{v}}$ in $\begin{smallmatrix} P_1 \\ \cdots \\ P_n \end{smallmatrix}$ (which can be effectively found whenever $\mathcal{P}$ is *e-a-p*). So $\Delta = \max_{\tilde{v} \in M} \{\Delta_{\tilde{v}}\}$ is an almost-period for all $\tilde{v} \in M$. This Δ is suitable as almost-period of v in $\begin{smallmatrix} P_1 \\ \cdots \\ P_n \end{smallmatrix}$ and thus as almost-period of u in $\begin{smallmatrix} P_1^{+c_1} \\ \cdots \\ P_n^{+c_n} \end{smallmatrix}$. ⌟

Remark 9. We shall work on theories of structures with the order relation and (at least) the successor function. Given a set of predicate $\mathcal{P}$ (not necessarily an *a-p* system), we shall consider its closure by translations $\mathcal{P}'$. The structures $\langle \omega; S, <, 0, \mathcal{P} \rangle$ and $\langle \omega; S, <, 0, \mathcal{P}' \rangle$ are interdefinable since any $P^{+c} \in \mathcal{P}$ is definable using $P \in \mathcal{P}$, S and $<$ by

$$\begin{array}{rll} P^{+c}(x) & \Leftrightarrow P(\underbrace{S \ldots S}_{c \text{ times}}\, x) & \text{if } c \geqslant 0, \\ & \Leftrightarrow \exists x' (\underbrace{S \ldots S}_{|c| \text{ times}}\, x' = x \wedge P(x')) & \text{if } c < 0. \end{array}$$
⌟

7. For instance, any almost-period for $u = \binom{11}{10}$ in $\begin{smallmatrix} P_1^{+1} \\ P_2^{-2} \end{smallmatrix}$ is an almost-period for $v = \binom{\cdot\,\cdot\,11}{10\,\cdot\,\cdot\,\cdot}$ in $\begin{smallmatrix} P_1 \\ P_2 \end{smallmatrix}$ and conversely.

3. Quantifiers Elimination and Almost-Periodicity

We now get to the following question: given a set of unary predicates $\mathcal{P} = \{P_i \mid i \in I\}$, does the theory of the structure $\langle \omega; S, <, 0, \mathcal{P} \rangle$ admit Quantifier Elimination?

As explained above (see Remark 9), we shall suppose that $\mathcal{P}$ is closed by translations.

Theorem 10. *Let $\mathcal{P}$ be a set of unary predicates defined on ω, closed by translations. Then the theory T of the structure $\langle \omega; S, <, 0, \mathcal{P} \rangle$ admits Quantifiers Elimination if and only if $\mathcal{P}$ is an almost-periodic system.*

Moreover, whenever this is the case, T is decidable if and only if $\mathcal{P}$ is an effectively almost-periodic system of decidable predicates.

PROOF. — We first show that if $\mathcal{P}$ is an *a-p* system, then T admits QE. Then we show the link between the decidability of T and the effectiveness of the almost-periodicity of $\mathcal{P}$. Finally, we show why the fact that T admits QE implies that $\mathcal{P}$ is an *a-p* system.

First step: *if P is an a-p system, then T admits QE.*

In order to proof the first step, it suffices to show that we are able to eliminate the quantifier from any formula of the form

$$\exists x \ \varphi(x, x_1, \ldots, x_l)$$

where φ is an conjunctive open formula and $x_1, \ldots, x_l$ are variables differing from x.

First, let us precise the notations and simplify the general form of conjunctive open formulæ.

By $\bar{x}$ we denote the variables $x_1, \ldots, x_l$ appearing in φ and differing from x (the quantified variable). By $\tau(\bar{x})$ (resp. $\psi(\bar{x})$) we denote a term (resp. a formula) in which only appear variables of $\bar{x}$.

Terms are expressions of the form S^n0, S^nx and S^nx_i, with $n \in \mathbb{N}$ and $x_i \in \bar{x}$.

An atomic formula is any expression of the form $\tau_1 = \tau_2$, $\tau_1 < \tau_2$ or $P(\tau_1)$ where τ_1 and τ_2 are terms and $P \in \mathcal{P}$.

Of course, $\tau_1 = \tau_2$, $\neg(\tau_1 = \tau_2)$ and $\neg(\tau_1 < \tau_2)$ are definable by positive formulæ using 'S' and '$<$'. Thus we may suppose φ only contains non-negated comparisons.

Without loss of generality, we may suppose that $\mathcal{P}$ is closed by negations. Indeed, using Remark 8, we see that $\mathcal{P}$ is *(e-)a-p* if and only if its closure by negations is *(e-)a-p*.

Thus we just have to take into account formulæ of the form $\tau_1 < \tau_2$ and $P(\tau)$.

However, we temporarily enlarge our language in order to use 'extended terms': we introduce terms like $S^{-n}0$ and $S^{-n}x_i$ where $n \in \mathbb{N}$ and $x_i \in \bar{x}$ (remember x is *not* a variable of $\bar{x}$). It is obvious that 'extended comparisons' are always equivalent to some usual comparison [8], and the same is true for predicates of $\mathcal{P}$ as this system is closed by translations.

Now, any formula φ can be written as a conjunction [9]

$$\bigwedge_{a \in A}(\tau_a(\bar{x}) < x) \ \wedge \ \bigwedge_{b \in B}(x < \sigma_b(\bar{x})) \ \wedge \ \bigwedge_{c \in C} P_c(x) \tag{1}$$

where A, B, C are finite indexing sets, τ_a and σ_b are extended terms for any $a \in A$ and $b \in B$, and $P_c \in \mathcal{P}$ for any $c \in C$. Notice that the terms τ_a (resp. σ_b) give lower (resp. upper) bounds on x.

We now show how to transform the formula $\exists x \varphi(x, \bar{x})$ into an equivalent open formula.

(a) Suppose we have no predicate P_c, that is, $C = \varnothing$.

Then we eliminate the quantifier exactly in the same way as when studying the structure $\langle \omega; S, <, 0 \rangle$: a solution for the formula (1) exists if and only if there is some 'space' between upper and lower bounds, that is if and only if these two conditions are both satisfied:

8. For instance, $S^{-n}x_1 < S^m x_2$ is equivalent to $x_1 < S^{n+m}x_2$.
9. Of course, atomic formulæ not containing x may be moved outside the action of the quantifier. Moreover, formulæ like $S^{n_1}x < S^{n_2}x$ are equivalent to $S^{n_1}0 < S^{n_2}0$.

- the maximum lower bound +1 (if any) is strictly lower than the minimum upper bound (if any),
- the minimum upper bound (if any) is strictly positive.

This can be expressed by the open formula

$$\bigwedge_{\substack{a\in A\\ b\in B}} (S\tau_a(\bar{x}) < \sigma_b(\bar{x})) \ \wedge \ \bigwedge_{b\in B} (S0 < \sigma_b(\bar{x})). \tag{2}$$

(b) Suppose $C \neq \varnothing$.

We now need to find some integer x between the upper and lower bounds such that $\bigwedge_{c\in C} P_c(x)$ is satisfied. Fortunately, the almost-periodicity of $\mathcal{P}$ tells us that there exists some Δ such that it suffices to test the Δ first possible values: either one of them satisfies $\bigwedge_{c\in C} P_c(x)$ or this formula is always false. Here, Δ is an almost-period for the word $\substack{1\\ \cdots\\ 1}$ in the word $\substack{P_1\\ \cdots\\ P_C}$.

Practically, given a lower bound $\tau_\alpha, \alpha \in A$, we test all the conditions on $\tau_\alpha + 1, \ldots, \tau_\alpha + \Delta$ *and* whether these values are positive. Since all the $\tau_\alpha + 1$ might be negative, we also test the condition on $0, \ldots, \Delta$ which would then be the Δ first possible values. So we get the following formula

$$\begin{aligned} &\bigvee_{\substack{i=1\ldots\Delta\\ \alpha\in A}} \Big(0 \leqslant S^i\tau_\alpha \ \wedge \ \bigwedge_{a\in A} \tau_a < S^i\tau_\alpha \\ &\qquad \wedge \ \bigwedge_{b\in B} S^i\tau_\alpha < \sigma_b \ \wedge \ \bigwedge_{j=1\ldots r} P_j(S^i\tau_\alpha)\Big) \\ &\vee \bigvee_{i=0\ldots\Delta-1} \Big(\bigwedge_{a\in A} \tau_a < S^i 0) \\ &\qquad \wedge \bigwedge_{b\in B} S^i 0 < \sigma_b \wedge \bigwedge_{j=1\ldots r} P_j(S^i 0)\Big). \end{aligned} \tag{3}$$

As we said before, any 'extended formula' appearing in (3) may be written back as a usual formula. We then have an open formula equivalent to the initial $\exists x\varphi(x, \bar{x})$, and this concludes the quantifier elimination.

Second step: *link between the decidability of T and the effectiveness of the almost-periodicity of P.*

When $\mathcal{P}$ is an *e-a-p* system, each Δ we need can be effectively obtained. If $\mathcal{P}$ is also a system of decidable predicates, then it is obvious we have a decidable theory: the satisfaction of a sentence can be checked just by computation of the quantifier-free equivalent form.

Conversely, if $\mathcal{P}$ is an *a-p* system and the theory of $\langle \omega; S, <, 0, \mathcal{P} \rangle$ is decidable, then $\mathcal{P}$ is *e-a-p*:

For any $p \in \mathcal{P}$, we may easily code the condition of the occurrence of a finite word $u = u_0 \dots u_l \in \{0,1\}^*$ at position x in the word W_P by the formula

$$\bigwedge_{j=0\dots l} (\neg)_j P(S^j x)$$

where $(\neg)_j$ means that the negation symbol is present if and only if $u_j =$ '0'.

Using this formulation, given a word

$$u = \begin{pmatrix} u_{10} & \dots & u_{1l} \\ u_{n0} & \dots & u_{nl} \end{pmatrix} \in (\{0,1\}^n)^*,$$

given some predicates $P_1, \dots, P_n \in \mathcal{P}$ and given a natural number $m > l$, we write the following sentence φ_m

$$\begin{aligned} &\forall x \Big(\bigvee_{i=0\dots m-l} \bigwedge_{\substack{j=0\dots l \\ k=1\dots n}} (\neg)_{kj} P_k(S^{i+j}x) \Big) \\ &\vee \ \forall x \Big(x > S^m 0 \to \neg \bigwedge_{\substack{j=0\dots l \\ k=1\dots n}} (\neg)_{kj} P_k(S^j x) \Big) \end{aligned} \tag{4}$$

which expresses that m is an almost-period for the word u in $\begin{smallmatrix} P_1 \\ \vdots \\ P_n \end{smallmatrix}$ (again, $(\neg)_{kj}$ means that the negation symbol is present if and only if $u_{kj} =$ '0').

We may effectively find an almost-period for u by deciding the sentences $\varphi_{l+1}, \varphi_{l+2}, \dots$ until we find a true one. This algorithm stops since $\mathcal{P}$ is *a-p* and since only finitely many natural numbers are not an almost-period for u.

Third step: *any set $\mathcal{P}$ of unary predicates such that the theory T of the structure $\langle \omega; S, <, 0, \mathcal{P}\rangle$ admits QE is an almost-periodic system.*

For a contradiction, suppose we have a word $u \in (\{0,1\}^n)^*$ of length l and some predicates $P_1, \ldots, P_n \in \mathcal{P}$ such that u appears infinitely often in $\begin{smallmatrix}P_1\\ \vdots\\ P_n\end{smallmatrix}$ but such that there exist arbitrarily large segments of $\begin{smallmatrix}P_1\\ \vdots\\ P_n\end{smallmatrix}$ not containing u.

Let $\varphi(x, y)$ be the formula expressing that $x < y$ and that u does not appear in the segment $[x, y]$ of $\begin{smallmatrix}P_1\\ \vdots\\ P_n\end{smallmatrix}$:

$$x < y \quad \wedge \quad \forall z\ (x < z \wedge S^l z < y) \ \rightarrow \neg \bigwedge_{\substack{j=0\ldots l-1\\ k=1\ldots n}} (\neg)_{kj} P_k(S^j z).$$

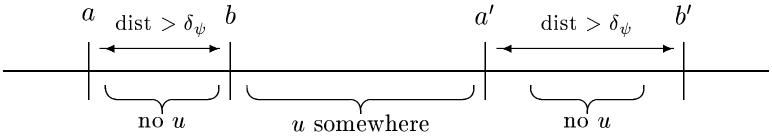
As the theory T admits QE, $\varphi(x, y)$ is equivalent to an open formula $\psi(x, y)$. The atomic subformulæ of ψ are of three possible types: those which contain neither x nor y, those which contain either x or y, and those which contain both x and y.

These last ones are necessarily comparisons. They are of the form $S^{n_1}x < S^{n_2}y$ or $S^{n_1}y < S^{n_2}x$. Let δ_ψ be the maximum of the $n_1 + n_2$'s. There are infinitely many couples (a, b) such that both $a + \delta_\psi < b$ and $\psi(a, b)$ are satisfied (otherwise we would have a contradiction with the choice of u). As ψ is built with finitely many atomic formulæ, it is possible to find two pairs of numbers (a, b) and (a', b') such that the following three conditions are satisfied:

(a) $a + \delta_\psi < b < a'$ and $a' + \delta_\psi < b'$,

(b) any atomic formula of ψ is satisfied on (a, b) if and only if it is satisfied on (a', b'),

(c) u does appear in the segment $[b, a']$ of $\begin{smallmatrix}P_1\\ \vdots\\ P_n\end{smallmatrix}$.



Thanks to condition (a) and to the choice of δ_ψ, any comparison appearing in ψ gets the same truth value on (a,b), on (a',b') and on (a,b').

Obviously, we have the same behaviour for atomic formulæ which contain neither x nor y.

Now, if y does not appear in an atomic formula ξ of ψ, then the truth value of ξ is the same on (a,b) and on (a,b').

Similarly, if ξ does not contain the variable x, then its truth value is the same on (a',b') and on (a,b'), and using condition (b), it is the same on (a,b) as well.

Since $\psi(a,b)$ is true and since any atomic formula of ψ is satisfied on (a,b) if and only if it is satisfied on (a,b'), $\psi(a,b')$ has to be true. However, this tell us u does not appear in the segment $[b,a']$ of $\substack{P_1\\ \vdots\\ P_n}$, but this contradicts condition (b). This concludes the proof of Theorem 10 □

4. Decidability, Index Function, Successor and Almost-Periodicity

In order to study structures like $\langle\omega; S, <, R, 0\rangle$ or $\langle\omega; +, <, R, 0, 1\rangle$, where R is a non-almost-periodic infinite predicate on $\mathbb{N}$, we shall consider larger structures, built on two domains, one of them being 'at the level of the elements of R', the other one being used to index these elements. The results of this section and of the next one aim to study these larger structures and to show, in certain cases, that their theories admit a property like QE, in this case, that they are 'partially existential' (see Definition 11). We shall then study the links with the initial structures in order to know about their decidability. However, the structures we are going to consider now are a bit more general than what we shall need later.

We first study the structures of the form $\langle\omega; S, <, R, 0\rangle$.

Let $R \subseteq \mathbb{N}$ be an infinite set.

We consider the following structure made of two domains: both domains are ω, but in order to avoid any ambiguity, we temporarily denote them

ω_1 and ω_2. On each domain ω_i ($i = 1, 2$), there is a successor function S_i, an order relation $<_i$ and a zero constant 0_i. When it is clear what is the domain we are working with, we shall use the symbols $S, <, 0$ instead of $S_1, <_1, 0_1$ or $S_2, <_2, 0_2$.

Let $X = \{x, x', x_1, x_2, \ldots\}$ be the set of variables whose values are in ω_1 and let $Y = \{y, y', y_1, y_2, \ldots\}$ be the set of variables whose values are in ω_2.

The domain ω_1 is used as the set which contains R. The elements of R can be enumerated, and this is done using the domain ω_2: we add to our language a unary function symbol '$R_{(\cdot)}$' whose domain is ω_2 and whose images are in ω_1, and that is interpreted as

$$R_{(\cdot)} : \omega_2 \to \omega_1 : n \mapsto R_n = \text{the } (n+1)^{\text{th}} \text{ element of } R.$$

The notion of 'distance' between two consecutive elements of R can be expressed using the unary predicates $I_k(y)$ defined on ω_2 by the formula $S_1^k R_y < R_{(S_2 y)}$ with $k \in \mathbb{N}$. The predicate $I_k(y)$ clearly means that the distance between the $(y+1)^{\text{th}}$ element of R and its successor in R is strictly greater than k. Let $\mathcal{I}_R$ designate the set of these predicates.

Let $\mathcal{P}$ be a system of almost-periodic predicates defined on ω_1 and suppose it is closed by negations and translations.

Using the function $R_{(\cdot)}$, any predicate defined on ω_1 — not just unary ones — gives rise to a predicate of the same arity defined on ω_2: if $X \subseteq \omega_1^n$, we denote by $X_R \subseteq \omega_2^n$ the predicate defined by

$$\forall b_1, \ldots, b_n \in \omega_2 \quad X_R(b_1, \ldots, b_n) \Leftrightarrow X(R_{b_1}, \ldots, R_{b_n}).$$

In particular, $\mathcal{P}$ gives rise to a set $\mathcal{P}_R$ of unary predicates defined on ω_2 (in general, these are not almost-periodic!).

Finally, given a set $\mathcal{Q}$ of (non necessarily unary) predicates defined on $R \subseteq \omega_1$ (that is, if $Q \in \mathcal{Q}$ is an n-ary predicate, then $Q \subseteq R^n$), we may equivalently [10] consider the sets $\mathcal{Q}$ and $\mathcal{Q}_R$. This will allow us to move these predicates in the 'index theory' (see later). In the following we shall denote the predicates of $\mathcal{Q}$ under their form $Q(R_{(\cdot)})$.

10. We have to show the interdefinability: for an n-ary predicate, we have $Q(x_1, \ldots, x_n) \Leftrightarrow \exists y_1, \ldots, y_n \ (x_1 = R_{y_1} \wedge \ldots \wedge x_n = R_{y_n} \wedge Q(x_1, \ldots, x_n)) \Leftrightarrow \exists y_1, \ldots, y_n \ (x_1 = R_{y_1} \wedge \ldots \wedge x_n = R_{y_n} \wedge Q_R(y_1, \ldots, y_n))$.

We consider the theories of the following two structures: first of all, the theory of $\langle\omega_1, \omega_2; S_1, S_2, <_1, <_2, 0_1, 0_2, R_{(\cdot)}, \mathcal{P}, \mathcal{Q}_R\rangle$ which will be abbreviated as $T_R^<$. Notice that the predicates of $\mathcal{P}_R$ and of $\mathcal{I}_R$ are definable in this theory.

Then we consider the so-called 'index theory' T_{Ind}, that is the theory of the structure $\langle\omega_2; S_2, <_2, 0_2, \mathcal{P}_R, \mathcal{Q}_R, \mathcal{I}_R\rangle$.

We shall see that any formula of $T_R^<$ which only contains variables of Y and constants of ω_2 is equivalent to a formula of T_{Ind} — the converse statement being trivially true. Notice that the function $R_{(\cdot)}$ does *not* belong to T_{Ind}, so that the predicates $\mathcal{P}_R$, $\mathcal{Q}_R$ and $\mathcal{I}_R$ have to be considered as predicates on ω_2 and not as predicates on images of $R_{(\cdot)}$.

Definition 11. *By '$T_R^<$* is an existential theory modulo T_{Ind}*' we mean that for any formula $\varphi(\bar{x}, \bar{y})$ of $T_R^<$, there exists an equivalent formula of the form $\exists\bar{x}', \bar{y}'\ \theta(\bar{x}, \bar{x}', \bar{y}, \bar{y}')$ where θ is a combination of* open *formulæ of $T_R^<$ and of* arbitrary *formulæ of T_{Ind}. Such a formula will be said to be* open modulo T_{Ind}.

Theorem 12. *Let R be an infinite set of natural numbers, $\mathcal{P}$ an almost-periodic system of unary predicates and $\mathcal{Q}$ a set of predicates defined on R. Then the theory $T_R^<$ is existential modulo T_{Ind}. Moreover, $T_R^<$ is decidable if and only if the system $\mathcal{P}$ is decidable and effectively almost-periodic and the theory T_{Ind} is decidable.*

Proof. — We shall show that in the present case, we are even able to eliminate the quantifiers acting on variables of X, that is, we may replace any formula $\varphi(\bar{x}, \bar{y})$ by a formula of the form $\exists\bar{y}'\ \ \theta(\bar{x}, \bar{y}, \bar{y}')$ where θ is open modulo T_{Ind}.

We first show that any formula of $T_R^<$ containing only variables of Y and the constant 0_2 is equivalent to a formula of T_{Ind}.

Then we prove that $T_R^<$ is existential modulo T_{Ind} and that it is effectively existential modulo T_{Ind} when $\mathcal{P}$ is *e-a-p*.

Then we roughly explain the decidability algorithm for sentences of $T_R^<$ when $\mathcal{P}$ is decidable and *e-a-p* and T_{Ind} is decidable.

Finally, the converse implication of Theorem 12 is immediate: for the effectiveness of the almost-periodicity of $\mathcal{P}$, it suffices to express with a sentence that Δ is an almost-period for a given u and to decide this sentence for an increasing sequence of Δ (see Theorem 10); moreover, since any sentence of T_{Ind} is a sentence of $T_R^<$, the decidability of T_{Ind} is induced by that of $T_R^<$.

First step: *any formula of $T_R^<$ containing only variables of Y and the constant 0_2 is equivalent to a formula of T_{Ind}.*

Once again, we use extended terms on ω_1 (see page 18). There are three kinds of terms in $T_R^<$:

Terms τ of ω_1 formed with variables and constants of ω_1	Terms σ of ω_2 formed with variables and constants of ω_2	Terms ν of ω_1 formed with variables and constants of ω_2
$\left.\begin{array}{c} S^n 0 \\ S^n x_i \end{array}\right\} \left\{\begin{array}{l} n \in \mathbb{Z} \\ x_i \in X \end{array}\right.$	$\left.\begin{array}{c} S^m 0 \\ S^m y_j \end{array}\right\} \left\{\begin{array}{l} m \in \mathbb{N} \\ y_j \in Y \end{array}\right.$	$\left.\begin{array}{c} S^n R_{(S^m 0)} \\ S^n R_{(S^m y_j)} \end{array}\right\} \left\{\begin{array}{l} n \in \mathbb{Z} \\ m \in \mathbb{N} \\ y_j \in Y \end{array}\right.$

Similarly, there are three kinds of atomic formulæ (we omit equality because it is definable by an open formula using $<$ and S):

Formulæ containing only vars. and csts. of ω_1	Formulæ containing only vars. and csts. of ω_2	Formulæ mixing vars. and csts. of ω_1 and ω_2
$\tau_1 < \tau_2$ $P(\tau_1), \quad P \in \mathcal{P}$	$\sigma_1 < \sigma_2$ $Q(R_{\sigma_1}, \ldots, R_{\sigma_k}), \quad Q \in \mathcal{Q}$	$\nu_1 < \nu_2$ $\tau_1 < \nu_1$ $\nu_1 < \tau_1$ $P(\nu_1), \quad P \in \mathcal{P}$

After simplifying the S^n's and using the fact that $\mathcal{P}$ is closed by translations and negations, we see that any open formula of $T_R^<$ is built out of the following atomic formulæ:

Formulæ containing only vars. and csts. of ω_1	Formulæ containing only vars. and csts. of ω_2	Formulæ mixing vars. and csts. of ω_1 and ω_2
	$S^m y_j \lesseqgtr 0$	
	$S^m y_j \lesseqgtr y_{j'}$	
$S^n x_i \lesseqgtr 0$	$S^n R_{(S^m y_j)} \lesseqgtr R_{(S^{m'} y_{j'})}$ $(*)$	$S^n R_{(S^m y_j)} \lesseqgtr 0$
$S^n x_i \lesseqgtr x_{i'}$	$S^n R_{(S^m y_j)} \lesseqgtr R_{(S^{m'} y_j)}$ $(**)$	$S^n R_{(S^m y_j)} \lesseqgtr x_i$
$P(x_i)$, $P \in \mathcal{P}$	$P(R_{(S^m y_j)})$, $P \in \mathcal{P}$	$R_{(S^m y_j)} \lesseqgtr S^n x_i$
	$Q(R_{(S^{m_1} y_{j_1})}, \ldots, R_{(S^{m_k} y_{j_k})})$, $Q \in \mathcal{Q}$	
	$I_k(S^m y_j)$	

(5)

The open formulæ of T_{Ind} are those which are built with any of the atomic formulæ of the central column, except $(*)$ and $(**)$.

In order to show that the formulæ of T_{Ind} are, up to equivalence, exactly those of $T_R^<$ which only contain variables of Y and the constant 0_2, it suffices to show what the formulæ of the form $(*)$ and $(**)$ are equivalent to.

First consider the formulæ $(*)$. Let ψ be the formula $S^n R_{(S^m y_j)} > R_{(S^{m'} y_{j'})}$. Thanks to the fact that the function $R_{(\cdot)}$ is strictly increasing, we have that

$$\psi \quad \Leftrightarrow \quad (S^m y_j > S^{m'} y_{j'}) \underbrace{\vee \bigvee_{i=0\ldots n-1} \left(S^{m+i} y_j = S^{m'} y_{j'} \wedge S^n R_{(S^m y_j)} > R_{(S^{m+i} y_j)} \right)}_{\text{if } n>0}, \tag{6}$$

because for any $i \geqslant n$, $S^n R_{(S^m y_j)} \leqslant R_{(S^{m+i} y_j)}$ so that the disjunction does not need to include any $i \geqslant n$. Notice that there is no formula $(*)$ in (6), but there are formulæ $(**)$. Similarly, if ψ is the formula $R_{(S^m y_j)} > S^n R_{(S^{m'} y_{j'})}$, then by transforming ψ into $\neg(R_{(S^m y_j)} < S^{n+1} R_{(S^{m'} y_{j'})})$, we get back to the previous situation.

Now, let us get rid of the formulæ $(**)$:

- If ψ is the formula $R_{(S^m y_j)} < R_{(S^{m'} y_j)}$, that is, $n = 0$, then ψ is equivalent to $S^m y_j < S^{m'} y_j$;

- If ψ is the formula $S^n R_{(S^m y_j)} > R_{(S^{m'} y_j)}$ with $m \geqslant m'$, then the formula is true since $R_{(\cdot)}$ is strictly increasing;

- Similarly, if ψ is the formula $S^n R_{(S^m y_j)} < R_{(S^{m'} y_j)}$ with $m \geqslant m'$, then ψ is false for the same reason;

- If ψ is the formula $S^n R_{(S^m y_j)} > R_{(S^{m'} y_j)}$ with $m < m'$, then we transform ψ into the equivalent formula $\neg(S^{n-1} R_{(S^m y_j)} < R_{(S^{m'} y_j)})$, and this leads us to the last possible case:

- If ψ is the formula $S^n R_{(S^m y_j)} < R_{(S^{m'} y_j)}$ with $m < m'$. Let $l = m' - m$. Using once again the monotonicity of the function $R_{(\cdot)}$, we have that

$$\begin{aligned} S^n R_{(S^m y_j)} &< R_{(S^{m+l} y_j)} \\ &\Leftrightarrow \text{“}n < R_{(S^{m+l} y_j)} - R_{(S^m y_j)}\text{”} \\ &\Leftrightarrow \text{“}n < \underbrace{R_{(S^{m+l} y_j)} - R_{(S^{m+l-1} y_j)}}_{\beta_l} \\ &\qquad + \underbrace{R_{(S^{m+l-1} y_j)} - R_{(S^{m+l-2} y_j)}}_{\beta_{l-1}} \\ &\qquad + \cdots + \underbrace{R_{(S^{m+1} y_j)} - R_{(S^m y_j)}}_{\beta_1}\text{”} \end{aligned}$$

Since the difference between two consecutive elements of R is a strictly positive natural number, the sum of the β_i's will be greater or equal to $n+1$ if and only if we may find natural lower bounds $\alpha_i \leqslant \beta_i$ such that the sum of these α_i's is equal to $n+1$. Thus we only have finitely many possibilities to test:

$$\begin{aligned} \cdots \Leftrightarrow\ &\text{“}\bigvee_{\substack{1 \leqslant \alpha_1, \ldots, \alpha_l \leqslant n+1 \\ \Sigma_i \alpha_i = n+1}} \bigwedge_{i=1\ldots l} \alpha_i \leqslant \beta_i\text{”} \\ \Leftrightarrow\ &\text{“}\bigvee_{\substack{1 \leqslant \alpha_1, \ldots, \alpha_l \leqslant n+1 \\ \Sigma_i \alpha_i = n+1}} \bigwedge_{i=1\ldots l} \alpha_i - 1 < R_{(S^{m+i+1} y_j)} - R_{(S^{m+i} y_j)}\text{”} \end{aligned}$$

$$\Leftrightarrow \bigvee_{\substack{1\leqslant\alpha_1,\ldots,\alpha_l\leqslant n+1\\ \Sigma_i\alpha_i=n+1}} \bigwedge_{i=1\ldots l} I_{\alpha_i-1}(S^{m+i}y_j)$$

and this last expression is a formula of $T_R^<$ which does not contain any formula of the form $(*)$ or $(**)$ except some I_k.

Second step: $T_R^<$ *is existential modulo* T_{Ind}.

We now show that:

(a) we are able to eliminate the quantifier from any formula of the form $\exists x\varphi\,(x,\bar{x},\bar{y})$ where φ is open modulo T_{Ind},

(b) any formula of the form $\neg\,(\exists\bar{y}\;\varphi(\bar{x},\bar{y},\bar{y}'))$, where φ is open modulo T_{Ind}, is equivalent to a formula of the form $\exists\tilde{y}\;\theta(\bar{x},\tilde{y},\bar{y}')$ where θ is open modulo T_{Ind}.

As a careful induction shows, this suffices in order to prove that the theory $T_R^<$ is existential modulo T_{Ind}.

Let us start with (a). Consider a formula $\exists x\;\varphi(x,\bar{x},\bar{y})$ with φ open modulo T_{Ind}. Like in the proof of Theorem 10, after we moved out of the action of the quantifier all sub-formulæ not containing the variable x (including any formulæ of T_{Ind}), it just remains predicates $P(x)$ and formulæ $x<\cdot$ and $\cdot<x$ giving upper and lower bounds on x. Using a completely similar reasoning as in Theorem 10, we get the same formula as in (2) and (3), and one checks immediately that these are expressible using formulæ of the list (5). Thus $\exists x\;\varphi(x,\bar{x},\bar{y})$ is indeed equivalent to some open formula (modulo T_{Ind}).

It remains to show (b). It would be interesting to apply once again a reasoning similar to that of Theorem 10, but the 'lower bound' conditions on the variable y might be expressed by formulæ $S^n x_i < S^{n'}R_{(S^m y)}$ or $S^n 0 < S^{n'}R_{(S^m y)}$. In order to apply the same method as in Theorem 10, we should be able to 'invert' the function $R_{(\cdot)}$. However, adding this inverse function to our language would considerably complicate the set of terms of $T_R^<\ldots$ Let us proceed differently.

If φ does not contain a sub-formula of the form

$$S^n\tau \lesseqgtr S^{n'}R_{(S^m y)} \qquad (\star)$$

with $\tau \in X \cup \{0\}$, then we get a formula open modulo T_{Ind} by moving out of the action of the quantifier any formula not containing a variable of $\bar{y}$. We show, using an induction on the number of formulæ $(\star)$, that we may move them out of the action of the '$\neg\exists\bar{y}$' by adding new existential quantifiers. Here is the (effective) process we may apply:

Suppose that the formula $\neg(\exists\bar{y}\ \varphi(\bar{x},\bar{y},\bar{y}'))$ is of the form

$$\neg\left(\exists\bar{y}\ \left(S^n\tau < S^{n'}R_{(S^m y)} \wedge \psi(\bar{x},\bar{y},\bar{y}')\right)\right) \tag{7}$$

where y is a variable of $\bar{y}$, $\tau \in X \cup \{0\}$ and where ψ is open in $T_R^<$ and may still contain sub-formulæ $(\star)$ (the case '$>$' is completely similar and is discussed afterwards). The trick will consist in "framing", if possible, the term $S^n\tau$ between terms made with a new variable of Y and to compare y to this new variable. That way, the 'bounding condition' will be split into the conjunction of a formula containing other variables than those of $\bar{y}$ (so that they will be moved out of the action of the $\neg\exists\bar{y}$) and comparisons between y and these new variables (and this will be a formula of T_{Ind}).

We need a $y^\star$ such that

$$S^{n'}R_{(S^m y^\star)} \leqslant S^n\tau < S^{n'}R_{(S^{m+1}y^\star)}.$$

Due to the monotonicity of $R_{(\cdot)}$, there exists such a $y^\star$ if and only if $S^{n'}R_{(S^m 0)} \leqslant S^n\tau$. If such a framing does not exist, then for any y, $S^n\tau < S^{n'}R_{(S^m y)}$ is true, so that there is no condition. Thus we may replace formula (7) by the equivalent formula

$$\begin{aligned}\exists y^\star \Big[&\Big(S^{n'}R_{(S^m y^\star)} \leqslant S^n\tau < S^{n'}R_{(S^{m+1}y^\star)} \vee S^n\tau < S^{n'}R_{(S^m 0)}\Big) \\ &\wedge\ \underbrace{\neg\Big(\exists\bar{y}\big(y^\star < y \vee S^n\tau < S^{n'}R_{(S^m 0)}\big) \wedge \psi(\bar{x},\bar{y},\bar{y}')\Big)}_{\psi'}\Big].\end{aligned} \tag{8}$$

Using the distributivity of $\wedge$ on $\vee$ and after a few elementary manipulations, we may write ψ' under the form

$$\begin{aligned}&\neg\Big(\exists\bar{y}\ \underbrace{(y^\star < y) \wedge \psi(\bar{x},\bar{y},\bar{y}')}_{\psi_1}\Big) \\ &\wedge\Big[\neg(S^n\tau < S^{n'}R_{(S^m 0)}) \vee \neg(\exists\bar{y}\psi(\bar{x},\bar{y},\bar{y}'))\Big].\end{aligned} \tag{9}$$

In this last formula, there is one less sub-formula $(\star)$ in ψ and in ψ_1 than in φ. If they do not contain any, then ψ and ψ_1 are equivalent to formulæ of T_{Ind} and (9) is open modulo T_{Ind}, as well as (8). Otherwise, we start this process again for ψ and ψ_1, and this process eventually stops since at each step we get rid of one formula $(\star)$.

Similarly, for a formula of the form

$$\neg\left(\exists \bar{y}\left(S^n\tau > S^{n'}R_{(S^m y)} \wedge \psi(\bar{x},\bar{y},\bar{y}')\right)\right),$$

arguing in the same way, we may replace it by

$$\begin{aligned}&\left(S^n\tau \leqslant S^{n'}R_{(S^m 0)}\right)\\ \vee\ &\exists y^\star\Big[S^{n'}R_{(S^m y^\star)} < S^n\tau \leqslant S^{n'}R_{(S^{m+1}y^\star)}\\ &\qquad \wedge \neg\big(\exists \bar{y}\ \underbrace{(y^\star > y \wedge \psi(\bar{x},\bar{y},\bar{y}'))}_{\psi'}\ \big)\Big].\end{aligned} \tag{10}$$

This time, if it is impossible to frame the term $S^n\tau$, then there is no y satisfying the bounding condition and the formula is true. Again, if (10) is not yet open modulo T_{Ind}, then we start this process again.

At the end, it just remains a formula $\exists \bar{y}^\star \theta(\bar{x},\bar{y}^\star,\bar{y}')$ where θ is open modulo T_{Ind}, and this proves point (b).

Third step: *$T_R^<$ is effectively existential modulo T_{Ind} when $\mathcal{P}$ is* e-a-p.

When $\mathcal{P}$ is decidable and *e-a-p*, the quantifier elimination with respect to the variables of X is effective, like in Theorem 10, as well as the above process. The theory $T_R^<$ is thus effectively existential modulo T_{Ind}.

Fourth step: *decidability algorithm for sentences of $T_R^<$ when $\mathcal{P}$ is decidable and* e-a-p *and T_{Ind} is decidable.*

Let us roughly explain the decidability algorithm of $T_R^<$ under the hypothesis that $\mathcal{P}$ is decidable and *e-a-p* and that T_{Ind} is decidable.

Given a sentence φ to be decided, we apply the result we just have shown to replace both φ and $\neg\varphi$ by equivalent formulæ $\exists \bar{y}\ \theta_\varphi(\bar{y})$ and $\exists \bar{y}'\ \theta_{\neg\varphi}(\bar{y}')$, where θ_φ and $\theta_{\neg\varphi}$ are open modulo T_{Ind}, $\bar{y} = (y_1, \ldots, y_n)$, and $\bar{y}' = (y'_1, \ldots, y'_m)$.

The decidability algorithm consists in alternatively enumerating all possible n-uples $\bar{a}$ and m-uples $\bar{a}'$ of natural numbers and to compute the truth values of $\theta_\varphi(\bar{a})$ and of $\theta_{\neg\varphi}(\bar{a}')$. These values may be effectively computed since any sub-formula of $\theta_\varphi(\bar{a})$ (resp. $\theta_{\neg\varphi}(\bar{a}')$) is now either open and variable-free or is a sentence of T_{Ind} — since the free variables have been replaced by the values of $\bar{a}$ (resp. $\bar{a}'$). In the first case, we only need to compute the truth value, and in the second case, the truth value of the sentence can be found as we supposed that T_{Ind} is decidable.

The algorithm stops as soon as it founds a n-uple $\bar{a}$ satisfying $\theta_\varphi(\bar{a})$ (φ is then true) or a m-uple $\bar{a}'$ satisfying $\theta_{\neg\varphi}(\bar{a}')$ (it is then false).

This concludes the proof of Theorem 12. □

Example 13. If $\mathcal{P}$ is a decidable and *e-a-p* system and if $\mathcal{Q}$ is a set of unary predicates such that the set $\mathcal{P}_R \cup \mathcal{Q}_R \cup \mathcal{I}_R$ forms a decidable and *e-a-p* system, then T_{Ind} is decidable by Theorem 10 (it even has effective QE) so that Theorem 12 applies. ⌟

Example 14. If R is a unary predicate such that the limit [11] of the differences of consecutive elements of R exists and is bounded, then Example 13 applies as soon as $\mathcal{P}$ is decidable and *e-a-p* and $\mathcal{Q}$ is a set of unary predicates such that $\mathcal{P}_R \cup \mathcal{Q}_R$ is decidable and *e-a-p*.

Proof. — The condition on R means that R is a ultimately periodic predicate.

Let $L = \lim_{y\in\omega} R_{(y+1)} - R_y$. This means that

$$\exists m \forall y > m : R_{(y+1)} - R_y = L.$$

Thus, beyond the m^{th} position, the symbols of the characteristic word of I_k will all be '0' when $k \geqslant L$, and will all be '1' when $k < L$. A word will be infinitely present in I_k if and only if either this word only contains '0' symbols and $k \geqslant L$ or only contains '1' and $k < L$. As a consequence, $\mathcal{P}_R \cup \mathcal{Q}_R \cup \mathcal{I}_R$ is *e-a-p* as soon as $\mathcal{P}_R \cup \mathcal{Q}_R$ is. ⌟

11. We also need to suppose that this limit is effectively calculable, that is, it is possible to determine the value of an m beyond which all differences are equal.

Example 15. If R is a unary predicate such that the limit [12] of the difference of consecutive elements of R 'exists' and is infinite, then Example 13 applies as soon as $\mathcal{P}$ is *e-a-p* and $\mathcal{Q}$ is a set of unary predicates such that $\mathcal{P}_R \cup \mathcal{Q}_R$ is decidable and *e-a-p*.

PROOF. — Indeed, we have

$$\forall k \exists m_k \forall y > m_k : R_{(y+1)} - R_y > k.$$

Thus the symbols of I_k all are '0' beyond the m_k^{th} position, and we may apply the same reasoning as in Example 14. ⌟

5. Decidability, Index Function, Addition and Almost-Periodicity

We now want to extend Presburger Arithmetic by considering structures of the form $\langle \omega; +, <, R, 0, 1 \rangle$, where R is an infinite predicate. We shall show that when R satisfies certain conditions (when R is 'sparse'— see Definition 18), the theory of some larger structure is existential modulo a sub-theory, just like in Section 4. In order to isolate certain variables appearing in inequalities, we shall have to temporarily extend our language, allowing sums with negative coefficients. These sums will be called 'operators'. Basically, the sparseness condition on R will allow us to re-use the 'framing trick' of Theorem 12.

5.1. Operators and Sparse Predicates: Definitions and Properties

Definition 16. *Let $R \subseteq \omega$ be a unary predicate, and let $R_{(\cdot)}$ denote the index function defined in the previous section. An* operator *(on R) is any expression A^R of the form $a_n R_{(S^n \cdot)} + \cdots + a_0 R_{(S^0 \cdot)}$ with $a_0, \ldots, a_n \in \mathbb{Z}$. We also call* operator *the associated function $A^R(\cdot) : \mathbb{N} \to \mathbb{Z} : y \mapsto a_n R_{(S^n y)} + \cdots + a_0 R_{(S^0 y)}$.*

12. The previous remark applies here as well...

Definition 17. *We say an operator A^R is*

- null *(on R) if its image is $\{0\}$,*
- positive *(on R) if the set $\{y \in \mathbb{N} \mid A^R y \leqslant 0\}$ is finite, and*
- negative *(on R) if the set $\{y \in \mathbb{N} \mid A^R y \geqslant 0\}$ is finite.*

This will be respectively denoted $A^R =_R 0$, $A^R >_R 0$ and $A^R <_R 0$.

Finally, two operators A^R and B^R are said to be equal *on R if their difference $A^R - B^R$, defined in the obvious way, is null on R.*

Definition 18. *A unary predicate R is* sparse *if for any operator A^R, the following two conditions are satisfied:*

P1. $A^R =_R 0$ or $A^R <_R 0$ or $A^R >_R 0$,

P2. If $A^R >_R 0$, then there exists a natural number Δ such that $A^R(S^\Delta y) - R_y > 0$ for all $y \in \mathbb{N}$.

A predicate R is said to be effectively sparse *whenever it is sparse and there exists an algorithm that decides, for any given operator A^R, which condition of P1 is satisfied, and, in case $A^R >_R 0$, that gives a natural number Δ satisfying condition P2.*

Examples of sparse predicates are given in Appendix B. Here are a few consequences of Definition 18.

Lemma 19. *Let R be a sparse predicate and let A^R be a positive operator on R. Then there exists a natural number Λ such that $A^R(S^\Lambda \cdot)$ is a* strictly increasing *positive function.*

Proof. — Using Definition 18.P2, we know that there exists a natural number Δ such that for any y, $A^R(S^\Delta y)$ is strictly positive and lowered by R_y.

Consider the operator $B^R = A^R(S\cdot) - A^R(\cdot)$. Since R is sparse, we have either $B^R <_R 0$ or $B^R =_R 0$ or $B^R >_R 0$. Let us show the first two cases are impossible.

Suppose $B^R =_R 0$. Then $A^R(\cdot)$ is constant, and so is $A^R(S^\Delta \cdot)$. But this contradicts the fact this last function is lowered by a strictly increasing integer function.

Similarly, suppose $B^R <_R 0$. So the set of non negative values of $A^R(S\cdot) - A^R(\cdot)$ is finite, and thus there exists a natural number N such that $A^R(S^N \cdot)$ is a strictly decreasing function. But the integer function $A^R(S^{\max\{\Delta,N\}} \cdot)$ is simultaneously strictly decreasing and lowered by a strictly increasing integer function. This gives a contradiction.

Now, $B^R >_R 0$ and this implies that there exists a natural number Δ' such that $B^R(S^{\Delta'} y) > R_y \geqslant 0$ for any y, i.e. $A^R(S^{\Delta'} y)$ is strictly increasing for any y. It suffices to take $\Lambda = \max\{\Delta, \Delta'\}$. Notice that Λ is effectively obtained whenever R is effectively sparse.

Remark 20. Re-using the 'growth operator' $A^R(S\cdot) - A^R(\cdot)$, we show that any positive operator is a 'superpolynomial' function: for any operator $A^R >_R 0$ and any polynomial $p(y) : \mathbb{N} \to \mathbb{Z}$, there exists a Δ such that for all $y > \Delta$, $A^R(y) > p(y)$. Let us show this by induction on the degree of p:

- for $\deg p = 0$, it is an immediate consequence of the lemma.
- for $\deg p = n + 1$, since the polynomial $p(y + 1) - p(y)$ has degree n and since $A^R(S\cdot) - A^R >_R 0$, using the induction hypothesis, there exists a Δ_n such that the growth of A^R is strictly bigger than that of p. Let $c = p(\Delta_n) - A^R(\Delta_n)$. We see that it suffices to take $\Delta = \Delta_n + \max\{0, c\}$. □

Considering the function $R_{(\cdot)}$ as an operator (which is trivially $>_R 0$), we get the result that any polynomial eventually lowers the function $R_{(\cdot)}$. ⌟

5.2. Decidability of the Addition

As before, we want to study the theory of a structure built on two domains. This time, it is the theory T_R^+ of

$$\langle \omega_1, \omega_2; +_1, <_1, 0_1, 1_1, \mathcal{E}, S_2, <_2, 0_2, R_{(\cdot)}, \mathcal{Q}_R \rangle$$

where $\mathcal{E} = \{\cdot \equiv_m c \mid c, m \in \mathbb{N}, c < m \neq 0, 1\}$ denotes the set of congruence predicates on ω_1 and $\mathcal{Q}_R$ is as in Section 4.

The variables whose values are in ω_1 are elements of the set

$$X = \{x, x', x_1, \ldots\},$$

and those whose values are in ω_2 will be elements of

$$Y = \{y, y', y_1, \ldots, z, z', z_1, \ldots\}.$$

Notice that the successor function is trivially definable in ω_1 by '$+_1 1_1$' so that it is useless to add S_1 to our language. However, we cannot define S_2 by a quantifier-free formula.

We also consider the sub-theory T_{Ind} of $\langle \omega_2; S_2, <_2, 0_2, \mathcal{E}_R, \mathcal{Q}_R, \mathcal{I}_R \rangle$ which is defined like in Theorem 12 (see page 24, where $\mathcal{P}$ has been replaced by $\mathcal{E}$, and both $\mathcal{Q}_R$ and $\mathcal{I}_R$ are as before).

Theorem 21. *Let R be a sparse predicate and let $\mathcal{Q}$ be a set of predicates defined on R. Then the theory T_R^+ is existential modulo T_{Ind}. Moreover, it is decidable if and only if R is both decidable and effectively sparse and T_{Ind} is decidable.*

PROOF. — The proof of this theorem follows the same structure as that of Theorem 12: we first show that we have QE with respect to quantifications on X; then we show that any formula $\neg(\exists \bar{y}\ \varphi(\bar{x}, \bar{y}, \bar{y}'))$, where φ is open modulo T_{Ind}, is equivalent to an existential formula (modulo T_{Ind}). The decidability algorithm of T_R^+ is similar to that of Theorem 12 (see page 30). Finally, for the 'only if' part of the proof, any sentence of T_{Ind} is a sentence of T_R^+ (and thus is decidable) and since R is sparse, its effectiveness will be obtained using sentences of T_R^+ (see the end of the proof).

First step: *QE with respect to quantifications on X.*

Let φ be an open formula modulo T_{Ind}. Let us show we are able to get rid of the quantifier of $\exists x\ \varphi(x, \bar{x}, \bar{y})$. The arguments we use are the same as in Presburger's proof.

The atomic formulæ containing the variable x are comparisons and congruences. We would like to apply the same reasoning as in Theorem 10 (with the congruences as system of almost-periodic predicates). However, the terms may be more complicated: the coefficient of x can be any natural number and the congruences may mix x and other variables.

Without loss of generality, we may suppose that φ is a conjunction of atomic formulæ, each of them containing the variable x. We convert equalities into inequalities and in each of them, we simplify the occurrences of x in order to keep x just on one side. We replace negated congruences by a (finite) disjunction of congruences

$$\tau(x,\bar{x},\bar{y}) \not\equiv_m c \Leftrightarrow \bigvee_{\substack{i<m \\ i\neq c}} \tau(x,\bar{x},\bar{y}) \equiv_m i,$$

and then split any congruence into a part containing x and a part containing the other variables (this last part being then moved out of the action of the quantifier):

$$nx + \tau(\bar{x},\bar{y}) \equiv_m c \;\Leftrightarrow\; \bigvee_{0\leqslant i<m} (nx \equiv_m i \;\wedge\; \tau(\bar{x},\bar{y}) \equiv_m c - i)\,.$$

Using the abbreviation 'nx' for '$\underbrace{x + \cdots + x}_{n\text{times}}$', the remaining atomic formulæ are either inequalities $nx + \tau_1(\bar{x},\bar{y}) \lesseqgtr \tau_2(\bar{x},\bar{y})$ or congruences $nx \equiv_m c$, where τ_1 and τ_2 are terms of T_R^+ and $n, m, c \in \mathbb{N}$ with $c < m \neq 0, 1$.

We extend our language in order to accept sums with negative coefficients. Denoting by $\sigma(\bar{x},\bar{y})$ the new terms, this allows us to restrict ourself to inequalities $nx \lesseqgtr \sigma(\bar{x},\bar{y})$ and congruences.

For any natural number $k > 0$, we have the equivalences

$$\begin{aligned} nx \lesseqgtr \sigma(\bar{x},\bar{y}) \quad &\Leftrightarrow \quad k\cdot nx \lesseqgtr k\cdot\sigma(\bar{x},\bar{y}) \\ nx \equiv_m c \quad &\Leftrightarrow \quad k\cdot nx \equiv_{k\cdot m} k\cdot c \end{aligned}$$

(the 'multiplication' by k has still to be considered as a repeated sum). For any such atomic formula ψ, let n_ψ be the coefficient of x. If we let l be the lowest common multiple of the n_ψ's, we can transform all these formulæ ψ so that the coefficients of x are always l. Denote by $\psi(lx,\bar{x},\bar{y})$ any of these new formulæ. It is clear there exists a solution x to the conjunction of the $\psi(lx,\bar{x},\bar{y})$ if and only if there exists a x' multiple of l satisfying the conjunction of the $\psi(x',\bar{x},\bar{y})$. So we may replace $\exists x\; \varphi(x,\bar{x},\bar{y})$ by the equivalent formula $\exists x\,(x \equiv_l 0 \wedge \bigwedge \psi(x,\bar{x},\bar{y}))$ in which x only appears with coefficients 1. Writing it using a canonical form, we have reduced φ to the form

$$\exists x \left(\bigwedge_{a=1\ldots A} \sigma_{1a} < x \;\; \wedge \bigwedge_{b=1\ldots B} x < \sigma_{2b} \;\; \wedge \bigwedge_{d=1\ldots D} x \equiv_{m_d} c_d \right). \tag{11}$$

We are exactly in the same situation as in Theorem 10: we have to satisfy lower and upper bounds, as well as some periodic predicates. Applying the same reasoning, we finally get a formula equivalent to (11) of the form (3) (on page 19).

Once again, we are able to transform extended terms back to classical terms of T_R^+ using the equivalences

$$\begin{aligned} \tau_1 - \tau_2 < \tau_1' - \tau_2' \quad &\Leftrightarrow \quad \tau_1 + \tau_2' < \tau_1' + \tau_2 \\ \tau_1 - \tau_2 \equiv_m 0 \quad &\Leftrightarrow \quad \tau_1 + (m-1) \cdot \tau_2 \equiv_m 0. \end{aligned}$$

This concludes the elimination of the quantifier $\exists x$.

Second step: *any formula $\neg\,(\exists \bar{y}\ \varphi(\bar{x},\bar{y},\bar{y}'))$, where φ is open modulo T_{Ind}, is equivalent to an existential formula (modulo T_{Ind}).*

Now, we show how to transform the formula $\neg\,(\exists \bar{y}\ \varphi(\bar{x},\bar{y},\bar{y}'))$ into an existential formula modulo T_{Ind}. Like in Theorem 12, we shall frame annoying terms in such a way that the quantifier $\neg\exists\bar{y}$ only acts on a formula of T_{Ind}. Again, we extend our language in order to accept sums with negative coefficients. We first need a few lemmas...

Notation. Let $y, \bar{z} \in Y$ be variables and let Δ be a natural number. We denote by $M_\Delta(\cdot)$ the following formula (according to its arity):

$$\begin{aligned} M_\Delta(y) \;&\Leftrightarrow\; y > \Delta \\ M_\Delta(y,\bar{z}) \;&\Leftrightarrow\; \bigwedge_{z\in\bar{z}} y > (S^\Delta z). \end{aligned}$$ ⌟

Lemma 22. *Let R be a sparse predicate and let A^R be a positive operator on R. Let $\tau(\bar{z})$ be a term of T_R^+ whose variables differ from y. Then there exists a natural number Δ such that*

$$\forall y, \bar{z}\left(M_\Delta(y,\bar{z}) \Rightarrow A^R y > |\tau(\bar{z})|\right).$$

Proof. — Let c be the sum of the absolute values of the coefficients appearing in the term $\tau(\bar{z})$ (including the constants of ω_1) and let d be the maximum exponent of the 'S_2' of $\tau(\bar{z})$ (with $d = 0$ if there is no S_2, for instance when $\bar{z} = \varnothing$). If we set $z' = \max\{z \in \bar{z}\}$, we see that $c \cdot R_{(S^d z')} \geqslant |\tau(\bar{z})|$. This shows that it suffices to prove the existence of a Δ such that $A^R(S^\Delta \cdot)$ is strictly increasing and such that $A^R(S^\Delta z) > c \cdot R_{(S^d z)}$ for any z (then, when $M_\Delta(y, \bar{z})$ is satisfied, we have that $A^R y > A^R(S^\Delta z') > c \cdot R_{(S^d z')} \geqslant |\tau(\bar{z})|$).

Lemma 19 gives us a Λ such that any $\Delta > \Lambda$ satisfies the first condition. It remains to show that Δ can be taken so that the second condition is fulfilled.

It suffices to show the result in the case $d = 0$: for $d \neq 0$, we suppose $\Delta > d$ and replace the condition $A^R(S^\Delta z) > c \cdot R_{(S^d z)}$ by $A^R(S^{\Delta - d} z) > c \cdot R_z$; this leads back to the case $d = 0$.

Using Definition 18.P2, we find a Δ_1 such that $A^R(S^{\Delta_1} z) - R_z > 0$ for any z. Since the operator $A^R(S^{\Delta_1} \cdot) - R_{(\cdot)}$ is positive, there exists a Δ_2 such that $\left[A^R(S^{\Delta_1} S^{\Delta_2} z) - R_{(S^{\Delta_2} z)}\right] - R_z > 0$, and the monotonicity of $R_{(\cdot)}$ implies that $A^R(S^{\Delta_1 + \Delta_2} z) - 2R_z > 0$. Continuing in the same way, we find $\Delta_1, \ldots, \Delta_c$ such that $A^R(S^{\Delta_1 + \cdots + \Delta_c} z) - cR_z > 0$. It suffices to take $\Delta = \Delta_1 + \cdots + \Delta_c$. Moreover, Δ can be found effectively when R is effectively sparse. □

Corollary 23. *Let A^R be a positive operator on R. Applied to the operator $A^R(S\cdot) - A^R(\cdot)$ which is also positive on R by Lemma 19, the previous lemma shows that for any given term τ, there exists a Δ such that the images by A^R of two consecutive natural numbers greater (at least by Δ) than the values of the variables of $\bar{z}$ are apart from each other at least by $|\tau(\bar{z})|$. This fact will be used later.*

The following lemma will allow inductions on the amount of variables of Y: it replaces any formula of T_R^+ by a disjunction of formulæ containing less variables of Y and of formulæ in which one variable of Y bounds the others.

Notation. Let $\bar{y}$ be the sequence $(y_1, \ldots, y_n)$ of variables of Y. For any $1 \leqslant i \leqslant n$, we denote by $\bar{y}_{\hat{i}}$ the reduced sequence $(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$. ⌟

Lemma 24. *Let $\bar{y} = (y_1, \ldots, y_n)$, let $\varphi(\bar{y})$ be an open formula* [13] *modulo T_{Ind} and let $\Delta_1, \ldots, \Delta_n$ be natural numbers. Then we may construct a formula $\psi(\bar{z})$ (open modulo T_{Ind}) with $\bar{z} = (z_1, \ldots, z_{n-1})$ and such that*

$$\exists \bar{y}\ \varphi(\bar{y}) \quad \Leftrightarrow \bigvee_{1 \leqslant i \leqslant n} \exists \bar{y}\ \big(M_{\Delta_i}(y_i, \bar{y}_{\not i}) \wedge \varphi(\bar{y})\big) \ \vee\ \exists \bar{z}\ \psi(\bar{z}).$$

Proof. — Suppose that $\exists \bar{y} \varphi(\bar{y})$ is satisfied but that the disjunction on $1 \leqslant i \leqslant n$ is false. Then for each i, $M_{\Delta_i}(y_i, \bar{y}_{\not i})$ is false, so that there exists a $j_i \neq i$ with $y_i \leqslant S^{\Delta_i} y_{j_i}$. By a simple counting argument, we see there must exist a 'cycle'. Suppose this cycle is $y_1 \leqslant S^{\Delta_1} y_2$, $y_2 \leqslant S^{\Delta_2} y_3, \ldots, y_k \leqslant S^{\Delta_k} y_1$ with $2 \leqslant k \leqslant n$. This implies $y_1 \leqslant S^{\Delta_1 + \cdots + \Delta_{k-1}} y_k$ and $y_k \leqslant S^{\Delta_k} y_1$. Thus there exists some $l \leqslant \Delta_1 + \cdots + \Delta_k$ such that $y_1 = S^l y_k$ or $y_k = S^l y_1$. In the general case, there exist some values of i and j such that $y_i = S^l y_j$ for some $0 \leqslant l \leqslant \Delta_1 + \cdots + \Delta_n$. Let $\Delta = \Delta_1 + \cdots + \Delta_n$. In order to satisfy the implication "$\Rightarrow$", it suffices to take as formula '$\exists \bar{z} \psi(\bar{z})$' the prenex form of

$$\bigvee_{\substack{0 \leqslant i \neq j \leqslant n \\ 0 \leqslant l \leqslant \Delta}} \exists \underbrace{\bar{y}_{\not i}}_{\bar{z}}\ \varphi(y_1, \ldots, y_{i-1}, S^l y_j, y_{i+1}, \ldots, y_n).$$

It is obvious that the implication "$\Leftarrow$" is also satisfied for this choice of ψ. □

Now, let us transform the formula $\neg(\exists \bar{y} \varphi)$ into an existential formula modulo T_{Ind}, when φ is open modulo T_{Ind}. We proceed by induction on the amount of variables of $\bar{y} = (y_1, \ldots, y_n)$. For $n = 0$, there is nothing to do. We show how to get from n to $n + 1$.

By Lemma 24, it suffices to find natural numbers $\Delta_1, \ldots, \Delta_n$ such that the formula

$$\bigwedge_{1 \leqslant i \leqslant n} \neg \big(\exists \bar{y}\ \big(M_{\Delta_i}(y_i, \bar{y}_{\not i}) \wedge \varphi(\bar{x}, \bar{y}, \bar{y}')\big)\big) \quad \wedge \quad \neg(\exists \bar{z}\ \psi(\bar{x}, \bar{y}', \bar{z}))$$

is equivalent to an open formula modulo T_{Ind}. Using the induction hypothesis, we may get rid of the $\neg(\exists \bar{z} \psi(\bar{z}))$ since $\bar{z}$ contains one less variable

13. For clearness reasons in this lemma, we shall not indicate unquantified variables... We should have written $\varphi(\bar{x}, \bar{y}, \bar{y}')$.

than $\bar{y}$. It remains to prove the following statement (in which we renamed the variables):

Lemma 25. *Given a variable y, a sequence of variables $\bar{z}$ and a formula $\varphi(\bar{x}, y, \bar{z}, \bar{y}')$ (open modulo T_{Ind}), it is possible to find a natural number Δ such that the formula*

$$\neg\,(\exists y \exists \bar{z}\,(M_\Delta(y, \bar{z}) \;\wedge\; \varphi(\bar{x}, y, \bar{z}, \bar{y}'))) \tag{12}$$

is equivalent to a formula of the form

$$\exists \tilde{y}\,(\psi_1(\bar{x}, \tilde{y}, \bar{y}') \;\wedge\; \neg\,(\exists \bar{z}\; \psi_2(\bar{x}, \tilde{y}, \bar{y}', \bar{z}))) \tag{13}$$

where $\tilde{y}$ is a sequence of new variables and ψ_1, ψ_2 are open modulo T_{Ind}.

Proof. — We proceed as in Theorem 12.

The two kind of sub-formulæ we have to treat are congruences and inequalities between terms containing y, $\bar{z}$, variables of X and constants of ω_1. Using extended terms, these can be written $A^R y + \tau(\bar{z}) + \sigma(\bar{x}) \equiv_m c$ and $A^R y + \tau(\bar{z}) \lesseqgtr \sigma(\bar{x})$ with $A^R >_R 0$.

Once again, we are going to proceed with an induction on the amount of sub-formulæ which are not open modulo T_{Ind} and contain y. Each step will be feasible if we satisfy some M_Δ-condition, and the maximum of all these Δ's will be the one we need in order to transform (12) into (13). Then we shall be able — with a possible renaming of linked variables — to move all the new quantifiers to the beginning of the formula.

Of course, if there is no problematic formula, we may commute the $\exists y$ with the $\exists \bar{z}$. Then, since any atomic formula containing y is open modulo T_{Ind}, we may move down the quantifier $\exists y$ on these sub-formulæ, and they remain open modulo T_{Ind}. So we immediately get a formula $\neg\,(\exists \bar{z}\; \psi(\bar{x}, \bar{y}', \bar{z}))$ on which we may apply our induction hypothesis. (In this case, we do not even need to use a M_Δ-condition.)

We now get rid of the congruence in

$$\neg\left(\exists y, \bar{z}\left(A^R y + \tau(\bar{z}) + \sigma(\bar{x}) \equiv_m c \;\wedge\; \varphi(\bar{x}, y, \bar{z}, \bar{y}')\right)\right)$$

(in order to simplify the notation, we move the formula $M_\Delta(y, \bar{z})$ into φ).

Let $A^R = a_n R_{(S^n \cdot)} + \cdots + a_0 R_{(S^0 \cdot)}$. We may replace the congruence $A^R y + \tau(\bar{z}) + \sigma(\bar{x}) \equiv_m c$ by the disjunction

$$\bigvee_{\substack{i_0 = 0 \ldots m-1 \\ \vdots \\ i_n = 0 \ldots m-1}} \left(\bigwedge_{j=0\ldots n} R_{(S^j y)} \equiv_m i_j \ \wedge \sum_{j=0\ldots n} a_j \cdot i_j + \tau(\bar{z}) + \sigma(\bar{x}) \equiv_m c \right).$$

Since y only appears in formulæ of T_{Ind} in this last expression, we need to eliminate one less problematic formula. This gives us the induction for congruences. (Again, we did not need to use the M_Δ-condition.)

Finally, we treat the inequality in

$$\neg \left(\exists y, \bar{z} \left(A^R y + \tau(\bar{z}) \lessgtr \sigma(\bar{x}) \wedge M_\Delta(y, \bar{z}) \wedge \varphi(\bar{x}, y, \bar{z}, \bar{y}') \right) \right) \tag{14}$$

(we shall precise later whether it is a '$<$' or a '$>$').

We want to frame the term $\sigma(\bar{x})$ (like in Theorem 12) using the image by A^R of a new variable $y^\star$. This will be possible if the minimal value m_A of the images of the operator A^R is lower or equal to $\sigma(\bar{x})$ (remember $A^R >_R 0$). The value of m_A can be computed whenever R is effectively sparse, since, by Lemma 19, we may compute a natural number Δ_1 such that $A^R(y)$ is strictly increasing for any $y > \Delta_1$.

Let $y^\star$ be a new variable satisfying

$$A^R(Sy^\star) > \sigma(\bar{x}) \geqslant A^R y^\star \ \vee \ \sigma(\bar{x}) < m_A. \tag{15}$$

The first problem we encounter is that the function $A^R(\cdot)$ is not monotonic. Thus, we might have a $y^\star$ satisfying (15) but which would not be extremal with respect to this property (see Figure 1); consequently, we would not be allowed to deduce from it implications like "$\forall y > y^\star, \forall y' > y^\star \left(y > y' \Rightarrow A^R y > A^R y' \right)$".

This problem can be avoided by setting additional conditions on Δ.

Let M_A be the maximal value of the images of A^R before it gets strictly increasing. This can be computed when R is effectively sparse (like m_A). Since $A^R(y)$ is strictly increasing for $y > \Delta_1$, there exists some $\Delta_2 > \Delta_1$ such that $A^R(y) > M_A$ when $y > \Delta_2$.

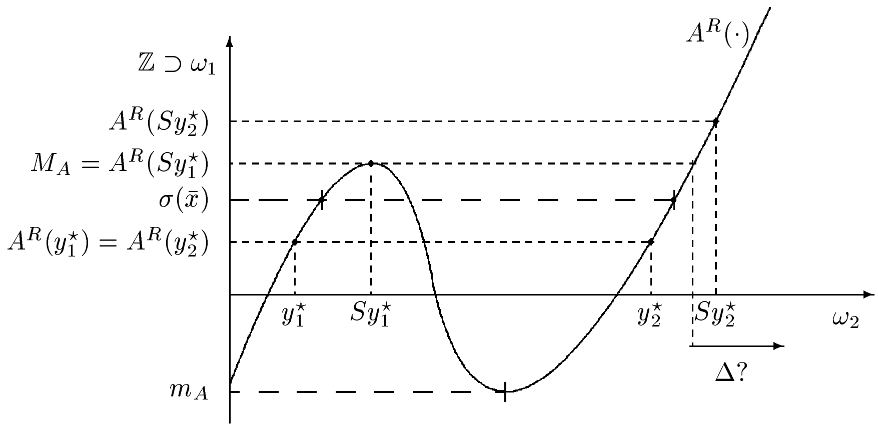


Figure 1: $y_1^\star$ and $y_2^\star$ satisfy (15).

Using Corollary 23, since $A^R(S\cdot) - A^R(\cdot)$ is positive on R, there exists a Δ_3 such that

$$\forall y, \bar{z} \left(M_{\Delta_3}(y, \bar{z}) \Rightarrow A^R(Sy) - A^R y > |\tau(\bar{z})| \right).$$

Let $\Delta = \max\{\Delta_2, \Delta_3\} + 1$. This ensures that for any $y > \Delta$ and any y',

$$y < y' \overset{\Delta_1}{\Rightarrow} A^R y < A^R y' \overset{\Delta_3}{\Rightarrow} A^R y + \tau(\bar{z}) < A^R y',$$

$$y > y' \overset{\Delta_2}{\Rightarrow} A^R y > A^R y' \overset{\Delta_3+1}{\Rightarrow} A^R y + \tau(\bar{z}) > A^R y'.$$

In particular, if we replace y' by $y^\star$ or $Sy^\star$, we get

$$y < y^\star \Rightarrow A^R y + \tau(\bar{z}) < A^R y^\star,$$

$$y > Sy^\star \Rightarrow A^R y + \tau(\bar{z}) > A^R(Sy^\star).$$

If $\sigma(\bar{x}) \geqslant m_A$, then the images of $y^\star$ 'frame' $\sigma(\bar{x})$ and these formulæ become

$$y < y^\star \Rightarrow A^R y + \tau(\bar{z}) < \sigma(\bar{x}),$$

$$y > Sy^\star \Rightarrow A^R y + \tau(\bar{z}) > \sigma(\bar{x}).$$

In this case, it remains to check whether $y = y^\star$ or $y = Sy^\star$ is a possible solution for the existence.

On the contrary, if $\sigma(\bar{x}) < m_A$, then whatever is the value of $y^\star$, we always have that $A^R y^\star > \sigma(\bar{x})$. In particular, $A^R 0 > \sigma(\bar{x})$ and thanks to the choice of Δ_3, $A^R y + \tau(\bar{z}) \geqslant A^R y - |\tau(\bar{z})| > \sigma(\bar{x})$ for any $y \geqslant 1$. Consequently, $A^R y + \tau(\bar{z}) > \sigma(\bar{x})$ is satisfied for any y such that $M_\Delta(y, \bar{z})$ is true (since $\Delta \geqslant 1$).

Summing up:

1) Suppose the inequality we work on is '$<$'. Then the formula

$$\neg\left(\exists y, \bar{z}\left(A^R y + \tau(\bar{z}) < \sigma(\bar{x}) \wedge M_\Delta(y, \bar{z}) \wedge \varphi(\bar{x}, y, \bar{z}, \bar{y}')\right)\right)$$

is equivalent to the formula

$$\begin{aligned}
&(\sigma(\bar{x}) < m_A) \vee \exists y^\star \Bigg[A^R y^\star \leqslant \sigma(\bar{x}) < A^R(Sy^\star) \\
&\quad \left.\wedge \neg\exists y, \bar{z}\,(M_\Delta(y, \bar{z}) \wedge y < y^\star \wedge \varphi(\bar{x}, y, \bar{z}, \bar{y}'))\right\}\psi_1 \\
&\quad \left.\begin{aligned}&\wedge \neg\exists \bar{z}\big[A^R y^\star + \tau(\bar{z}) < \sigma(\bar{x}) \wedge M_\Delta(y^\star, \bar{z}) \wedge \varphi(\bar{x}, y^\star, \bar{z}, \bar{y}') \\
&\qquad \vee A^R(Sy^\star) + \tau(\bar{z}) < \sigma(\bar{x}) \wedge M_\Delta(Sy^\star, \bar{z}) \wedge \varphi(\bar{x}, Sy^\star, \bar{z}, \bar{y}')\big]\Bigg]\end{aligned}\right\}\psi_2.
\end{aligned}$$

2) On the other hand, if we work on '$>$', then the formula

$$\neg\left(\exists y, \bar{z}\left(A^R y + \tau(\bar{z}) > \sigma(\bar{x}) \wedge M_\Delta(y, \bar{z}) \wedge \varphi(\bar{x}, y, \bar{z}, \bar{y}')\right)\right)$$

is equivalent to

$$\begin{aligned}
&\exists y^\star \Big[\left(A^R y^\star \leqslant \sigma(\bar{x}) < A^R(Sy^\star) \vee \sigma(\bar{x}) < m_A\right) \\
&\quad \left.\wedge \neg\exists y, \bar{z}\,(M_\Delta(y, \bar{z}) \wedge y > Sy^\star \wedge \varphi(\bar{x}, y, \bar{z}, \bar{y}'))\right\}\psi_1 \\
&\quad \left.\begin{aligned}&\wedge \neg\exists \bar{z}\big[A^R y^\star + \tau(\bar{z}) < \sigma(\bar{x}) \wedge M_\Delta(y^\star, \bar{z}) \wedge \varphi(\bar{x}, y^\star, \bar{z}, \bar{y}') \\
&\qquad \vee A^R(Sy^\star) + \tau(\bar{z}) < \sigma(\bar{x}) \wedge M_\Delta(Sy^\star, \bar{z}) \wedge \varphi(\bar{x}, Sy^\star, \bar{z}, \bar{y}')\big]\Big]\end{aligned}\right\}\psi_2.
\end{aligned}$$

In each case, the formula ψ_1 contains one less annoying formula and the formula ψ_2 does not contain the variable y. We may thus finish the work by induction. Notice that a value of Δ can be effectively computed when R is effectively sparse. Using the maximum of the Δ's we found on each step of the induction, we finally proved Lemma 25.

Third step: *if R is sparse and T_R^+ is decidable, then R is effectively sparse.*

Like in Theorem 10, this will come from the fact that the conditions can be expressed by sentences of our language.

Let A^R be an operator $a_n R_{(S^n\cdot)} + \cdots + a_0 R_{(S^0\cdot)}$ with $a_i \in \mathbb{Z}$ $(i = 0 \dots n)$. We know $A^R \lesseqqgtr_R 0$. Here is the algorithm that allows us to find out which possibility is the good one.

Let J^+ denote the set of i's such that $a_i \geqslant 0$ and J^- the set of i's such that $a_i < 0$. We have

$$A^R(y) \lesseqqgtr 0 \Leftrightarrow \sum_{i \in J^+} a_i R_{(S^i y)} \lesseqqgtr \sum_{i \in J^-} |a_i| R_{(S^i y)}.$$

We have $A^R =_R 0$ if and only if for any y, $A^R(y) = 0$. This can be expressed by the sentence of T_R^+

$$\forall y \sum_{i \in J^+} a_i R_{(S^i y)} = \sum_{i \in J^-} |a_i| R_{(S^i y)} \tag{16}$$

and the hypothesis tells us that it can be decided. Similarly, $A^R >_R 0$ if and only if finitely many y are such that $A^R(y) \leqslant 0$. This can be expressed by the sentence of T_R^+

$$\exists y \forall y' \left(y' > y \Rightarrow \sum_{i \in J^+} a_i R_{(S^i y')} > \sum_{i \in J^-} |a_i| R_{(S^i y')} \right) \tag{17}$$

which is also decidable. Finally, $A^R <_R 0$ if and only if the previous two sentences are false. Thus we have an algorithm deciding 18.P1.

Suppose $A^R >_R 0$. The deciding algorithm for 18.P2 works in the following way: it decides the following sentences for an increasing sequence of Δ

$$\forall y \sum_{i \in J^+} a_i R_{(S^{\Delta+i} y)} > \sum_{i \in J^-} |a_i| R_{(S^{\Delta+i} y)} + R_{(y)} \tag{18}$$

and gives the value of Δ as soon as such a sentence is true. It will eventually stop since we supposed R is sparse and $A^R >_R 0$.

This concludes the proof of Theorem 21. □

6. Back to $\langle \omega; <, R\rangle$ and $\langle \omega; +, R\rangle$

6.1. Decidability of the Theories of $\langle \omega; <, R\rangle$ and $\langle \omega; +, R\rangle$ for certain R

We now want to study the theories of the structures $\langle \omega; S, <, R, 0\rangle$ and $\langle \omega; +, R, 0, 1\rangle$ (where R only plays a role of predicate). These will be abbreviated by $\langle \omega; <, R\rangle$ and $\langle \omega; +, R\rangle$.

In order to use our previous results, we shall naturally look at the set $\mathcal{R}$ of sparse predicates on which congruences are almost-periodic, that is such that $\mathcal{E}_R$ is *a-p*.

When $R \in \mathcal{R}$, we may see the theory of $\langle \omega; +, R\rangle$ as a sub-theory of the theory T_R^+ (see its definition on page 34, with $\mathcal{Q}_R = \varnothing$). This gives the following result:

Corollary 26. *Let $R \in \mathcal{R}$. The theory of $\langle \omega; +, R\rangle$ is decidable if and only if R is effectively sparse* (e-s) *and $\mathcal{E}_R$ is* e-a-p.

It is then interesting to remember Remark 20 that shows that a sparse predicate is 'superpolynomial', and to compare this corollary with Büchi's famous theorem [2] that says that $\langle \omega; +, R\rangle$ is undecidable whenever R is polynomial.

Proof. — First consider a predicate $R \in \mathcal{R}$ such that R is *e-s* and $\mathcal{E}_R$ is *e-a-p*. Since R is *e-s*, we may find a natural number Δ such that for any y, $R_{S^{\Delta+1}y} - R_{S^{\Delta}y} > y$. In particular, $\lim_{y\to\infty} R_{Sy} - R_y = \infty$. By remark [14] 15 and Theorem 10, the theory T_{Ind} is decidable. Theorem 21 now applies and shows T_R^+ is decidable. Since any sentence of $\langle \omega; +, R\rangle$ is a sentence of T_R^+, we get the result we wanted.

14. In this case, the 'effective limit' condition mentioned in Footnote 11 is satisfied.

In order to show the converse implication, it suffices to show that formulæ (4) on page 20 (expressing the effectiveness of the almost-periodicity of $\mathcal{E}_R$) and (16–18) on page 44 (expressing the effectiveness of the sparseness of R) are expressible in $\langle\omega;+,R\rangle$. Notice that the variable 'x' appearing in formula (4) should be a variable $y \in Y$ in T_R^+, and that the predicates 'P_k' are predicates of $\mathcal{E}_R$ or of $\mathcal{I}_R$.

The predicate R is clearly decidable (since for any natural m, '$R(S^m0)$' is a sentence of $\langle\omega;+,R\rangle$), so for each natural number m, we may effectively find the m^{th} element of R. We may add to our language a new constant symbol 'R_m' for each m.

Moreover, congruences are definable in $\langle\omega;+,R\rangle$ by

$$\tau(\bar{x}) \equiv_m c \quad \Leftrightarrow \quad \exists x'(\tau(\bar{x}) = \underbrace{x' + \cdots + x'}_{m\text{times}} + c).$$

Now, we show that any formula of T_R^+ whose free variables are variables of X may be written as an equivalent [15] formula of $\langle\omega;+,R\rangle$. This 'partial interdefinability' will suffice in order to convert (4) and (16–18) to formulæ of $\langle\omega;+,R\rangle$.

By the usual induction argument on the construction of formulæ, it suffices to show how to convert

$$\exists y \bigwedge_{i=1\dots n} \varphi_i(\bar{x}, y, \bar{y})$$

(where all the quantified variables of the φ_i's are in X) into formulæ not containing y anymore (although they may contain many new variables of X). The atomic formulæ [16] of φ_i that contain y are of the following types:

$$S^n y \underset{(1)}{\lesseqgtr} S^{n'} y', \quad A^R y + B^R \bar{y} + \tau(\bar{x}) \underset{(2)}{\lesseqgtr} 0, A^R y + B^R \bar{y} + \tau(\bar{x}) \underset{(3)}{\equiv_m} c. \quad (19)$$

We use the fact that y represents the index of an element of R, replace the quantification $\exists y(\cdots)$ by a quantification on a new variable x and add the condition $x \in R$, i.e. we get a formula of the form $\exists x\,(R(x) \wedge \ldots)$.

15. In the sense that the subset of ω_1 defined by the first formula coincides with the subset of ω defined by the second one.
16. Once again, we use extended formulæ with operators, etc.

The replacement of formulæ (1), (2) and (3) by equivalent formulæ with x is possible thanks to the fact the 'successor in R' operation is definable in $\langle\omega;+,R\rangle$: suppose y corresponds to '$x_0 \in R$'; then Sy will correspond to the x_1 which satisfies

$$x_0 < x_1 \;\wedge\; R(x_1) \;\wedge\; \forall x' \,(x_0 < x' < x_1 \Rightarrow \neg\,(R(x'))) \,.$$

Continuing that way, we easily convert $S^n y$ to a variable of X for any natural number n.

Consequently, wishing to replace $\exists y(\ldots)$ by $\exists x\,(R(x) \wedge \ldots)$, formula (1) becomes

$$\begin{aligned}\exists x_0 < \ldots < x_n \Big[x_0 = x \wedge \bigwedge_{j=0\ldots n} R(x_j) \\ \wedge\; \forall x' \;\big[(x_0 < x' < x_n \wedge \bigwedge_{j=0\ldots n} x' \neq x_j) \Rightarrow \neg\,(R(x'))\,\big] \wedge x_n \lesseqgtr S^{n'} y' \Big].\end{aligned}$$

In the same way, for formulæ (2) and (3), with $A^R = \sum_{j=0\ldots n} a_j R_{S^j\cdot}$, we introduce new variables $x_0, \ldots, x_n$ and replace (2) (and similarly (3)) by

$$\begin{aligned}\exists x_0 < \ldots < x_n \Big[x_0 = x \;\wedge\; \bigwedge_{j=0\ldots n} R(x_j) \\ \wedge\; \forall x' \big[(x_0 < x' < x_n \;\wedge\; \bigwedge_{j=0\ldots n} x' \neq x_j) \Rightarrow \neg\,(R(x'))\,\big] \\ \wedge\; \sum_{j=0\ldots n} a_j x_j + B^R \bar{y} + \tau(\bar{x}) \lesseqgtr 0 \Big].\end{aligned}$$

In particular, the sentences (4) and (16–18) may be written as sentences of $\langle\omega;+,R\rangle$. The algorithms we gave on pages 20 and 44 still apply here, and this proves that $\mathcal{E}_R$ is *e-a-p* and R is *e-s*. □

Remark 27. The proof of Corollary 26 shows that for any n, definable subsets of ω^n in $\langle\omega;+,R\rangle$ are exactly definable subsets of ω_1^n in T_R^+. We could not much more extend this notion of 'partial interdefinability' since a subset of $\omega_1^{n_1} \times \omega_2^{n_2}$ would be meaningless in $\langle\omega;+,R\rangle$ for $n_2 \geqslant 1$. ⌟

When R is a predicate such that $\mathcal{I}_R \cup \mathcal{E}_R$ is almost-periodic, a result similar to Corollary 26 can be given for the theory of $\langle\omega;<,R,\mathcal{E}\rangle$:

Corollary 28. *Let $R \in \mathcal{R}$. The theory of $\langle \omega; <, R, \mathcal{E} \rangle$ is decidable if and only if $\mathcal{I}_R \cup \mathcal{E}_R$ is* e-a-p.

The proof consists on one hand in moving into the theory $T_R^<$ described on page 24 and in using Theorem 12, and on the other hand in noticing that the method we used in the previous proof in order to transform formulæ (1) and (2) still works [17] in this case. □

Example 29. We show in Examples 33, 36 and 37 (in Appendix B) that the predicates R_{c^n}, $R_{n!}$ and R_{Fib} (where c is a natural number > 1) are effectively sparse predicates. For any m, the rests of c^n modulo m form a periodic sequence (m being a period); those of $n!$ are constant for any $n \geqslant m$; and those of Fibonacci sequence also form a periodic sequence. Indeed, they satisfy the same induction rule, so that two consecutive elements determine the rest of the sequence. Since any segment of length $m^2 + 2$ of the sequence of the rests contains (at least) two equal pairs of consecutive elements, we see that the period is lower than or equal to $m^2 + 2$ (it can be effectively found by computing the first $m^2 + 2$ Fibonacci numbers).

As a consequence, the systems $\mathcal{E}_{R_{c^n}}$, $\mathcal{E}_{R_{n!}}$ and $\mathcal{E}_{R_{Fib}}$ are periodic, and thus *e-a-p*.

Corollaries 26 and 28 do apply, showing the decidability of the theories of the structures $\langle \omega; <, \mathcal{E}_{R_{c^n}} \rangle$, $\langle \omega; +, R_{c^n} \rangle$, $\langle \omega; <, \mathcal{E}_{R_{n!}} \rangle$, $\langle \omega; +, R_{n!} \rangle$, $\langle \omega; <, \mathcal{E}_{R_{Fib}} \rangle$ and $\langle \omega; +, R_{Fib} \rangle$. ⌟

6.2. A Structure whose Theory is Undecidable but whose Relations are Decidable

We now show a quite peculiar structure: all of its definable relations are decidable [18], but the theory of this structure is undecidable [19].

Let W be the non-effectively almost-periodic word we build in Appendix A. We define a predicate R_W from the sparse predicate $R_{n!}$ in such a way

17. Except that operators are replaced by terms '$S^n R_{S^l y}$'.
18. Recall that a relation is decidable if there exists an algorithm (we might not know it and it depends on the relation) which decides, for any n-uple of natural numbers, whether this n-uple belongs or not to the relation.
19. There is no 'global' algorithm deciding any sentence.

that the elements of $\mathcal{E}_R$ are almost everywhere equal to W.

Let R_W be the predicate defined by

$$x \in R_W \;\Leftrightarrow\; \exists y \,\Big((x = y! \wedge \neg\,(W(y))) \;\vee\; (x = 1 + y! \wedge W(y)) \Big).$$

Since the unary relations 'being a factorial number' and 'belonging to W' are decidable (since W is computable), this predicate is decidable as well. By Example 36 on page 56, it is also effectively sparse. Moreover, for any natural number $m > 1$, the sequence of the rests modulo m of the elements of R_W coincides almost everywhere with the characteristic word of W, since for any $j > m$, $R_j \equiv_m 0 \Leftrightarrow W(j)$. This shows $\mathcal{E}_{R_W}$ forms an almost-periodic system [20] but which is *not e-a-p*. By Example 15, we do not have to worry for the almost-periodicity of the system of predicates $\mathcal{I}_R$.

Since $R \in \mathcal{R}$, Corollary 26 tells us that the theory of $\langle \omega; +, R_W \rangle$ is undecidable.

However, the definable sets (or relations) of $\langle \omega; +, R_W \rangle$ are the same as those of T_R^+. But T_R^+ is an existential theory. Thus any n-ary relation $\varphi(\bar{x})$ is decidable: a deciding algorithm *exists*, and is described on page 30. However, it requires us to know an existential version of both $\varphi(\bar{x})$ and $\neg\,(\varphi(\bar{x}))$. This means that these formulæ do exist, but there is no *global* algorithm giving them.

Appendices

A. A Non-Effectively Almost-Periodic Word

We want to show that there exists a decidable [21] almost-periodic word which is not effectively almost-periodic. We shall consider a machine M (say a Turing machine) whose entries are natural numbers and for which the halting problem is undecidable (that is, there is no algorithm which decides, for any natural number n, whether M gives or not an answer (after a finite time) for the input n). Then, we shall use this machine

20. The fact this forms an *a-p system* comes from the fact the elements of $\mathcal{E}_{R_W}$ are almost everywhere equal.

21. In the sense that we are able to find the ith symbol of this word for any $i \in \mathbb{N}$.

to construct an *a-p* word, and show that if this word was *e-a-p*, then, by 'analyzing' this word, we would be able to decide the halting problem of M.

First, let us show a general way of constructing *a-p* words.

Let w be a finite word on $\{0,1\}$. We denote by $\overline{w}$ the word obtained by changing all symbols '0' into '1' and conversely. Also, when w_1 and w_2 are words, we denote by w_1w_2 the word obtained by concatenation.

We define by induction a product $\times$ on $\{0,1\}^+$ (the set of non-empty finite words on $\{0,1\}$) in the following way:

$$\begin{array}{rl} u \times 0 = & u \\ u \times 1 = & \overline{u} \\ u \times (v_1 v_2) = & (u \times v_1)(u \times v_2) \end{array} \qquad u, v_1, v_2 \in \{0,1\}^+ .$$

One easily checks that this product is associative. Moreover, given any words $u \in \{0,1\}^+, v \in \{0,1\}^*$, the word $u \times (0v)$ begins with u. This implies that if we consider a sequence of words $u_0, u_1, \ldots \in 0\{0,1\}^*$, infinitely [22] many of them being of length at least 2, then the sequence $u_0, u_0 \times u_1, u_0 \times u_1 \times u_2, \ldots$ 'converges' [23] to an infinite word of $\{0,1\}^\omega$.

Example 30. Thue-Morse word is obtained as 'limit' of the following iteration:

- $t_0 =$ '0'
- t_i is made out of t_{i-1} by replacing each symbol '0' with '01' and each '1' with '10'.

So Thue-Morse word starts with 0110100110010110... It is easy to show that it is also the limit of the product $01 \times 01 \times 01 \times \cdots$ ⌟

Lemma 31. *For any words* $u_0, u_1, \ldots \in 0\{0,1\}^*$, *the word* $W = u_0 \times u_1 \times \cdots$ *is almost-periodic.*

22. If there are only finitely many words of length at least 2, then we obviously get a finite word.
23. This means that there exists a unique word $W \in \{0,1\}^\omega$ such that any word of the sequence is an initial subword of W.

PROOF. — First suppose only finitely many u_i contain a symbol '1'. Then W is a periodic word. Indeed, let n be such that $u_i \in \{0\}^+$ for all $i > n$. Then

$$\begin{aligned} u_0 \times u_1 \times \cdots &= (u_0 \times \cdots \times u_n) \times (000 \cdots) \\ &= (u_0 \times \cdots \times u_n)(u_0 \times \cdots \times u_n)(u_0 \times \cdots \times u_n) \cdots \end{aligned}$$

Now suppose there are infinitely many u_i containing a symbol '1'. Given a word $v \in \{0,1\}^*$ appearing in $u_0 \times u_1 \times \cdots$, let us show that there exists an almost-period Δ_v. We know that v appears in $u_0 \times \cdots \times u_n$ for some n. Let $p > n$ be a number such that u_p contains a '1'. Since v appears in $u_0 \times \cdots \times u_{p-1}$, v and $\overline{v}$ appear in $w = u_0 \times \cdots \times u_p$. Further finite words (and thus W itself) are concatenations of w and $\overline{w}$, so that v appears in any of their segment of length $2\,|w|$. This shows $2\,|w|$ is suitable as almost-period for v. □

We now give an algorithm that constructs a decidable non-effectively almost-periodic word W.

Let M be a machine whose halting problem is undecidable and whose entries are natural numbers (see [4, §3.8] for the existence of such a machine). Denote by $t(n)$ the answer time of M for the entry $n \in \mathbb{N}$ (with the convention $t(n) = \infty$ if the machine does not answer).

Without loss of generality, we may suppose that for all $n \in \mathbb{N}$, $t(n) \geqslant n$. (If M does not satisfy this condition, it suffices to transform it into a machine M' that loops n times before transmitting the data 'n' to M.)

We construct the word W using the program P described on Figure 2, page 52: this program uses natural variables i and k and finite string variables $u_0, u_1, \ldots$ on $\{0,1\}^*$. The program P does the following work: for increasing k's, it initializes the variable u_k and looks for the values of $i \leqslant k$ such that $t(i) = k$. Once such an i is found, a new value is assigned to the variable u_i, depending on the actual values of $u_i, \ldots, u_k$.

The program also uses a string variable w which always gets the value of $u_0 \times \cdots \times u_k$.

While P is running, the value of w tends to a word of $\{0,1\}^\omega$. Indeed, it is clear that at any time, all the variables u_j ($j = 1 \ldots k$) begin with a '0'.

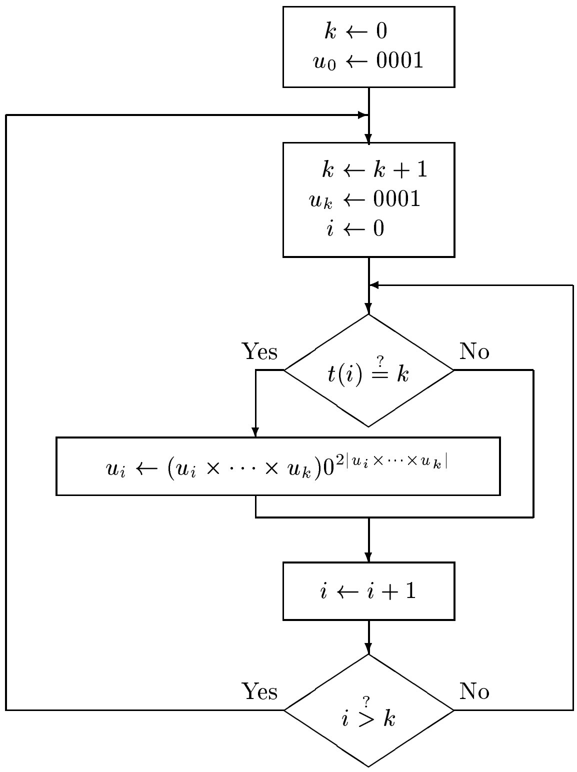


Figure 2: The program P.

When k is incremented, the new value of w is $w' = www\overline{w}$, and w is an initial segment of w'. Similarly, when the value of a variable u_i is changed into $u'_i = (u_i \times \cdots \times u_k)0^{|u_i \times \cdots \times u_k|}$, the new value of w contains the old one as initial segment: let $x = u_0 \times \cdots \times u_{i-1}$. The old value of w is

$$\underbrace{u_0 \times \cdots \times u_{i-1}}_{x} \times u_i \times u_{i+1} \times \cdots \times u_k$$

and the new one is

$$x \times ((u_i \times \cdots \times u_k)0^{|u_i \times \cdots \times u_k|}) \times u_{i+1} \times \cdots \times u_k$$

$$= \Big(\underbrace{(x \times (u_i \times \cdots \times u_k))}_{w} (x \times 0^{|u_i \times \cdots \times u_k|}) \Big) \times u_{i+1} \times \cdots \times u_k$$

$$= \big(w \underbrace{\bar{x} \cdots \bar{x}}_{|u_i \times \cdots \times u_k| \text{ times}} \big) \times u_{i+1} \times \cdots \times u_k,$$

and as the values of $u_{i+1}, \ldots, u_k$ all begin with '0', we have the result we wanted.

It is clear that this program will change the initial value of a variable u_i at most once (when $k = t(i)$). Let $\tilde{u}_i$ be the value of the variable u_i after it has been re-assigned, or its initial value if it is never re-assigned. Clearly, $\tilde{u}_n =$ '0001' if and only if $t(n) = \infty$. Let W be the infinite word $\tilde{u}_0 \times \tilde{u}_1 \times \cdots$. This word is well defined as all $\tilde{u}_i$ begin with a '0', and by Lemma 31, W is almost-periodic. Moreover, W is the limit of the values taken by the variable w, and this shows that W is decidable [24].

The hypothesis on the undecidability of the halting problem for M implies that it is impossible to know, after P worked some time, which unchanged variables will be re-assigned later. This is what we shall use in order to show W is not effectively almost-periodic: supposing W is *e-a-p* and using an induction on i, we will have an algorithm that gives a time after what we are sure that the value of u_i is $\tilde{u}_i$.

Before proceeding with this part, let us show a result on the particular form of the values that the variables u_i can get.

Lemma 32. *Let x be the value of a variable u_i at some time. Then $\bar{x}$ does not appear in the word xx.*

Proof. — Let Ω be the subset of $\{0,1\}^*$ built inductively by the rules

- $0001 \in \Omega$,
- If $x \in \Omega$, then $x0^{2|x|} \in \Omega$,
- If $x, y \in \Omega$, then $x \times y \in \Omega$.

24. In order to decide whether the n^{th} symbol of W is a '1', it suffices to let the program P run until $|w| \geqslant n$. This will eventually happen since $|w|$ is strictly increasing with k.

We show our lemma for Ω (all the possible values of the variable u_i being in Ω.)

- Obviously, $\overline{0001} = 1110$ does not appear in 00010001,
- Also, for any word x, $\overline{x0^{2|x|}} = \overline{x}1^{2|x|}$ does not appear in $x0^{2|x|}x0^{2|x|}$,
- Finally, consider a word $x \times y$ with $x, y \in \Omega$. Thanks to the associativity of $\times$, we may suppose y cannot be factorized as a product of words of Ω. So
  - if $y = 0001$, then $\overline{x \times y} = \overline{xxx}x$ cannot appear in $(x \times y)(x \times y) = xxx\overline{x}xxx\overline{x}$ using our induction hypothesis,
  - if $y = z0^{2|z|}$ with $z \in \Omega$, then $x \times y = (x \times z)x^{2|z|}$ and $\overline{x \times y} = \overline{x \times z}\overline{x}^{2|z|}$. Since $|x \times z| = |x|\,|z| = \frac{1}{2}\left|x^{2|z|}\right|$, the word $\overline{x \times y}$ cannot appear in $(x \times y)^2$, otherwise $\overline{x}$ would have to appear in xx, which is impossible using our induction hypothesis. □

It remains to prove that W cannot be *e-a-p*. Suppose it is and let us come to the algorithm that would allow us to compute $\tilde{u}_i$ for any $i \in \mathbb{N}$.

First start with $i = 0$. Since 0001 appears in W, it appears infinitely often. We find an almost-period Δ for 0001 in W. Our algorithm starts the program P until $k = \Delta$. At that time, we are sure that $u_0 = \tilde{u}_0$: any further modification of u_0 would force u_0 (and thus W) to contain a sequence of symbols '0' of length at least 2Δ — since $|u_0 \times \cdots \times u_k| \geqslant 2k$ — but this would contradict the fact Δ is an almost-period for 0001.

Now, we want to compute $\tilde{u}_i$. Using our induction hypothesis, we first compute $\tilde{u}_0, \ldots, \tilde{u}_{i-1}$. We set $W_i = \tilde{u}_0 \times \cdots \times \tilde{u}_{i-1}$ and look for an almost-period Δ for $\overline{W_i}$ in W. Since the word u_i will contain a '1', $\overline{W_i}$ will appear in W and thus will appear infinitely often. We start P until $k = \Delta + i$. At that time, the value of u_i must be equal to $\tilde{u}_i$: any further modification of u_i would force it to contain a sequence of '0' of length at least 2Δ so that W would contain at least 2Δ concatenated copies of W_i. However, the previous lemma tells us that $\overline{W_i}$ cannot appear in these 2Δ copies so that we contradict the fact that Δ is an almost-period for $\overline{W_i}$ in W. □

B. Some Examples of Sparse Predicates

Example 33. Let $R = R_{2^n}$ be the set of powers of two. Then any operator A^R can be written under the form $A^R(y) = a_n 2^{(S^n y)} + \cdots + a_0 2^{(S^0 y)} = (2^n a_n + \cdots + 2a_1 + a_0)2^y = c2^y$. Thus R is effectively sparse, since $A^R \lesseqgtr_R 0 \Leftrightarrow c \lesseqgtr 0$ and when $c > 0$, it suffices to take $\Delta = 1$ in order to satisfy condition P2: $\forall y A^R(Sy) - 2^y = (2c-1)2^y > 0$.

Of course, this example extends to any predicate of the form $R_{c^n} = \{c^n \mid n \in \mathbb{N}\}$ where c is a natural number > 1. ⌟

Counter-example 34. Contrarily to what Semenov writes in [10, p. 410, ex1], this example does not extend to predicates R such that the rational sequence $\frac{R_{j+1}}{R_j}$ is periodic. Indeed, let R be the set defined by the induction $R_0 = 1$, $R_{2n+1} = 4R_{2n}$, $R_{2n+2} = 2R_{2n+1}$ and let A^R be the operator $R_{(S\cdot)} - 3R_{(\cdot)}$. We immediately check $A^R(y) > 0$ if y is even, and $A^R(y) < 0$ if y is odd. This contradicts condition 18.P1.

This example also shows that a subset of a sparse predicate does not need to be sparse itself — our counter-example being a subset of the set of powers of two. ⌟

Example 35. Let us show that any set R such that $\lim_{j\to\infty} \frac{R_{j+1}}{R_j} = \infty$ is sparse. We know that for any natural number N there exists a Λ such that $\forall j > \Lambda : \frac{R_{j+1}}{R_j} > N$. Let A^R be the operator $\sum_{i=0\ldots n} a_i R_{(S^i\cdot)}$ and let us show that $A^R \lesseqgtr_R 0 \Leftrightarrow a_n \lesseqgtr 0$. Suppose $a_n > 0$ (the argument is similar if $a_n < 0$), let $N = 1 + \sum_{i=0\ldots n-1} |a_i|$ and let Λ have a value corresponding to N. Then for any $y > \Lambda$,

$$\begin{aligned} A^R(y) &\geqslant a_n R_{S^n y} - \sum_{i=0\ldots n-1} |a_i|\, R_{S^i y} \\ &\geqslant R_{S^n y} - \sum_{i=0\ldots n-1} |a_i|\, R_{S^{n-1} y} \\ &\geqslant \left(N - \sum_{i=0\ldots n-1} |a_i|\right) R_{S^{n-1} y} > 0 \end{aligned}$$

so that condition 18.P1 is fulfilled. For condition P2, taking $N = 2 + \sum_{i=0\ldots n-1} |a_i|$, the same computation tells us that for any y, $A^R(S^\Lambda y) -$

$R_y > 0$. When the limit is effective, that is, when there exists an algorithm giving Λ as a function of N, then R is effectively sparse.

This time, any (infinite) subset R' of a predicate R of this type remains sparse, since clearly $\lim_{j\to\infty} \frac{R'_{j+1}}{R'_j} = \infty$. Moreover, if R is effectively sparse, then so is R', because the same algorithm still works. ⌟

Example 36. The predicate $R_{n!} = \{n! \mid n \in \mathbb{N}\}$ is effectively sparse (it suffices to take $\Lambda = N$ in the previous example) and so is any subset of it. ⌟

Example 37. Let R_{Fib} be the set of Fibonacci numbers $\{1, 2, 3, 5, 8, 13, \ldots\}$. These are recursively defined by $R_0 = 1, R_1 = 2, R_{n+2} = R_n + R_{n+1}$. One also shows that they are obtained as images of the function

$$R_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+2}\right).$$

Notice that this implies that for any n, $\frac{\alpha^{n+2}}{\sqrt{5}} - 1 < R_n < \frac{\alpha^{n+2}}{\sqrt{5}} + 1$ where $\alpha = \frac{1+\sqrt{5}}{2}$.

Using the induction rule, any operator A^R can be written as

$$\begin{aligned} A^R(y) &= a_n R_{(S^n y)} + \cdots + a_0 R_y \\ &= (a_n + a_{n-1}) R_{(S^{n-1} y)} + (a_n + a_{n-2}) R_{(S^{n-2} y)} \\ &\quad + a_{n-3} R_{(S^{n-3} y)} + \cdots + a_0 R_y \\ &= \cdots = c_1 R_{Sy} + c_2 R_y. \end{aligned} \tag{20}$$

If $c_1 = 0$, then we may effectively find out which condition P1 is satisfied and what is the Δ of P2, just like in Example 33.

Otherwise, we have that

$$\begin{aligned} A^R(y) \lesseqgtr 0 &\Leftrightarrow c_1 R_{Sy} \lesseqgtr -c_2 R_y \\ &\Leftrightarrow \frac{R_{Sy}}{R_y} \lesseqgtr \frac{-c_2}{c_1}. \end{aligned} \tag{21}$$

Since $\lim_{y\to\infty} \frac{R_{Sy}}{R_y} = \alpha$, we have $A^R \lesseqgtr_R 0 \Leftrightarrow \alpha \lesseqgtr \frac{-c_2}{c_1}$, and this determines P1.

Suppose $A^R >_R 0$. In order to satisfy P2, we need to find a Δ such that for any y,

$$0 < A^R(S^\Delta y) - R_y.$$

It suffices to find a Δ such that $R_0 = 1 < A^R(\Delta)$ and such that the growth of $A^R(S^\Delta \cdot)$ is greater than that of $R_{(\cdot)}$. The first condition is

$$1 < c_1 R_{\Delta+1} + c_2 R_\Delta$$

and using the bounds $\frac{\alpha^{n+2}}{\sqrt{5}} - 1 < R_n < \frac{\alpha^{n+2}}{\sqrt{5}} + 1$, it suffices to find Δ such that

$$\begin{aligned} & 1 < c_1 \frac{\alpha^{\Delta+3}}{\sqrt{5}} - |c_1| + c_2 \frac{\alpha^{\Delta+2}}{\sqrt{5}} - |c_2| \\ \Leftrightarrow\ & \sqrt{5}\,(1 + |c_1| + |c_2|) \ < \ c_1 \alpha^{\Delta+3} + c_2 \alpha^{\Delta+2} \\ \Leftrightarrow\ & \frac{\sqrt{5}\,(1 + |c_1| + |c_2|)}{c_1 \alpha^3 + c_2 \alpha^2} \ < \ \alpha^\Delta. \end{aligned} \tag{22}$$

(we may divide by $c_1\alpha^3 + c_2\alpha^2$ without changing the signs since $A^R >_R 0$, so that $c_1\alpha + c_2 > 0$.) Now, in order to have that $A^R(\Delta) > R_0$, it suffices to have

$$\Delta > \log_\alpha \frac{\sqrt{5}\,(1 + |c_1| + |c_2|)}{c_1 \alpha^3 + c_2 \alpha^2}. \tag{23}$$

The second condition on Δ requires that for any y,

$$\begin{aligned} & A^R(S^{\Delta+1} y) - A^R(S^\Delta y) > R_{Sy} - R_y \\ \Leftrightarrow\ & A^R(S^{\Delta-1} y) > R_{S^{-1} y} \\ \Leftrightarrow\ & c_1 R_{S^\Delta y} + c_2 R_{S^{\Delta-1} y} > R_{S^{-1} y} \end{aligned} \tag{24}$$

and using the same bounds as before, it suffices to have

$$\begin{aligned} & c_1 \frac{\alpha^{\Delta+y+2}}{\sqrt{5}} - |c_1| + c_2 \frac{\alpha^{\Delta+y+1}}{\sqrt{5}} - |c_2| > \frac{\alpha^{y+1}}{\sqrt{5}} + 1 \\ \Leftrightarrow\ & \alpha^\Delta \alpha^{y+1}(c_1\alpha + c_2) > \alpha^{y+1} + \sqrt{5}\,(|c_1| + |c_2| + 1)\,. \end{aligned} \tag{25}$$

Again, we know that $c_1\alpha + c_2 > 0$, and thus the condition becomes

$$\Delta > \log_\alpha \left(\frac{1}{(c_1\alpha + c_2)} \Big(1 + \frac{\sqrt{5}}{\alpha^{y+1}}\,(|c_1| + |c_2| + 1) \Big) \right). \tag{26}$$

Since the function $\frac{1}{\alpha^{y+1}}$ is decreasing, the right member of this inequality gets its maximum value for $y = 0$.

The value of Δ we need to find can be taken as the first natural number greater than both (23) and (26) (computed in $y = 0$), and this concludes the proof that R_{Fib} is an effectively sparse predicate. ⌟

References

[1] P.T. Bateman, C.G. Jockusch, and A.R. Woods. Decidability and Undecidability of Theories with a Predicate for the Primes. *Journal of Symbolic Logic*, Vol. 58, no. 2, pp. 672–687, 1993.

[2] J.R. Büchi. Weak second-order arithmetic and finite automata. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, Vol. 6, pp. 66–92, 1960.

[3] A. Cobham. On the base-dependance of sets of numbers recognizable by finite-automata. *Math. Systems Theory*, Vol. 3, pp. 186–192, 1969.

[4] H.B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, New York, 1973.

[5] C. H. Langford. On inductive relations. *Bull. Amer. Math. Soc.*, Vol. 33, pp. 599–607, 1927.

[6] Th. Lavendhomme and A. Maes. Note on the undecidability of $\langle \omega; +, P_{m,r} \rangle$. *This Cahier*, pages 61–68, 2000.

[7] C. Michaux and R. Villemaire. Presburger arithmetic and recognizability of sets of natural numbers by automata: New proofs of Cobham's and Semenov's theorems. *Annals of Pure and Applied Logic*, Vol. 77, pp. 251–277, 1996.

[8] M. Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. *Comptes rendus du 1er Congrès des Mathématiciens des Pays Slaves*, pages 92–101, 1929.

[9] A.L. Semenov. The Presburger nature of predicates that are regular in two number systems. *Siberrian Math. Journal*, Vol. 18, no. 2, pp. 289–299, 1977.

[10] A.L. Semenov. On certain extensions of the arithmetic of addition of natural numbers. *Math USSR Izv.*, Vol. 15, no. 2, pp. 401–418, 1980. English translation of *Izv. Akad. Nauk SSSR Ser. Mat.*, Vol. 43, no. 5, pp. 1175–1195, 1979.

[11] A.L. Semenov. Logical theories of one-place functions on the set of natural numbers. *Math USSR Izv.*, Vol. 22, no. 3, pp. 587–618, 1984. English transl. of *Izv. Akad. Nauk SSSR Ser. Mat.*, Vol. 47, no. 3, pp. 623–658, 1983.

Université de Mons–Hainaut
Institut de Mathématique et d'Informatique
Bâtiment "Le Pentagone"
Avenue du Champ de Mars 6
7000 Mons
Belgium
e-mail : `maesa@sun1.umh.ac.be`

Cahiers du Centre de logique
Volume 11

Note on the Undecidability of $\langle \omega; +, P_{m,r} \rangle$

by

T. Lavendhomme and A. Maes [1]
Université de Cergy-Pontoise, *Université de Mons-Hainaut*

1. Introduction

Let P be the set of prime numbers. In 1993, P.T. Bateman, C.G. Jockusch and A.R. Woods [1] showed that the linear case of Schinzel's Hypothesis (H) implies the undecidability of the first order theory of $\langle \omega; +, P \rangle$.

Bateman, Jockusch and Woods gave two proofs of their result: the first one shows that the multiplication is definable in $\langle \omega; +, P \rangle$ by defining the range of a polynomial g of degree two; the second one shows that a 'substructure [2]' isomorphic to $\langle \omega; +, \cdot \rangle$ is definable in $\langle \omega; +, P \rangle$. The definability argument is closely related to the existence of sequences (of quadratic growth in the first case and of linear growth in the second case) of consecutive elements of P.

In 1996, M. Boffa [2] extended the first proof in order to show that (H) also implies the undecidability of the first order theory of $\langle \omega; +, P_{m,r} \rangle$.

1. Aspirant du Fonds national belge de la recherche scientifique.
2. More precisely, the domain of this structure is the quotient of a definable set by a definable equivalence relation, and the graphs of the functions $+$ and $\cdot$ are defined on the classes of this equivalence relation.

Here, $P_{m,r}$ is the set $\{p \in P \mid p \equiv r (\mathrm{mod}\, m)\}$ where m and r are natural numbers, $m > 2$, $r < m$ and r is prime to m. His argument consists in modifying the proof of Bateman & al. by carefully choosing another polynomial g (of degree $\phi(m)$, where ϕ is Euler's totient function). As a consequence, the considered sequences of elements of $P_{m,r}$ grow faster and their elements are no more consecutive in $P_{m,r}$, but are separated by a fixed number of elements of $P_{m,r}$.

In order to show Boffa's result again, we modify a preliminary lemma of [1]. This will enable us to remain as close as possible to the two original proofs of Bateman & al. In particular, we keep polynomials of degree two and sequences of quadratic or linear growth.

2. Basic Definitions and Preliminary Lemma

Definition 1. *Let P be the set of prime numbers. Let $m > 2$ and $r < m$ be two mutually prime natural numbers.*

We set $P_{m,r} = \{p \in P \mid p \equiv r (\mathrm{mod}\, m)\}$.

Conjecture: Schinzel's Hypothesis (H). *Let $f_1(x), \ldots, f_n(x)$ be irreducible polynomials over $\mathbb{Z}$, each having a positive leading coefficient. Suppose that there is no prime p which divides $f_1(x) \cdot f_2(x) \cdots f_n(x)$ for all $x \in \omega$. Then there exist infinitely many $x \in \omega$ such that $f_1(x), f_2(x), \ldots, f_n(x)$ are all prime.*

The linear case of (H) is the case in which the $f_i(x)$ are all linear and have the same leading coefficient.

The following technical lemma is based on Lemma 1 of [1].

Lemma 2. *Assume the linear case of (H). Let m and r be as before, let $b_0, b_1, \ldots, b_n$ be an increasing sequence of natural numbers, and assume that:*

(a) *there is no prime p such that*

$$\{b_i (\mathrm{mod}\, p), 0 \leqslant i \leqslant n\} = \{0, 1, \ldots, p-1\};$$

(*b*) $b_i \equiv b_j (\bmod m)$ *for all* $0 \leqslant i, j \leqslant n$.

Then there are infinitely many natural numbers a *such that* $a + b_0, a + b_1, \ldots, a + b_n$ are consecutive *in* $P_{m,r}$.

PROOF. — We follow the proof of [1, Lemma 1].

Let $a_1 < \cdots < a_s$ be the integers between b_0 and b_n which are not of the form b_i for any $i \leqslant n$. We shall show that there exist infinitely many integers a such that

$$a + a_i \notin P; \tag{1}$$

$$a + b_j \in P_{m,r}. \tag{2}$$

Let p_i denote the ith prime, and let $c = mp_1 p_2 \ldots p_{(\delta+1+s)}$ with $\delta = \max\{b_n, p_k\}$ where p_k is the greatest prime factor of m. We want to use (H) for the following polynomials:

$$f_j(x) = cx + a_0 + b_j, \tag{3}$$

where a_0 will be chosen later.

We shall then choose a to be $cx + a_0$.

In order to ensure (2) (using (b) and the fact that $m|c$), we require that:

$$a_0 + b_0 \equiv r \pmod{m}. \tag{4}$$

In order to ensure (1) (with the same argument as in [1]), we require that a_0 satisfies

$$a_0 \equiv -a_i \pmod{p_{(\delta+1+i)}} \qquad (i = 1, 2, \ldots, s). \tag{5}$$

To ensure that Schinzel's hypothesis applies for the $f_j(x)$, we also require that:

$$a_0 \not\equiv -b_j \pmod{p_i} \qquad (j = 0, 1, \ldots, n) \tag{6}$$

for each $i = 1, 2, \ldots, s$ such that p_i is not a prime factor of m.

To prove that Schinzel's hypothesis applies for the $f_j(x)$, we may restrict attention, as in [1], to primes $p \leqslant p_{(\delta+1+s)}$.

If p is a prime factor of m, note that $p|c$. So

$$\prod_{j=0}^{n} f_j(x) \equiv \prod_{j=0}^{n} (a_0 + b_j) \pmod p$$

and so, by condition (b),

$$\prod_{j=0}^{n} f_j(x) \equiv \prod_{j=0}^{n} (a_0 + b_0) \pmod p.$$

(4) ensures us that this is not $\equiv 0 \pmod p$. Therefore

$$p \nmid \prod_{j=0}^{n} f_j(x).$$

If $p = p_{(\delta+1+i)}$ for some $i = 1, 2, \dots, s$ or if $p = p_i$ for $i = 1, 2, \dots, s$ such that p_i is not a prime factor of m, the proofs are similar as in [1, Lemma 1], using, respectively, (5) and (6).

Note that (6) has a solution by hypothesis (a), therefore the Chinese Remainder Theorem gives an a_0 satisfying (4), (5) and (6). □

3. First Proof of Theorem 3

Theorem 3. *Assume the linear case of (H). Then the first order theory of* $\langle \omega; +, P_{m,r} \rangle$ *is undecidable. ([2])*

Recall the following theorem (Büchi):

> *for any polynomial $g(x)$ of degree at least two with coefficients from ω, the multiplication is definable in $\langle \omega; +, \mathrm{Im}(g) \rangle$ and therefore, by Church's Theorem, $\langle \omega; +, \mathrm{Im}(g) \rangle$ is undecidable.*

PROOF of Theorem 3. — The polynomial $g(x) = m(x^2+x)$, easily, satisfies the hypotheses of Büchi's theorem. So it suffices to prove that $\mathrm{Im}(g)$ is definable in $\langle \omega; +, P_{m,r} \rangle$. The proof is similar as in Theorem 1 of [1].

Define $T(a,b)$ to hold iff

$$(\exists n \geqslant 1)[a+g(n)=b \\ \wedge \quad a+g(0),a+g(1),\ldots,a+g(n) \text{ are consecutive in } P_{m,r}].$$

T is definable in $\langle\omega;+,P_{m,r}\rangle$ because $T(a,b)$ holds if and only if $a, a+2m$ and b are all in $P_{m,r}$, $a+m$ is not in $P_{m,r}$, and $(e-d)-(d-c)=2m$ whenever c,d,e are consecutive in $P_{m,r}\cap[a,b]$ with $c<d<e$.

The end of the proof is identical as in [1, Theorem 1], defining $\mathrm{Im}(g)$ by:

$$k\in \mathrm{Im}(g) \iff (\exists a)\ T(a,a+k) \vee k=0$$

and using Lemma 2. □

4. Second Proof of Theorem 3

We now want to prove Theorem 3 like in [1, §4]. This will result from the following theorem (similar to [1, Theorem 2]):

Theorem 4. *Let $m\neq 0$ be a natural number and let R be a predicate satisfying the following property for any $d,l\in\omega$ with $l>m$:*

$(\star)$ *There exist infinitely many numbers a such that $a, a+d,\ldots,a+(l-1)d$ are consecutive elements of R* if and only if *$m|d$ and $p|d$ for all primes $p\leqslant l$.*

Then the first-order theory of the structure $\langle\omega;+,R\rangle$ is undecidable.

Proof of Theorem 3 using Theorem 4. — It suffices to show that $P_{m,r}$ satisfies $(\star)$.

The 'if' part is an immediate consequence of Lemma 2 (with $b_i=i\cdot d$). The 'only if' part is obvious since $P_{m,r}\subset P$ and that $a,a+d,\ldots,a+(l-1)d$ have to be congruent mod m. □

Proof of Theorem 4. — It suffices to show that the predicate $D(x,y)$ used in [1, Theorem 2] is definable by a first order formula and to continue with the same proof.

Recall that $D(x,y)$ is the predicate

$$x > 0 \ \wedge \ (\exists n \geqslant 1)\,(\exists l \geqslant 1)\Big(\pi_n | x \ \wedge \ l < p_{n+1} \ \wedge \ y = (l-1)x\Big)$$

where $\pi_n{=}p_1 \cdots p_n$.

We may define a predicate $D'(x,y)$ by the first-order formula:

$$\begin{aligned} & x > 0 \ \wedge \ y \geqslant \underbrace{x + \cdots + x}_{m-1 \text{ times}} \\ & \wedge \ \forall r \exists a \Big[\Big(a \leqslant s < a + y \ \wedge \ R(s)\Big) \\ & \qquad\qquad \Rightarrow \ \Big(R(s+x) \ \wedge \ \forall t\big(0 < t < x \Rightarrow \neg R(s+t)\big)\Big)\Big]\Big]. \end{aligned}$$

This predicate means that y is a multiple of $x > 0$, say $y/x = l - 1$ with $l \geqslant m$, and that there are infinitely many a such that $a, a+x, \ldots, a+(l-1)x = a+y$ are consecutive elements of R.

Using $(\star)$, this predicate is equivalent to

$$x > 0 \ \wedge \ m|x \ \wedge \ (\exists n \geqslant 1)\,(\exists l \geqslant m)\Big(\pi_n | x \wedge l < p_{n+1} \wedge y = (l-1)x\Big).$$

We see that there are two families of elements missing from D' in order to get D. These are:

(a) the pairs (x,y) such that

$$x > 0 \ \wedge \ (\exists n \geqslant 1)\,(\exists 1 \leqslant l < m)\Big(\pi_n | x \ \wedge \ l < p_{n+1} \ \wedge \ y = (l-1)x\Big),$$

(b) the pairs (x,y) such that

$$x > 0 \ \wedge \ m \nmid x \ \wedge \ (\exists n \geqslant 1)\,(\exists l \geqslant m)\Big(\pi_n | x \wedge l < p_{n+1} \wedge y = (l-1)x\Big).$$

We may define the first family by testing whether, for some $l = 1 \ldots m-1$, we have

$$x > 0 \ \wedge \ (\exists n \geqslant 1)\Big(\pi_n \mid x \ \wedge \ l < p_{n+1} \ \wedge \ y = (l-1)x\Big).$$

Denote by h the computable function $h(l)=\min\{i : l<p_{i+1}\}$. Then the previous formula is equivalent to

$$x>0 \ \wedge \ \pi_{h(l)}|x \ \wedge \ y=(l-1)x.$$

This allows us to define the first family by a finite disjunction.

We now look at the second family. Suppose m factorizes as $p_{\alpha_1}^{\beta_1}\cdots p_{\alpha_u}^{\beta_u}$.

Then we have the following equivalences:

$$x \not\equiv_m 0 \ \wedge \ (\exists n\geqslant 1)\,(\exists l\geqslant m)\Big(\pi_n|x\wedge l<p_{n+1}\wedge y=(l-1)x\Big)$$

$$\Updownarrow \ (1)$$

$$\begin{aligned}&x\not\equiv_m 0 \ \wedge \ x\equiv_{p_{\alpha_1}\cdots p_{\alpha_u}} 0\\ &\wedge\ (\exists n\geqslant 1)\,(\exists l\geqslant m)\Big(\pi_n|mx \ \wedge \ l<p_{n+1} \ \wedge \ my=(l-1)mx\Big)\end{aligned}$$

$$\Updownarrow \ (2)$$

$$x\not\equiv_m 0 \ \wedge \ x\equiv_{p_{\alpha_1}\cdots p_{\alpha_u}} 0 \ \wedge \ D'(mx,my).$$

Indeed, (1) can be deduced from the fact that ($m<p_{n+1}$ and $\pi_n|x$) implies $p_{\alpha_1}\cdots p_{\alpha_u}|x$, so that no new factors are added when we replace x and y by mx and my. Thus any l and n acceptable to D' are still acceptable to D, i.e. we only add pairs of D. The definition of D' gives (2).

As a conclusion, we may define $D(x,y)$ by the first-order formula

$$\begin{aligned}D(x,y)\quad\Leftrightarrow\quad &D'(x,y)\ \vee\ (x\equiv_{p_{\alpha_1}\cdots p_{\alpha_u}} 0\wedge D'(mx,my))\\ &\vee \bigvee_{l=1\ldots m-1}(x>0\ \wedge\ \pi_{h(l)}|x\ \wedge\ y=(l-1)x).\end{aligned}$$

□

References

[1] P.T. Bateman, C.G. Jockusch, and A.R. Woods. Decidability and Undecidability of Theories with a Predicate for the Primes. *Journal of Symbolic Logic*, Vol. 58, no. 2, pp. 672–687, 1993.

[2] M. Boffa. More on an undecidability result of Bateman, Jockusch and Woods. *Journal of Symbolic Logic*, Vol. 63, no. 1, p. 50, 1998.

T. Lavendhomme
Rue François Vander Elst, 54
1950 Kraainem
Belgium
e-mail : kestemont@mark.ucl.ac.be

A. Maes
Université de Mons–Hainaut
Institut de Mathématique et d'Informatique
Bâtiment "Le Pentagone"
Avenue du Champ de Mars 6
7000 Mons
Belgium
e-mail : maesa@sun1.umh.ac.be

Cahiers du Centre de logique
Volume 11

On functions, computables by Turing machines

by

M. Margenstern and L. Pavlotskaïa
Université de Metz, *Institut d'ingénierie énergétique de Moscou*

1. Abstract

The notion of a function, computable by a Turing machine on a given set of words is defined. It is proved that this notion is very sensitive to the definition of a computation, in particular for universal Turing machines. Indeed, it is proved that there are universal machines which cannot compute any function on any set. An example is given of a machine which possesses this property and which cannot be replaced by a machine computing a function on a set of configurations where the former machine halts.

2. Introduction

A Turing machine is an algorithm which transforms sequences of symbols or words written on the cells of its tape. A cell which does not contain any symbol is said *empty.* This situation is marked by a particular symbol called *blank* symbol, or short, *the blank.* Applying that algorithm to a word written on consecutive cells of the tape, *i.e* to a word which contains no blanks, defines a function on the set of all these words, provided that the position of the machine head on the tape and its state are fixed when the computation starts.

For such computations, the notions of prefix or suffix of the initial word, of blank and of empty tape are natural. From that, a certain number of properties ensue which are specific to those transformations.

Each function defined in that way can be computed by a Turing machine which extends its computation only in the following case: the cells scanned during the computation which were out of the initial word, were empty at the time when the computation has started; also the set of words to which the algorithm is applied is recursive.

In this paper, we study the question whether computation still possess the above properties when it is allowed that the finite sequence of symbols written on the tape at the initial time may contain blanks. In particular, we study properties of universal Turing machines in this point of view.

It is thus proved that the set of configurations used by a universal Turing machine for modelling other computations may be non-recursive. In this paper, we introduce the notion of machine computing a function on a set with the help of the empty tape. We prove that there are such machines for which the corresponding set of configurations is not recursively enumerable. Applying this result to a particular universal Turing machine, we obtain that there is a universal Turing machine which does not compute a function on any set. Moreover, this machine cannot be replaced by another machine which would compute a function on a set of configurations on which the former machine could be applied.

3. Conventions and Notations

We begin by recalling definitions on the computation of a Turing machine.

First, define the notion of *configuration*. This notion is intuitively clear. The configuration consists of what is on the machine tape at a given time together with the supplementary information of the position of the machine head and its state while reading the scanned cell. It is more delicate to formalize this notion: we first need to define the *initial* configuration, or *starting* configuration — we shall distinguish between both notions — then, step by step, to define the *current* configuration starting from the previous current configuration, the first current one being a starting configuration.

Assume the tape to be infinite only on the right side. Assume that there is a one-to-one mapping from $\mathbb{N}$ onto the cells of the tape. The integer thus associated to a cell is called its *address*. Most often, the cell with x as address will be called *cell x*. Assume that the address of the leftmost cell is 0, that cell $x+1$ is the immediate right neighbour of cell x, and that cell $x-1$ is the immediate left neighbour of cell x. Assume we have also at our disposal a function $\sigma(x,t)$ which gives the content of cell x at time t. When time t is already known from the context, we shall write simply $\sigma(x)$ instead of $\sigma(x,t)$.

It is assumed that all possible values of $\sigma(x,t)$ belong to a fixed finite set, characteristic of the considered machine. This set is called the machine *alphabet* and is denoted by A in the sequel. Choose one letter among those of A to be called the *blank* denoted by $\sqcup$.

Consider now the smallest integer N such that at initial time 0, $\sigma(x) = \sqcup$ holds for all $x \geqslant N$. The *support* of the *starting* configuration is constituted of word $\sigma(0)\ldots\sigma(N-1)$, possibly empty. The starting configuration is defined by giving the smallest integer K such that at the initial time the address of any cell belonging to the support including the scanned cell is less than K, the address x of the scanned cell, and the state under which the machine is at initial time 0. The starting configuration can be encoded as a word of the form uq_iv where $u, v \in A^*$, $|uv| = N$, $|u| = x$ and $q_i \in Q$, Q denoting the finite set of the machine states $\{q_0, q_1, \ldots, q_s\}$.

By convention, state q_0 is the final state: when the machine is under that state it halts. Call q_1 the *initial* state. Call a starting configuration *initial* if the machine state at the initial time is q_1.

The *current configuration* of the machine at time t is defined by induction on t. For $t = 0$, it is the starting configuration. Assume the current configuration to be encoded by uq_iv at time t, with q_i as the *current* state at that time. Denote $v(t)$ the address of the cell scanned at time t. Notice that $v(t) = |u|$. Let $q_i\xi M\eta q_j$ be the instruction to be performed at time $t+1$. The current configuration at time $t+1$ is encoded by $u_1q_jv_1$ with

$|u_1| = v(t+1)$ and $\sigma(x, t+1) = \sigma(x,t)$ for all x except $v(t)$, and for $x = v(t)$ we have: $\sigma(x,t) = \xi$ and $\sigma(x, t+1) = \eta$. Integer $|uv|$ is the *current bound* of the configuration at time t. Denote that current bound by $\ell(t)$.

Call a current configuration *final* if and only if the current state is final state q_0. There is no current configuration after a final configuration.

Say that a starting configuration is a *halting configuration* if and only if the sequence of the following current configurations is finite and if the last configuration of the sequence is a final one and if, during the computation, the machine head scanned any cell with as address an integer less than the starting bound. This latter condition is formally written as:

$$\forall x\,(x < \ell(0) \Rightarrow \exists t\,(t \leqslant t_f \wedge v(t) = x)).$$

where t_f is the time when the final state is reached.

4. Universal Machines on a Set

Let Z denote a Turing machine. Denote by Z_{end} the set of its halting configurations.

Let $\mathcal{G}$ a fixed Gödel numbering of machine Z configurations.

Definition 1.
Call machine Z universal *if and only if set $\mathcal{G}(Z_{end})$ is m-complete.*

If machine Z is universal in that sense, see [1], there is a total recursive function f with two arguments such that the set of $U_n = \{x\ :\ f(n,x) \in \mathcal{G}(Z_{end})\}$ contains all recursively enumerable sets. Say then that f is an *encoding function* for machine Z. Set $K_{Z,f} = \mathcal{G}^{-1}(\mathrm{Im} f)$.

Definition 2. *Say that universal machine Z is* universal on set E of configurations *if and only if there is an encoding function f for Z such that $E = K_{Z,f}$.*

Theorem 3.
Any universal machine Z is universal on a set containing Z_{end}.

PROOF. — Let f be an encoding function for Z. Then, $K_{Z,f} \cap Z_{end}$ is infinite. There is a total recursive function d which enumerates elements of $\mathcal{G}(Z_{end})$. Accordingly, let c, ℓ and r be a recursive encoding of couples of integers: if $u = c(n,x)$, then $n = \ell(u)$, $x = r(u)$ and conversly. Let $\varphi(u) = f(\ell(u), r(u))$. As $\operatorname{Im}\varphi = \operatorname{Im} f$ and $\operatorname{Im}\varphi \cap \operatorname{Im} d$ is infinite, there are two total recursive increasing functions h and k such that for all natural number n, $\varphi(h(n)) = d(k(n))$ and $\varphi(h(n)) < \varphi(h(n+1)$. Define then φ^* by:

$$\varphi^*(n) = \begin{cases} \varphi(n) & \text{if } n \notin \operatorname{Im} h \\ d(j) & \text{if } n = h(j). \end{cases}$$

In these conditions, $\mathcal{G}(Z_{end}) = \operatorname{Im} d \subset \operatorname{Im}\varphi^*$. Let then $f^*(n,x) = \varphi^*(c(n,x))$. Function f^* is then an encoding function for Z. Indeed, let

$$U_n^* = \{x \ : \ f^*(n,x) \in \mathcal{G}(Z_{end})\}.$$

It is not difficult to prove that $U_n = U_n^*$: φ and φ^* take identical values when $n \notin \operatorname{Im} h$ and when this is not the case, $n = h(j)$ for some j. Consequently, $\varphi^*(n) = d(j)$ and $\varphi(n) = d(k(j))$. These latter two values are thus simultaneously in $\mathcal{G}(Z_{end})$. Moreover, as $K_{Z,f^*} = \mathcal{G}^{-1}(\operatorname{Im} f^*)$, which contains Z_{end} since $\operatorname{Im} f^* \supset \operatorname{Im} d$, set K_{Z,f^*} satisfies definition 2. □

Consider now a partial recursive function f whose domain, say $\operatorname{dom} f$, is m-complete. It is possible to construct a Turing machine F on alphabet $\{0,1\}$ such that $F_{end} \supset \{q_1 01^n \ : \ n \in \operatorname{dom} f\}$, that $F_{end} \cap \{q_1 01^n \ : \ n \notin \operatorname{dom} f\} = \varnothing$ and such that F applied to $q_1 01^n$ for $n \in \operatorname{dom} f$ computes $q_* 01^{f(n)}$ where q_* is the halting state of machine F.

Let us denote by $s(n)$ the space used by machine F to compute $f(n)$. Notice that s is a partial recursive function, that its domain of definition is the same as for f and that $s(n) > n$ for any natural number n where f is defined (else, if $s(n_0) \leqslant n_0$ for some natural number n_0, F provides the same result for all $n \geqslant n_0$ and thus, F cannot compute f).

Let us construct machine F_1 as described in Figure 1 (page 74).

Then,

$$F_{1,end} \supset \{\alpha_1 01^n 0 x_1 \ldots x_r \ : \ n \in \operatorname{dom} f, x_i \in \{0,1\}, r + n + 2 \leqslant s(n)\}$$

<table>
<tr><th></th><th>0</th><th>1</th><th>ϵ_0</th></tr>
<tr><td>α_1</td><td>$R\alpha_2$</td><td></td><td></td></tr>
<tr><td>α_2</td><td>$\epsilon_0 L\alpha_3$</td><td>R</td><td></td></tr>
<tr><td>α_3</td><td>Sq_1</td><td>L</td><td></td></tr>
<tr><td>q_i</td><td colspan="2">F</td><td>$0R\beta_i$</td></tr>
<tr><td>β_i</td><td>$\epsilon_0 Lq_i$</td><td>$\epsilon_0 Lq_i$</td><td></td></tr>
</table>

In this table, the action performed by the machine under state **i** while reading x in the scanned cell is denoted $yM\mathbf{j}$ where $y =$ output symbol, $M =$ move performed by the machine head, $\mathbf{j} =$ new state of the head. Move M is R, L or S according to whether the head goes to the right cell, the left one or stays on the same cell. If $y = x$, y is omitted. Same convention for the new state.

Figure 1: Table of machine F_1.

and

$$F_{1,end} \cap \{\alpha_1 01^n 0x_1 \dots x_r \ : \ n \notin \mathrm{dom}\, f, x_i \in \{0,1\}, r \in \mathbb{N}\} = \varnothing.$$

Consider the mapping which associates configuration $uq_iv\epsilon_0$ of machine F_1 to configuration uq_iv of machine F. This defines a function θ through function $\mathcal{G}$ such that $x \in \mathcal{G}(F_{end}) \Leftrightarrow \theta(x) \in \mathcal{G}(F_{1\,end})$. This proves that machine F_1 is universal.

Theorem 4. *There is a non-recursive set of codes on which machine F_1 is universal.*

PROOF. — Let χ_1 be an encoding of all configurations of F_1. As F_1 is universal, there is a total recursive function θ such that if $V_n = \{x \ : \theta(n,x) \in \chi_1(F_{1\,\mathrm{end}})\}$, then the set of V_n's contains all recursively enumerable sets. Let θ^* be the function obtained from θ as in the proof of theorem 3. Let $K = \{\theta^*(n,x)\}$. K is a set of configuration encodings and, by construction of θ^*, machine F_1 is universal on K. Let us prove that

K is a non recursive recursively enumerable set. It is obvious that K is a recursively enumerable set as a range of a partial recursive function.

Assume K recursive.

Let then $q_1 01^u$ be a configuration on which machine F halts. In that case, $\alpha_1 01^u \in F_{1\,end}$. Thus, there are natural numbers n and x such that $\theta^*(n,x) = \chi_1(\alpha_1 01^u) \in \mathcal{G}(F_{1\,end})$ since K contains $\mathcal{G}(F_{1\,end})$ according to theorem 3. Consequently, $\chi_1(\alpha_1 01^u) \in K$.

Assume now that F does not halt on $q_1 01^u$. Machine F_1 neither halts on $\alpha_1 01^u 0$. Assume that $\chi_1(\alpha_1 01^u) = \theta^*(n,x)$ for some couple n, x. As $\chi_1(\alpha_1 01^u) \notin \mathcal{G}(F_{1\,end})$, $\theta^*(n,x) = \theta(n,x)$ by construction of θ^*.

In the proof of the universality of machine F_1, function θ can be taken as equal to $\chi_1(\alpha_1 01^{g(n,x)}0)$, where $g(n,x)$ is associated to the m-completeness of dom f. Thus, $\theta^*(n,x)$ is the encoding of some configuration $\alpha_1 01^v 0$. Now, we assumed that $\theta^*(n,x) = \chi_1(\alpha_1 01^u)$. As χ_1 is a one-to-one mapping, there is a contradiction.

Consequently, if F_1 does not halt on $\alpha_1 01^u$, $\chi_1(\alpha_1 01^u) \notin K$. Then if K were recursive, the halting of F on configurations $q_1 01^u$ would be decidable, which is impossible. □

5. Machines Computing a Function on a Set

Consider now a Turing machine Z on alphabet A. Fix a word w of A^* and observe the computation of Z starting from the configuration $q_1 w$. Say that the machine *goes out* from this word if the sequence of computations leads to a first time t for which $v(t) = |w|$. If the machine does not halt before such a time or if its computation is not interrupted before it, that time can either be finite or infinite.

It then ensues that the computation of Z starting from configuration $q_1 w$ leads to any one of the following five possibilities:

- the machine computation leads to a configuration of the form uq_j with $|w| = |u|$, which is the case of an exit;
- the machine computation endlessly goes on within the space defined by

word w: for all time t, $v(t) < |w|$; this is a case when the computation time of the machine on w is infinite;

- configuration q_1w is a halting configuration; in that case, say that the computation of Z on w is *determined*;
- the machine computation halts and the head did not visit all cells with an address less than $|w|$;
- the machine computation is interrupted under state q_j on a cell containing x_i because the program of Z contains no instruction associated to couple q_jx_i.

Each of these possibilities is recursively enumerable and no other one can occur. Among these possibilities, call *undetermined* any one of them which is different from the case of a *determined* computation.

Give a special status to the first one and the third one of the above listed cases by setting:

Definition 5. *Let Z be a Turing machine on alphabet A. Call* domain of applicability *of Z the set A_Z of words w of A^* such that the machine computation starting from configuration q_1w leads to a configuration of the form uq_j with $|w| = |u|$.*

In the case when the computation on a word of the applicability domain goes out of that word, there are several ways of going on the computation. As in the classical definitions, one can consider that the machine meets a cell containing the blank, which boils down to take $uq_j\sqcup$ as a starting configuration. One can also consider that a cell is appended to the tape or that a letter is appended to word uq_j. This time, that letter or cell content can be any letter of alphabet A. Intuitively, the computation of Z on w has been *extended*. In order to make this notion more precise, if w is a finite or infinite word on A and k is a natural number with $0 \leqslant k \leqslant |w|$, denote by $w|_k$ the unique word u of length k such that $w = uv$, for some word v on A. It is now possible to state the following definition:

Definition 6. *Let Z be a Turing machine on alphabet A and $w \in A_Z$. Call* extension of the computation on w *any finite or infinite word y on A such that for all k with $0 \leqslant k \leqslant |y|$, $w(y|_k) \in A_Z$.*

This allows us to introduce the notion of computing a function on a set by a Turing machine.

Definition 7. *Let Z be a Turing machine on alphabet A and E be a set of finite words on A. Say that Z* computes a function *on E if and only if set E satisfies to the following conditions:*

(*a*) $E \subset A_Z$;

(*b*) *if the computation of Z on w is undetermined, w has at most one extension;*

(*c*) *for all $w \in A_Z$:*

- *a either $w \in E$;*
- *or there is $y \in E$ such that w is a non empty prefix of y;*
- *or there is $y \in E$ such that y is a strict prefix of w.*

This definition coincides with the traditional one of the computation of a partial recursive function by a Turing machine. Indeed, the rôle of set E is then played by the configurations encoding natural numbers or n-tuples of natural numbers. The definition we suggest here deals with the status which should be given to the undetermined computations which do not lead to a halting configuration. In that case, the domains of definition of computed functions as well as the functions themselves may have properties different from the classical ones.

Theorem 8. *Whatever set E is, machine F_1 computes no function on E.*

PROOF. — Let E be a set and assume that F_1 computes a function on E. Then $E \subset A_{F_1}$ and E has properties (b) and (c) of definition 7. Let m be such that F_1 does not halt on $q_1 01^m 0$ and that F_1 does not cycle during that computation (this can always be assumed). Then $q_1 01^m 0 \in A_{F_1}$ and for all $x \in \{0,1\}^*$, $q_1 01^m 0x \in A_{F_1}$. If $q_1 01^m 0x$ extends a word of E, this is also the case for all $q_1 01^m 0xy$ for all $y \in \{0,1\}^*$. This is a contradiction with the uniqueness of extension (point (b) on above definition). For the same reason, $q_1 01^m 0x$ cannot be extended by a word of E and it cannot neither extend a word of E. □

Let $S_n = \{01^n0x \ : \ |x| + n + 2 = s(n), x \in \{0,1\}^*\}$ for $n \in \mathrm{dom}\, f$ and set $S = \bigcup_{n\in\mathrm{dom}\, f} S_n$. Notice that, by construction of F_1, $S \subset A_{F_1}$.

Lemma 9. *If G is a machine computing a function on E with $A_G \subset \{0,1\}^*$, then $S \not\subset A_G$.*

PROOF. — Assume G to be a machine computing a function on E with $A_G \subset \{0,1\}^*$ and $S \subset A_G$. There is a shortest x in word length such that:

- either the computation of G on 01^n0x is determined,
- or 01^n0y with $|y| > |x|$ has several extensions in A_G.

If the computation on 01^n0x is determined, we have:

$$01^n0xw \notin A_G \tag{1}$$

for $|w| \geqslant 1$.

If F_1 does not halt on 01^n0y for some y such that $|y| \leqslant |x|$, then it does not halt: otherwise, the machine would halt on any word $01^n0xz \in S_n$ with $|z| \geqslant 1$. Since $S \subset A_G$ is assumed, that would be a contradiction to (1).

If 01^n0x has a unique extension in A_G, consider 01^n0y with $y = x\epsilon$, where $\epsilon \in A$, the machine alphabet, and ϵ is chosen in a such way that 01^n0y is not the restriction of the extension of 01^n0x in A_G. Then, by uniqueness of the extension in A_G, the computation of G on 01^n0y is not determined. If the computation of F_1 on 01^n0y would ultimately halt, then that computation would be determined on 01^n0yz for some z with $|z| \geqslant 1$. But then, as $S \subset A_G$, 01^n0yz would be in A_G, a contradiction with the uniqueness of the extension of 01^n0x in A_G.

If F_1 halts on 01^n0, there is an x such that $01^n0x \in S$ and so, the computation of G on 01^n0x has a unique extension in A_G. Necessarily, $|x|$ is unique with this property.

If F_1 does not halt on 01^n0, it does not halt neither on any 01^n0y, whatever y is in A^*. Now, there is a shortest x such that either the computation of G is determined on x or 01^n0y, with $|y| > |x|$ has several extensions in A_G.

As A_G is recursive by construction, it can be seen that those computations make it possible to decide the halting of F_1 on words 01^n0, which is impossible. □

As a first corollary of the lemma, we get the following result:

Theorem 10. *There is no machine G with $A_G \subset \{0,1\}^*$ which would compute a function an a set E containing $\{\alpha_1 01^n 0x\}_{end} = F_{1\,end} \cap \{\alpha_1 01^n 0x\}$.*

PROOF. — Indeed, $S \subset \{\alpha_1 01^n 0x\}_{end}$. If G computes a function on E containing $\{\alpha_1 01^n 0x\}_{end}$, we get that $S \subset A_G$ since $E \subset A_G$, by definition of the notion of computation on E. Then $S \subset A_G$, which is a contradiction with the lemma. □

As a second corollary, we get:

Theorem 11. *There is no machine G with $A_G \subset \{0,1\}^*$ which would compute a function on a set E and such that G_{end} would contain $\{\alpha_1 01^n 0x\}_{end}$.*

PROOF. — If G would exist with $G_{end} \supset \{\alpha_1 01^n 0x\}_{end}$, it could be inferred that $G \subset A_G$ as far as $S \subset \{\alpha_1 01^n 0x\}_{end}$. □

6. Computations with the Empty Tape

Let us turn back to the notion of computing a function on a set by a Turing machine.

Definition 12. *Say that set E is a* functional domain *if there is a Turing machine which computes a funtion on E.*

The following result can then be proved:

Theorem 13. *Any recursively enumerable functional domain is recursive.*

PROOF. — Let Z be a Turing machine computing a function on functional domain E assumed to be recursively enumerable. By construction, A_Z, the applicability domain of Z, is recursive.

By construction, if $x \notin A_Z$, then $x \notin E$. Consider now $x \in A_Z$. By enumerating the words of E, we get either a $y \in E$ such that $y = x$, or a $y \in E$ such that x is a prefix of y or such that y is a prefix of x. □

Definition 14. *Let Z be a Turing machine on alphabet A. Say that Z* computes a funtion on E with the empty tape *if and only if Z computes a function on E and if any word in A_Z with a prefix in E is of the form $xx_0{}^n$ where $x \in E$ and $x_0 \in A$ is a fixed symbol considered as the blank.*

The following result can then be proved:

Theorem 15. *There is a Turing machine Z and a non recursively enumerable set E such that Z computes a function on E with the empty tape.*

Proof. — Consider machine F, constructed in §4 which transforms the starting configuration $q_1 01^n$ into the final configuration $q_0 01^{f(n)}$, where f is a partial recursive function whose domain of definition is m-complete. As in §4, let function $s(n)$ denote the space of the computation of F on data n. Take as E the set constituted of the union of the set of 01^n0's for $n \notin \mathrm{dom} f$ and of the set of $01^n0^{s(n)-n-2}1$'s for $n \in \mathrm{dom}\, f$. Set E is not recursively enumerable: otherwise, enumerating its elements would provide a decision algorithm for membership to $\mathrm{dom}\, f$ which is not a recursive set.

Starting from machine F, one constructs machine F_2 whose program is given in Figure 2 (page 81).

Notice that, contrary to machine F_1, machine F_2 has a different behaviour according to the symbol which it meets beyond the right end of the initial word: either 1 or 0. The machine writes, correspondingly, ϵ_1 or ϵ_0. When later it again meets ϵ_0, the machine behaves as machine F_1. If then it meets ϵ_1, it halts.

Notice that A_{F_2} contains $\alpha_1 01^n$ for all n. Notice also that the latter set also contains $\alpha_1 01^n 0^k$ for all k if $n \notin \mathrm{dom}\, f$, for all k with $1 \leqslant k \leqslant s(n)$ if $n \in \mathrm{dom}\, f$. In that latter case, A_{F_2} also contains $\alpha_1 01^n 0^k$. Moreover, A_{F_2} contains precisely the just indicated elements. It is then clear that $E \subset A_{F_2}$ and that E also satisfies properties (c) of definition 7. Besides, if

	0	1	ϵ_0	ϵ_1
α_1	$R\alpha_2$			
α_2	$\epsilon_0 L\alpha_3$	R		
α_3	Sq_1	L		
q_i	F		$0R\beta_i$	$oS\alpha_*$
β_i	$\epsilon_0 Lq_i$	$\epsilon_1 Lq_i$		

Figure 2: Table of machine F_2.

$x \in A_{F_2}$ has a strict prefix in E, it then ensues from the above definition of A_{F_2} that x is of the form $\alpha_1 01^n 0^k$. Consequently, here 0 plays the rôle of blank. Thus, machine F_2 computes a function on E with the empty tape. □

That result can be reformulated in a different way, from "the point of view of the function".

Let Z be a Turing machine computing a function on set E with the empty tape. Fix a numbering $\delta(x)$ of all the words on alphabet X. Define E_{end} as the set of words of E on which the computation of the machine is either determined or has a finite extension. Consider function $f_{Z,\delta}(n)$ defined as follows:

- if $n \in \delta(E_{\text{end}})$, then $f_{Z,\delta}(n)$ is defined and $f_{Z,\delta}(n) = \delta(w)$ where w is the result of the computation of Z on $\delta^{-1}(n)$;
- if $n \notin \delta(E_{\text{end}})$, then $f_{Z,\delta}(n)$ is not defined.

This definition of a function computed by a Turing machine coincides with the traditional definition of the computation of a partial recursive function by a Turing machine. We can state:

Theorem 16. *For any partial recursive function φ, there is a Turing machine Z such that machine Z computes φ on the set of configurations $q_1 01^n 0$ with the empty tape.*

	0	1	ϵ_0
α_1	$R\alpha_2$		
α_2	$\epsilon_0 L\alpha_3$	R	
α_3	Sq_1	L	
q_i	M		$0R\beta_i$
β_i	$\epsilon_0 Lq_i$		

Figure 3: Table of machine M_2.

PROOF. — Let φ be a partial recursive function. Denote by M a Turing machine on $\{0,1\}$ in the classical sense, transforming any configuration $q_1 01^n 0$ where $n \in \operatorname{dom}\varphi$ into configuration $q_0 01^{\varphi(n)} 0$ and which does not halt its computation on any configuration $q_1 01^n 0$ such that $n \notin \operatorname{dom}\varphi$, where q_1 is the initial state of M and q_0 its final state. Such a machine is easily obtained through the construction performed in [2].

Let then M_2 be the machine whose program is given by Figure 3.

It is then clear that the applicatibility domain of machine M_2 is the set of configurations of the form $q_1 01^n 00^k 1$. The machine computation is interrupted when the head reads symbol 1 under any one of states β_i, as there is no corresponding instruction. For $f_{Z,\delta}$, the computation gives φ by taking any recursive encoding of the words on $\{0,1\}$ as well as its inverse. □

Let us now turn back to machine F_2 of theorem 15. If we take the following mapping on E for numbering δ:

$$\delta(q_1 01^n 0) = n$$
$$\delta(q_1 01^n 0^{s(n)-n-2} 0) = n, \text{if } n \in \operatorname{dom} f$$

then F_2 computes f on E with the empty tape. Notice that on this set, the just defined numbering δ has no recursive inverse on $\mathbb{N}$. However, δ has an inverse on set E_{end}, which can be extended to a recursive function.

Acknowledgements

This work has been realized thanks to the funding of the cooperation between both authors by Metz University Institute of Technology. Both authors also acknowledge INTAS project 97–1259 for helping them to enhance their cooperation.

References

[1] M.D. Davis. A note on universal Turing machines. In C. Shannon and J. McCarthy, editors, *Automata Studies*, pages 167–175. Princeton University Press, 1956.

[2] S.C. Kleene. *Introduction to Metamathematics.* van Nostrand, New York, 1952.

M. Margenstern
GIFM - Université de Metz
I.U.T. de Metz
France
e-mail: margens@iut.univ-metz.fr

L. Pavlotskaïa
Institut d'ingénierie énergétique de Moscou
Russia
e-mail: vm@mpei-rt.msk.su

Cahiers du Centre de logique
Volume 11

On Extensions of Presburger Arithmetic

by

F. Point [1]
Université de Mons-Hainaut

1. Introduction

Let $R = (r_n)$ be a strictly increasing sequence of natural numbers with $r_0 = 1$. We denote by the same letter, the unary predicate $R(x)$ expressing that x belongs to the sequence R; let S be the successor function on R i.e. $S(r_n) = r_{n+1}$. We will consider two kinds of extensions of Presburger arithmetic, namely $\langle \mathbb{N}, +, R \rangle$ and $\langle \mathbb{N}, <, R \rangle$ and under certain conditions on R, we will prove relative quantifier elimination results.

Let n be a natural number, there exists $k \geqslant 0$ such that $r_k \leqslant n < r_{k+1}$; write $n = x_k \cdot r_k + q_k$ with x_k, q_k in $\mathbb{N}$ and $q_k < r_k$. We associate in this way a finite sequence of non zero natural numbers $x_{k_m} 0^{i_m} x_{k_{m-1}} \dots x_{k_0}$ to n, where $x_{k_m} = x_k$ and 0^{i_j} denotes a sequence of i_j zeros i_j natural number. Let us denote the set of those finite sequences by $L(R)$. Let V_R be the unary function sending n to r_{k_0}, the least r_j appearing in that decomposition of n in base R (obtained using the Euclidean algorithm), let λ_R be the unary function sending n to r_k, the largest r_j appearing in that decomposition i.e. $r_k \leqslant n < r_{k+1}$ and finally let f_R be the function

1. Senior Research Associate, F.N.R.S.

sending each natural number n to the n^{th} element r_n of the sequence R.

For R a sparse unary predicate, A.L. Semenov proves that the theory $\mathrm{Th}\langle \mathbb{N}, +, R, S\rangle$ is existential (see [12, Theorem 3]). His proof is syntactic, here we will give a model theoretic proof, using the same strategy as van den Dries when proving that $\langle \mathbb{N}, +, \dot{-}, <, 0, 1, \cdot/n; n \in \omega - \{0\}, \lambda_2\rangle$ admits quantifier elimination (see [16] and [15]). The condition of (almost) sparsity on a sequence R in order to get $\mathrm{Th}\langle \mathbb{N}, +, R, S\rangle$ existential is not necessary as shows the example of the sequence $R = (n + 2^n)$ which is not sparse but such that $\mathrm{Th}\langle \mathbb{N}, +, (n + 2^n), S\rangle$ is existentially bi-interpretable with $\mathrm{Th}\langle \mathbb{N}, +, n \mapsto 2^n\rangle$, the later theory being existential (see [13]).

Then, we will discuss under which conditions a sequence R is (almost) sparse. However, we will always assume that the sequence (r_{n+1}/r_n) has a limit. If this limit is infinite, then the sequence is sparse as Semenov shows (see [12]). We will show that if R satisfies first the A. Bertrand conditions which imply that asymptotically it looks like the sequence of powers of a real number θ (strictly greater than 1) and second a linear recurrence whose characteristic polynomial is the minimal polynomial of θ, then it is sparse and we will point out why it may fail to be sparse if it does not satisfy a linear recurrence. Now, let us discuss a difference between sparse and non sparse sequences.

Recall that a Pisot number is a real algebraic number strictly greater than 1 whose algebraic conjugates are of modulus strictly smaller than 1. If the sequence R satisfies the A. Bertrand conditions and a linear recurrence whose associated polynomial is the minimal polynomial of a Pisot number, then the theory of $\langle \mathbb{N}, +, V_R\rangle$ is decidable (see [10] and in [9, Theorem 2]). Is it true that for all effectively sparse sequences R such that the theory of $\langle \mathbb{N}, +, R\rangle$ is decidable, we have that the theory of $\langle \mathbb{N}, +, V_R\rangle$ is decidable?

If $L(R)$ contains all finite words of the form $10^{n_1}10^{n_2}10^{n_3}\dots$ with $n_i \geqslant n$, for some fixed natural number $n(*)$, then $\mathrm{Th}\langle \mathbb{N}, +, V_R, f_R\rangle$ is undecidable (see [7] and Proposition 8). A corollary of this undecidability result is the undecidability of $\mathrm{Th}\langle \mathbb{N}, +, V_{(n+2^n)}\rangle$, which contrasts with the decidability of $\mathrm{Th}\langle \mathbb{N}, +, V_{(2^n)}\rangle$ (see [5], [10] and [3]). (Unlike the above results the decidability proof of this theory uses automata theory and its quantifier complexity is still unknown.) Note that if the sequence R satisfies the A. Bertrand conditions, then $L(R)$ has the above property $(*)$.

Finally, we will examine the theory of $\langle \mathbb{N}, <, R, S\rangle$ which has also been considered by A.L. Semenov (see [12, Theorem 2]). Here, again by model theoretic means, we will obtain that, under certain conditions over R, this theory is existential.

2. $\langle \mathbb{N}, +, \boldsymbol{R}\rangle$

First, let us set up some notations. We will work in the following language: $L = \{+, \dot{-}, <, 0, 1, \cdot/n; n \in \omega - \{0\}, \lambda_R, S, S^{-1}\}$, where

$$\begin{array}{rcl} x \dot{-} y & = & 0 \quad \text{if } x \leqslant y, \\ & = & x - y \quad \text{otherwise}, \\ x/n & = & y \text{ iff } \bigvee_{k=0}^{n-1} x = n.y + k, \text{ where } n.y = \underbrace{y + \cdots + y}_{n \text{ times}} \\ \lambda_R(x) & = & r_n \text{ iff } r_n \leqslant x < r_{n+1} \text{ and } \lambda_R(0) = 0, \\ S(x) & = & x \text{ iff } r_n < x < r_{n+1}, S(0) = 0, \text{ and} \\ S(r_n) & = & r_{n+1},\, n \geqslant 0, \text{ and } S^{-1}(r_n) = r_{n-1},\, n > 0, \\ S^{-1}(1) & = & 1. \end{array}$$

In our proof, we will use the following two well known results of Presburger (see for instance [8, §3.2 Theorems 32A and 32E]):

(a) $\mathrm{Th}\langle \mathbb{N}, +, \dot{-}, <, 0, 1, ./n; n \in \mathbb{N}\rangle$ is decidable and admits quantifier elimination in $L_P = \{+, \dot{-}, <, 0, 1, ./n; n \in \mathbb{N}\}$.

(b) $\mathrm{Th}\langle R, <, 1, S\rangle$ admits quantifier elimination and so does $\mathrm{Th}\langle R, <, 1, S, S^{-1}\rangle$ where $S^{-1}(x) = y$ iff $(x = 1 \wedge y = 1) \vee (x > 1 \wedge Sy = x)$. Let T_d be a universal axiomatisation of this theory.

First, let us recall the notion of a sparse predicate. We need the notion of an operator A on R, it is any expression of the form $a_n S^n y + \cdots + a_0 S^0 y$, where $R(y)$ and $a_i \in \mathbb{Z}$, $0 \leqslant i \leqslant n$. The predicate R is sparse if for any operator A on R, the following holds:

(a) either $A = 0$ for all y in R, or $A >_{pp} 0$, or $A <_{pp} 0$, where $A >_{pp} 0$ (respectively $A <_{pp} 0$) means that for all but finitely many y in R, $A(y) > 0$ (respectively $A(y) < 0$).

(b) If $A >_{pp} 0$, then there exists a natural number Δ such that for all y in R, $A(S^{\Delta}y) - y > 0$.

By an abuse of notation, we will consider "terms" of the form $\sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i$, where $n_i \in \mathbb{N} - \{0\}$, $m_i \in \mathbb{Z}$, $m_0 \neq 0$.
We will interpret $x + (-1) \cdot y$ by $x \dot{-} y$ if $x \geqslant y$ and by 0, otherwise.
We have $\displaystyle\sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i = \sum_{i \geqslant 0} \left(\frac{m_i}{n_i} \cdot S^{-i}(x) - q_i \cdot \frac{m_i}{n_i} \right)$,
where $0 \leqslant q_i < n_i$, $S^{-i}(x) \equiv q_i \pmod{n_i}$ and $\frac{m_i}{n_i}$ is the quotient of m_i by n_i. We will write down a set T of axioms which amount to put the pair of conditions above on operators A except that we will replace "$A = 0$" by "$A =_{pp} 0$". We will call a sequence satisfying those conditions almost sparse. We will show that T admits quantifier elimination in the language L and in case T is effective, T will be decidable.

Let T be the following set of axioms:

(A1) T_P, a set of universal axioms in the language $\{+, \dot{-}, <, 0, 1\}$ for the theory of abelian, discretely ordered, simplifiable semi-groups with neutral element 0 and 1 as the least element strictly greater than 0, plus the following set of axioms:

$$\text{for each } n,\ n \in \omega - \{0\}, \forall x \forall y \Big(x/n = y \Leftrightarrow \bigvee_{k=0}^{n-1} x = n \cdot y + k \Big).$$

Let us abbreviate the atomic formulas $\lambda_R(x) = x$ by $R(x)$, $x = n \cdot (x/n)$ by $x \equiv 0 \pmod{n}$ and $\displaystyle\prod_j n_j \cdot t \gtreqqless \sum_{i \geqslant 0} \Big(\prod_{j \neq i} n_j \Big) m_i \cdot t_i$ by $\displaystyle t \gtreqqless \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot t_i$.

(A2) $R(1) \wedge \forall x((x \geqslant 1 \wedge R(x)) \rightarrow (x < S(x) \wedge \forall y((y > x \wedge R(y)) \rightarrow y \geqslant S(x))) \wedge R(S(x)))$.

(A3) $\forall x[S^{-1}(S(x)) = x \wedge ((x = 1 \wedge S^{-1}(1) = 1) \vee ((x > 1 \wedge R(x)) \rightarrow (S^{-1}(x) < x \wedge S(S^{-1}(x)) = x)))]$.

(A4) $\forall x((\lambda_R(0) = 0 \wedge S(0) = 0 \wedge S^{-1}(0) = 0) \wedge$
$(x \geqslant 1 \rightarrow (\lambda_R(x) \leqslant x < S(\lambda_R(x)) \wedge \lambda_R(\lambda_R(x)) = \lambda_R(x) \wedge$

$\forall y(\lambda_R(x) < y < S(\lambda_R(x)) \rightarrow$
$(\lambda_R(y) = \lambda_R(x) \wedge S(y) = y \wedge S^{-1}(y) = y))))).$

(A5) For each natural number $n \neq 0$, we add the following axiom scheme in order to ensure that the sequence R is periodic, modulo n. For some constant $k(n)$, $\forall x \forall y(x > k(n) \wedge y > k(n) \wedge R(x) \wedge R(y) \rightarrow (\bigwedge_{z \in E} S^z x \equiv \ell_z (\operatorname{mod} n) \rightarrow \bigwedge_{z \in E} S^{m+z}(y) \equiv \ell_z (\operatorname{mod} n)))$ where $S^0(x) = x$, E varies over finite subsets of integers and m is a positive integer that might depend on n and E.

(A6) Let $n_i \in \mathbb{N} - \{0\}$, $m_i \in \mathbb{Z}$, $m_0 \neq 0$, and suppose $\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(x) > 0$ for $x > k(\mathbf{m}, \mathbf{n})$, where $\mathbf{m}$ and $\mathbf{n}$ stand respectively for the finite sequences (m_i), (n_i) then there exists $t \in \mathbb{N}$, such that:

(i) $\forall x(x > k(\mathbf{m}, \mathbf{n}) \wedge R(x) \rightarrow$
$(S^{t-1}(x) \leqslant \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(x) < S^t(x) \wedge$
$S^{t-1}(x) \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i < S^t(x)) \quad \vee$
$(\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(x) = S^{t-1}(x) \wedge \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i > 0 \wedge$
$S^{-i}(x) \equiv d_i \pmod{n_i} \wedge$
$S^{t-2}(x) \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i < S^{t-1}(x)) \quad \vee$
$(\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i \leqslant 0 \wedge S^{-i}(x) \equiv d_i (\operatorname{mod} n_i) \wedge$
$S^{t-1}(x) \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i < S^t(x))).$
Moreover,

(ii) $\forall x((x > k(\mathbf{m}, \mathbf{n}) \wedge R(x) \wedge S^{t-1}(x) = \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(x)) \rightarrow$
$(\forall y(y > k(\mathbf{m}, \mathbf{n}) \wedge R(y) \wedge S^{t-1}(y) = \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(y)))).$

(iii) there exists j_0 such that for all $j > j_0$, and there exists $k' > k$ such that

(iiia) $\forall x(x > k'(\mathbf{m}, \mathbf{n}) \wedge R(x) \rightarrow$
$S^{t-1}(x) \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i + S^{-j}(x) < S^t(x)).$

$$\begin{aligned}&\text{(iiib)}\ (\forall x((x > k'(\mathbf{m},\mathbf{n}) \wedge R(x) \wedge S^{t-1}(x) < \textstyle\sum_{i\geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(x)) \rightarrow \\ &\quad S^{t-1}(x) < \textstyle\sum_{i\geqslant 0} m_i \cdot S^{-i}(x)/n_i - S^{-j}(x) < S^t(x))) \quad \wedge \\ &\quad (\forall x((x > k'(\mathbf{m},\mathbf{n}) \wedge R(x) \wedge S^{t-1}(x) = \textstyle\sum_{i\geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(x)) \rightarrow \\ &\quad S^{t-2}(x) \leqslant \textstyle\sum_{i\geqslant 0} m_i \cdot S^{-i}(x)/n_i - S^{-j}(x) < S^{t-1}(x))).\end{aligned}$$

First, we will show that T is model-complete. This will follow from the three following Lemmas.

Lemma 1. *Let $\mathcal{M}$ be a model of T. Then, for each a, b, c in $\mathcal{M}$, and each pair of non zero distinct natural numbers n, m, we have:*
$(a/n)/m = a/(n \cdot m)$ and $a/n = (a \cdot m)/(m \cdot n)$, if

$$\begin{aligned} a > b,\ a/n \dot{-} b/n &= (a \dot{-} b)/n \text{ if } a - a/n \cdot n \geqslant b - b/n \cdot n, \\ &= (a \dot{-} b)/n + 1, \text{ otherwise,} \\ a/n + b/n &= (a+b)/n - 1 \text{ if } (a - a/n \cdot n) + (b - b/n \cdot n) \geqslant n, \\ &= (a+b)/n, \text{ otherwise,}\end{aligned}$$

suppose $b \geqslant m$, $b \equiv k \pmod m$, $0 \leqslant k < m$,
if $a + n \cdot b/m = c$, then $b = (mc \dot{-} ma)/n + k$ and
if $a \dot{-} n \cdot b/m = c$, then $b = (ma \dot{-} mc)/n + k$.

Let $\mathcal{A}, \mathcal{B}$ be models of T with $\mathcal{A} \subset \mathcal{B}$ and let b belong to $R(\mathcal{B}) - \mathcal{A}$.
Set $\mathcal{A}_- = \{a \in \mathcal{A} : a < b\}$, $\mathcal{A}_+ = \{a \in \mathcal{A} : b < a\}$.

Lemma 2. *For all $z \in \mathbb{Z} - \{0\}$, $\mathcal{A}_- < S^z b < \mathcal{A}_+$.*

Proof. — By the way of contradiction, suppose that there exists a in $\mathcal{A}$ such that $b < a < S(b)$, since $\mathcal{A}$ is closed under λ_R, we have $b < \lambda_R(a) \leqslant a < S(b)$ which contradicts axiom (A4).

In the same way, we have that if $S^{-1}(b) < a < b$, for some a in $\mathcal{A}$, then $S^{-1}(b) < \lambda_R(a) < b$. □

Lemma 3. *Let $n_i \in \mathbb{N} - \{0\}$, $m_i \in \mathbb{Z} - \{0\}$ and suppose $\sum_{i\geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(b) > 0$ and $S^{-i}(b) \equiv d_i (\operatorname{mod} n_i)$.*

First, let $a \in \mathcal{A}_+$, then:

$$\begin{aligned}
\lambda_R(a + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i) &= \lambda_R(a), \textit{if } S(\lambda_R(a)) - a \in \mathcal{A}_+, \\
&= S(\lambda_R(a)) \textit{if } S(\lambda_R(a)) - a \in \mathcal{A}_- \\
\lambda_R(a - \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i) &= \lambda_R(a), \textit{if } a - \lambda_R(a) \in \mathcal{A}_+, \\
&= S^{-1}(\lambda_R(a)) \textit{if } a - \lambda_R(a) \in \mathcal{A}_-.
\end{aligned}$$

Second, suppose that $a \in \mathcal{A}_-$.

Either $S^{t-1}(b) < \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(b) < S^t(b)$, then

$$\lambda_R(a + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a) = S^{t-1}(b),$$

Or $S^{t-1}(b) = \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(b)$, then

if $\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i \geqslant 0$, then $\lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a) = S^{t-2}(b)$,

$$\begin{aligned}
\lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i + a) &= S^{t-2}(b) \textit{ if in addition } a < \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i, \\
&= S^{t-1}(b) \textit{ if in addition } a \geqslant \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i,
\end{aligned}$$

if $\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i < 0$, then :

$$\begin{aligned}
\lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i + a) &= S^{t-1}(b), \\
\lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a) &= S^{t-1}(b) \textit{ if in addition } a > -\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i, \\
&= S^{t-2}(b) \textit{ if in addition } a < -\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i.
\end{aligned}$$

PROOF. — First note that axiom (A6) implies that $\lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i) = S^t(b)$ for some integer t. So, by Lemma 2, $\mathcal{A}_- < \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < \mathcal{A}_+$.

(a) Let $a \in \mathcal{A}_+$ and
suppose that $S(\lambda_R(a)) - a \in \mathcal{A}_+$, then $\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < S(\lambda_R(a)) - a$
and so $\lambda_R(a) \leqslant a < a + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < S(\lambda_R(a))$,
suppose that $S(\lambda_R(a)) - a \in \mathcal{A}_-$, then $S(\lambda_R(a)) - a < \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i$
and so $S(\lambda_R(a)) < \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i + a$.
We have that $S^{-j_0}(S\lambda_R(a))$ belongs to $\mathcal{A}_+$ and that, by axiom (A6), $S^{-j_0}(S\lambda_R(a)) + S(\lambda_R(a)) < S^2(\lambda_R(a))$,
so $a + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < S(\lambda_R(a)) + S^{-j_0}(S\lambda_R(a)) < S^2\lambda_R(a)$,
suppose that $a - \lambda_R(a) \in \mathcal{A}_+$, then $\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < a - \lambda_R(a)$
i.e. $\lambda_R(a) < a - \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < a < S(\lambda_R(a))$,
suppose that $a - \lambda_R(a) \in \mathcal{A}_-$, then $a - \lambda_R(a) < \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i$,
i.e. $a - \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < \lambda_R(a)$.

By axiom (A6), $S^{-1}\lambda_R(a) + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < a$
(indeed $\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < S^{-j_0-1}(a)$).

(b) Let $a \in \mathcal{A}_-$.
First, assume that $S^{t-1}(b) < \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^i(b)$. So by axiom (A6),
$\lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i) = S^{t-1}(b)$, so $S^{s-1}(b) \leqslant a + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i$.
We have that $\mathcal{A}_- < S^{-j_0}(b)$ by Lemma 2.

By axiom (A6) (iiia), $S^{-j_0}(b) + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i < S^i(b)$.
So, $\lambda_R(a + \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i) = S^{t-1}(b)$. By a similar argument and axiom (A6) (iiib), we have that $\lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a) = S^{t-1}(b)$.
Second, assume that $S^{t-1}(b) = \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(b)$.
Then, as in axiom (A6) we have two subcases.

First, suppose that $\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i \geqslant 0$.

Then, $S^{t-2}(b) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{t-2}(b)) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i)$ and by axiom (A6) (iiib) we have $S^{t-2}(b) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a)$.

If $a < \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i$, then $S^{t-2}(b) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i + a)$ and if $a \geqslant \sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i$, then $S^{s-1}(b) \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i + a$ and by axiom (A6) (iiia), and the same reasoning as before, $\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i + a < S^t(b)$.

Second, suppose that $\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i < 0$.

Then, $S^{t-1}(b) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i)$. So as before, using axiom (A6) (iiia), $S^{t-1}(b) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i + a)$. If $a \leqslant -\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i$, then $S^{t-1}(b) = \lambda_R(\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a)$ and if $a > -\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot d_i$, then $\sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a < S^{s-1}(b)$ and by axiom (A6) (iiib) $S^{t-2}(b) \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(b)/n_i - a$. □

Corollary 4.

(a) *Let a belong to $\mathcal{A}_+$, then $S(a \pm b/n \cdot m) = a \pm b/n \cdot m$ if $a \pm b/n \cdot m$ does not belong to R i.e. if $\lambda_R(a \pm b/n \cdot m) \neq a + b/n \cdot m$ which is always the case if $m \neq 0$.*

(b) *Let a belong to $\mathcal{A}_-$, then unless $a + b/n \cdot m = S^{-t}(b)$, $S(a + b/n \cdot m) = a + b/n \cdot m$ and unless $b/n \cdot m - a = S^{-t}(b)$, $S(b/n \cdot m - a) = b/n \cdot m - a$.*

(c) *Let $\mathcal{A}_- < c < \mathcal{A}_+$ suppose c does not belong to R, then $c/n \cdot m$ does not belong to R.*

PROOF:

(a) $\lambda_R(a \pm b/n \cdot m)$ is either equal to $\lambda_R(a)$ or $S(\lambda_R(a))$ or $S^{-1}(\lambda_R(a))$.

(c) By axiom (A6), $\lambda_R(c/n \cdot m) = S^t(c)$ for some c, but since c does not belong to R, $S(c) = c$. □

Proposition 5. *T is model-complete.*

PROOF. — Using Robinson criterium, it suffices to show that given $\mathcal{A}$, $\mathcal{B}$ two models of T with $\mathcal{A} \subset \mathcal{B}$, then $\mathcal{A} \subset_{ec} \mathcal{B}$. Since T is universal, it suffices to consider the extensions $\mathcal{B}$ of the form $\langle \mathcal{A}, b \rangle$ with $b \in \mathcal{B} - \mathcal{A}$.

We show that $\mathcal{A} \subset_{ec} \mathcal{B}$ in the following two cases:

(a) $R\langle \mathcal{A}, b \rangle = R\langle \mathcal{A} \rangle$,

(b) $R(b)$.

Using Lemmas 2 and 3, one can show that this suffices to ensure that, in all cases, $\mathcal{A} \subset_{ec} \mathcal{B}$. Indeed, suppose b does not belong to R, first consider terms of the form $t = a \pm b/n \cdot m$, with $a \in \mathcal{A}$ and $m, n \in \mathbb{N} - \{0\}$ (note that $S(b/n \cdot m) = b$ (see Corollary 4)). Either t belongs to R and so we are back to case (b) since by Lemma 1, $\langle \mathcal{A}, t \rangle = \langle \mathcal{A}, b \rangle$ or for some a' in $\mathcal{A}$, $n', m', \lambda_R(a' \pm t/n' \cdot m')$ does not belong to $\mathcal{A}$, but t does not belong to R. By Lemma 1, $\langle \mathcal{A}, a' \pm t/n' \cdot m' \rangle = \langle \mathcal{A}, t \rangle = \langle \mathcal{A}, b \rangle$. So, we may as well assume that $\lambda_R(t) \notin \mathcal{A}$ and $t \notin R$. Set $t_1 = \lambda_R(t)$ and consider $\langle \mathcal{A}, t_1 \rangle$. By Lemma 2, the $S^z(t_1)$, $z \in \mathbb{Z}$, are all in the same cut and $\lambda_R(a \pm t_1/n \cdot m)$ only depends on the cut t_1 is in and where a is with respect to this cut. But $\lambda_R(a + t_1/n \cdot m) \leqslant a + t_1/n \cdot m < a + t/n \cdot m < a + S(t_1)/n \cdot m$, and if $a > t_1$, $\lambda_R(a - S(t_1)/n \cdot m) \leqslant a - S(t_1)/n \cdot m < a - t/n \cdot m < a - t_1/n \cdot m$, so $\lambda_R(a \pm t/n \cdot m)$ either belongs to $\mathcal{A}$ or it is equal to $S^z(t_1)$ for a certain z and similarly for $\lambda_R(a \pm t/n \cdot m + \sum_i m_i \cdot S^i(t_1)/n_i)$ (see Lemma 3). So we see that $R\langle \mathcal{A}, t_1, t \rangle = R\langle \mathcal{A}, t_1 \rangle$ and replacing $\mathcal{A}$ by $\langle \mathcal{A}, t_1 \rangle$, we are in case (a).

(a) Since T_P is model-complete, by the existential Frayne theorem, there exists an embedding f of $\langle \mathcal{A}, b \rangle$ in an ultrapower of $\mathcal{A}$ which is the identity on $\mathcal{A}$ and respects L_P (see [6, 4.3.13]). Let us show that f is an L-morphism. Let $t(a, b)$ be an L-term where b is occuring and let $a \in \mathcal{A}$. By hypothesis, $\lambda_R(t(a, b))$ belongs to $\mathcal{A}$. We have by axiom (A4), $\lambda_R(t(a,b)) \leqslant t(a,b) < S(\lambda_R(t(a,b)))$, since f respects the order we have $f(\lambda_R(t(a,b))) \leqslant f(t(a,b)) < f(S(\lambda_R(t(a,b))))$, since f is the identity on $\mathcal{A}$, we have $\lambda_R(t(a,b)) \leqslant f(t(a,b)) < S(\lambda_R(t(a,b)))$. So, by axiom (A4), $\lambda_R(f(t(a,b))) = \lambda_R(t(a,b))$. Consider $S(t(a,b))$, suppose that $\lambda_R(t(a,b)) \neq t(a,b)$ so $\lambda_R(t(a,b)) = f(\lambda_R(t(a,b))) \neq f(t(a,b))$, by definition of S, $S(t(a,b)) = t(a,b)$. So $f(S(t(a,b))) = f(t(a,b)) = S(f(t(a,b)))$. If $\lambda_R(t(a,b)) = t(a,b)$, $t(a,b)$ belongs to $\mathcal{A}$, and so $f(S(t(a,b))) = S(t(a,b)) = S(f(t(a,b)))$.

(b) Since T_d is model-complete, by the existential Frayne theorem, there exists an embedding g of $R(\langle \mathcal{A}, b\rangle)$ in an ultrapower $R(\mathcal{A})^\alpha/U$, where U is a α-regular ultrafilter on α, and $\alpha = \max\{|L|, |\mathcal{B}|\}$, which is the identity on $R(\mathcal{A})$ and respects $\{1, <, S, S^{-1}\}$ (see [6, Corollary 4.3.13]).

Using the embedding g, we define an image $b^\star$ of b in $\mathcal{A}^\alpha/U$ and then we will define an embedding of $\langle \mathcal{A}, b\rangle$ sending b to $b^\star$, $S^z(b)$ to $S^z(b^\star)$, $z \in \mathbb{Z} - \{0\}$, and which restriction to $\mathcal{A}$ will be the identity. The element $b^\star$ will be the realization of the following type $p(x) = \{S^z x \equiv \ell_{j_z} (\text{mod } m)$, $R(\mathcal{A}_-) < x < R(\mathcal{A}_+) : S^z b \equiv \ell_{j_z} (\text{mod } m)$, m is a natural number, $\ell_{j_z} \in \{0, \ldots, m-1\}$, $z \in \mathbb{Z} - \{0\}\}$, where $\mathcal{A}_- = \{a \in \mathcal{A} : a < b\}$ and $\mathcal{A}_+ = \{a \in \mathcal{A} : b < a\}$. By axiom (A5), this type is finitely satisfiable by an element of the form $S^n(g(b))$, for some n, which is in the same cut as b with respect to $\mathcal{A}$ (see Lemma 2). Moreover, since g is a morphism for S and S^{-1}, we have that $S^{n\pm1}(g(b)) = S^n(g(S^{\pm1}(b)))$.

If $n = m$, then $b/n.n = b - k$, for some $0 \leqslant k < n$ and since b and $b^\star$ satisfies the same congruences, we have $b^\star/n.n = b^\star - k$ with the same k. From Lemma 3, we see that we may define $(\Sigma_i m_i.S^i(b)/n_i \pm a)^\star$ as $\Sigma_i m_i.S^i(b^\star)/n_i \pm a$, moreover, since for any z, $S^z(b)$ and $S^z(b^\star)$ satisfy the same congruences, if we apply the function λ_R to it, we obtain a term of the same form where b is possibly replaced by $S^t(b)$. The fact that $\Sigma_i m_i.S^i(b)/n_i \pm a = S^{t'}(b)$ for some t' is enforced by axiom (A6) and the fact that $b > \mathcal{A}_-$. Since $b^\star$ is also strictly greater than $\mathcal{A}_-$, we will also have that $\Sigma_i m_i.S^i(b^\star)/n_i \pm a = S^{t'}(b^\star)$. □

Proposition 6. *Suppose that $\mathcal{N}$ satisfies T, then T axiomatizes $\mathrm{Th}(\mathcal{N})$, T is complete, admits quantifier elimination and is decidable if T is recursive.*

PROOF. — Note that T is a universal set of axioms such that $\mathcal{N}$ embeds in any model of T. We have shown that T is model-complete. This entails that $\mathcal{N}$ is a prime model for T, that T is complete and axiomatises $\mathcal{N}$ and finally that T admits quantifier elimination. The theory $\mathrm{Th}(\mathcal{N})$ will be decidable if T is recursive. □

Corollary 7. *The definable functions are given by finitely many L-terms.*

Proof. — This follows directly, using compactness, from the fact that T is universal and admits quantifier elimination. □

Let us recall the A. Bertrand conditions on a sequence R (see [2]). A. Bertrand has caracterized the sets $L(R)$ (see Introduction) stable by

(B1) left translations i.e. if $x_n x_{n-1} \dots x_h \in L(R)$ then $x_n x_{n-1} \dots x_h 0 \dots 0 \in L(R)$ and

(B2) left truncations i.e. if $x_n x_{n-1} \dots x_h \in L(R)$ then for $m < n$, $x_m x_{m-1} \dots x_h \in L(R)$.

It is shown that those conditions imply the existence of a real number $\theta > 1$, such that $L(R) = L(\theta)$, where $L(\theta)$ is the language associated with the θ-shift and such that $\lim \left(\frac{r_k}{\theta^k}\right)$ is a non zero real number and so $L(R)$ consists of finite words written in a finite alphabet (indeed $\lim r_{n+1}/r_n$ is equal to θ). Note that if in addition the θ-expansion of 1 is almost periodic, then $L(\theta)(= L(R))$ is rational which implies that the sequence R satisfies a linear recurrence (see [14]).

Proposition 8. *Let R be a sequence such that $L(R)$ contains all finite words of the form $10^{n_1}10^{n_2}10^{n_3}\dots$ with $n_i \geqslant n$, for some fixed natural number n $(\star)$. Then, $\mathrm{Th}\langle \mathbb{N}, +, V_R, f_R\rangle$ is undecidable. If R satisfies the A. Bertrand conditions, then R has this property $(\star)$.*

Proof. — The proof of the undecidability of $\mathrm{Th}(\mathbb{N}, +, V_R, f_R)$ is along the same lines that the one for $\mathrm{Th}(\mathbb{N}, +, V_2, n \to 2^n)$ with the difference that in this case not all sequences of 0's and 1's belong to $L(R)$. For x belonging to R, one interprets x belongs to y as "1" in the x position belongs to the normal representation of y in base R and then for x any natural number, x belongs to y iff $f_R(x)$ belongs to y. Then one may define the set of multiples of $x \leqslant z$ as the smallest set v of natural numbers containing 0 and closed by the following relation: if $u < z$ and if u belongs to v, then $u+x$ belongs to to v. Then, if $x \geqslant n$, x divides y, denoted by "$x \mid y$", iff y belongs to the set of mutiples of x less or equal to y.

Suppose that R is a sequence satisfying the A. Bertrand conditions. As in proof of Proposition 15 in [4], we use the criterium of Parry for a sequence w to belong to $L(\theta)$ (see [11]). Let $e(1)$ be the θ-expansion of 1. Then $e(1)$ begins with a letter e strictly greater than 0 and does not end with an infinite sequence of 0's (see [4, section 2.2]). Let $e(1) = e0^{n-1}e' \dots$, with $e, e' \neq 0$, $n \geqslant 1$. Then all sequences of the form $10^{n_1}10^{n_2}10^{n_3}\dots$, with $n_i \geqslant n$, $i > 0$, belongs to $L(R)$. □

Corollary 9. $\mathrm{Th}\langle \mathbb{N}, +, V_{(n+2^n)}\rangle$ *is undecidable.*

Proof. — In order to apply the preceeding Proposition, we directly check that $L(R)$ contains all finite words of the form $10^{n_1}10^{n_2}10^{n_3}\dots$ with $n_i \geqslant n$. (We could have checked the A. Bertrand conditions). We will show that this is the case for $n = 2$. It suffices to check that $2^n + n < 2^n + n + 2^{n-2} + n - 2 + \ldots + 2^3 + 3 < 2^{n+1} + n + 1$. We have $2^{n+1} + n + 1 = 2.2^n + n + 1$, so we have to verify that $2^{n-2} + n - 2 + \cdots + 2^3 + 3 < 2^n + 1$. By induction on n, we suppose that $2^{n-4} + n - 4 + \cdots + 2^3 + 3 < 2^{n-2} + 1$. Write $2^n = 4.2^{n-2}$, we have $n - 2 < 2^{n-2}$ and so get the result. □

Now, we will address the question: for which sequence R, does $\mathcal{N}$ satisfy T ? First, we will assume that in addition our sequence R satisfies one of the hypotheses:

(A) either, there exists $\theta > 1$ such that $\lim_{k\to\infty} \frac{r_k}{\theta^k} = \tau$, where $\tau \in \mathbb{R}^+ - \{0\}$,

(B) or, $\lim_{t\to\infty} \frac{r_k}{r_{k-1}} = \infty$.

As we have seen above, the hypothesis on our sequence that asymptotically it looks like a sequence of powers of θ is implied by the A. Bertrand conditions on R. Moreover, if the sequence R satisfies a linear recurrence, then axiom (A5) holds in the corresponding theory T. Either, we will require that the sequence R is such that $L(R)$ satisfies the A. Bertrand conditions (A) and (B) and that R satisfies a linear recurrence whose characteristic polynomial is the minimal polynomial of θ, let us call these hypotheses $(A)^\star$, or we will assume hypothesis (B). (For a discussion of hypothesis $(A)^\star$, see [4] after Theorems 2 and 5). Note that if the sequence

R satisfies (A)* and if in addition θ is a Pisot number, then decidability of $\mathrm{Th}(\mathcal{N})$ can be proved by automata theory (see [10] and [9, Theorem 2]).

Examples of Sparse Sequences (see [12])

(a) $R =$ Fibonaci sequence i.e. $r_n = r_{n-1} + r_{n-2}$, $r_0 = 1$ et $r_1 = 1$ and the corresponding $\theta = (1+\sqrt{5})/2$.

(b) $R = (2^n)$.

(c) $R = (n!)$.

Example of a Non Sparse Sequence

Let $R = (2^n + n)$. In this case, one may define the function $n \mapsto 2^n$ as follows. Let (n, m) belong to $\mathbb{N}^2$. Then $(n, m) = (n, 2^n)$ iff there exists r in R such that $(n = 2r - Sr + 1$ and $m = Sr - r - 1)$. The theory $(\mathbb{N}, +, n \mapsto 2^n)$ has been proved to be decidable and it admits quantifier elimination in $\{+, -, <, 0, 1, 2^x, \ell(x), \lambda_2(x), D_n; n \in \mathbb{N} - \{0\}\}$ where $2^{\ell(x)} = \lambda_2(x)$ and $\ell(2x) = \ell(x) + 1$ (see [13, Theorem 2] and [7]). One can check directly that axiom (A5) is satisfied. But axiom (A6) (iii) does not hold (see Lemma 11 below).

Set $x = 2^k + k$ and $t(x) = 2x - S(x) + 1 = k$. Then $t(S^n x) = t(x) + n$. We have $x < y$ iff $t(x) < t(y)$. But the only way to insure that if $\mathcal{A}_- < b < \mathcal{A}_+$, then $t(b)$ does not belong to $\mathcal{A}$ is to require that if $t(x) < u < t(y)$, then there exists z in R with $x < z < y \wedge u = t(z)$.

Now, in case of hypothesis (B) i.e. if $\lim_{k\to\infty} \dfrac{r_k}{r_{k-1}} = \infty$, we will rewrite axiom (A6) of the theory T in a more explicit way. Let hypothesis (B)* be hypothesis (B) plus the fact that R satisfies axiom (A5).

First, we note that $\displaystyle\sum_{i\geqslant 0}^{i\leqslant N} \frac{m_i}{n_i} \cdot S^{-i}(x) > 0$ for x large enough, iff $m_0 > 0$.

This implies, in particular, that if $\displaystyle\sum_{i\geqslant 0}^{i\leqslant N} \frac{m_i}{n_i} \cdot S^{-i}(x) = 0$ for x large enough, then all $m_i = 0$. Moreover, $\sum_{i\geqslant 0}^{i\leqslant N} \frac{m_i}{n_i} \cdot S^{-i}(x) > 0$ for x large enough, implies that $\sum_{i\geqslant 0}^{i\leqslant N} \frac{m_i}{n_i} \cdot S^{-i}(x) > \sum_{i\geqslant 0} |m_i|$, for x large enough and so that $\sum_{i\geqslant 0}^{i\leqslant N} m_i \cdot S^{-i}(x)/n_i > 0.(\star)$

PROOF. — Take k large enough such that first $r_k > \frac{2.n_0}{|m_0|} \cdot \sum_{i\geqslant 0} |m_i|$ and then such that $\frac{r_{k-N}}{r_{k-N+1}} < \min\left\{\frac{|m_0|}{2.n_0} \cdot \left(\sum_{i\geqslant 0}\left|\frac{m_i}{n_i}\right|\right)^{-1}, 1\right\}$.
If $m_0 > 0$, then $\sum_{i\geqslant 1}^{i\leqslant N} \frac{m_i}{n_i} \cdot r_{k-i} > \frac{m_0}{n_0} \cdot r_k$.
On the other hand, if $m_0 > 0$ and $\sum_{i\geqslant 0}^{i\leqslant N} m_i \cdot r_{k-i}/n_i > 0$, we get a contradiction. Indeed, $\sum_{i\geqslant 1}^{i\leqslant N} m_i/n_i \cdot \frac{r_{k-i}}{r_k} > \frac{m_0}{n_0}$ which is a contradiction.

Now let us assume that $\sum_{i\geqslant 0}^{i\leqslant N} m_i \cdot r_{k-i}/n_i > 0$, for k large enough. So by the above $m_0 > 0$. We have $\frac{m_0}{n_0} r_k + r_k \left(\sum_{i\geqslant 1}^{i\leqslant N} \frac{m_i}{n_i} \cdot \frac{r_{k-i}}{r_k}\right) > 0$. By hypothesis on r_k and the quotients $\frac{r_{k-N}}{r_{k-N+1}}$, we get the result. The last assertion follows from: let $r_{k-i} \equiv d_i \pmod{n_i}$ with $0 \leqslant d_i < n_i$, then $\sum_{i\geqslant 0}^{i\leqslant N} m_i \cdot S^{-i}(x)/n_i = \sum_{i\geqslant 0}^{i\leqslant N} \frac{m_i}{n_i} \cdot S^{-i}(x) - \sum_{i\geqslant 0}^{i\leqslant N} \frac{m_i}{n_i} \cdot d_i$.

Axiom (A6)

Let $n_i \in \mathbb{N} - \{0\}$, $m_i \in \mathbb{Z}$, $m_0 \neq 0$, and suppose $\sum_{i \geqslant 0} \frac{m_i}{n_o} \cdot S^{-i}(x) > 0$ for $x > k(\mathbf{m}, \mathbf{n})$. In this case this implies that $m_0 > 0$.

(i) Suppose that $\frac{m_0}{n_0} > 1$ then there is $k(\mathbf{m}, \mathbf{n}) = k \in \mathbb{N}$,

$$\forall x (x > k(\mathbf{m}, \mathbf{n}) \wedge R(x) \rightarrow x \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i < S(x)).$$

Set $t = 0$.
Suppose that $m_0/n_0 < 1$ then there $k(\mathbf{m}, \mathbf{n}) = k \in \mathbb{N}$,

$$\forall x (x > k(\mathbf{m}, \mathbf{n}) \wedge R(x) \rightarrow S^{-1}(x) \leqslant \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i < x).$$

Set $t = -1$.

(ii) There exists $k'(\mathbf{n}, \mathbf{m}) > k(\mathbf{n}, \mathbf{m})$, with $k(\mathbf{n}, \mathbf{m})$ as above such that:

$$\begin{aligned} \forall x (x > k'(\mathbf{n}, \mathbf{m}) \wedge R(x) \rightarrow S^{-t}(x) &< S^{-1}(x) + \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i \\ &< S^{-t+1}(x)). \end{aligned}$$

(iii) There exists $k'(\mathbf{n}, \mathbf{m}) > k(\mathbf{n}, \mathbf{m})$, with $k(\mathbf{n}, \mathbf{m})$ as above such that:

$$\begin{aligned} \forall x (x > k'(\mathbf{n}, \mathbf{m}) \wedge R(x) \rightarrow S^{-t}(x) &< \sum_{i \geqslant 0} m_i \cdot S^{-i}(x)/n_i - S^{-1}(x) \\ &< S^{-t+1}(x))). \end{aligned}$$

Now, let us show that this axiom (A6) holds under the hypothesis that $\lim_{n \to \infty} r_n/r_{n-1} = \infty$. Let x belong to R and suppose that $x \equiv q (\operatorname{mod} n_0)$ with $0 \leqslant q < n_0$,

Now, suppose that $\frac{m_0}{n_0} > 1$.

We want to show that

$$x \leqslant m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i \tag{1}$$

and

$$m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i < S(x). \tag{2}$$

(1) is equivalent to $0 \leqslant m_0 \cdot x/n_0 - x + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ or equivalently that $0 \leqslant (m_0 - n_0) \cdot x/n_0 - q + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$. This follows from $(\star)$.

(2) is equivalent to $0 \leqslant S(x) - m_0 \cdot x/n_0 - \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ which holds by $(\star)$.

Suppose that $\dfrac{m_0}{n_0} = 1$. First suppose that $\sum_{i \geqslant 1} \dfrac{m_i}{n_i} \cdot S^{-i}(x) > 0$. (Note that this condition is equivalent to $m_{i_0} > 0$, where i_0 is the smallest i such that $m_i \neq 0$).

We want to show that

$$x \leqslant m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i \tag{3}$$

and

$$m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i < S(x). \tag{4}$$

(3) is equivalent to $0 \leqslant m_0 \cdot x/n_0 - x + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ or equivalently that $0 \leqslant q + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$. This follows from $(\star)$.

(4) is equivalent to $0 \leqslant S(x) - m_0 \cdot x/n_0 - \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ which holds by $(\star)$.

Second suppose that $\sum_{i \geqslant 1} \dfrac{m_i}{n_i} \cdot S^{-i}(x) < 0$.

We want to show that

$$S^{-1}(x) \leqslant m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i \tag{5}$$

and

$$m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i < x. \tag{6}$$

(5) is equivalent to $0 \leqslant m_0 \cdot x/n_0 - S^{-1}(x) + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ which holds by $(\star)$.

(6) is equivalent to $0 \leqslant x - m_0 \cdot x/n_0 - \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ or equivalently that $0 \leqslant q - \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$. This follows from $(\star)$.

Suppose that $\dfrac{m_0}{n_0} < 1$.

We want to show that

$$S^{-1}(x) \leqslant m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i \tag{7}$$

and

$$m_0 \cdot x/n_0 + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i < x. \tag{8}$$

(7) is equivalent to $0 \leqslant m_0 \cdot x/n_0 - S^{-1}(x) + \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ which holds by $(\star)$.

(8) is equivalent to $0 \leqslant x - m_0 \cdot x/n_0 - \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ or equivalently that $0 \leqslant (n_0 - m_0) \cdot x/n_0 - q - \sum_{i \geqslant 1} m_i \cdot S^{-i}(x)/n_i$ where $0 \leqslant q < n_0$. This follows from $(\star)$.

One can show similarly that (i) and (iii) hold since in the discussion above we only use the value of the coefficient of the term in x.

Now we will examine in case of hypothesis (A) i.e. in case there exists $\theta > 1$ such that $\lim_{k \to \infty} \dfrac{r_k}{\theta^k} = \tau$, where $\tau \in \mathbb{R}^+ - \{0\}$, the meaning of axiom (A6). We will examine terms of the form $\sum_{i \geqslant 0} m_i \cdot r_{k-i}/n_i$. We have $r_{k-i}/n_i = r_{k-i} \cdot \dfrac{1}{n_i} - \dfrac{d_i}{n_i}$, with $0 \leqslant d_i < n_i$. Set $q_i = \dfrac{m_i}{n_i}$ and let

$u = \sum_{i \geqslant 0} q_i \cdot \theta^{-i}$, $\overline{u} = \sum_{i \geqslant 0} |q_i| \cdot \theta^{-i}$. The first part of axiom (A6) requires that if (9) and (10) below hold for k_0 large enough, then it holds for all $k > k_0$.

$$r_{k-s-1} \leqslant \sum_{i \geqslant 0} m_i \cdot r_{k-i}/n_i \tag{9}$$

and

$$\sum_i m_i \cdot r_{k-i}/n_i < r_{k-s}. \tag{10}$$

Then if (9) and (10) hold and look for j such that the following holds: (where t may be replaced by $t-1$ in case $\sum_{i \geqslant 0} \frac{m_i}{n_i} \cdot S^{-i}(x) = S^{t-1}(x)$).

$$r_{k-t-1} \leqslant \sum_{i \geqslant 0} m_i \cdot r_{k-i}/n_i + q \cdot r_{k-j} \tag{11}$$

and

$$\sum_i m_i \cdot r_{k-i}/n_i + q \cdot r_{k-j} < r_{k-t}, \quad \text{with } q > 0. \tag{12}$$

$$r_{k-t-1} \leqslant \sum_{i \geqslant 0} m_i \cdot r_{k-i}/n_i - q \cdot r_{k-j} \tag{13}$$

and

$$\sum_{i \geqslant 0} m_i \cdot r_{k-i}/n_i - q \cdot r_{k-j} < r_{k-t}, \text{with } q > 0. \tag{14}$$

Note that in anycase, (9) implies (11) and that (10) implies (14).

Let $\eta = \theta^{-k} \cdot \sum_{i \geqslant 0} |q_i| \cdot d_i \cdot \theta^t$ and note that $\eta = 0$ iff $r_{k-i} \equiv 0 \pmod{n_i}$, for all i.

(9) is implied by: $\theta^{-1} \cdot (\tau + \epsilon) \leqslant (u \cdot \tau - \overline{u} \cdot \epsilon) \cdot \theta^t - \eta$ which is equivalent to $\epsilon \cdot (\theta^{-1} + \overline{u} \cdot \theta^t) \leqslant \tau \cdot (u \cdot \theta^t - \theta^{-1}) - \eta$.

(10) is implied by: $(\tau \cdot u + \epsilon \cdot \overline{u}) \cdot \theta^t + \eta < (\tau - \epsilon)$, which is equivalent to $\epsilon \cdot (\overline{u} \cdot \theta^t + 1) < \tau \cdot (1 - u \cdot \theta^t) - \eta$.

We are able to satisfy those two inequalities whenever t has been chosen such that $1 > u \cdot \theta^t > \theta^{-1}$ or equivalently that $u \cdot \theta^t < 1 < u \cdot \theta^{t+1}$. Choose t minimal in absolute value such that $1 \leqslant u \cdot \theta^{t+1}$ and so $u \cdot \theta^t < 1$. Now, for k large enough (10) will hold. Indeed, first choose k_0 such that $\eta < 1/2 \cdot \tau(1-u\theta^t)$, then choose $\epsilon < 1/2 \cdot \tau(1-u\theta^t)/(\overline{u} \cdot \theta^s + 1)$ and finally choose $k_1 \geqslant k_0$ such that for all $k \geqslant k_1$, $\tau - \epsilon < r_k/\theta^k < \tau + \epsilon$.

Since s is such that $u \cdot \theta^t < 1 \leqslant u \cdot \theta^{t+1}$, there exists j large enough such that: $(u + q \cdot \theta^{-j}) \cdot \theta^t < 1 \leqslant u \cdot \theta^{t+1} < (u + q \cdot \theta^{-j}) \cdot \theta^{t+1}$ and so (11) and (12) also hold whenever (9) and (10) hold.

If t is such that $u \cdot \theta^t < 1 < u \cdot \theta^{t+1}$, then there exists j large enough such that: $1 < (u - q \cdot \theta^{-j}) \cdot \theta^{t+1}$ and so $(u - q \cdot \theta^{-j}) \cdot \theta^t < u \cdot \theta^t < 1 < (u - q \cdot \theta^{-j}) \cdot \theta^{t+1} < u \cdot \theta^{t+1}$ and so (13) will hold whenever (9) and (10) hold.

So we see that for inequalities (9) and (13), we have to distinguish the cases whether there is a t such that $u \cdot \theta^t < 1 < u \cdot \theta^{t+1}$, or $u = \theta^{-(t+1)}$.

(a) There is a t such that $u \cdot \theta^t < 1 < u \cdot \theta^{t+1}$. So if k_0 has been chosen such that $\eta = \theta^{-k_0} \cdot \sum_{i \geqslant 0} |q_i| \cdot d_i \cdot \theta^t < 1/2\tau \cdot (u \cdot \theta^t - \theta^{-1})$ and $k_1 > k_0$ such that for all $k \geqslant k_1$, $\tau - \epsilon < r_k/\theta^k < \tau + \epsilon$ with $\epsilon \leqslant 1/2 \cdot \tau \cdot (u \cdot \theta^t - \theta^{-1})/(\theta^{-1} + \overline{u} \cdot \theta^t)$, then (9) will hold for all $k \geqslant k_1$ and also, by the discussion above, (13) for k and j large enough.

(b) There is t such that $u = \theta^{-t-1}$.

First assume that the sequence $\left(\frac{r_k}{\theta^k}\right)$ is constant for k large enough. So for such k, (9) is equivalent to: $\theta^{-1} \cdot \tau \leqslant \theta_0^t(u \cdot \tau - \theta^{-k} \cdot \sum_{i \geqslant 0} d_i \cdot q_i)$ i.e. $0 \leqslant -\theta^{-k} \cdot \sum_{i \geqslant 0} d_i \cdot q_i$. So we have to take into account the congruences that each r_{k-i} satisfies modulo n_i.

If $\displaystyle\sum_{i \geqslant 0} \frac{d_i}{n_i} \cdot m_i \leqslant 0$, then (9) holds. If $\displaystyle\sum_{i \geqslant 0} \frac{d_i}{n_i} \cdot m_i > 0$ then we have

$$r_{k-t-2} \leqslant \sum_{i \geqslant 0} m_i \cdot r_{k-i}/n_i < r_{k-t-1},$$

for k large enough such that the sequence is constant and $\eta \leqslant \tau(\theta - 1) \cdot \theta^{-2}$.

Second, assume that in addition the sequence satisfies a linear recurrence whose characteristic polynomial is the minimal polynomial of θ. Then $\sum_{i\geqslant 0} q_i \cdot \theta^{-i} = \theta^{-t-1}(= u)$ implies that $\sum_{i\geqslant 0} q_i \cdot \theta^{-i+n} = \theta^{n-t-1}$ with $n \geqslant \max\{i, t+1\}$ and so $\sum_{i\geqslant 0} q_i \cdot r_{k-i} = r_{k-t-1}$, for k large enough.

Therefore, if $\sum_{i\geqslant 0} \frac{d_i}{n_i} \cdot m_i \leqslant 0$, then (9) holds since

$$\sum_{i\geqslant 0} m_i \cdot r_{k-i}/n_i = \sum_{i\geqslant 0} \frac{m_i}{n_i} \cdot r_{k-i} - \sum_{i\geqslant 0} m_i \cdot \frac{d_i}{n_i} = r_{k-t-1} - \sum_{i\geqslant 0} m_i \cdot \frac{d_i}{n_i}.$$

If $\sum_{i\geqslant 0} \frac{d_i}{n_i} \cdot m_i > 0$, then we will show that $r_{k-t-2} \leqslant \sum_{i\geqslant 0} m_i \cdot r_{k-i}/n_i < r_{k-t-1}$. The right inequality is clearly satisfied, for the left one, we choose k large enough such that $\tau-\epsilon < r_k/\theta^k < \tau+\epsilon$ with $\epsilon < 1/2\cdot\tau\cdot(\theta-1)/(\theta+1)$ and $\eta \cdot \theta < 1/2 \cdot \tau \cdot (\theta - 1)/(\theta + 1)$.

Let us examine conditions (13) in this case where $u = \theta^{-t-1}$.

First suppose that $\sum_{i\geqslant 0} \frac{d_i}{n_i} \cdot m_i = 0$. Let $a > 0$, then we put a condition on a in order to get that $r_{k-t-1} - a = \sum_{i\geqslant 0} \frac{m_i}{n_i} \cdot r_{k-i} - a > r_{k-(t+2)}$. Dividing both sides of the inequality by $\theta^{k-(t+1)}$, we get $\theta^{-1} \cdot r_{k-(t+2)}/\theta^{k-(t+2)} < (r_{k-(t+1)}-a)/\theta^{k-(t+1)}$. This is implied by $\theta^{-1}\cdot(\tau+\epsilon) < (\tau-\epsilon)-a/\theta^{k-(s+1)}$ i.e. $a/\theta^{k-(t+1)} < (\tau - \epsilon) - \theta^{-1} \cdot (\tau + \epsilon)$.

Choose $\epsilon = 1/2 \cdot \tau \cdot (\theta - 1)/(\theta + 1)$. Then we choose k_0 such that for all $k > k_0$, $\tau - \epsilon < r_k/\theta^k < \theta + \epsilon$. The condition on a is that $a/\theta^{k-(t+1)} < 1/2 \cdot \tau \cdot (1 - \theta^{-1})$. Now if a is in the form r_{k-j}, then j has to be chosen such that $\theta^{-j+(t+1)} < 1/4 \cdot (1 - \theta^{-1})$.

If $\sum_{i\geqslant 0} \frac{d_i}{n_i} \cdot m_i \neq 0$, then $\sum_{i\geqslant 0} m_i \cdot r_{k-i}/n_i = \sum_{i\geqslant 0} \frac{m_i}{n_i} \cdot r_{k-i} - \sum_{i\geqslant 0} m_i \cdot \frac{d_i}{n_i}$. Then, instead of considering a, consider $a + \sum_{i\geqslant 0} m_i \cdot \frac{d_i}{n_i}$.

If $\sum_{i\geqslant 0} m_i \cdot \dfrac{d_i}{n_i} > 0$, then the discussion is as above, if $\sum_{i\geqslant 0} m_i \cdot \dfrac{d_i}{n_i} < 0$, then we have to consider the finitely many cases where $0 < a \leqslant \sum_{i\geqslant 0} m_i \cdot \dfrac{d_i}{n_i}$.

When the sequence does not satisfy a linear recurrence, then we will show although conditions (9) up to (12) can be satisfied if θ is not a root of unity, (13) may not be satisfied as will show the example of the sequence $(2^k + k)$.

Suppose that the sequence does not necessarily satisfy a linear recurrence. Then we have to show that if $r_{k-s-1} \leqslant \sum_{i\geqslant 0} m_i \cdot r_{k-i}/n_i < r_{k-s}$, for k large enough, then it holds for all $k+t$, where $t > 0$. We have shown above that the only case to consider is when $u = \theta^{-s-1}$. We have two cases either for all $t > 0$, $r_{k+t-(s+1)} \leqslant \sum_{i\geqslant 0} m_i \cdot r_{k+t-i}/n_i < r_{k+t-s}$, or there exists $t' > 0$ and $s' \neq s$ such that $r_{k+t'-(s'+1)} \leqslant \sum_{i\geqslant 0} m_i \cdot r_{k+t'-i}/n_i < r_{k+t'-s'}$. We have seen that these inequalities hold if s' has been chosen such that $u \cdot \theta^{s'} < 1 < u \cdot \theta^{s'+1}$, or $u = \theta^{-(s'+1)}$. In this later case, we have $\theta^{s'-s} = 1$, which contradicts the hypothesis on θ not being a root of unity. In the former case where $u \cdot \theta^{s'} < 1 < u \cdot \theta^{s'+1}$, we can meet conditions (9) up to (14) as the above discussion shows, for $k > k_0 + t'$. So now suppose that for all k large enough, we have $r_{k-s-1} \leqslant \sum_{i\geqslant 0} m_i \cdot r_{k-i}/r_{k-s}$. Then the problem is that for $a > t(r_k)$, where $t(r_k)$ is a non constant term in r_k, we may have $r_{k-s-2} \leqslant \sum_{i\geqslant 0} m_i \cdot r_{k-i}/n_i - a < r_{k-s-1}$, but for $0 < a < t(r_k)$, $r_{k-s-1} \leqslant \sum_{i\geqslant 0} m_i \cdot r_{k-i}/n_i - a < r_{k-s}$.

So, we have shown the following Proposition.

Proposition 10. *Let R be a sequence satisfying the A. Bertrand conditions. Let θ be a real number strictly greater than 1 such that $L(R)$ is equal to $L(\theta)$. Suppose in addition that the θ-development of 1 is almost periodic (so R satisfies a linear recurrence) and that the minimal polynomial of θ is the characteristic polynomial of this linear recurrence. Then R is an almost sparse sequence (and R satisfies axiom $(A6)$).*

Lemma 11. *The sequence* $(2^k + k)$ *is not sparse.*

PROOF. — Let $r_k = 2^k + k$, $\theta = 2$ and $\tau = 1$. Let $q_i = \frac{m_i}{n_i}$ and $2^{k-i} + k - i \equiv d_i \pmod{n_i}$. Suppose that $\sum_{i\geqslant 0} q_i \cdot 2^{-i} = 2^{-s-1}$ for some s and $\sum_i q_i - 1 \geqslant 0$, (take for instance $q_2 = 2$, and if $i \neq 2$, $q_i = 0$, so $q_2 \cdot 2^{-2} = 2^{-1}$, $s = 0$ and $q_2 - 1 > 0$).

Then, $2^{k-s-1}+k-s-1 \leqslant \sum_{i\geqslant 0} q_i \cdot (2^{k-i}+k-i) - \sum_{i\geqslant 0} q_i \cdot d_i < 2^{k-s}$ is equivalent to: $2^{k-s-1} + k - s - 1 \leqslant 2^k \sum_{i\geqslant 0} q_i \cdot 2^{-i} + \sum_{i\geqslant 0} q_i \cdot k - \sum_{i\geqslant 0} q_i \cdot (i + d_i) < 2^{k-s}$, $k - s - 1 \leqslant \sum_{i\geqslant 0} q_i \cdot k - \sum_{i\geqslant 0} q_i \cdot (i + d_i) < 2^{k-s-1}$, which holds for k large enough iff $\sum_{i\geqslant 0} q_i \cdot (i + d_i) - (s + 1) \leqslant k(\sum_{i\geqslant 0} q_i - 1)$ iff $\sum_i q_i - 1 \geqslant 0$.

If $1 - \sum_i q_i = 0$, then the requirement is that $s + 1 \geqslant \sum_i q_i \cdot (i + d_i)$.

Suppose now that $1 - \sum_i q_i < 0$, we are going to examine the condition that if, for $k > k_0$, $2^{k-s-1} + k - s - 1 < \sum_{i\geqslant 0} q_i \cdot (2^{k-i} + k - i - d_i) < 2^{k-s}$, then, for some j depending only on k_0,
$2^{k-s-1} + k - s - 1 < \sum_{i\geqslant 0} q_i \cdot (2^{k-i} + k - i - d_i) - 2^{k-j} - k + j < 2^{k-s}$ holds, equivalently, $2^{k-s-1} + k - s - 1 \leqslant$
$2^k \sum_{i\geqslant 0} q_i \cdot 2^{-i} - 2^k \cdot 2^{-j} + (\sum_{i\geqslant 0} q_i) \cdot k - \sum_{i\geqslant 0} q_i(d_i + i) + j < 2^{k-s}$. But, if k is large enough, the inequality: $2^k \cdot 2^{-j} + \sum_{i\geqslant 0} q_i(i + d_i) - j - (s + 1) \leqslant$ $k \cdot \left(\sum_{i\geqslant 0} q_i - 1 - q\right)$ never holds. Suppose, we replace $2^{k-j} - k + j$ by some $a > 0$. We get: $2^{k-s-1} + k - s - 1 \leqslant \sum_{i\geqslant 0} q_i \cdot (2^{k-i} + k - i - d_i) - a$, equivalently that $a \leqslant (\sum_{i\geqslant 0} q_i - 1)k - \sum_{i\geqslant 0} q_i(d_i + i) + s + 1$. So the fact that this inequality holds depends on the relative position of a and k. □

Proposition 12. *If the sequence R satisfies either hypotheses $(A)^*$ or $(B)^*$ above, then $\mathcal{N}$ is a model of the corresponding theory T.*

Proof. — Let us first show that axiom (A5) is satisfied if our sequence satisfies a linear recurrence. Let us fix a natural number n. The numbers of r-tuples $(\ell_{1_i}, \dots, \ell_{r_i})$ which are the values of an element of R in $\mathbb{Z}/q_1\mathbb{Z} \times \cdots \times \mathbb{Z}/q_r\mathbb{Z}$ is finite. By hypothesis, the value of the n^{th} element of R depends only on the k preceeding ones for a fixed k. So necessarily after a sequence of length $\leqslant k^{n^r}$, we will have a repetition of a sequence of some k consecutive r-tuples.

The fact that it satisfies axiom (A6) follows from Proposition 10. □

Remark: To put necessary and sufficient conditions on a sequence R in order to get decidability of the structure $\langle \mathbb{N}, +, R\rangle$ seems to depend on non trivial number theoretic results as the following results of Bateman, Jockusch and Woods illustrate: $\mathrm{Th}\langle \mathbb{N}, +, \mathcal{P}\rangle$, where $\mathcal{P}$ denotes the set of prime numbers, is undecidable under the linear Schintzel hypothesis (see Theorem 1 in [1]). Under the same hypothesis, they also prove that $\mathrm{Th}\langle \mathbb{N}, +, V_2, (2^n)_{n\in\mathcal{P}}\rangle$ is decidable (see Corollary 1 in [1]).

3. $\langle \mathbb{N}, <, R\rangle$

We are going to consider the following expansion of Presburger arithmetic: $\mathcal{N} = \langle \mathbb{N}, <, 0, 1, \pm 1, \lambda_R, S, S^{-1}\rangle$, where $+1$ denotes the successor function on the natural numbers and -1 the predecessor function i.e. if $n > 0$, then $n - 1$ is defined as the greatest natural number $< n$ and we define $0 - 1$ as equal to 0. Since R is cofinal in $\mathbb{N}$ and since $r_0 = 1$, for each non zero natural number x, there exists n such that $r_n \leqslant x < r_{n+1}$, then we define $\lambda_R(x) = r_n$, $\lambda_R(0) = 0$. If $r_n < x < r_{n+1}$, then $S(x) = x$, also $S(0) = 0$, and $S(r_n) = r_{n+1}$, $n \geqslant 0$, and $S^{-1}(r_n) = r_{n-1}$, $n > 0$, $S^{-1}(1) = 1$. If $\lambda_R(x) = x$, we will use the following abreviation $R(x)$. Let L be the following language $\{<, 0, \pm 1, \lambda_R, S, S^{-1}\}$.

The following structures, namely $\langle \mathbb{N}, <, 0, \pm 1\rangle$ and $\langle R, <, 1, S, S^{-1}\rangle$ have the same theory which admits quantifier elimination. Let T_d be a universal axiomatisation of this theory of discrete linear orders with a smallest element, a successor and a predecessor functions (see [8, § 3.2]).

We define the following predicates $I_{\overline{n},\overline{m},\overline{n}',\overline{m}'}(a,b)$ as follows:

$$\exists c(a < c < b \wedge R(c) \wedge \bigwedge_i S^{n_i}(c) = c \pm m_i \wedge \bigwedge_j S^{n'_j}(c) \neq c \pm m'_j),$$

where $\overline{n}, \overline{n}'$ (respectively $\overline{m}, \overline{m}'$) range over the finite sequences of integers (respectively natural numbers). In order to get an expansion $\mathcal{N}$ with an existential theory, we are going to put a condition on the tuples (a,b) for which either $I_{\overline{n},\overline{m},\overline{n}',\overline{m}'}$ holds or does not hold (see axiom (A5) below). This condition is justified by the following Proposition.

Proposition 13. *Suppose there exists a predicate $I_{\overline{n},\overline{m},\overline{n}',\overline{m}'} = I$ such that $\mathbb{N}$ satisfies $\neg I(n_k, m_k)$ and $I(m_k, n_{k+1})$, $k \in \mathbb{N}$, $n_k < m_k < n_{k+1}$ with $|m_k - n_k|$ increasing and $|R \cap [n_k m_k]|$ unbounded. Then the formula $\varphi(u,v) = (u < v \wedge \neg I(u,v))$ is not equivalent in $\mathrm{Th}(\mathcal{N})$ to an existential formula.*

PROOF. — By the way of contradiction, suppose that $\varphi(u,v)$ is equivalent to an existential formula $\psi(u,v) = \exists z_1 \ldots \exists z_m \theta(u,v,z_1,\ldots,z_m)$, where $\theta(u,v,z_1,\ldots,z_m)$ is a conjunction of basic formulas $\tau(u,v,z_1,\ldots,z_m)$.

We assume that among those basic formulas there is "$u < v$". A term $t(x)$ is of the form $\ldots S^{n_2}(\lambda_R(S^{n_1}(\lambda_R(x+m_1)) + m_2))\ldots$, where m_i and n_i belong to $\mathbb{Z}$. We say that x is connected to y if θ implies that $(x < y$ and $t_1(x) > t_2(y))$ or vice-versa with x in place of y, where t_1, t_2 are terms. By hypothesis, we have $\varphi(n_k, m_k)$, let $z_1(k), \ldots, z_m(k)$ elements of $\mathbb{N}$ such that $\theta(n_k, m_k, z_1(k), \ldots, z_m(k))$ holds. We claim that because of the fact that $n_k < m_k < n_{k+1}$ with $|m_k - n_k|$ increasing and $|R \cap [n_k m_k]|$ unbounded, no z_i can be connected to both u and v. If z_i is connected to u, define $z_{i,\mathrm{new}}(k) = z_i(k)$ and if z_i is connected to v, define $z_{i,\mathrm{new}}(k)$ by $z_i(k+1)$. Now, we have that $\mathcal{N}$ satisfies $\psi(n_k, m_{k+1})$ since we have $\theta(n_k, m_{k+1}, z_{1,\mathrm{new}}(k), \ldots, z_{m,\mathrm{new}}(k))$, which contradicts the fact that $I(m_k, n_{k+1})$. □

Example

Let$R = (2n; n \in \mathbb{N}) \cup (2^k + 1; k \in \mathbb{N})$,
$I(a,b) = \exists c(a < c < b \wedge R(c) \wedge S(c) = c+1)$ and
$n_k = 2^k + 2$, $m_k = 2^{k+1} - 1$.

Let T be the following theory:

(a) T_d in $\{0, \pm 1, <\}$.

(b) T_d in $\{1, S^{\pm 1}, <\}$.

(c) $\forall x((\lambda_R(0) = 0 \wedge S(0) = 0 \wedge S^{-1}(0) = 0) \wedge (x \geqslant 1 \rightarrow ((\lambda_R(x) \leqslant x < S(\lambda_R(x)) \wedge \lambda_R(\lambda_R(x)) = \lambda_R(x) \wedge \forall y(\lambda_R(x) < y < S(\lambda_R(x)) \rightarrow \lambda_R(y) = \lambda_R(x) \wedge S(y) = y \wedge S^{-1}(y) = y))))$.

(d)$_n R(r_n)$.

(e) For each sequences $\overline{n}, \overline{m}, \overline{n}', \overline{m}'$, and predicate $I = I_{\overline{n},\overline{m},\overline{n}',\overline{m}'}(\cdot,\cdot)$, there exists $n(I)$ and $m(I)$ such that we have either the axiom: $(\forall a \forall b\, ((n(I) < a < b \wedge \neg I(a,b)) \rightarrow (b \leqslant S^{m(I)}(\lambda_R(a))))$, or $(\forall a \forall b\, ((n(I) < a < b) \rightarrow \neg I(a,b)))$.

Proposition 14. *T is model-complete.*

Proof. — The proof of this Proposition follows the same pattern as the proof of Proposition 5. Using Robinson criterium, it suffices to show that given $\mathcal{A}$, $\mathcal{B}$ two models of T with $\mathcal{A} \subset_L \mathcal{B}$, then $\mathcal{A} \subset_{ec} \mathcal{B}$. Let $\varphi(a_1, \ldots, a_n)$ be an existential formula with parameters in $\mathcal{A}$ of the form $\exists b_1 \cdots \exists b_s \theta(a, b)$, where θ is a quantifier-free L-formula. The terms occuring in θ are of the form $\ldots S^{z_2}(\lambda_R(b_j + m_1)) + m_2)) \ldots$, $1 \leqslant j \leqslant s$, where $b_j \in \mathcal{B} - \mathcal{A}$, m_i, $z_i \in \mathbb{Z}$. We have two cases either every sub-term $c_1, \ldots, c_r$ which belongs to R, belongs to $R(\mathcal{A})$ or there is at least one which does not. Consider $\langle \mathcal{A}, c_1, \ldots, c_r \rangle$ with $c_1, \ldots, c_r \in \mathcal{B} - \mathcal{A}$ in the following two cases:

(a) $R\langle \mathcal{A}, c_1, \ldots, c_r \rangle = R\langle \mathcal{A} \rangle$,

(b) $\bigwedge_{i \leqslant s} R(c_i)$ and $c_{s+1}, \ldots, c_r$ do not belong to R.

(a) Since T_d is universal, $\langle \mathcal{A}, c_1, \ldots, c_r \rangle$ satisfies T_d and since it is model-complete, by the existential Frayne theorem, there exists an embedding f of $\langle \mathcal{A}, c_1, \ldots, c_r \rangle$ in an ultrapower of $\mathcal{A}$ which is the identity on $\mathcal{A}$ and respects $\{0, \pm 1, <\}$ (see [6], Corollary 4.3.13). Let us show that f is an L-morphism. Let $t(a, \mathbf{c})$ be an L-term where at least one of $c_1, \ldots, c_r$ is occuring and let $a \in \mathcal{A}$. By hypothesis, $\lambda_R(t(a, \mathbf{c}))$ belongs to $\mathcal{A}$. We have that $\lambda_R(t(a, \mathbf{c})) \leqslant t(a, \mathbf{c}) < S(\lambda_R(t(a, \mathbf{c})))$.

Since f respects the order, we have $f(\lambda_R(t(a,\mathbf{c}))) \leqslant f(t(a,\mathbf{c})) < f(S(\lambda_R(t(a,\mathbf{c}))))$, since f is the identity on $\mathcal{A}$, we have $\lambda_R(t(a,\mathbf{c})) \leqslant f(t(a,\mathbf{c})) < S(\lambda_R(t(a,\mathbf{c})))$. So $\lambda_R(f(t(a,\mathbf{c}))) = \lambda_R(t(a,\mathbf{c}))$. Consider $S(t(a,\mathbf{c}))$, suppose that $\lambda_R(t(a,\mathbf{c})) \neq t(a,\mathbf{c})$ so $\lambda_R(t(a,\mathbf{c})) = f(\lambda_R(t(a,\mathbf{c}))) \neq f(t(a,\mathbf{c}))$, by definition of S, $S(t(a,\mathbf{c})) = t(a,\mathbf{c})$. So $f(S(t(a,\mathbf{c}))) = f(t(a,\mathbf{c})) = S(f(t(a,\mathbf{c})))$. If $\lambda_R(t(a,\mathbf{c})) = t(a,\mathbf{c})$, $t(a,\mathbf{c})$ belongs to $\mathcal{A}$, and so $fS(t(a,\mathbf{c})) = S(t(a,\mathbf{c})) = S(f(t(a,\mathbf{c})))$. Let t_1, t_2 belong to $\mathcal{B}$. Suppose that $I_{n,m,n',m'}(t_1,t_2)$. This is equivalent to $I_{n,m,n',m'}(\lambda_R(t_1), \lambda_R(t_2)+1)$ with $\lambda_R(t_1)$, $\lambda_R(t_2)$ belonging to $\mathcal{A}$ since there are no new elements of R in $\langle \mathcal{A}, c_1, \ldots, c_r\rangle$.

(b) In this case, we will show that $\langle \mathcal{A}, c_1, \ldots, c_s\rangle$ embeds in an ultrapower of $\mathcal{A}$ by an L-morphism and then we will use (a) to embed $\langle\langle \mathcal{A}, c_1, \ldots, c_s\rangle, c_{s+1}, \ldots, c_r\rangle$ in an (iterated) ultrapower of $\mathcal{A}$ and so $\mathcal{A}$ will satisfy $\varphi(a_1, \ldots, a_n)$. Since T_d is model-complete, by the existential Frayne theorem, there exists an embedding g of $R(\langle a, c_1, \ldots, c_s\rangle)$ in an ultrapower $R(\mathcal{A})^\alpha/U$, where U is a α-regular ultrafilter on α, and $\alpha = |\mathcal{B}|$, which is the identity on $R(\mathcal{A})$ and respects $\{1, <, S, S^{-1}\}$ (see [6], Corollary 4.3.13).

For each $c_1, \ldots, c_s \in \mathcal{B} - \mathcal{A}$, let $\mathcal{A}_{c_i,-} = \{a \in \mathcal{A} : a < c_i\}$, $\mathcal{A}_{c_i,+} = \{a \in \mathcal{A} : c_i < a\}$. As in Lemma 1, we get: for all $z \in \mathbb{Z} - \{0\}$, $\mathcal{A}_{c_i,-} < S^z c_i < \mathcal{A}_{c_i,+}$.

First, we have to identify which c_i's determine the same cuts in $\mathcal{A}$; by renaming them, we may assume that $c_1 < c_2 < \cdots < c_s$ and we also assume that $S^{-n'} c_{i+1}$ is not in the same archimedean class as $S^n c_i$. (If it is not the case e.g. if $S^{-n'} c_{i+1} = S^n c_i + m$, then replace $S^{n'}(S^n c_i + m)$). We partition $\{1, \ldots, s\} = J_1 \cup \cdots \cup J_h$ according to the fact that i, $i' \in J_j$ iff c_i, $c_{i'}$ belong to the same cut in $\mathcal{A}$.

Now we are going to modify the images $g(c_i)$, $1 \leqslant i \leqslant s$, in $\mathcal{A}^\alpha/U$ in order that it satisfies the same type as $c_1, \ldots, c_s$ do over $\mathcal{A}$. Let $p(x_1, \ldots, x_m)$ be the set of all formulas of the form $\bigwedge_{i,j} S^{n_{i,j}}(x_j) = x_j \pm m_{ij} \neq x_j \pm m'_{ij}$, where $n_{ij}, n'_{ij} \in \mathbb{Z}$, m_{ij}, $m'_{ij} \in \mathbb{N} - \{0\}$, satisfied by $c_1, \ldots, c_s$ in $\mathcal{B}$. We want to show that this type is finitely satisfiable in $\mathcal{A}$ in order to get a solution in $\mathcal{A}^\alpha/U$. Consider a formula τ in $p(x_1, \ldots, x_m)$ of the form $\bigwedge_j \bigwedge_{(i,t \in J_j)} (S^{n_{it}}(x_t) = x_t \pm m_{it} \wedge S^{n'_{it}}(x_t) \neq x_t \pm m'_{it})$ and set for $t \in J_j$, $\mathcal{A}_{c_t,-} = \mathcal{A}_{J_j,-}$ and $\mathcal{A}_{c_t,+} = \mathcal{A}_{J_j,+}$. Let $I_{j,t}$ be the predicates corresponding to the formula τ, $1 \leqslant j \leqslant h$, $t \in J_j$. For each of those predicates, the theory T contains the following axiom: $(\forall a \forall b\ ((n(I) < a < b \wedge \neg I_{j,t}(a,b)) \rightarrow (b < S^{n(I_{j,t})}(\lambda_R(a)))$. For each j, $1 \leqslant j \leqslant h$, let $(a_{\ell,j})_{i \in I}$ be a cofinal sequence in $\mathcal{A}_{J_{j,-}}$. We have that $\mathcal{A}$ satisfies $I_{j,t}(a_{\ell,j}, S^{n(I_{j,t})}(a_{\ell,j}))$. Therefore in $\mathcal{A}$, we have $(a_t)_{t \in J_j}$ such that $\mathcal{A}$ satisfies $\tau((a_t)_{t \in J_j})$ and $a_{\ell,j} < a_{n_1} < \cdots < a_{n_j} < S^{n(I_{j,t})}(a_{\ell,j})$ with $J_j = \{n_1, \ldots, n_j\}$. So $\mathcal{A}^\alpha/U$ satisfies $\tau(([a_{t,\ell}]_U)_{t \in J_j})$ with $\mathcal{A}_{J_{j,-}} < (([a_{t,\ell}]_U)_{t \in J_j}) < \mathcal{A}_{J_{j,+}}$. So the type $p(x_1, \ldots, x_m)$ is finitely satisfiable in $\mathcal{A}^\alpha/U$ by an element which lies in the right cut with respect to $\mathcal{A}$. □

Corollary 15. *T is complete and modulo T, any L-formula is equivalent to an existential L-formula. Moreover if T is recursive, then T is decidable.*

Proof. — $\mathcal{N}$ is is a prime model of T and so T being model-complete, T is complete. Since T is model-complete, any L-formula is equivalent in T, to an existential L-formula. The last assertion follows from the completeness of T. □

Corollary 16. (See also [13]). *Let R be a sequence such that $\lim_{n \to \infty} r_{n+1} - r_n = +\infty$, then $\mathrm{Th}\langle \mathbb{N}, <, R, I_{\overline{n}, \overline{m}, \overline{n'}, \overline{m'}} \rangle$, where $\overline{n}, \overline{m}, \overline{n'}, \overline{m'}$ vary over the set of finite sequences of natural numbers, is model-complete.*

Proof. — It suffices to show that under the above hypothesis on R, $\langle \mathbb{N}, <, R \rangle$ is a model of a theory T of the form given above.
Choose $n(I) = n_0$ such that $\forall n > n_0\ (r_{n+1} - r_n) > max\{\overline{m}, \overline{m}'\}$. So, the following axioms hold in that structure:
$\forall a \forall b\ (n(I) < a < b \wedge \neg I_{\overline{0}, \overline{0}, \overline{n}', \overline{m}'}(a,b)\ \rightarrow b \leqslant S(\lambda_R(a)))$
$\forall a \forall b\ (n(I) < a < b \rightarrow \neg I_{\overline{m}, \overline{n}, \overline{m}', \overline{n}'}(a,b))$, if $\{\overline{m}, \overline{n}\} \neq \{\overline{0}\}$. □

Acknowledgements

We would like to thank M. Boffa, C. Frougny, A. Maes and C. Michaux for interesting discussions, we like to thank Carol Wood and the Logic group of Wesleyan University for their hospitality during the spring 1996 and the opportunity to give a talk on this paper.

References

[1] P.T. Bateman, C.G. Jockusch, and A.R. Woods. Decidability and Undecidability of Theories with a Predicate for the Primes. *Journal of Symbolic Logic*, Vol. 58, no. 2, pp. 672–687, 1993.

[2] A. Bertrand-Mathis. Comment réécrire les nombres entiers dans une base qui n'est pas entière. *Acta Math. Hungr.*, Vol. 58, no. 2, pp. 672–687, 1993.

[3] V. Bruyère, G. Hansel, C. Michaux, and R. Villemaire. Logic and p-recognizable sets of integers. *Bull. Belg. Math. Soc.*, Vol. 1, no. 191-238, , 1994.

[4] V. Bruyère and F. Point. On the Cobham-Semenov Theorem. *Theory Comput. Systems*, Vol. 30, pp. 197–220, 1997.

[5] J.R. Büchi. Weak second-order arithmetic and finite automata. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, Vol. 6, pp. 66–92, 1960.

[6] C.C. Chang and H.J. Keisler. *Model Theory*, volume 73 of *Studies in Logic and the Fondations of Mathematics*. North-Holland, 1973.

[7] G. Cherlin and F. Point. On extensions of Pressburger arithmetic. In *Proc. 4th Easter Model Theory Conference, Gross Köris, Seminarberichte 86*, pages 17–34, Humboldt University zu Berlin, 1986.

[8] H.B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, New York, 1973.

[9] C. Frougny and B. Solomyak. On representations of integers in linear numeration systems. L.i.t.p. report, Paris 6/7, 1994/1995.

[10] B.R. Hodgson. Décidabilité par automate fini. *Ann. Sci. Math. Québec*, Vol. 7, pp. 39–57, 1983.

[11] W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hungar.*, Vol. 11, pp. 401–406, 1960.

[12] A.L. Semenov. On certain extensions of the arithmetic of addition of natural numbers. *Math USSR Izv.*, Vol. 15, no. 2, pp. 401–418, 1980. English translation of *Izv. Akad. Nauk SSSR Ser. Mat.*, Vol. 43, no. 5, pp. 1175–1195, 1979.

[13] A.L. Semenov. Logical theories of one-place functions on the set of natural numbers. *Math USSR Izv.*, Vol. 22, no. 3, pp. 587–618, 1984. English transl. of *Izv. Akad. Nauk SSSR Ser. Mat.*, Vol. 47, no. 3, pp. 623–658, 1983.

[14] J. Shallit. Numeration Systems. In *Linear Recurrences, and Regular Sets, I.C.A.L.P.92, Wien,* Vol. 623 of *Lecture Notes in Computer Science*, pages 89–100. Springer, 1992.

[15] L. van den Dries. On expansions of $\mathbb{Q}$ and $\mathbb{Z}$. Manuscript, 1/12/1985.

[16] L. van den Dries. The field of reals with a predicate for powers of two. *Manuscripta Math.*, pages 187–195, 1985.

Université de Mons–Hainaut
Institut de Mathématique et d'Informatique
"Le Pentagone"
Avenue du Champ de Mars 6
7000 Mons
Belgium
e-mail : point@sun1.umh.ac.be

Added in proof: An expanded version of the first part of this article will appear in the Journal of Symbolic Logic.