Infinite and Bi-infinite Words with Decidable Monadic Theories

Dietrich Kuske[1], Jiamou Liu[2], and Anastasia Moskvina[2]

1 **Technische Universität Ilmenau, Germany**
2 **Auckland University of Technology, New Zealand**

Abstract

We study word structures of the form $(D, \leq, P)$ where D is either $\mathbb{N}$ or $\mathbb{Z}$, $\leq$ is a linear ordering on D and $P \subseteq D$ is a predicate on D. In particular we show:

(a) The set of recursive ω-words with decidable monadic second order theories is Σ_3-complete.
(b) We characterise those sets $P \subseteq \mathbb{Z}$ that yield bi-infinite words $(\mathbb{Z}, \leq, P)$ with decidable monadic second order theories.
(c) We show that such "tame" predicates P exist in every Turing degree.
(d) We determine, for $P \subseteq \mathbb{Z}$, the number of predicates $Q \subseteq \mathbb{Z}$ such that $(\mathbb{Z}, \leq, P)$ and $(\mathbb{Z}, \leq, Q)$ are indistinguishable.

Through these results we demonstrate similarities and differences between logical properties of infinite and bi-infinite words.

1998 ACM Subject Classification F.4.1

Keywords and phrases infinite words, bi-infinite words, monadic second order logic

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

The decision problem for logical theories of linear structures and their expansions has been an important question in theoretical computer science. Büchi in [2] proved that the monadic second order theory of the linear ordering $(\mathbb{N}, \leq)$ is decidable. Expanding the structure $(\mathbb{N}, \leq)$ by unary functions or binary relations typically leads to undecidable monadic theories. Hence many works have been focusing on structures of the form $(\mathbb{N}, \leq, P)$ where P is a unary predicate. Elgot and Rabin [5] showed that for many natural unary predicates P, such as the set of factorial numbers, the set of powers of k, and the set of kth powers (for fixed k), the structure $(\mathbb{N}, \leq, P)$ has decidable monadic second order theory; on the other hand, there are structures $(\mathbb{N}, \leq, P)$ whose monadic theory is undecidable [3]. Numerous subsequent works further expanded the field [13, 4, 10, 11, 9, 8].

1. Semenov generalised periodicity to a notion of "almost periodicity". While periodicity implies that certain patterns are repeated through a fixed period, almost periodicity captures the fact that certain patterns occur before the expiration of some period. This led him to consider "recurrent structures" within an infinite word. Such a recurrent structure is captured by a certain function, which he called "indicator of recurrence". In [10], he provided a full characterisation: $(\mathbb{N}, \leq, P)$ has decidable monadic theory if and only if P is recursive and there is a recursive indicator of recurrence for P.
2. Rabinovich and Thomas generalised periodicity to a notion of "uniform periodicity". Such a uniform periodicity condition is captured by a *homogeneous set* which exists by Ramsey's theorem. More precisely, a k-homogeneous set for $(\mathbb{N}, \leq, P)$ partitions the

Conference title on which this volume is based on.
Editors: Billy Editor and Bill Editors; pp. 1–15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

natural numbers into infinitely many finite segments that all have the same k-type. A uniformly homogeneous set specifies an ascending sequence of numbers that ultimately becomes k-homogeneous for any $k > 0$. In [9], Rabinovich and Thomas provided a full characterisation: $(\mathbb{N}, \leq, P)$ has a decidable monadic theory if and only if P is recursive and there is a recursive uniformly homogeneous set.

Note that a recursive uniformly homogeneous set describes *how to* divide $(\mathbb{N}, \leq, P)$ such that the factors all have the same k-type. If P is recursive, this implies that the recurring k-type can be computed. A weakening of the existence of a recursive uniformly homogeneous set is therefore the requirement that one can compute a k-type such that $(\mathbb{N}, \leq, P)$ *can, in some way,* be divided. Nevertheless, Rabinovich and Thomas also showed that the monadic second order theory of $(\mathbb{N}, \leq, P)$ is decidable if and only if P is recursive and there is a "recursive type-function" (see below for precise definitions).

This paper has three general goals: The first is to compare these characterisations in some precise sense. The second is to investigate the above results in the context of *bi-infinite words*, which are structures of the form $(\mathbb{Z}, \leq, P)$. The third is to compare the logical properties of infinite words and bi-infinite words. More specifically, the paper discusses:

(a) In Section 4, we analyze the recursion-theoretical bound of the set of all computable predicates $P \subseteq \mathbb{N}$ where $(\mathbb{N}, \leq, P)$ has a decidable monadic theory. The second characterisation by Rabinovich and Thomas turns out to be a Σ_5-statement. In contrast, the characterisation by Semenov and the 1st characterisation by Rabinovich and Thomas both consist of Σ_3 statements, and hence deciding if a given $(\mathbb{N}, \leq, P)$ has decidable monadic theory is in Σ_3. We show that the problem is in fact Σ_3-complete. Hence these two characterisations are optimal in terms of their recursion-theoretical complexity.

(b) In Section 5, we then investigate which of the three characterisations can be lifted to bi-infinite words, i.e., structures of the form $(\mathbb{Z}, \leq, P)$ with $P \subseteq \mathbb{Z}$. It turns out that this is nicely possible for Semenov's characterisation and for the second characterisation by Rabinovich and Thomas, but not for their first one.

(c) If the monadic second order theory of $(\mathbb{N}, \leq, P)$ is decidable, then P is recursive. For bi-infinite words of the form $(\mathbb{Z}, \leq, P)$, this turns out not to be necessary. In Section 6, we actually show that every Turing degree contains a set $P \subseteq \mathbb{Z}$ such that the monadic second order theory of $(\mathbb{Z}, \leq, P)$ is decidable.

(d) The final Section 7 investigates how many bi-infinite words are indistinguishable from $(\mathbb{Z}, \leq, P)$. It turns out that this depends on the periodicity properties of P: if P is periodic, there are only finitely many equivalent bi-infinite words, if P is recurrent and non-periodic, there are $2^{\aleph_0}$ many, and if P is not recurrent, then there are $\aleph_0$ many.

2 Preliminaries

2.1 Words

We use $\mathbb{N}$, $\widetilde{\mathbb{N}}$ and $\mathbb{Z}$ to denote the set of natural numbers (including 0), negative integers (not containing 0), and integers, respectively. A *finite word* is a mapping $u\colon \{0, 1, \ldots, n-1\} \to \{0, 1\}$ with $n \in \mathbb{N}$, it is usually written $u(0)u(1)u(2)\cdots u(n-1)$. The set of positions of u is $\{0, 1, \ldots, n-1\}$, its length $|u|$ is n. The unique finite word of length 0 is denoted ε. The set of all (resp. non-empty) finite words is $\{0, 1\}^*$ (resp. $\{0, 1\}^+$). An ω*-word* is a mapping $\alpha\colon \mathbb{N} \to \{0, 1\}$; it is usually written as the sequence $\alpha(0)\alpha(1)\alpha(2)\cdots$. Its set of positions is $\mathbb{N}$; $\{0, 1\}^\omega$ is the set of ω-words. An ω^**-word* is a mapping $\alpha\colon \widetilde{\mathbb{N}} \to \{0, 1\}$; it is usually written as the sequence $\cdots\alpha(-3)\alpha(-2)\alpha(-1)$. Its set of positions is $\widetilde{\mathbb{N}}$ and $\{0, 1\}^{\omega^*}$ is the set of ω^*-words. Finally, a *bi-infinite word* ξ is a mapping from $\mathbb{Z}$ into $\{0, 1\}$, written as the sequence $\cdots\xi(-2)\xi(-1)\xi(0)\xi(1)\xi(2)\cdots$ (this notation has to be taken with care since, e.g., the bi-infinite words $\xi_i\colon \mathbb{Z}\to\{0,1\}\colon n\mapsto(|n|+i) \bmod 2$ with $i\in\{0,1\}$ are both described as $\cdots 0101010\cdots$, but they are different). The set of positions of a bi-infinite word is $\mathbb{Z}$. When saying "word", we mean "a finite, an ω-, an ω^*- or a bi-infinite word", "infinite word" means "ω- or ω^*-word".

The concatenation uv of two finite words u,v has its usual meaning. More generally, and in a similar way, we can also concatenate a finite or ω^*-word u and a finite or ω-word v giving rise to some word uv. Similarly, we can concatenate infinitely many finite words u_i giving an ω-word $u_0u_1u_2\cdots$, an ω^*-word $\cdots u_{-2}u_{-1}u_0$, and a bi-infinite word $\cdots u_{-2}u_{-1}u_0u_1u_2\cdots$ (where the position 0 is the first position of u_0). As usual, u^ω denotes the ω-word $uuuu\cdots$ for $u\in\{0,1\}^+$, analogously, $u^{\omega^*}=\cdots uuu$.

Let w be some word and i,j be two positions with $i\le j$. Then we write $w[i,j]$ for the finite word $w(i)w(i+1)\cdots w(j)\in\{0,1\}^+$. A finite word u is a *factor* of w if $u=w[i,j]$ for some i,j or if u is the empty word ε. The set of factors of w is $F(w)$. If w is an ω- or a bi-infinite word, then $w[i,\infty)$ is the ω-word $w(i)w(i+1)w(i+2)\cdots$. If w is an ω^*- or a bi-infinite word, then $w(-\infty,i]$ is the ω^*-word $\cdots w(i-2)w(i-1)w(i)$. A bi-infinite word β is *recurrent* if for all $u\in F(\beta)$ and all $i\in\mathbb{Z}$, $u\in F(\beta[i,\infty))\cap F(\beta(-\infty,i])$.

Let u be some finite word. Then u^R is the *reversal* of u, i.e., the finite word of length $|u|$ with $u^R(i)=u(|u|-i-1)$ for all $0\le i<|u|$. The reversal of an ω-word (resp. ω^*-word) α is the ω^*-word (resp. ω-word) α^R with $\alpha^R(i)=\alpha(-i-1)$ for all positions i. Finally, the reversal of a bi-infinite word ξ is the bi-infinite word ξ^R with $\xi^R(i)=\xi(-i)$ for all $i\in\mathbb{Z}$.

2.2 Logic

With any word w, we associate a relational structure $M_w=(D,\le,P)$ where $D\subseteq\mathbb{Z}$ is the set of positions of w, $\le$ is the restriction of the natural linear order on $\mathbb{Z}$ to D, and $P=\{n\in D\mid w(n)=1\}=w^{-1}(1)$. Structures of this form are called *labeled linear orders*. The word w is *recursive* (resp. *recursively enumerable*) if so is the set P.

We use the standard logical system over the signature of labeled linear orders. Hence first order logic FO has relational symbols $\le$ and P. The monadic second order logic MSO extends FO by allowing unary second order variables $X,Y,\dots$, their corresponding atomic predicates (e.g. $X(y)$), and quantification over set variables. By Sent, we denote the set of sentences of the logic MSO. For a word w and an MSO-sentence φ, we write $w\models\varphi$ for "the sentence φ holds in the relational structure M_w". The MSO-*theory* of the word w is the set $\mathsf{MTh}(M)$ of all MSO-sentences φ that are true in w.

▶ **Example 2.1.** Let $n\in\mathbb{N}$ and consider the following formula:

$$\varphi(x,y)=\exists X\colon\forall z\colon(X(z)\Leftrightarrow z=x\vee(x<z\wedge X(z-n)))\wedge X(y)$$

If w is a word with positions i,j, then $w\models\varphi(i,j)$ if and only if $i\le j$ and $n\mid j-i$.

With any MSO-formula φ, we associate its *quantifier rank* $\mathsf{qr}(\varphi)\in\mathbb{N}$: the atomic formulas have quantifier rank 0; $\mathsf{qr}(\varphi_1\wedge\varphi_2)=\mathsf{qr}(\varphi_1\vee\varphi_2)=\max\{\mathsf{qr}(\varphi_1),\mathsf{qr}(\varphi_2)\}$; $\mathsf{qr}(\neg\varphi)=\mathsf{qr}(\varphi)$; and $\mathsf{qr}(\exists X\colon\varphi)=\mathsf{qr}(\forall X\colon\varphi)=\mathsf{qr}(\varphi_1)+1$ where X is a first- or second-order variable.

▶ **Definition 2.2.** Let $k\in\mathbb{N}$. Two words w_1 and w_2 are k*-equivalent* (denoted $w_1\equiv_k w_2$) if $w_1\models\varphi$ iff $w_2\models\varphi$ for all MSO-sentences φ with $\mathsf{qr}(\varphi)\le k$. Equivalence classes of this equivalence relation are called k*-types*. The words w_1 and w_2 are MSO-*equivalent* (denoted $w_1\equiv w_2$) if $w_1\equiv_k w_2$ for all $k\in\mathbb{N}$. Equivalence classes of $\equiv$ are called *types*.

Let $k\ge 2$ and u,v be two words with $u\equiv_k v$. If u is finite, then it satisfies the sentence $(\exists x\forall y\colon x\le y)\wedge(\exists x\forall y\colon x\ge y)$. Consequently, also v is finite. Analogously, u is an ω-word iff v is an ω-word etc. We will therefore speak of a "k-type of finite words" when we mean a k-type that contains some finite word (and analogously for ω-, ω^*-, bi-infinite words etc).

Often, we will use the following known results without mentioning them again. They follow from the well-understood relation between MSO and automata (cf. [15, 6]).

▶ **Theorem 2.3.** *1. Let $k \geq 2$.*

- *For any ω-word (ω^*-word) α, there exist finite words x and y with $xy \equiv_k x$ ($yx \equiv_k x$), $yy \equiv_k y$ and $\alpha \equiv_k xy^\omega$ ($\alpha \equiv_k y^{\omega^*}x$). Any such pair (x, y) is* a representative of the k-type of α.
- *For any bi-infinite word ξ, there exist finite words x, y and z with $xy \equiv_k yz \equiv_k y$, $xx \equiv_k x$, $zz \equiv_k z$, and $\xi \equiv_k x^{\omega^*}yz^\omega$. Any such triple (x, y, z) is* a representative of the k-type of ξ.

2. The following sets are decidable:

- $\{\varphi \in \mathsf{Sent} \mid \forall u \in \{0,1\}^* \colon u \models \varphi\}$ *and* $\{(u, \varphi) \mid u \in \{0,1\}^*, \varphi \in \mathsf{Sent}, u \models \varphi\}$
- $\{(u, v, \varphi) \mid u, v \in \{0,1\}^*, v \neq \varepsilon, \varphi \in \mathsf{Sent}, uv^\omega \models \varphi\}$
- $\{(u, v, w, \varphi) \mid u, v, w \in \{0,1\}^*, u, w \neq \varepsilon, \varphi \in \mathsf{Sent}, u^{\omega^*}vw^\omega \models \varphi\}$
- $\{(u, v, k) \mid u, v \in \{0,1\}^*, k \in \mathbb{N}, u \equiv_k v\}$. *This means in particular that it is decidable whether u and v represent the same k-type of finite words.*
- *Similarly, it is decidable whether two pairs of finite words represent the same k-type of ω-words (of ω^*-words, resp). It is also decidable whether two triples of finite words represent the same k-type of bi-infinite words.*

3. If $u, v \in \{0,1\}^* \cup \{0,1\}^{\omega^*}$ and $u', v' \in \{0,1\}^* \cup \{0,1\}^\omega$ with $u \equiv_k v$ and $u' \equiv_k v'$, then $uu' \equiv_k vv'$. From representatives of the k-types of u and v, one can compute a representative of the k-type of uv.

4. If $u_i, v_i \in \{0,1\}^+$ with $u_i \equiv_k v_i$ for all $i \in \mathbb{Z}$, then we have

$$u_0u_1 \cdots \equiv_k v_0v_1 \cdots, \text{ and } \quad \cdots u_{-1}u_0 \equiv_k \cdots v_{-1}v_0, \text{ and } \quad \cdots u_{-1}u_0u_1 \cdots \equiv_k \cdots v_{-1}v_0v_1 \cdots$$

5. If u is a finite or ω^*-word and v is a finite or ω-word such that $\mathsf{MTh}(u)$ *and* $\mathsf{MTh}(v)$ *are both decidable, then* $\mathsf{MTh}(uv)$ *is decidable [12].*

2.3 Recursion theoretic notions

This paper makes use of standard notions in recursion theory; the reader is referred to [14] for a thorough introduction. We assume a canonical effective enumeration $\Phi_0, \Phi_1, \Phi_2, \ldots$ of all partial recursive functions on the natural numbers. The set W_e is the domain $\mathsf{dom}(\Phi_e)$ and is the *eth recursively enumerable set*. Let TOT be the set $\{e \in \mathbb{N} \mid \Phi_e \text{ is total}\}$ and REC be the set $\{e \in \mathbb{N} \mid W_e \text{ is decidable}\}$.

A set $A \subseteq \mathbb{N}$ belongs to the level Π_2 of the arithmetical hierarchy if there exists a decidable set $P \subseteq \mathbb{N}^{m+n+1}$ such that A is the set of natural numbers a satisfying $\forall x_1, \ldots, x_m \exists y_1, \ldots y_n \colon P(a, \bar{x}, \bar{y})$. A set $B \subseteq \mathbb{N}$ is Π_2-*hard* if, for every $A \in \Pi_2$, there exists a m-reduction from A to B; the set B is Π_2-*complete* if, in addition, $B \in \Pi_2$. Similarly, $A \subseteq \mathbb{N}$ belongs to Σ_3 if there exists a decidable set $P \subseteq \mathbb{N}^{\ell+m+n+1}$ such that A is the set of natural numbers a satisfying $\exists x_1, \ldots, x_\ell \forall y_1, \ldots, y_m \exists z_1, \ldots z_n \colon P(a, \bar{x}, \bar{y}, \bar{z})$. The notions Σ_3-hard and Σ_3-complete are defined similarly. For our purposes, it is important that the set TOT is Π_2-complete and the set REC is Σ_3-complete [14].

3 When is the MSO-theory of an ω-word decidable?

In this section, we recall the answers by Semenov [10] and by Rabinovich and Thomas [9]. Semenov defined a form of *"periodic words"* in which words from certain regular sets recur.

▶ **Definition 3.1.** Let α be some ω-word. An *indicator of recurrence* for α is a function $\mathrm{rec}\colon \mathsf{Sent} \to \mathbb{N} \cup \{\top\}$ such that, for every MSO-sentence φ, the following hold:

- if $\mathrm{rec}(\varphi) = \top$, then $\forall k\, \exists j \geq i \geq k\colon \alpha[i,j] \models \varphi$
- if $\mathrm{rec}(\varphi) \neq \top$, then $\forall j \geq i \geq \mathrm{rec}(\varphi)\colon \alpha[i,j] \models \neg\varphi$

▶ **Theorem 3.2** (Semenov's Characterisation [10])**.** *Let α be an ω-word. Then* $\mathsf{MTh}(\alpha)$ *is decidable if and only if the ω-word α is recursive and there exists a recursive indicator of recurrence for α.*

Note that an ω-word can have many recursive indicators of recurrence: if rec is such an indicator, then so is $\varphi \mapsto 2 \cdot \mathrm{rec}(\varphi)$.

Two other characterisations are given by Rabinovich and Thomas in [9]. The idea is to decompose an infinite word into infinitely many finite sections all of which (except possibly the first one) have the same k-type.

▶ **Definition 3.3.** Let $\alpha \in \{0,1\}^\omega$, $u, v \in \{0,1\}^+$, $k \in \mathbb{N}$, and $H \subseteq \mathbb{N}$ be infinite.

- The set H is a *k-homogeneous factorisation of α into (u,v)* if $\alpha[0, i-1] \equiv_k u$ and $\alpha[i, j-1] \equiv_k v$ for all $i, j \in H$ with $i < j$. The set H is *k-homogeneous for α* if it is a k-homogeneous factorisation of α into some finite words (u,v).
- Let $H = \{h_i \mid i \in \mathbb{N}\}$ with $h_0 < h_1 < \dots$. The set H is *uniformly homogeneous for α* if, for all $k \in \mathbb{N}$, the set $\{h_i \mid i \geq k\}$ is k-homogeneous for α.

As with indicators of recurrence, any ω-word has many uniformly homogeneous sets: the existence of at least one follows by a repeated and standard application of Ramsey's theorem, and there are infinitely many since any infinite subset of a uniformly homogeneous set is again uniformly homogeneous.

▶ **Theorem 3.4** (1st Rabinovich-Thomas' Characterisation [9])**.** *Let α be an ω-word. Then* $\mathsf{MTh}(\alpha)$ *is decidable if and only if the ω-word α is recursive and there exists a recursive uniformly homogeneous set for α.*

Suppose $h_0 < h_1 < h_2 < \dots$ is an enumeration of some uniformly homogeneous set for α. This sequence determines finite words u_k and v_k such that $w \equiv_k u_k(v_k)^\omega$, $u_k v_k \equiv_k u_k$, and $v_k v_k \equiv_k v_k$: simply set $u_k = \alpha[0, h_k - 1]$ and $v_k = \alpha[h_k, h_{k+1} - 1]$. If the ω-word α is recursive, we can therefore, from $k \in \mathbb{N}$, compute a representative of the k-type of α.

▶ **Definition 3.5.** Let α be some ω-word and $\mathrm{tp}\colon \mathbb{N} \to \{0,1\}^+ \times \{0,1\}^+$. The function tp is a *type-function* if, for all $k \in \mathbb{N}$, α has a k-homogeneous factorisation into $\mathrm{tp}(k) = (u,v)$.

Let tp be a type-function for the ω-word α and let $k \in \mathbb{N}$. Then there exists a k-homogeneous factorisation H of α into $\mathrm{tp}(k) = (u,v)$. Let $H = \{h_0 < h_1 < h_2 < \dots\}$. Then we have $\alpha = \alpha[0, h_0 - 1]\, \alpha[h_0, h_1 - 1]\alpha[h_1, h_2 - 1] \cdots \equiv_k uv^\omega$. Furthermore, $v \equiv_k \alpha[h_0, h_2 - 1] = \alpha[h_0, h_1 - 1]\alpha[h_1, h_2 - 1] \equiv_k vv$. Consequently, $\mathrm{tp}(k)$ is a representative of the k-type of α.

▶ **Theorem 3.6** (2nd Rabinovich-Thomas' Characterisation [9])**.** *Let α be an ω-word. Then* $\mathsf{MTh}(\alpha)$ *is decidable if and only if α has a recursive type-function.*

Note that, differently from Thm. 3.4 this theorem does not mention that α is recursive. But this recursiveness is implicit: Let tp be a recursive type-function and $k \in \mathbb{N}$. Then one can write a FO sentence of quantifier-depth $k+2$ expressing that $\alpha(k) = 1$. Let $\mathrm{tp}(k+2) = (u,v)$. Then $\alpha \equiv_{k+2} uv^\omega$ implies $\alpha(k) = uv^k(k)$, hence $\alpha(k)$ is computable from k.

4 How hard is it to tell if the MSO-theory of an ω-word is decidable?

In this section, we determine the recursion-theoretical complexity of the question whether the MSO-theory of a recursive ω-word is decidable. Technically, we will consider the following two sets:

$$\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}} = \{e \in \mathsf{REC} \mid \mathsf{MTh}(\mathbb{N}, \leq, W_e) \text{ is decidable}\} \qquad \mathsf{UndecTh}^{\mathsf{MSO}}_{\mathbb{N}} = \mathsf{REC} \setminus \mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$$

Recall that $W_e \subseteq \mathbb{N}$ denotes the e^{th} recursively enumerable set.

But first note the following: Let α be some recursive word. Then, by Büchi's and McNaughton's theorems, $\mathsf{MTh}(\alpha)$ is decidable iff the set of deterministic parity automata accepting α is decidable. Recall that "the deterministic parity automaton no. n accepts α" (where we assume any computable enumeration of all deterministic parity automata) is a Boolean combination of Σ_2-statements, cf. [15, Prop. 5.3]. It follows that $e \in \mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ if and only if the following holds:

$$\exists f \in \mathsf{TOT}\, \forall n\colon \Phi_f(n) = 1 \Leftrightarrow \text{the deterministic parity automaton no. } n \text{ accepts } (\mathbb{N}, \leq, W_e)$$

Hence $\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ belongs to Σ_4. The following lemma improves this by one level in the arithmetical hierarchy:

▶ **Lemma 4.1.** *The set* $\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ *belongs to* Σ_3.

We present two proofs of this lemma, one based on the first Rabinovich-Thomas characterisation, the second one based on the Semenov characterization.

Proof. (based on Thm. 3.4) Let α be some recursive ω-word. Recall that a set $H \subseteq \mathbb{N}$ is infinite and recursive if there exists a total computable and strictly monotone function f such that $H = \{f(n) \mid n \in \mathbb{N}\}$. Now consider the following:

$$\begin{aligned}\exists e\, \forall k, i, j, i', j'\colon\ & e \in \mathsf{TOT} \wedge\ (i < j \Rightarrow \Phi_e(i) < \Phi_e(j)) \wedge \\ & (k \leq i < j \wedge k \leq i' < j' \Rightarrow \alpha[\Phi_e(i), \Phi_e(j) - 1] \equiv_k \alpha[\Phi_e(i'), \Phi_e(j') - 1])\end{aligned}$$

It expresses that there exists a total recursive function (namely Φ_e) that is strictly monotone. Its image then consists of the numbers $\Phi_e(0) < \Phi_e(1) < \Phi_e(2) < \dots$ The last line expresses that this image is uniformly homogeneous for α. Hence this statement says that there exists a recursive uniformly homogeneous set for α, i.e., that $\mathsf{MTh}(\alpha)$ is decidable by Thm. 3.4.

From $k, i, i', j, j' \in \mathbb{N}$ with $k \leq i < j$, and $k \leq i' < j'$ we can compute the finite words $\alpha[\Phi_e(i), \Phi_e(j) - 1]$ and $\alpha[\Phi_e(i'), \Phi_e(j') - 1]$ since α is recursive. Hence it is decidable whether $\alpha[\Phi_e(i), \Phi_e(j) - 1] \equiv_k \alpha[\Phi_e(i'), \Phi_e(j') - 1]$. The whole statement is in Σ_3 as $\mathsf{TOT} \in \Pi_2$. ◀

Proof. (based on Thm. 3.2) We enumerate the set Sent of MSO-sentences in any effective way as $\varphi_0, \varphi_1, \dots$. Let $e \in \mathsf{TOT}$. Then the function $\mathrm{rec}\colon \mathsf{Sent} \to \mathbb{N}\colon \varphi_i \mapsto \Phi_e(i)$ is an indicator of recurrence for the ω-word α if and only if the following holds for all $\varphi \in \mathsf{Sent}$

$$(\mathrm{rec}(\varphi) \neq \top \Rightarrow \forall k \geq j \geq \mathrm{rec}(\varphi)\colon \alpha[j, k] \models \neg\varphi)\ \wedge\ (\mathrm{rec}(\varphi) = \top \Rightarrow \forall j \exists \ell \geq k \geq j\colon \alpha[k, \ell] \models \varphi)$$

Given the definition of rec, this is equivalent to requiring (for all $i \in \mathbb{N}$)

$$(\Phi_e(i) \neq \top \Rightarrow \forall k \geq j \geq \Phi_e(i)\colon \alpha[j, k] \models \neg\varphi_i)\ \wedge\ (\Phi_e(i) = \top \Rightarrow \forall j \exists \ell \geq k \geq j\colon \alpha[k, \ell] \models \varphi_i)$$

If α is recursive, this is a Π_2-statement. Prefixing it with $\exists e \in \mathsf{TOT}\, \forall i$ yields a Σ_3-statement that expresses the existence of a recursive indicator of recurrence. ◀

▶ Remark. From the 2nd characterisation by Rabinovich and Thomas (Thm. 3.6), we can only infer that $\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ is in Σ_5: Let α be some recursive ω-word and $u, v \in \{0,1\}^+$. Then, by the proof of [9, Prop. 7], there exists a k-homogeneous factorisation of α into (u, v), if the following Σ_3-statement $\varphi(u, v)$ holds: $\exists x \forall y \exists z, z'\colon (\alpha[0, x-1] \equiv_k u \wedge y < z < z' \wedge \alpha[x, z-1] \equiv_k \alpha[z, z'-1] \equiv_k v)$. Hence the function $\mathrm{tp}\colon \mathbb{N} \to \{0,1\}^+ \times \{0,1\}^+$ is a type-function if the Π_4-statment $\forall k \in \mathbb{N}\colon \varphi(\mathrm{tp}(k))$ holds. Consequently, there is a recursive type-function if we have $\exists e\colon e \in \mathsf{TOT} \wedge \forall k\colon \varphi(\Phi_e(k))$ which is a Σ_5-statement.

The above raises the natural question whether these characterisations are "optimal". Namely, if one can separate $\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ from $\mathsf{UndecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ using a simpler statement. We now prepare a negative answer to this last question (which is an affirmative answer to the optimality question posed first).

We now construct an m-reduction from the set REC to any separator of $\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ and $\mathsf{UndecTh}^{\mathsf{MSO}}_{\mathbb{N}}$: Let $e \in \mathbb{N}$. One can compute $f \in \mathbb{N}$ such that Φ_f is total and injective and $\{\Phi_f(i) \mid i \in \mathbb{N}\} = \{2a \mid a \in W_e\} \cup (2\mathbb{N}+1)$. For $i \in \mathbb{N}$, set $x_i = 2^{\Phi_f(i)} \times \prod_{0 \le j \le i}(2j+1)$ and consider the ω-word $\alpha_e = 10^{x_0}10^{x_1}10^{x_2}\cdots$. Since Φ_f is total, this ω-word is recursive.

▶ **Lemma 4.2.** *Let $e \in \mathbb{N}$. The* MSO*-theory of the ω-word α_e is decidable if and only if the e^{th} recursively enumerable set W_e is recursive, i.e., $e \in$* REC.

Proof. First suppose that the MSO-theory of α_e is decidable. For $a \in \mathbb{N}$, we have $a \in W_e$ iff there exists $i \ge 0$ with $2a = \Phi_f(i)$ iff there exists $i \ge 0$ such that 2^{2a} is the greatest power of 2 that divides x_i. Consequently, $a \in W_e$ if the ω-word α_e satisfies

$$\exists x, y \in P\colon\ (x < y \wedge \forall z\colon (x < z < y \Rightarrow z \notin P)) \ \wedge\ \left(2^{2a} \mid y - x - 1 \wedge 2^{2a+1} \not| y - x - 1\right) \quad (1)$$

Recall that $n \mid y - x - 1$ is expressible by an MSO-formula. Since validity in α_e of the resulting MSO-sentence is decidable, the set W_e is recursive.

Conversely, let W_e be recursive. To show that the MSO-theory of α_e is decidable, let φ be some MSO-sentence. Let $k = \mathsf{qr}(\varphi)$ be the quantifier-rank of φ. To decide whether $\alpha_e \models \varphi$, we proceed as follows:

- Using standard semigroup arguments, compute $\ell > 0$ such that $0^\ell \equiv_k 0^{2\ell}$ and determine $a, b \in \mathbb{N}$ with $\ell = 2^a(2b+1)$.
- Compute $i \ge b$ such that $\Phi_f(j) > a$ for all $j > i$: to this aim, first determine $A = \{n \le a \mid n \in W_e \text{ or } a \text{ odd}\}$ which is possible since W_e is decidable. Then compute the least $i \ge b$ such that $A \subseteq \{\Phi_f(j) \mid j \le i\}$. Since Φ_f is injective, $\Phi_f(j) > a$ for all $j > i$.
- Decide whether $10^{x_0}10^{x_1}\dots10^{x_i}(10^\ell)^\omega$ satisfies φ which is possible since this ω-word is ultimately periodic.

Let $j > i$. Then $\Phi_f(j) > a$ and $j > i \ge a$ imply that x_j is a multiple of ℓ. Thus $0^{x_j} \equiv_k 0^\ell$. We therefore obtain $\alpha_e \equiv_k 10^{x_1}10^{x_2}\cdots10^{x_i}(10^\ell)^\omega$. Hence the above algorithm is correct. ◀

Lemmas 4.2 and 4.1 imply that the problem of deciding whether a recursive ω-word has a decidable MSO-theory is Σ_3-complete:

▶ **Theorem 4.3.**
- $\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ *is in Σ_3.*
- *Any set containing $\mathsf{DecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ and disjoint from $\mathsf{UndecTh}^{\mathsf{MSO}}_{\mathbb{N}}$ is Σ_3-hard.*

Remark. Thm. 3.4 also holds for the weaker logics FO and FO+MOD that extends FO by modulo-counting quantifiers [9]. Consequently, Lemma 4.1 also holds, *mutatis mutantis*, for these logics.

Conversely, Lemma 4.2 also holds for FO+MOD since (1) is easily expressible in this logic. To also handle FO, replace the definition of x_i by $x_i = \Phi_f(j)$. A similar argument as in Lemma 4.2 proves that W_e is recursive iff the ω-word α_e obtained this way has a decidable FO-theory. Thus, Thm. 4.3 also holds for the logics FO and FO+MOD.

5 When is the MSO-theory of a bi-infinite word decidable?

In this section, we investigate whether the characterisations from Theorems 3.2, 3.4, and 3.6 can be lifted from ω- to bi-infinite words.

5.1 A characerization à la Semenov

▶ **Definition 5.1.** Let ξ be a bi-infinite word. A pair of functions $(\mathrm{rec}_\leftarrow, \mathrm{rec}_\rightarrow)$ with $\mathrm{rec}_\leftarrow, \mathrm{rec}_\rightarrow \colon \mathsf{Sent} \to \mathbb{Z} \cup \{\top\}$ is an *indicator of recurrence for* ξ if for any $\varphi \in \mathsf{Sent}$:

- if $\mathrm{rec}_\leftarrow(\varphi) = \top$, $\forall k \in \mathbb{Z}\, \exists i \le j \le k \colon \xi[i,j] \models \varphi$; otherwise, $\forall i \le j \le \mathrm{rec}_\leftarrow(\varphi) \colon \xi[i,j] \models \neg\varphi$
- if $\mathrm{rec}_\rightarrow(\varphi) = \top$, $\forall k \in \mathbb{Z}\, \exists j \ge i \ge k \colon \xi[i,j] \models \varphi$; otherwise, $\forall j \ge i \ge \mathrm{rec}_\rightarrow(\varphi) \colon \xi[i,j] \models \neg\varphi$

A bi-infinite word ξ "consists" of an ω^*-word $\xi_\leftarrow$ and an ω-word $\xi_\rightarrow$. Then, roughly speaking, an indicator of recurrence for the *bi-infinite* word ξ consists of a pair of indicators of recurrence, one for $\xi_\leftarrow$ and one for $\xi_\rightarrow$. Therefore, the following is similar to Thm. 3.2.

▶ **Theorem 5.2.** *Let ξ be a bi-infinite word. Then $\mathsf{MTh}(\xi)$ is decidable if and only if ξ has a recursive indicator of recurrence and the bi-infinite word ξ is recursive or recurrent.*

This theorem is an immediate consequence of Propositions 5.3 and 5.4 below. If ξ is non-recurrent, there is a finite word u that has a leftmost or a rightmost occurrence in ξ, say at a position $x \in \mathbb{Z}$. Then x is definable in MSO. Consequently, also the position 0 is definable. This allows one to reduce the decidability of $\mathsf{MTh}(\xi)$ to the decidability of both $\mathsf{MTh}(\xi(-\infty,-1])$ and $\mathsf{MTh}(\xi[0,\infty))$. Hence Prop. 5.3 is a consequence of Thm. 3.2.

▶ Proposition 5.3. Let ξ be a non-recurrent bi-infinite word. Then $\mathsf{MTh}(\xi)$ is decidable if and only if ξ has a recursive indicator of recurrence and the bi-infinite word ξ is recursive.

▶ Proposition 5.4. Let ξ be a recurrent bi-infinite word. Then $\mathsf{MTh}(\xi)$ is decidable if and only if ξ has a recursive indicator of recurrence.

Proof. First suppose $\mathsf{MTh}(\xi)$ is decidable. We have to construct a recursive indicator of recurrence $(\mathrm{rec}_\leftarrow, \mathrm{rec}_\rightarrow)$ for ξ. Let $\varphi \in \mathsf{Sent}$. Set $\mathrm{rec}_\leftarrow(\varphi) = \mathrm{rec}_\rightarrow(\varphi) = \top$ if there exist integers $i \le j$ with $\xi[i,j] \models \varphi$, otherwise set $\mathrm{rec}_\leftarrow(\varphi) = \mathrm{rec}_\rightarrow(\varphi) = 0$.

It remains to be shown that these functions are recursive and that they form an indicator of recurrence. Regarding the recursiveness, note that there are $i \le j$ with $\xi[i,j] \models \varphi$ iff $\xi \models \exists x, y \colon x \le y \wedge \varphi_{x,y}$ where $\varphi_{x,y}$ is obtained form φ by restricting all quantifiers to the interval $[x,y]$. Since $\mathsf{MTh}(\xi)$ is decidable, the functions $\mathrm{rec}_\leftarrow$ and $\mathrm{rec}_\rightarrow$ are recursive.

Next we show that $(\mathrm{rec}_\leftarrow, \mathrm{rec}_\rightarrow)$ is an indicator of recurrence for ξ: If $\mathrm{rec}_\leftarrow(\varphi) = \top$, then (by the definition of $\mathrm{rec}_\leftarrow$) there are $i \le j$ with $\xi[i,j] \models \varphi$. Since ξ is recurrent, it follows that there are arbitrary small and large integers $a \le b$ with $\xi[a,b] = \xi[i,j] \models \varphi$. If, in the other case, $\mathrm{rec}_\leftarrow(\varphi) = 0$, then there are no integers $i \le j$ with $\xi[i,j] \models \varphi$, in particular, there are no integers $i \le j \le \mathrm{rec}_\leftarrow(\varphi)$ with $\xi[i,j] \models \varphi$.

Conversely, suppose $(\mathrm{rec}_\leftarrow, \mathrm{rec}_\rightarrow)$ is a recursive indicator of recurrence for ξ. Then, for $\varphi \in \mathsf{Sent}$, we can decide whether there are integers $i \le j$ with $\xi[i,j] \models \varphi$ (since ξ is recurrent, this is the case if and only if $\mathrm{rec}_\leftarrow(\varphi) = \top$). In [1, Thm. 3.1(2)] and in [10, 7], it is stated that then $\mathsf{MTh}(\xi)$ is decidable (a proof can be extracted from [6, Section IX.6]). ◀

Thm. 5.2 connects the decidability of the MSO theory of a recurrent bi-infinite word ξ with a decidability question on its set of factors $F(\xi)$. It follows that, if $\mathsf{MTh}(\xi)$ is decidable, then $F(\xi)$ is decidable. We now show that the converse implication does not hold.

▶ **Lemma 5.5.** *A set of finite words F containing at least one non-empty word is the factor set of a recurrent bi-infinite word if and only if it satisfies the following conditions:*

(a) If $uvw \in F$, then $v \in F$.
(b) For any $u, w \in F$, there is a word $v \in F$ such that $uvw \in F$

Proof. Necessity of (a) and (b) is obvious. So suppose $F \subseteq \{0,1\}^*$ contains at least one non-empty word u and satisfies (a) and (b). We construct a bi-infinite recurrent word ξ such that $F(\xi) = F$. Since F is non-empty, (b) implies that F is infinite. Let $F = \{u_i \mid i \in \mathbb{N}\}$. Inductively, we define two sequences $(x_i)_{i>0}$ and $(y_i)_{i>0}$ of words from F such that, for all $i \in \mathbb{N}$, the finite word $w_i = u_i x_i u_{i-1} x_{i-1} \dots u_1 x_1 u_0 y_1 u_1 y_2 u_2 \dots y_i u_i$ belongs to F.

Let $i > 0$ and suppose we already defined the words x_j and y_j for $j < i$ such that $w_{i-1} \in F$. Then, by (b), there exists $x_i \in F$ such that $u_i x_i w_{i-1} \in F$. Again by (b), there exists $y_i \in F$ such that $u_i x_i w_{i-1} y_i u_i \in F$. Now set $\xi = \cdots u_3 x_3\, u_2 x_2\, u_1 x_1\, u_0\, y_1 u_1\, y_2 u_2\, y_3 u_3 \cdots$. Let $v \in \{0,1\}^*$ be some factor of ξ. Then there is $i \in \mathbb{N}$ such that v is a factor of w_i. Since $w_i \in F$, condition (a) implies $v \in F$. Hence $F(\xi) = F$.

Now let $v \in F(\xi) = F$. By (b), there are infinitely many $i \in \mathbb{N}$ such that v is a factor of u_i. Hence ξ is recurrent. ◀

▶ **Theorem 5.6.** *There exists a recurrent bi-infinite word ξ whose set of factors is decidable, but* $\mathsf{MTh}(\xi)$ *is undecidable.*

Proof. Let $f\colon \mathbb{N} \to \mathbb{N}$ be some recursive and total function such that $\{f(i) \mid i \in \mathbb{N}\}$ is not recursive. Let $F \subseteq \{0,1\}^*$ be the set of all finite words u with the following property: If $10^{2i+1}10^{2j}1$ is a factor of u, then $j = f(i)$. This set is clearly recursive, contains a non-empty word, and satisfies conditions (a) and (b) from Lemma 5.5. Hence there exists a bi-infinite word ξ with $F(\xi) = F$. For $j \in \mathbb{N}$, consider the following sentence:

$$\exists x < y\colon P(x) \wedge P(y+2j) \wedge \neg 2 \mid y - x - 1 \wedge \forall z\colon (x < z < y + 2j \wedge P(z) \to z = y)$$

It expresses that the language $1(00)^*010^{2j}1$ contains a factor of ξ. But this is the case iff it contains a factor of some word from F iff there exists $i \in \mathbb{N}$ with $j = f(i)$. Since this is undecidable, the MSO-theory of ξ is undecidable by Thm. 5.2. ◀

5.2 A characerization à la Rabinovich-Thomas I

We return to the question when the MSO-theory of a recurrent bi-infinite word is decidable. We will see that Thm. 3.4 naturally extends to *recursive* bi-infinite words. We will then demonstrate that it does not extend to non-recursive bi-infinite words.

▶ **Definition 5.7.** Let $\xi \in \{0,1\}^{\mathbb{Z}}$, $u, v, w \in \{0,1\}^+$, $k \in \mathbb{N}$, and let $H_\leftarrow = \{h_i^- \mid i \in \mathbb{N}\}$ and $H_\to = \{h_i^+ \mid i \in \mathbb{N}\}$ with $h_0^- > h_1^- > \dots$ and $h_0^+ < h_1^+ < \dots$.

- The pair $(H_\leftarrow, H_\to)$ is a *k-homogeneous factorisation of ξ into (u, v, w)* if
  - $\xi[i, j-1] \equiv_k u$ for all $i, j \in H_\leftarrow$ with $i < j$,
  - $\xi[i, j-1] \equiv_k v$ for all $i \in H_\leftarrow$ and $j \in H_\to$ with $i < j$ and
  - $\xi[i, j-1] \equiv_k w$ for all $i, j \in H_\to$ with $i < j$.
- The pair $(H_\leftarrow, H_\to)$ is *k-homogeneneous for ξ* if it is a k-homogeneous factorisation of ξ into some finite words (u, v, w).
- The pair $(H_\leftarrow, H_\to)$ is *uniformly homogeneous for ξ* if, for all $k \in \mathbb{N}$, the pair $(\{h_i^- \mid i \geq k\}, \{h_i^+ \mid i \geq k\})$ is k-homogeneous for ξ.

Let ξ be a bi-infinite word split into an ω^*-word $\xi_\leftarrow$ and an ω-word $\xi_\to$. As for any ω-word, there exists a uniformly homogeneous set $H_\to$ for $\xi_\to$. Symmetrically, there exists a set $H_\leftarrow \subseteq \widetilde{\mathbb{N}}$ that is "uniformly homogeneous" for $\xi_\leftarrow$. Then the pair $(H_\leftarrow, H_\to)$ is a uniformly homogeneous pair for $\xi = \xi_\leftarrow \xi_\to$.

▶ **Lemma 5.8.** *Let ξ be a recursive bi-infinite word with a decidable* MSO*-theory. Then the* MSO*-theories of $\xi_\leftarrow = \xi(-\infty,-1]$ and of $\xi_\rightarrow = \xi[0,\infty)$ are both decidable.*

Proof. We handle the cases of recurrent and non-recurrent words separately.

First let ξ be non-recurrent. Then some word $u \in F(\xi)$ has a leftmost or a rightmost occurrence, at some position $x \in \mathbb{Z}$ which is definable in FO. Hence, also the positions -1 and 0 are definable. Hence the MSO-theories of $\xi_\leftarrow$ and of $\xi_\rightarrow$ can be reduced to that of ξ and are therefore decidable.

Now let ξ be recurrent. By Thm. 5.2, ξ has a recursive indicator of recurrence $(\mathrm{rec}_\leftarrow, \mathrm{rec}_\rightarrow)$. Define the functions $f, g\colon \mathsf{Sent} \to \mathbb{N} \cup \{\top\}$ as follows:

$$f(\varphi) = \begin{cases} \top & \text{if } \mathrm{rec}_\leftarrow(\varphi) = \top \\ 0 & \text{if } \mathrm{rec}_\leftarrow(\varphi) \geq 0 \\ |\mathrm{rec}_\leftarrow(\varphi)| - 1 & \text{otherwise} \end{cases} \quad \text{and} \quad g(\varphi) = \begin{cases} \top & \text{if } \mathrm{rec}_\rightarrow(\varphi) = \top \\ 0 & \text{if } \mathrm{rec}_\rightarrow(\varphi) < 0 \\ \mathrm{rec}_\rightarrow(\varphi) & \text{otherwise} \end{cases}$$

Exploiting the properties of $\mathrm{rec}_\leftarrow$ and $\mathrm{rec}_\rightarrow$, it is then routine to check that f, g are indicators of recurrences for the two ω-words $\xi_\leftarrow^R$ and $\xi_\rightarrow$. Note that $\xi_\leftarrow^R$ and $\xi_\rightarrow$ are recursive ω-words. Hence, by Thm. 3.2, the MSO-theories of $\xi_\leftarrow^R$ and of $\xi_\rightarrow$ are both decidable. ◀

▶ **Theorem 5.9.** *A recursive bi-infinite word ξ has a decidable* MSO*-theory if and only if there exists a recursive uniformly homogeneous pair for ξ.*

Proof. Suppose $\mathsf{MTh}(\xi)$ is decidable. By Lemma 5.8, the MSO-theories of $\xi_\leftarrow^R = \xi(-\infty,-1]^R$ and of $\xi_\rightarrow = \xi[0,\infty)$ are both decidable. Consequently, by Thm. 3.4, there are recursive uniformly homogeneous factorisations $H_\leftarrow^R, H_\rightarrow \subseteq \mathbb{N}$ for $\xi_\leftarrow^R$ and $\xi_\rightarrow$ into (x^R, y^R) and (y', z), respectively. Deleting, if necessary, the minimal element from $H_\leftarrow^R$, we can assume $0 \notin H_\leftarrow^R$. We set $H_\leftarrow = \{-n \mid n \in H_\leftarrow^R\} \subseteq \widetilde{\mathbb{N}}$ and show that $(H_\leftarrow, H_\rightarrow)$ is a uniformly homogeneous pair for ξ: Let $H_\leftarrow = \{h_i^- \mid i \in \mathbb{N}\}$ and $H_\rightarrow = \{h_i^+ \mid i \in \mathbb{N}\}$ such that $h_0^- > h_1^- > \dots$ and $h_0^+ < h_1^+ < \dots$.

- Let $j > i \geq k$. Then $\xi[h_i^- + 1, h_j^-] = \xi_\leftarrow[h_i^- + 1, h_j^-] = (\xi_\leftarrow^R[-h_j^-, -h_i^- - 1])^R \equiv_k y^R$.
- Let $i, j \geq k$. Then $\xi[h_i^-, h_j^+ - 1] = \xi_\leftarrow[h_i^- + 1, 0]\, \xi_\rightarrow[0, h_j^+ - 1] \equiv_k xy'$
- Let $j > i \geq k$. Then $\xi[h_i^+, h_j^+ - 1] = \xi_\rightarrow[h_i^+, h_j^+ - 1] \equiv_k z$.

Hence the pair $(\{h_i^- \mid i \geq k\}, \{h_i^+ \mid i \geq k\})$ is a k-homogeneous factorisation of ξ into (y^R, xy', z). Since k is arbitrary, $(H_\leftarrow, H_\rightarrow)$ is uniformly homogeneous for ξ. Since these two sets are clearly recursive, this proves the first implication.

Conversely, suppose there exists a recursive uniformly homogeneous pair $(H_\leftarrow, H_\rightarrow)$ for ξ. Then the sets $H_\leftarrow^R = \{|n| \mid n \in H_\leftarrow \cap \widetilde{\mathbb{N}}\}$ and $H_\rightarrow \cap \mathbb{N}$ are recursive and uniformly homogeneous for $\xi_\leftarrow^R$ and $\xi_\rightarrow$, resp. Since $\xi_\leftarrow$ and $\xi_\rightarrow$ are both recursive, we can apply Thm. 3.4. Hence the infinite words $\xi_\leftarrow$ and $\xi_\rightarrow$ both have decidable MSO-theories. Since $\xi = \xi_\leftarrow \xi_\rightarrow$, the MSO-theory of ξ is decidable. ◀

We next show that we cannot hope to extend Thm. 5.9 to non-recursive words:

▶ **Theorem 5.10.** *There exists a recurrent r.e. bi-infinite word ξ with decidable* MSO*-theory such that there is no r.e. uniformly homogeneous pair for ξ.*

Proof. We prove this theorem by constructing a recurrent bi-infinite word ξ such that the set $F(\xi)$ of factors is $\{0,1\}^*$. Hence ξ has decidable MSO-theory by Thm. 5.2.

There is a computable function $f\colon \mathbb{N}^2 \to \mathbb{N}$ such that the following hold:

- $\Phi_{f(e,s)}$ is total and $W_{f(e,s)} \subseteq \{0, 1, \dots, s\}$ for any $e, s \in \mathbb{N}$.
- $W_e = \bigcup_{s \in \mathbb{N}} W_{f(e,s)}$ for any $e \in \mathbb{N}$.

In the following, we fix the function f and write $W_{e,s}$ for $W_{f(e,s)}$. Furthermore, we fix some recursive enumeration $u_0, u_1, \dots$ of the set $\{0,1\}^+$ of non-empty finite words.

Construction

By induction on $s \in \mathbb{N}$, we construct tuples

$$t_s = (w_s, m_{0,s}, m_{1,s}, \ldots, m_{s,s}, P_s) \in \{0,1\}^* \times \mathbb{N}^{s+1} \times 2^{\{0,\ldots,s\}} \qquad \text{such that}$$

- $m_{i,s} + |u_i| \le m_{i+1,s}$ for all $0 \le i < s$ and $m_{s,s} + |u_s| \le |w_s|$ (in particular, $|w_s| > s$),
- $w_s[m_{i,s}, m_{i,s} + |u_i| - 1] = u_i$ for all $0 \le i \le s$, and
- for all $e \in P_s$, there exist $a, b \in W_e$ with $a < b < |w_s|$ and $w_s[a, b-1] \in 1^*$.

Set $w_0 = u_0$, $m_{0,0} = 0$, and $P_0 = \emptyset$. Then the inductive invariant holds for the tuple $t_0 = (w_0, m_0, P_0)$.

Now suppose the tuple t_s has been constructed. Let H_{s+1} denote the set of indices $0 \le e \le s+1$ with $e \notin P_s$ such that $W_{e,s}$ contains at least two numbers $a > b \ge m_{e,s}$. In the construction of the tuple t_{s+1}, we distinguish two cases:

- 1st case: $H_{s+1} = \emptyset$. Then set $w_{s+1} = w_s u_{s+1}$, $m_{i,s+1} = m_{i,s}$ for $0 \le i \le s$, $m_{s+1,s+1} = |w_s|$, and $P_{s+1} = P_s$. Since the inductive invariant holds for the tuple t_s, it also holds for the newly constructed tuple t_{s+1}.
- 2nd case: $H_{s+1} \ne \emptyset$. Let e_{s+1} be the minimal element of H_{s+1} and let a_{s+1} and b_{s+1} be the minimal elements of $W_{e_{s+1},s}$ satisfying $m_{e,s} \le a_{s+1} < b_{s+1}$. Then set
  - $w_{s+1} = w_s[0, a_{s+1}-1]\, 1^{b_{s+1}-a_{s+1}}\, w_s[b_{s+1}, |w_s|-1]\, u_{e_{s+1}} u_{e_{s+1}+1} \ldots u_{s+1}$ (in other words, the words $u_{e_{s+1}}$ up to u_{s+1} are appended to w_s and the positions between a_{s+1} and $b_{s+1}-1$ are set to 1).
  - $m_{i,s+1} = \begin{cases} m_{i,s} & \text{if } i < e_{s+1} \\ |w_s u_{e_{s+1}} u_{e_{s+1}+1} \ldots u_{i-1}| & \text{if } e_{s+1} \le i \le s+1 \end{cases}$
  - $P_{s+1} = P_s \cup \{e_{s+1}\}$

  The first two conditions of the inductive invariant are obvious. Regarding the last one, let $e \in P_{s+1}$. If $e \ne e_{s+1}$, then $e \in P_s$ and therefore there exist $a, b \in W_e$ with $a < b < |w_s| < |w_{s+1}|$ such that $w_s[a, b-1] \in 1^*$. Note that any position in w_s that carries 1 also carries 1 in w_{s+1}. Hence $w_{s+1}[a, b-1] \in 1^*$ as well. It remains to consider the case $e = e_{s+1}$. But then, by the very construction, $a_{s+1} < b_{s+1}$ belong to $W_{e_{s+1},s} \subseteq W_e$ and satisfy $w_{s+1}[a_{s+1}, b_{s+1}-1] \in 1^*$.

This finishes the construction of the sequence of tuples t_s.

Verification

Let $\xi_\to$ be the ω-word with $\xi_\to(i) = 1$ iff there exists $s \in \mathbb{N}$ with $w_s(i) = 1$. Since the tuple t_{s+1} is computable from the tuple t_s, the word $\xi_\to$ is clearly recursively enumerable.

Furthermore, let $u \in \{0,1\}^+$. Then there exists $e \in \mathbb{N}$ with $u = u_e$. Note that $m_{e,s} \le m_{e,s+1}$ for all $e, s \in \mathbb{N}$. Furthermore, $m_{e,s} < m_{e,s+1}$ iff $H_{s+1} \ne \emptyset$ and $e_{s+1} \le e$. Since the numbers $e_{s'+1}$ for $s' \in \mathbb{N}$ (if defined) are mutually distinct, there exists $s \in \mathbb{N}$ such that $e_{t+1} > e$ and therefore $m_{e,s} = m_{e,t}$ for all $t \ge s$. Consequently, $\xi_\to[m_{e,s}, m_{e,s} + |u_e| - 1] = w_s[m_{e,s}, m_{e,s} + |u_e| - 1] = u_e = u$. This means that $F(\xi_\to) = \{0,1\}^*$. It follows that $\xi_\to$ is recurrent.

Claim 1. If W_e is infinite, then $e \in \bigcup_{s \in \mathbb{N}} P_s$.

Proof of Claim 1. By contradiction, suppose this is not the case. Let $e \in \mathbb{N}$ be minimal with W_e infinite and $e \notin \bigcup_{s \in \mathbb{N}} P_s$. Since W_e is infinite, we get $e \in H_{s+1}$ for almost all $s \in \mathbb{N}$. By minimality of e, there is $s \in \mathbb{N}$ with $e = \min H_{s+1}$. But then $e_{s+1} = e$ and $e \in P_{s+1}$. **q.e.d.**

Claim 2. No recursively enumerable set W is uniformly homogeneous for the ω-word $\xi_\to$.

Proof of Claim 2. Suppose W is recursively enumerable and uniformly homogeneous for $\xi_\to$. Then W is infinite and there exists $e \in \mathbb{N}$ with $W = W_e$. By claim 1, there exists $s \in \mathbb{N}$ with $e \in P_s$. Hence there are $a, b \in W_e$ with $w_s[a, b-1] \in 1^*$ and therefore $\xi_\to[a, b-1] = w_s[a, b-1]$. There are $d > c > b$ in W_e such that $\xi_\to[c, d-1] \notin 1^*$. But then $\xi_\to[a, b-1]$ and $\xi_\to[c, d-1]$ do not have the same 1-type. Hence the set W_e is not 1- and therefore not uniformly homogeneous for $\xi_\to$. **q.e.d.**

Finally, let $\xi_\leftarrow$ be the reversal of $\xi_\to$ and consider the bi-infinite word $\xi = \xi_\leftarrow \xi_\to$. By Thm. 5.2, $\mathsf{MTh}(\xi)$ is decidable since ξ is recurrent and contains every finite word as a factor. Finally, suppose $(H_\leftarrow, H_\to)$ is uniformly homogeneous for ξ. Then $H_\to \cap \mathbb{N}$ is uniformly homogeneous for $\xi_\to$. By claim 2, this set cannot be recursively enumerable. Hence $(H_\leftarrow, H_\to)$ is not recursively enumerable either. ◀

5.3 A characerization à la Rabinovich-Thomas II

We next extend the 2nd characterisation by Rabinovich and Thomas to bi-infinite words. Differently from the 1st characterization, this also covers non-recursive bi-infinite words.

▶ **Definition 5.11.** Let ξ be some bi-infinite word and $\mathrm{tp}\colon \mathbb{N} \to \{0,1\}^+ \times \{0,1\}^+ \times \{0,1\}^+$. The function tp is a *type-function for* ξ if, for all $k \in \mathbb{N}$, the bi-infinite word ξ has a k-homogeneous factorisation into $\mathrm{tp}(k)$.

▶ **Theorem 5.12.** *Let ξ be a bi-infinite word. Then* $\mathsf{MTh}(\xi)$ *is decidable if and only if ξ has a recursive type-function.*

Proof. First suppose that $\mathsf{MTh}(\xi)$ is decidable. We have to construct a recursive type-function $\mathrm{tp}\colon \mathbb{N} \to (\{0,1\}^+)^3$. To this aim, let $k \in \mathbb{N}$. Then one can compute a finite sequence $\varphi_1, \ldots, \varphi_n$ of MSO-sentences of quantifier-rank k such that, for all finite words u and v, we have $u \equiv_k v$ if and only if $\forall 1 \le i \le n\colon u \models \varphi_i \iff v \models \varphi_i$. For finite words u, v, and w, consider the following statement:

$$\begin{aligned} \exists H_\leftarrow, H_\to\colon \quad & \forall y \exists x, z\colon (x < y < z \wedge H_\leftarrow(x) \wedge H_\to(z)) \\ \wedge \quad & \forall x, y\colon (x < y \wedge H_\leftarrow(x) \wedge H_\leftarrow(y) \to \xi[x, y-1] \equiv_k u) \\ \wedge \quad & \forall x, y\colon ((H_\leftarrow(x) \wedge H_\to(y) \wedge x < y \to \xi[x, y-1] \equiv_k v) \\ \wedge \quad & \forall x, y\colon (x < y \wedge H_\to(x) \wedge H_\to(y) \to \xi[x, y-1] \equiv_k w) \end{aligned}$$

This statement holds for a bi-infinite word ξ iff ξ has a k-homogeneous factorisation into (u, v, w). Using $\varphi_1, \ldots, \varphi_n$, the statements $\xi[x, y-1] \equiv_k u$ etc. can be expressed as MSO-formulas with free variables x and y. Since $\mathsf{MTh}(\xi)$ is decidable, we can decide (given k, u, v, and w) whether ξ has a k-homogeneous factorisation into (u, v, w). Since some k-homogeneous factorisation always exist, this allows to compute, from k, a tuple $\mathrm{tp}(k)$ such that ξ has a k-homogeneous factorisation into $\mathrm{tp}(k)$; tp is the wanted type function.

Conversely suppose that tp is a recursive type-function for ξ. To show that $\mathsf{MTh}(\xi)$ is decidable, let $\varphi \in \mathsf{Sent}$ be any MSO-sentence. Let k denote the quantifier-rank of φ. First, compute $\mathrm{tp}(k) = (u, v, w)$. Then $\xi \models \varphi$ iff $u^{\omega^*} v w^\omega \models \varphi$ which is decidable since this bi-infinite word is ultimately periodic on the left and on the right. ◀

6 How complicated are bi-infinite words with decidable MSO-theories?

By Thm. 5.2, non-recurrent bi-infinite words with decidable MSO-theory are recursive. In this section, we will show in a strong sense that this does not hold for recurrent bi-infinite words: there are "arbitrarily complicated" bi-infinite words with decidable MSO-theories.

▶ **Definition 6.1.** Let $L \subseteq \{0,1\}^*$ be a language. A word $u \in L$ is *left-determined in L* if for any $k \in \mathbb{N}$ there is exactly one word $vu \in L$ with $|v| = k$. Similarly, u is *right-determined in L* if for any $k \in \mathbb{N}$ there is exactly one word $uv \in L$ with $|v| = k$. The word $u \in L$ is *determined in L* if it is both left- and right-determined.

Intuitively, a word $w \in L$ is left-determined (right-determined) in L if it can be extended on the left (right) in a unique way.

▶ **Lemma 6.2.** *Let ξ be a recurrent bi-infinite word. The following are equivalent:*

(1) ξ is periodic
(2) $F(\xi)$ contains a determined word
(3) $F(\xi)$ contains a right-determined word
(3') $F(\xi)$ contains a left-determined word

Proof. For (1)→(2), let $\xi = u^{\omega^*}u^\omega$ be a periodic word. Then u is determined in $F(\xi)$. The direction (2)→(3) is trivial by the very definition.

For (3)→(1), suppose u is a right-determined word in $F(\xi)$. Choose $i < j$ such that $\xi[i, i + |u| - 1] = \xi[j, j + |u| - 1] = u$ (such a pair $i < j$ exists since ξ is recurrent). With $p = j - i$, we claim $\xi(n) = \xi(n + p)$ for all $n \in \mathbb{Z}$: First let $n \geq j + |u|$. Then $\xi[i, n]$ and $\xi[j, n + p]$ are two words from $F(\xi)$ that both start with u. We have $|\xi[i, n]| = n - i - 1 = n + p - j - 1 = |\xi[j, n + p]|$. Since u is right-determined, this implies $\xi[i, n] = \xi[j, n + p]$ and therefore $\xi(n) = \xi(n + p)$. Consequently, $\xi[j + |u|, \infty) = \xi[j + |u|, j + |u| + p]^\omega$. Next let $n < j + |u|$. Since ξ is recurrent, there is $k < n$ with $\xi[k, k + |u| - 1] = u$. Since u is right-determined, this implies $\xi[k, \infty) = \xi[j + |u|, \infty) = \xi[j + |u|, j + |u| + p]^\omega$ and therefore in particular $\xi(n) = \xi(n + p)$. The implications (2)→(3')→(1) are shown analogously. ◀

Lemma 6.2 states that a recurrent non-periodic bi-infinite word does not contain any left-determined or right-determined factor, and thus can be extended in both directions (left and right) in at least two ways. This observation allows to prove the following:

▶ **Lemma 6.3.** *Let ξ be a recurrent non-periodic bi-infinite word. For any set $A \subseteq \mathbb{N}$, there is a recurrent bi-infinite word ξ_A such that $F(\xi) = F(\xi_A)$, $(A, F(\xi)) \leq_T \xi_A$, and $\xi_A \leq_T (A, F(\xi))$.*

Proof. Let $w_0, w_1, \ldots$ be the enumeration of $F(\xi)$ in length-lexicographic order. Note that this is recursive in $F(\xi)$. There is also an effective enumeration of all pairs of words of the same length, say $(\ell_0, r_0), (\ell_1, r_1), \ldots$. Now let $A \subseteq \mathbb{N}$ be arbitrary. We will construct a sequence of tuples $t_s = (u_s, v_s, x_s, y_s) \in (\{0,1\}^*)^4$ such that, for all $s \in \mathbb{N}$, the finite word

$$\begin{aligned} z_s &= w_s y_s v_s \, z_{s-1} \, u_s x_s w_s \\ &= w_s y_s v_s \, w_{s-1} y_{s-1} v_{s-1} \ldots w_0 y_0 v_0 \; u_0 x_0 w_0 \ldots u_{s-1} x_{s-1} w_{s-1} \, u_s x_s w_s \end{aligned}$$

belongs to $F(\xi)$ (the bi-infinite word ξ_A will be the "limit" of these words).

To start with $s = 0$ note the following: since ξ is recurrent and $w_0 \in F(\xi)$, the bi-infinite word ξ contains a factor of the form $w_0 x w_0$. Set $y_0 = x$ and $u_0 = v_0 = x_0 = \varepsilon$.

For the induction step, assume that we constructed the tuple t_s and that z_s is a factor of ξ. Since ξ is recurrent but not periodic, the word z_s is not right-determined in $F(\xi)$ by Lemma 6.2. Hence there are two distinct finite words u and u' of the same length such that $z_s u, z_s u' \in F(\xi)$. For (u, u'), choose the first such pair in the effective enumeration $(\ell_i, r_i)_{i \in \mathbb{N}}$. If $s \in A$, then set $u_{s+1} = u$, otherwise set $u_{s+1} = u'$. Now the word $z_s u_{s+1}$ is a factor of ξ. Since ξ is recurrent, there is $x_{s+1} \in \{0,1\}^*$ such that $z_s u_{s+1} x_{s+1} w_{s+1}$ is a factor of ξ – choose x_{s+1} length-lexicographically minimal among all possible such words.

To choose v_{s+1} and y_{s+1}, we proceed symmetrically to the left: $z'_s = z_s u_{s+1} x_{s+1} w_{s+1}$ is a factor of ξ that is not left-determined. Hence there exists a pair of distinct words v and v' of the same length with $vz'_s, v'z'_s \in F(w)$. Choose this pair minimal in the effective enumeration. If $s \in A$, then set $v_{s+1} = v$, otherwise set $v_{s+1} = v'$. Now there is $y_{s+1} \in \{0,1\}^*$ with $w_{s+1} y_{s+1} v_{s+1} z'_s \in F(\xi)$ since ξ is recurrent. Choosing y_{s+1} length-lexicographically minimal completes the construction of the tuple t_{s+1} and therefore the inductive construction of all the tuples t_s. Now set $\xi_A = \cdots w_1 y_1 v_1\, w_0 y_0 v_0\, u_0 x_0 w_0\, u_1 x_1 w_1 \cdots$. Observe the following:

- If $u \in F(\xi)$, then there exists $s \in \mathbb{N}$ such that $u \in F(z_s)$. Hence $F(\xi) \subseteq F(\xi_A)$.
- Let $u \in F(\xi_A)$. There exists $s \in \mathbb{N}$ such that $u \in F(z_s)$. In particular, $F(\xi_A) \subseteq F(\xi)$. Since z_s is a factor of ξ, there are infinitely many $i \in \mathbb{N}$ such that z_s (and therefore u) is a factor of w_i. Hence the word ξ_A is recurrent.

Since the above describes how to compute the bi-infinite word ξ_A using the oracles A and $F(w)$, we get $\xi_A \leq_T (A, F(\xi))$.

It remains to be shown that $A \leq_T (\xi_A, F(\xi))$ holds: To determine whether $s \in A$ suppose we already know which of the natural numbers $i < s$ belong to A. Then the construction of ξ_A above allows to build t_s using the oracle $F(\xi)$. Now construct t_{s+1} assuming $s \in A$ again using the oracle $F(\xi)$. If the resulting word z_{s+1} is an initial segment of ξ_A, then $s \in A$. Otherwise, $s \notin A$. ◀

From this lemma and Thm. 5.2, we get immediately that indeed, every decidable theory of some recurrent bi-infinite word is represented in every Turing-degree:

▶ **Theorem 6.4.** *Let ξ be a recurrent non-periodic bi-infinite word and* **a** *a Turing-degree above the degree of* MTh(ξ)*. Then* **a** *contains a bi-infinite word ξ_A with* MTh(ξ_A) = MTh(ξ)*.*

7 How many indistinguishable bi-infinite words are there?

If α and β are MSO-equivalent ω-words, then $\alpha = \beta$. In this final section we study this question for bi-infinite words. Shift-equivalence and period will be important notions in this context: two bi-infinite words ξ and ζ are *shift-equivalent* if there is $p \in \mathbb{N}$ with $\xi(n) = \zeta(n+p)$ for all $n \in \mathbb{Z}$. Furthermore, the period of the bi-infinite word ξ is the least natural number $p > 0$ with $\xi(n) = \xi(n+p)$ for all $n \in \mathbb{Z}$ – clearly, the period need not exist. To count the number of MSO-equivalent bi-infinite words, we need a characterisation when two bi-infinite words are MSO-equivalent.

▶ **Theorem 7.1.** *[6, Chp. 9, Thm. 6.1] Two bi-infinite words ξ and ζ are* MSO*-equivalent if and only if one of the following conditions is satisfied:*

1. *ξ and ζ are shift-equivalent.*
2. *ξ and ζ are recurrent and have the same set of factors.*

This characterisation is the central ingredient in the proof of the following result:

▶ **Theorem 7.2.** *Let ξ be a bi-infinite word.*

(a) If ξ is periodic, then the cardinality of the type of ξ is finite and equals the period of ξ.
(b) If ξ is non-recurrent, then the cardinality of the type of ξ is $\aleph_0$.
(c) If ξ is recurrent and non-periodic, then the cardinality of the type of ξ is $2^{\aleph_0}$.

Proof. (a) Let p be the period of ξ. Since p is minimal, there are precisely p distinct bi-infinite words that are shift-equivalent with ξ. Since shift-equivalent words are MSO-equivalent, the type of ξ contains at least p elements. It remains to be shown that no further MSO-equivalent word exists. So let ζ be some MSO-equivalent word. Then ζ is p-periodic since ξ (and therefore ζ) satisfies $\forall x\colon (P(x) \Leftrightarrow P(x+p))$ and does not satisfy $\forall x\colon (P(x) \Leftrightarrow P(x+q))$ for any $1 \leq q < p$. Furthermore $u = \xi[1,p]$ is a factor of ξ and therefore of ζ of length p. Hence $\zeta = u^{\omega^*} u^{\omega}$.

(b) This claim follows immediately from Thm. 7.1.

(c) This follows from Thm. 6.4 as there are $2^{\aleph_0}$ Turing-degrees above any Turing-degree. ◀

References

1 Bès, A., Rabinovich, A.: Decidable expansions of labelled linear orderings. Volume 7(2), 2011. Logical Methods in Computer Science.

2 Büchi, J.: On a decision method in restricted second order arithmetic, Methodology and Philosophy of Science (Proc. 1960 Internat. Congr .), Stanford Univ. Press, Stanford, Calif., 1962, 1-11.

3 Büchi, J., Landweber, L.: Definability in the monadic second-order theory of successor. Journal of Symbolic Logic 34, 166–170 (1969).

4 Carton, O., Thomas, W.: The monadic theory of morphic infinite words and generalizations. Information and Computation. 176(1), 51–56 (2002).

5 Elgot, C., Rabin, M.: *Decidability and undecidability of extensions of second (first) order theory of (generalized) successor.* J.Symb. Logic, 31(2), 169-181, 1996

6 Perrin, D., Pin, J.: Infinite words: automata, semigroups, logic and games. Pure and Applied Mathematics Series, vol. 141. Elsevier, Amsterdam, 2004.

7 Perrin, D., Schupp, P.E.: Automata on the Integers, Recurrence Distinguishability, and the Equivalence and Decidability of Monadic Theories. Proceedings of the Symposium on Logic in Computer Science, 301-304, 1986.

8 Rabinovich, A.: On decidability of monadic logic of order over the naturals extended by monadic predicates. Information and Computation Volume 205, Issue 6, June 2007, 870-889.

9 A. Rabinovich, W. Thomas: Decidable theories of the ordering on natural numbers with unary predicates, Computer Science Logic Lecture Notes in Computer Science Volume 4207, 2006, 562-574.

10 Semenov, A.: Logical theories of one-place functions on the set of natural numbers. Mathematics of the USSR - Izvestia, vol. 22, pp 587-618, 1984

11 Semenov, A.: Decidability of monadic theories. Mathematical foundations of computer science, Proc. 11th Symp., Praha/Czech. 1984, Lect. Notes Comput. Sci. 176, 162-175 (1984).

12 Shelah, S.: The monadic theory of order. Annals of Mathematics, vol. 102, 379–419, (1957).

13 Siefkes, D.: Decidable extensions of monadic second order successor arithmetic. In Automatentheorie und formale Sprachen, 441–472. Mannheim: Bibliogr. Inst. 1969.

14 Soare, R.: Recursively enumerable sets and degrees: A Study of Computable Functions and Computably Generated Sets, Perspectives in Mathematical Logic, Springer, (1987).

15 Thomas, W.: Languages, automata, and logic. In: Handbook of Formal Languages (G. Rozenberg, A. Salomaa, eds.), Vol 3, 389–455, Springer, (1997).