RAIRO-Theor. Inf. Appl. 47 (2013) 201–214
DOI: 10.1051/ita/2013035

Available online at:
www.rairo-ita.org

DECIDABILITY OF THE HD0L ULTIMATE PERIODICITY PROBLEM

Fabien Durand

Abstract. In this paper we prove the decidability of the HD0L ultimate periodicity problem.

Mathematics Subject Classification. 68Q45, 03B25.

1. Introduction

1.1. The HD0L ultimate periodicity problem

In this paper we prove the decidability of the following problem:

Input: Two finite alphabets A and B, an endomorphism $\sigma : A^* \to A^*$, a word $w \in A^*$ and a morphism $\phi : A^* \to B^*$.

Question: Do there exist two words u and v in B^*, with v non-empty, such that the sequence $(\phi(\sigma^n(w)))_n$ converges to uv^ω (that is, it is ultimately periodic)?

(The convergence of the sequence $(\phi(\sigma^n(w)))_n$ meaning that $(|\phi(\sigma^n(w))|)_n$ goes to $+\infty$ and that $(\phi(\sigma^n(www\ldots)))_n$ converges in $B^{\mathbb{N}}$ endowed with the usual product topology.) We will refer to it as the *HD0L ultimate periodicity problem*. Observe that it is slightly more general than the classical statement where it is assumed in the input that the sequence $(\phi(\sigma^n(w)))_n$ converges.

Theorem 1.1. *The HD0L ultimate periodicity problem is decidable.*

This result was announced in [8]. While we were finishing this paper, I. Mitrofanov put on Arxiv [20] another solution to this problem.

Keywords and phrases. HD0L – periodicity – decidability – return words.

[1] Université de Picardie Jules Verne, Laboratoire Amiénois de Mathématiques Fondamentales et Appliquées, CNRS-UMR 7352, 33 rue Saint Leu, 80039 Amiens Cedex 01, France. fabien.durand@u-picardie.fr

This problem was open for about 30 years.

In 1986, positive answers were given independently for D0L systems (or purely substitutive sequences) in both [10, 24], and, for automatic sequences (which are particular HD0L sequences) in [11]. Other proofs have been given for the D0L case in [12] and for automatic sequences in [1].

Recently in [8] the primitive case has been solved.

In [14] an equivalent statement of the HD0L ultimate periodicity problem in terms of recognizable sets of integers and abstract numeration systems, is given. In fact, J. Honkala already gave a positive answer to this question in [11] but in the restricted case of the usual integer bases, that is, for k-automatic sequences or constant length substitutive sequences. Recently, in [3], a positive answer was given for a (large) class of numeration systems including, for instance, the Fibonacci numeration system.

Let us point out that the characterisation of recognizable sets of integers for abstract numeration systems in terms of substitutions given in [19] (see also [16]), together with Theorem 1.1, provides a decision procedure to test whether a recognizable set of integers in some abstract numeration system is a finite union of arithmetic progressions.

1.2. Organisation of the paper

In Section 2 are the classical definitions.

In Section 3 we prove the HD0L ultimate periodicity problem for substitutive sequences. These sequences are such that $(\sigma^n(w))_n$ converges. This avoids testing the existence of the limit. Indeed, there are examples where $(\phi(\sigma^n(w)))_n$ converges and $(\sigma^n(w))_n$ does not: for σ and ϕ, defined by $\sigma(a) = cb$, $\sigma(b) = ba$, $\sigma(c) = ab$, $\phi(a) = \phi(c) = 0$ and $\phi(b) = 1$, the sequence $(\sigma^n(a))_n$ does not converge but $(\phi(\sigma^n(a)))_n$ does (to the Thue–Morse sequence).

Under these assumptions the proof could be sketched as follows. First we recall some primitivity arguments about matrices and substitutions. The ”best or easiest situation” is when we deal with growing substitutions and codings (letter-to-letter morphisms). It is known that we can always consider we are working with codings (see [2,4,6,22]). In [13] it is shown this can be algorithmically realised. We propose a different algorithm using the proof of [4] where we replace some (non-algorithmic) arguments (Lem. 2, Lems 3 and 4 of this paper) with algorithmic ones.

We treat growing and non-growing substitutions separately. For growing substitutions we look at their primitive components and we use the decidability result established in [8] about periodicity for primitive substitutions. Indeed, these primitive components should generate periodic sequences. Hence, we check it is the case (if not, then the sequence is not ultimately periodic). From there, Lemma 3.10 allows us to conclude.

For the non-growing case we use a result of Pansiot [23] saying that we can either consider we are in the growing case or there are longer and longer periodic words with the same period in the sequence. We again conclude with Lemma 3.10.

In Section 4 we show how to use the substitutive case to solve the general HD0L case. This concludes the proof of Theorem 1.1.

1.3. Questions and comments

We did not compute the complexity of the algorithm provided by our proof of the HD0L ultimate periodicity problem. Looking at Proposition 3.3 and the results in [8] that we use here, our approach provides a high complexity.

Our result is for one-dimensional sequences. What can be said about multidimensional sequences generated by substitution rules or self-similar tilings? It seems hopeless to generalise our method to tilings, although the main and key result we use to solve the HD0L ultimate periodicity problem (that is, the main result in [7], see [8]) has been generalised to higher dimensions by N. Priebe in [25] (see also [26]). However observe that in [17] the author gives a polynomial time algorithm to know whether or not a Number Decision Diagram defines a Presburger definable set (see also [21] where it was first proven but with a much greater complexity). From this result and [5,28,29] it is decidable to know whether a multidimensional automatic sequence (or fixed point of a multidimensional "uniform" substitution) has a certain type of periodicity (see [17,21]). From [9] this type of periodicity is equivalent to a block complexity condition.

2. Words, morphisms, substitutive and HD0L sequences

In this section we recall classical definitions and notation. Observe that the notion of substitution we use below could be slightly different from other definitions in literature.

2.1. Words and sequences

An *alphabet* A is a finite set of elements called *letters*. Its cardinality is $|A|$. A *word* over A is an element of the free monoid generated by A, denoted by A^*. Let $x = x_0x_1 \dots x_{n-1}$ (with $x_i \in A$, $0 \leq i \leq n-1$) be a word, its *length* is n and is denoted by $|x|$. The *empty word* is denoted by ϵ, $|\epsilon| = 0$. The set of non-empty words over A is denoted by A^+. The elements of $A^{\mathbb{N}}$ are called *sequences*. If $x = x_0x_1 \dots$ is a sequence (with $x_i \in A$, $i \in \mathbb{N}$) and $I = [k, l]$ an interval of $\mathbb{N}$ we set $x_I = x_kx_{k+1} \dots x_l$ and we say that x_I is a *factor* of x. If $k = 0$, we say that x_I is a *prefix* of x. The set of factors of length n of x is written $\mathcal{L}_n(x)$ and the set of factors of x, or the *language* of x, is denoted by $\mathcal{L}(x)$. The *occurrences* in x of a word u are the integers i such that $x_{[i,i+|u|-1]} = u$. If u has an occurrence in x, we also say that u *appears* in x. When x is a word, we use the same terminology with similar definitions.

The sequence x is *ultimately periodic* if there exist a word u and a non-empty word v such that $x = uv^\omega$, where $v^\omega = vvv\dots$. In this case v is called a *word period* and $|v|$ is called a *length period* of x. It is *periodic* if u is the empty word. A word u is *recurrent* in x if it appears in x infinitely often. The sequence x is *uniformly* *recurrent* if all words in its language appear infinitely often in x and with bounded gaps.

2.2. Morphisms and matrices

Let A and B be two alphabets. Let σ be a *morphism* from A^* to B^*. When $\sigma(A) \subset B$, we say σ is a *coding*. We say σ is *erasing* if there exists $b \in A$ such that $\sigma(b)$ is the empty word. Such a letter is called an *erased letter* (w.r.t. σ). If $\sigma(A)$ is included in B^+, it induces by concatenation a map from $A^{\mathbb{N}}$ to $B^{\mathbb{N}}$. This map is also denoted by σ. With the morphism σ is naturally associated its *incidence matrix* $M_\sigma = (m_{i,j})_{i\in B, j\in A}$ where $m_{i,j}$ is the number of occurrences of i in the word $\sigma(j)$.

Let σ be an endomorphism. We say it is *primitive* whenever its incidence matrix is primitive that is, when it has a power with positive coefficients). We denote by $\mathcal{L}(\sigma)$ the set of words that have an occurrence in some image of σ^n for some $n \in \mathbb{N}$. We call it the language of σ.

2.3. Substitutions and substitutive sequences

We say that an endomorphism $\sigma : A^* \to A^*$ is *prolongable on* $a \in A$ if there exists a word $u \in A^+$ such that $\sigma(a) = au$ and, moreover, if $\lim_{n\to+\infty} |\sigma^n(a)| = +\infty$. Prolongable endomorphisms are called *substitutions*.

We say a letter $b \in A$ is *growing* (w.r.t. σ) if $\lim_{n\to+\infty} |\sigma^n(b)| = +\infty$. We say σ is *growing* whenever all letters of A are growing.

Since for all $n \in \mathbb{N}$, $\sigma^n(a)$ is a prefix of $\sigma^{n+1}(a)$ and because $(|\sigma^n(a)|)_n$ tends to infinity with n, the sequence $(\sigma^n(aaa\ldots))_n$ converges (for the usual product topology on $A^{\mathbb{N}}$) to a sequence denoted by $\sigma^\omega(a)$. The endomorphism σ being continuous for the product topology, $\sigma^\omega(a)$ is a fixed point of σ: $\sigma(\sigma^\omega(a)) = \sigma^\omega(a)$. A sequence obtained in this way (by iterating a prolongable substitution) is said to be *purely substitutive* (w.r.t. σ). If $x \in A^{\mathbb{N}}$ is purely substitutive and $\phi : A^* \to B^*$ is a morphism then the sequence $y = \phi(x)$ is said to be a *morphic sequence* (w.r.t. (σ, ϕ)). When ϕ is a coding, we say y is *substitutive* (w.r.t. (σ, ϕ)). In these cases, when σ is primitive, y is uniformly recurrent (see [27]).

2.4. D0L and HD0L sequences

A *D0L system* is a triple $G = (A, \sigma, u)$ where A is a finite alphabet, $\sigma : A^* \to A^*$ is an endomorphism and u is a word in A^*. An *HD0L system* is a 5-tuple $G = (A, B, \sigma, \phi, u)$ where (A, σ, u) is a D0L system, B is a finite alphabet and $\phi : A^* \to B^*$ is a morphism. If it converges, the limit of $(\phi(\sigma^n(uuu\ldots))_n$ is called a *HD0L sequence*.

It is clear that substitutive sequences are HD0L sequences. We will show in the last section that HD0L sequences are substitutive sequences. Nevertheless, as the initial data are not the same, it is not enough to solve the ultimate periodicity problem for substitutive sequences. Indeed, if $(\sigma^n(uuu\ldots))_n$ does not converge it seems difficult to decide whether $(\phi(\sigma^n(uuu\dots))_n$ converges. We leave this question as an open problem.

3. Ultimate periodicity of substitutive sequences

In this section we prove the decidability of the HD0L ultimate periodicity problem for substitutive sequences.

In the sequel $\sigma : A^* \to A^*$ is a substitution prolongable on a, $\phi : A^* \to B^*$ is a morphism, $y = \sigma^\omega(a)$ and $x = \phi(y)$ is a sequence of $B^{\mathbb{N}}$. We have to find an algorithm which decides whether x is ultimately periodic or not.

3.1. Primitivity assumption and sub-substitutions

We recall that the HD0L ultimate periodicity problem has already been solved in the primitive case.

Theorem 3.1 [8]**.** *The HD0L ultimate periodicity problem is decidable in the context of primitive substitutions. Moreover, a word period can be explicitly computed.*

Proof. The first part is Theorem 26 in [8]. The second part can be easily deduced from the proof of this theorem. □

The following lemma shows that it is decidable to check that a nonnegative matrix is primitive.

Lemma 3.2 [15]**.** *The $n \times n$ nonnegative matrix M is primitive if and only if M^{n^2-2n+2} has positive entries.*

From Lemma 3.2, Sections 4.4 and 4.5 in [18] we deduce the following proposition.

Proposition 3.3. *Let $M = (m_{i,j})_{i,j\in A}$ be a matrix with non-negative coefficients. There exist three positive integers $p \neq 0$, q, l, where $q \le l-1$, and a partition $\{A_i; 1 \le i \le l\}$ of A such that*

$$M^p = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_q \\ A_{q+1} \\ A_{q+2} \\ \vdots \\ A_l \end{array} \begin{array}{c} \begin{array}{cccccccc} A_1 & A_2 & \dots & A_q & A_{q+1} & A_{q+2} & \dots & A_l \end{array} \\ \left(\begin{array}{cccccccc} M_1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ M_{1,2} & M_2 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{1,q} & M_{2,q} & \dots & M_q & 0 & 0 & \dots & 0 \\ M_{1,q+1} & M_{2,q+1} & \dots & M_{q,q+1} & M_{q+1} & 0 & \dots & 0 \\ M_{1,q+2} & M_{2,q+2} & \dots & M_{q,q+2} & 0 & M_{q+2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ M_{1,l} & M_{2,l} & \dots & M_{q,l} & 0 & 0 & \dots & M_l \end{array}\right) \end{array},$$

where the matrices M_i have only positive entries or are equal to zero. Moreover, the partition and p can be algorithmically computed.

The next three corollaries are consequences of Proposition 3.3.

The following corollary will be helpful to change the representation of x (in terms of (σ, ϕ)) to a more convenient representation.

Corollary 3.4. *Let $\tau : A^* \to A^*$ be an endomorphism whose incidence matrix has the form of M^p in Proposition 3.3. Then, for all $b \in A$ and all $j \geq 1$, the letters which have an occurrence in $\left(\tau^{|A|}\right)^j (b)$ or $\left(\tau^{|A|}\right)^{j+1} (b)$ are the same.*

Proof. Let us take the notation of Proposition 3.3. Let $A^{(0)}$ (resp. $A^{(1)}$) be the set of letters b belonging to some A_i where M_i is the null matrix (resp. is not the null matrix).

The conclusion is a consequence of the following two remarks. Let $b \in A_i$. From the shape of the incidence matrix of τ we get:

- If b belongs to $A^{(1)}$, then all letters occurring in $\tau(b)$ occur in $\tau^n(b)$ for all n.
- If b belongs to $A^{(0)}$, then all letters occurring in $\tau(b)$ belong to some A_j with $j > i$.

This achieves the proof. □

Corollary 3.5. *It is decidable whether a given letter is growing for a given endomorphism.*

Proof. Let τ be an endomorphism. Let us take the notation of Proposition 3.3. Let $A^{(0)}$ (resp. $A^{(1)}$) be the set of letters b belonging to some A_i where M_i is the null matrix (resp. the 1×1 matrix $[1]$). The letters belonging to $A \setminus (A^{(0)} \cup A^{(1)})$ are growing.

Let $b \in A_i \cap A^{(1)}$ for some i. Then, from Corollary 3.4, b is non-growing (w.r.t. τ) if and only if all letters occurring in $\tau^{p|A|}(b)$, except b, are erased with respect to $\tau^{p|A|}$. Let A' be the set of such non-growing letters.

Let $b \in A_i \cap A^{(0)}$ for some i. Then b is non-growing if and only if all letters occurring in $\tau^{p|A|}(b)$ are erased with respect to $\tau^{p|A|}$ or belong to A'.

Moreover, from Proposition 3.3 and Corollary 3.4 we can decide whether a letter is erased w.r.t. $\tau^{p|A|}$. This achieves the proof. □

In what follows we keep the notation of Proposition 3.3. We will say that $\{A_i; 1 \leq i \leq l\}$ is a *primitive component partition of A (with respect to M)*, the A_i being the *primitive components*. If i belongs to $\{q+1, \ldots, l\}$ we will say that A_i is a *principal primitive component of A (with respect to M)*.

Let $\tau : A^* \to A^*$ be a substitution whose incidence matrix has the form of M^p in Proposition 3.3. Let $i \in \{q+1, \ldots, l\}$. We denote τ_i the restriction $\tau_{/A_i} : A_i^* \to A^*$ of τ to A_i^*. Because $\tau_i(A_i)$ is included in A_i^* we can consider that τ_i is an endomorphism of A_i^* whose incidence matrix is M_i. When it defines a substitution, we say it is a *sub-substitution of* τ. Moreover the matrix M_i has positive coefficients which implies that the substitution τ_i is primitive.

A non-trivial primitive endomorphism always has some power that is a substitution. For non-primitive endomorphisms we have the following corollary.

Corollary 3.6. *Let $\tau : A^* \to A^*$ be an endomorphism whose incidence matrix has the form of M^p in Proposition 3.3. Then, with the notation of Proposition 3.3, there exists $k \leq |A|^{|A|}$ satisfying: for all $i \geq q+1$ such that M_i is neither a null matrix nor the 1×1 matrix $[1]$, the endomorphism τ_i^k is a (primitive) substitution for some letter in A_i.*

Proof. We only have to check there exists $k \leq |A|^{|A|}$ such that for all $i \geq q+1$ there exists a letter $b \in A_i$ satisfying $\tau_i^k(b) = bu$ for some non-empty word u.

Let $i \geq q+1$ and $c \in A_i$. There exist $k_i \geq 1$ and $j \geq 0$, with $k_i + j \leq |A|$, such that $\tau_i^j(c)$ and $\tau_i^{k_i+j}(c)$ start with the same letter b. That is to say, $\tau_i^{k_i}(b) = bu$ for some u. To conclude, it suffices to take $k = k_{q+1} \dots k_l$. □

The following lemma is easy to establish.

Lemma 3.7. *Let $x = \phi(\sigma^\omega(a))$. If $x = uv^\omega$, where v is not the empty word, then each sub-substitution σ' of σ such that $\mathcal{L}(\sigma') \subset \mathcal{L}(\sigma^\omega(a))$ verifies $\phi(\mathcal{L}(\sigma')) \subset \mathcal{L}(v^\omega)$.*

Proof. Let σ' be a sub-substitution of σ. Its incidence matrix being primitive, there exists an uniformly recurrent sequence z such that $\mathcal{L}(\sigma') = \mathcal{L}(z)$ (see [27]). Thus, the words of $\mathcal{L}(\sigma')$ appear infinitely often in $\sigma^\omega(a)$. Finally, for all $w \in \mathcal{L}(\sigma')$, $\phi(w)$ should occur in v^ω. □

3.2. Reduction of the problem

It may happen, as for σ defined by $a \mapsto ab$, $b \mapsto a$, $c \mapsto c$, that some letter of the alphabet, here the alphabet is $\{a, b, c\}$, does not appear in $\sigma^\omega(a)$. It is preferable to avoid this situation. Corollary 3.4 enables us to avoid this algorithmically. We explain this below. Indeed, from Proposition 3.3, taking a power of σ (that can be algorithmically found) if needed, we can suppose

(P1) the incidence matrix of σ has the form of M^p in Proposition 3.3.

Then, consider $\sigma^{|A|}$ instead of σ. Hence, σ will continue to satisfy (P1) and, from Corollary, 3.4 we have

(P2) for all $b \in A$ and all $j \geq 1$ the letters which have an occurrence in $\sigma^j(b)$ or $\sigma^{j+1}(b)$ are the same.

Notice that, as σ is a substitution, taking a power of σ instead of σ will change neither y nor x. It will not be the case when we will deal with endomorphisms which are not substitutions.

Let A' be the set of letters appearing in $\sigma^\omega(a)$. From (P2) it can be checked that $\sigma(A')$ is included in A'^* and that the set of letters appearing in $\sigma(a)$ is A'. Thus σ', the restriction of σ to A', defines a substitution prolongable on a satisfying $\sigma'^\omega(a) = \sigma^\omega(a)$ such that all letters of A' have an occurrence in $\sigma'^\omega(a)$ and all letters of A' occur in $\sigma'(a)$. Hence we can always suppose σ and a satisfy the following condition.

(P3) The set of letters occurring in $\sigma^\omega(a)$ is A.

When we work with morphic sequences it is much simpler to handle non-erasing substitutions and even better to suppose that ϕ is a coding. Such a reduction is possible as shown in [4].

Theorem 3.8. *Let x be a morphic sequence. Then, x is substitutive with respect to a non-erasing substitution.*

This result was previously proven in [6,22] (see also [2,4]). It was shown that it could be algorithmically done in [13]. In the sequel we give another algorithm.

The proof of J. Cassaigne and F. Nicolas is short and inspired by [7], in particular its second part which is clearly algorithmic. Whereas the first part (Lemma 2, Lemma 3 and Lemma 4 of [4]) is not because it uses the fact that from any sequence of integers, we can extract a subsequence that is either constant or strictly increasing. They use these lemmas to show the key point of their proof: we can always suppose that ϕ and σ fulfill the following:

$$|\phi(\sigma(a))| > |\phi(a)| > 0 \text{ and } |\phi(\sigma(b))| \geq |\phi(b)| \text{ for all } b \in A. \tag{3.1}$$

Below we show that this can be algorithmically realised. This provides another algorithm for Theorem 3.8.

First let us show that σ can be supposed to be non-erasing. As we explained before, there is no restriction to suppose σ satisfies (P1), (P2) and (P3).

As σ satisfies (P2), each letter e is either erased or, for all l, $\sigma^l(e)$ is not the empty word. Let A' be the set of non-erased letters and A'' the set of erased letters. Let ψ be the morphism that sends the elements of A'' to the empty word and that is the identity for the other letters. Then, we define σ' to be the unique endomorphism defined on A' satisfying $\psi \circ \sigma = \sigma' \circ \psi$. Observe that σ' is easily algorithmically definable and prolongable on a. Moreover we have $\sigma\psi = \sigma$. Let $z = \sigma'^\omega(a)$. Then, $\psi(y) = z$ and $\sigma(z) = y$.

Notice that σ' is non-erasing. Indeed, if $\sigma'(a') = \epsilon$ for some $a' \in A'$, then $\psi(\sigma(a')) = \sigma'(a') = \epsilon$. Hence $\sigma(a') = b_1 \dots b_k$ where the b_i's belong to A''. Then $\sigma^2(a') = \epsilon$. But, from Property (P2), $\sigma^2(a')$ is not the empty word.

Thus we can also consider

(P4) σ is non-erasing.

Consequently, from (P2), $|\phi \circ \sigma(\sigma(a))| > |\phi(\sigma(a))| > |\phi(a)|$, otherwise $\phi(\sigma^\omega(a))$ would not be an infinite sequence. Hence, replacing ϕ with $\phi \circ \sigma$ if needed, we can suppose ϕ and σ are such that $|\phi(\sigma(a))| > |\phi(a)| > 0$.

Moreover, we claim that $\sigma^2(b) = \sigma(b)$ for all non-growing letters $b \in A$. Let b be a non-growing letter. As σ is non-erasing we necessarily have $|\sigma^2(b)| \geq |\sigma(b)|$. Suppose $|\sigma^2(b)| > |\sigma(b)|$. Then, the letters occurring in $\sigma^2(b)$ and $\sigma(b)$ being the same, we would have $|\sigma^n(b)| \geq n+1$ for all n, and, b would not be a growing letter. Consequently, $|\sigma^2(b)| = |\sigma(b)|$. Let $\sigma(b) = b_1 b_2 \dots b_l$. Then, $|\sigma(b_i)| = 1$ for all i, and, from the shape of the incidence matrix of σ, $\sigma(b_i) = b_i$ for all i.

Therefore, replacing ϕ with $\phi \circ \sigma$ if needed, we can suppose $|\phi(\sigma^n(b))| \geq |\phi(b)|$ for all non-growing letters b and all n.

Again, replacing σ with σ^k, where $k = \max_{a \in A} |\phi(a)|$, if needed, we can suppose (3.1) holds for σ and ϕ.

Hence, together with the argument of the proof of Theorem 3.8 we obtain the algorithm we are looking for. This is summarised in the following theorem (first proved in [13]).

Theorem 3.9. *There exists an algorithm that given ϕ and σ compute a coding φ and a non-erasing substitution τ, prolongable on a, such that $x = \varphi(z)$ where $z = \tau^\omega(a)$.*

Thus, in the sequel we suppose ϕ is a coding and σ is a non-erasing substitution. We end this section with a technical lemma checking the ultimate periodicity.

Lemma 3.10. *Let $t \in A^{\mathbb{N}}$, φ be a coding defined on A^*, $z = \varphi(t)$, and, u and v be non-empty words. Then, $z = uv^\omega$ iff and only if for all recurrent words $B = b_1 b_2 \dots b_{2|v|} \in \mathcal{L}(t)$, where the b_i's are letters, there exist $r_B \in \{0, 1, 2\}$, s_B and p_B such that*

1. *$\varphi(B) = s_B v^{r_B} p_B$ where s_B is a suffix of v and p_B a prefix of v, and,*
2. *for all recurrent words $BB' \in \mathcal{L}(t)$, where B' is a word of length $2|v|$, $p_B s_{B'}$ is equal to v or the empty word.*

Proof. The proof is left to the reader. □

3.3. The case of substitutive sequences with respect to growing substitutions

In the sequel we suppose σ is a growing substitution. From Corollary 3.5 it is decidable to know whether we are in this situation.

We recall that from the previous section we can suppose ϕ is a coding and that σ satisfies (P1), (P2), (P3) and (P4).

Lemma 3.11. *Let u and v be two words. It is decidable to check whether or not $\mathcal{L}(u^\omega)$ is equal to $\mathcal{L}(v^\omega)$.*

Lemma 3.12. *The set of recurrent letters in $\sigma^\omega(a) = c_0 c_1 \dots$ is algorithmically computable. Moreover there is a computable i such that all letters occurring in $c_i c_{i+1} \dots$ are recurrent.*

Proof. Let $\sigma(a) = au$. Then, $\sigma^\omega(a) = au\sigma(u)\sigma^2(u)\dots$. Thus, from (P2), a letter is recurrent if and only if it appears in $\sigma(u)$. Moreover, all letters occurring in $\sigma(u)\sigma^2(u)\dots$ are recurrent. □

Lemma 3.13. *The set of recurrent words of length n in $\sigma^\omega(a) = c_0 c_1 \dots$ is algorithmically computable. Moreover there is a computable i such that all words of length n occurring in $c_i c_{i+1} \dots$ are recurrent.*

Proof. Let $n \in \mathbb{N}$. Let w_0 be the prefix of length n of $\sigma^n(a)$. Let $w_1, \ldots, w_{j_1}$ be the words of length n appearing in $\sigma(w_0)$. Then we do the same for w_1. We obtain some new words of length n: $w_{j_1+1}, \ldots, w_{j_2}$. We proceed similarly with w_2, w_3 and so on, until all the w_i are handled and no new words appear. At this point, the set A' of all collected words is the set of all words of length n occurring in $\sigma^\omega(a)$.

It remains to find the words in A' that are recurrent in $\sigma^\omega(a)$.

Consider A' as a new alphabet and $\sigma_n : A'^* \to A'^*$ the endomorphism defined, for all $(a_1 \ldots a_n)$ in A', by

$$\sigma_n((a_1 \ldots a_n)) = (b_1 \ldots b_n)(b_2 \ldots b_{n+1}) \ldots (b_{|\sigma(a_1)|} \ldots b_{|\sigma(a_1)|+n-1})$$

where $\sigma(a_1 \ldots a_n) = b_1 \ldots b_k$. Let $\sigma^\omega(a) = c_0 c_1 \ldots$, with $c_i \in A$, $i \geq 0$. It is easy to check that σ_n is prolongable on $c = (c_0 c_1 \ldots c_{n-1})$ and that

$$\sigma_n^\omega(c) = (c_0 \ldots c_{n-1})(c_1 \ldots c_n)(c_2 \ldots c_{n+1}) \ldots .$$

For details, see Section V.4 in [27]. Thus a word w of length n is recurrent in $\sigma^\omega(a)$ if and only if (w) (which is a letter of A') is recurrent in $\sigma_n^\omega(c)$.

We achieve the proof using Lemma 3.12 □

Theorem 3.14. *The HD0L ultimate periodicity problem is decidable for substitutive sequences w.r.t. growing substitutions. Moreover, some u and v in the description of the problem can be computed.*

Proof. In this proof we suppose σ is growing. Let us use the notation of Proposition 3.3. From Corollary 3.6, taking a power of σ (less than $|A|^{|A|}$) if needed, we can suppose that for all $i \geq q+1$ the endomorphism $\sigma_i : A_i^* \to A_i^*$ defines a primitive sub-substitution w.r.t. some letter $a_i \in A_i$. We recall that all sub-substitutions are primitive. We notice that, in the growing case, there is at least one sub-substitution.

Observe that for all $i \geq q+1$ and $b \in A_i$, the word $\sigma^n(b) = \sigma_i^n(b)$ is recurrent in $\sigma^\omega(a)$. Thus, to check the periodicity of x, we start checking with Theorem 3.1 that, for all $i \geq q+1$, the sequence $\phi(\sigma_i^\omega(a_i))$ is periodic. We point out that when the language is periodic then a word period $w(\sigma_i)$ can be computed. If for some σ_i the sequence $\phi(\sigma_i^\omega(a_i))$ is not periodic then x cannot be ultimately periodic. Indeed, suppose $x = uv^\omega$. As longer and longer words occurring in $\phi(\sigma_i^\omega(a_i))$ occurs in x, the uniform recurrence would imply that $\phi(\sigma_i^\omega(a_i)) = v^\omega$.

Then, we check that all the languages $\mathcal{L}(w(\sigma_i)^\omega)$ are equal using Lemma 3.11. From Lemma 3.7, if this check fails, then x is not periodic.

Hence we suppose this is the case: There exists a word v that is algorithmically given by Theorem 3.1 such that $\phi((w(\sigma_i))^\omega) = \mathcal{L}(v^\omega)$ for all i. Consequently, we should check whether u exists such that $x = uv^\omega$.

We conclude using Lemmas 3.13 and 3.10. □

3.4. The case of substitutive sequences with respect to non-growing substitutions

In the sequel we suppose that σ is a non-growing substitution. From Corollary 3.5 it is decidable to know whether we are in this situation. We recall that from the previous section we can suppose ϕ is a coding and that σ satisfies (P1), (P2), (P3) and (P4).

Lemma 3.15 ([23], Théorème 4.1)**.** *The substitution σ satisfies exactly one of the following two statements.*

1. *The length of words (occurring in $\sigma^\omega(a)$) consisting of non-growing letters is bounded.*
2. *There exists a growing letter $b \in A$, occurring in $\sigma^\omega(a)$, such that $\sigma(b) = vbu$ (or ubv) with $u \in C^* \setminus \{\epsilon\}$ where C is the set of all non-growing letters w.r.t. σ.*

Moreover, in situation (1) the sequence $\sigma^\omega(a)$ can be algorithmically defined as a substitutive sequence w.r.t. a growing substitution.

Lemma 3.16. *It is decidable to know whether σ satisfies (1) or (2) of Lemma 3.15.*

Proof. It can be easily algorithmically checked whether we are in situation (2) of Lemma 3.15. Thus it is decidable to know whether we are in situation (1) of Lemma 3.15. □

Theorem 3.17. *The HD0L ultimate periodicity problem is decidable for substitutive sequences w.r.t. non-erasing substitutions. Moreover, some u and v in the description of the problem can be computed.*

Proof. From Theorem 3.14, Lemma 3.15 and Lemma 3.16 it remains to consider that σ satisfies (2) in Lemma 3.15: Let b be a letter occurring in $\sigma^\omega(a)$ such that $\sigma(b) = vbu$ (or ubv) with $u \in C^* \setminus \{\epsilon\}$ where C is the set of all non-growing letters w.r.t. σ. Then, for all n, $\sigma^{n+1}(b) = \sigma^n(v)bu\sigma(u)\ldots\sigma^n(u)$. As the sequence $(|\sigma^n(u)|)_n$ is bounded, there exist i and j, $i < j$, such that $\sigma^i(u) = \sigma^j(u)$. Let $u' = \sigma^i(u)\sigma^{i+1}(u)\ldots\sigma^{j-1}(u)$. Then, we get $\mathcal{L}(u'^\omega) \subset \mathcal{L}(\sigma)$. We conclude using Lemmas 3.13 and 3.10. □

This provides the proof for Theorem 1.1 for substitutive sequences.

Theorem 3.18. *Suppose that the sequence x is substitutive with respect to (σ, ϕ). Then, it is decidable whether x is ultimately periodic: $x = uv^\omega$ for some u and non-empty v. Moreover, we can compute such u and v.*

4. Ultimate periodicity of HD0L sequences

In this section finish providing the proof for the Theorem 1.1: We solve the HD0L ultimate periodicity problem. We use the notation introduced in the input of the problem. We recall that in the previous section we prove this theorem for a special case of HD0L sequences: the substitutive sequences. These sequences are very convenient as, by definition, there is no problem with the existence of the limit in the statement of the HD0L ultimate periodicity problem. We gave, in Section 1.2, an example of an HD0L sequence where the sequence $(\sigma^n(a))_n$ does not converge but $(\phi(\sigma^n(a)))_n$ does.

Thus, in the general case, it would be convenient (but not necessary) to be able to decide the existence of the limit. As we did not succeed in solving this decidability problem, we leave this question as an open problem. We proceed in a different way.

Let us consider the input of the HD0L ultimate periodicity problem.

Lemma 4.1. *Let $a \in A$. Suppose σ satisfies* (P1) *and* (P2)*. Then, it is decidable whether:*

1. $(|\phi(\sigma^n(a))|)_n$ *tends to* 0,
2. $(|\phi(\sigma^n(a))|)_n$ *tends to infinity.*

Moreover, if $(|\phi(\sigma^n(a))|)_n$ does not tend to infinity then it is bounded.

Proof. Let A' be the set of letters occurring in $\sigma(a)$. We prove the decidability of (1). From (P2), for all $n \geq 1$, the set of letters occurring in $\sigma^n(a)$ is A'. Then, $(|\phi(\sigma^n(a))|)_n$ tends to 0 if and only if $\phi(a')$ is the empty word for all $a' \in A'$.

We prove the decidability of (2). Let us consider the notation of Proposition 3.3 for σ. As σ satisfies (P1) we can suppose $p = 1$.

Suppose a belongs to A_l. Then $(|\phi(\sigma^n(a))|)_n$ tends to infinity if and only if M_l is neither the 1×1–matrix $[1]$ nor the null matrix, and, there exists a letter $b \in A_l$ such that $\phi(b)$ is not the empty word. Thus for such a letter the problem is decidable. Moreover, if $(|\phi(\sigma^n(a))|)_n$ does not tend to infinity, then it is bounded.

Now we proceed by a finite induction. Suppose the problem is decidable for all letters in $\cup_{n+1\leq j\leq l} A_j$. We show it is decidable for all letters in $\cup_{n\leq j\leq l} A_j$.

Suppose a belongs to A_n. If M_n is the null matrix, then we conclude with our induction hypothesis.

Suppose M_n is the 1×1–matrix $[1]$. Then, $(|\phi(\sigma^n(a))|)_n$ tends to infinity if and only if there is a letter a' in $A' \setminus \{a\}$ such that $(|\phi(\sigma^n(a'))|)_n$ does not tend to zero. Hence the decidability is deduced from (1). Moreover, if $(|\phi(\sigma^n(a))|)_n$ does not tend to infinity, then it is bounded.

Suppose M_n is neither the 1×1–matrix $[1]$ nor the null matrix. Then, $(|\phi(\sigma^n(a))|)_n$ tends to infinity if and only if there exists a letter in A' such that $\phi(a')$ is not empty. Moreover, if $(|\phi(\sigma^n(a))|)_n$ does not tend to infinity, then it goes to 0 and thus is bounded. □

Let us conclude with the HD0L ultimate periodicity problem.

Let us first suppose that σ satisfies (P1) and (P2).

Let $w = w_0 \dots w_{|w|-1}$ where the w_i's belong to A. As we want to test the ultimate periodicity, from Lemma 4.1, we can suppose $(|\phi(\sigma^n(w_0))|)_n$ tends to infinity. Consequently we can suppose $w = w_0$. We set $a = w_0$.

Let j_0 be the smallest integer less or equal to $|A| + 1$ such that $\sigma^{j_0}(a)$ and $\sigma^{j_0+n_0}(a)$ start with the same first letter for some n_0 verifying $j_0 + n_0 \leq |A| + 1$. Such integers exist from the pigeon hole principle. We also assume n_0 is the smallest such integer. Let a_i be the first letter of $\sigma^{j_0+i}(a)$, $0 \leq i \leq n_0 - 1$. Notice that if $(|\phi(\sigma^{j_0+i+kn_0}(a_i))|)_k$ tends to infinity then $(\phi(\sigma^{j_0+i+kn_0}(a_i)))_k$ converges in $B^{\mathbb{N}}$. From Lemma 4.1 it is decidable to know whether $(|\phi(\sigma^{j_0+i+kn_0}(a_i))|)_k$ tends to infinity. Let Λ be the set of such a_i's. Then, the set of accumulation points in $B^{\mathbb{N}}$ of $(\phi(\sigma^n(w)))_n$ is computable: it is the set of the infinite sequences $\lim_{k\to+\infty} \phi(\sigma^{j_0+i+kn_0}(a_i))$ where a_i belongs to Λ.

Consequently, $(\phi(\sigma^n(w)))_n$ converges to an ultimately periodic sequence if and only if there exist $u, v \in B^*$ such that for all $0 \leq i \leq n_0 - 1$, $\lim_{k\to+\infty} \phi(\sigma^{j_0+i+kn_0}(a_i)) = uv^\omega$. Thus to decide whether $(\phi(\sigma^n(w)))_n$ converges to an ultimately periodic sequence, we first have to check (using Thm. 3.18) that for all $0 \leq i \leq n_0 - 1$, $\lim_{k\to+\infty} \phi(\sigma^{j_0+i+kn_0}(a_i)) = u_i v_i^\omega$, for some computable $u_i, v_i \in B^*$. Then, we check whether the sequences $u_i v_i^\omega$ are equal (which can be algorithmically realised).

Let then σ be an arbitrary morphism. From Proposition 3.3 we can suppose that σ^p satisfies (P1) and (P2) for some computable $p > 0$. Then, we proceed as before for the couples $(\sigma^p, \phi\circ\sigma^i)$, $0 \leq i \leq p-1$: We test their ultimate periodicity and then we compare the results to finally decide.

Acknowledgements. The author thanks the referees for their valuable and relevant comments. He also thanks the ANR program SubTile for its financial support.

References

[1] J.-P. Allouche, N. Rampersad and J. Shallit, Periodicity, repetitions, and orbits of an automatic sequence. *Theoret. Comput. Sci.* **410** (2009) 2795–2803.

[2] J.-P. Allouche and J.O. Shallit, *Automatic Sequences, Theory, Applications, Generalizations*. Cambridge University Press (2003).

[3] J.P. Bell, E. Charlier, A.S. Fraenkel and M. Rigo, A decision problem for ultimately periodic sets in non-standard numeration systems. *Internat. J. Algebra Comput.* **9** (2009) 809–839.

[4] J. Cassaigne and F. Nicolas, Quelques propriétés des mots substitutifs. *Bull. Belg. Math. Soc. Simon Stevin* **10** (2003) 661–676.

[5] A. Cerný and J. Gruska, Modular trellises. *The Book of L*, edited by G. Rozenberg and A. Salomaa. Springer-Verlag (1986) 45–61.

[6] A. Cobham, On the Hartmanis-Stearns problem for a class of tag machines. In *IEEE Conference Record of 1968 Ninth Annual Symposium on Switching and Automata Theory.* Also appeared as IBM Research Technical Report RC-2178, August 23 (1968) 51–60.

[7] F. Durand, A characterization of substitutive sequences using return words. *Discrete Math.* **179** (1998) 89–101.

[8] F. Durand, HD0L ω-equivalence and periodicity problems in the primitive case (To the memory of G. Rauzy). *J. Unif. Distrib. Theory* **7** (2012) 199–215.
[9] F. Durand and M. Rigo, Multidimensional extension of the Morse-Hedlund theorem. *Eur. J. Comb.* **34** (2013) 391–409.
[10] T. Harju and M. Linna, On the periodicity of morphisms on free monoids. *RAIRO ITA* **20** (1986) 47–54.
[11] J. Honkala, A decision method for the recognizability of sets defined by number systems. *RAIRO ITA* **20** (1986) 395–403.
[12] J. Honkala, Cancellation and periodicity properties of iterated morphisms. *Theoret. Comput. Sci.* **391** (2008) 61–64.
[13] J. Honkala, On the simplification of infinite morphic words. *Theoret. Comput. Sci.* **410** (2009) 997–1000.
[14] J. Honkala and M. Rigo, Decidability questions related to abstract numeration systems. *Discrete Math.* **285** (2004) 329–333.
[15] R.A. Horn and C.R. Johnson, *Matrix analysis*. Cambridge University Press (1990).
[16] P. Lecomte and M. Rigo, Abstract numeration systems. In *Combinatorics, automata and number theory*, Cambridge Univ. Press. *Encyclopedia Math. Appl.* **135** (2010) 108–162.
[17] J. Leroux, A polynomial time presburger criterion and synthesis for number decision diagrams. In *20th IEEE Symposium on Logic In Computer Science (LICS 2005)*, *IEEE Comput. Soc.* (2005) 147–156.
[18] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press (1995).
[19] A. Maes and M. Rigo, More on generalized automatic sequences. *J. Autom. Lang. Comb.* **7** (2002) 351–376.
[20] I. Mitrofanov, A proof for the decidability of HD0L ultimate periodicity. arXiv:1110.4780 (2011).
[21] A. Muchnik, The definable criterion for definability in Presburger arithmetic and its applications. *Theoret. Comput. Sci.* **290** (2003) 1433–1444.
[22] J.-J. Pansiot, Hiérarchie et fermeture de certaines classes de tag-systèmes. *Acta Informatica* **20** (1983) 179–196.
[23] J.-J. Pansiot, Complexité des facteurs des mots infinis engendrés par morphismes itérés. In *ICALP84*, *Lect. Notes Comput. Sci.* Vol. 172, edited by J. Paredaens. Springer-Verlag (1984) 380–389.
[24] J.-J. Pansiot, Decidability of periodicity for infinite words. *RAIRO ITA* **20** (1986) 43–46.
[25] N. Priebe, Towards a characterization of self-similar tilings in terms of derived Voronoĭ tessellations. *Geom. Dedicata* **79** (2000) 239–265.
[26] N. Priebe and B. Solomyak, Characterization of planar pseudo-self-similar tilings. *Discrete Comput. Geom.* **26** (2001) 289–306.
[27] M. Queffélec, *Substitution dynamical systems–spectral analysis*. In *Lect. Notes Math.*, Vol. 1294. Springer-Verlag (1987).
[28] O. Salon, Suites automatiques à multi-indices. In *Séminaire de Théorie des Nombres de Bordeaux* (1986-1987) 4.01–4.27.
[29] O. Salon, Suites automatiques à multi-indices et algébricité. *C. R. Acad. Sci. Paris* **305** (1987) 501–504.

Communicated by J. Kari.
Received November 14, 2011. Accepted February 1, 2013.