Characteristic morphisms of generalized episturmian words

Michelangelo Bucci, Aldo de Luca, and Alessandro De Luca

Dipartimento di Matematica e Applicazioni "R. Caccioppoli"

Università degli Studi di Napoli Federico II

Via Cintia, Monte S. Angelo, I-80126 Napoli, Italy

{micbucci,aldo.deluca,alessandro.deluca}@unina.it

Abstract

In a recent paper with L. Q. Zamboni, the authors introduced the class of ϑ*-episturmian* words. An infinite word over A is standard ϑ-episturmian, where ϑ is an involutory antimorphism of A^*, if its set of factors is closed under ϑ and its left special factors are prefixes. When ϑ is the reversal operator, one obtains the usual standard episturmian words. In this paper, we introduce and study ϑ*-characteristic morphisms*, that is, morphisms which map standard episturmian words into standard ϑ-episturmian words. They are a natural extension of standard episturmian morphisms. The main result of the paper is a characterization of these morphisms when they are injective. In order to prove this result, we also introduce and study a class of biprefix codes which are *overlap-free*, i.e., any two code words do not overlap properly, and *normal*, i.e., no proper suffix (prefix) of any code-word is left (right) special in the code. A further result is that any standard ϑ-episturmian word is a morphic image, by an injective ϑ-characteristic morphism, of a standard episturmian word.

Introduction

The study of combinatorial and structural properties of finite and infinite words is a subject of great interest, with many applications in mathematics, physics, computer science, and biology (see for instance [2, 14]). In this framework, *Sturmian words* play a central role, since they are the aperiodic infinite words of minimal "complexity" (see [2]). By definition, Sturmian words are on a binary alphabet; some natural extensions to the case of an alphabet with more than two letters have been given in [9, 12], introducing the class of the so-called *episturmian words*.

Several extensions of standard episturmian words are possible. For example, in [10] a generalization was obtained by making suitable hypotheses on the lengths of palindromic prefixes of an infinite word; in [8, 5, 4, 6] different extensions were introduced, all based on the replacement of the *reversal operator* R by an arbitrary *involutory antimorphism* ϑ of the free monoid A^*. In particular, the so called ϑ*-standard* and *standard* ϑ*-episturmian* words were studied. An infinite word over A is standard ϑ-episturmian if its set of factors is closed under ϑ and its left special factors are prefixes.

In this paper we introduce and study *ϑ-characteristic morphisms*, a natural extension of standard episturmian morphisms, which map all standard episturmian words on an alphabet X to standard ϑ-episturmian words over some alphabet A. When $X = A$ and $\vartheta = R$, one obtains the usual standard episturmian morphisms (cf. [9, 12, 11]). Beside being interesting by themselves, such morphisms are also a powerful tool for constructing nontrivial examples of standard ϑ-episturmian words and for studying their properties.

In Section 2 we introduce ϑ-characteristic morphisms and prove some of their structural properties (mainly concerning the images of letters). In Section 3 our main results are given. A first theorem is a characterization of injective ϑ-characteristic morphisms such that the images of the letters are unbordered ϑ-palindromes. The section concludes with a full characterization (cf. Theorem 3.13) of all injective ϑ-characteristic morphisms, to whose proof Section 5 is dedicated. This result, which solves a problem posed in [4], is very useful to construct nontrivial examples of ϑ-characteristic morphisms and then of standard ϑ-episturmian words. Moreover, one has a quite simple procedure to decide whether a given injective morphism is ϑ-characteristic.

In Section 4 we study some properties of two classes of codes: the *overlap-free codes*, i.e., codes whose any two elements do not overlap properly, and the *normal codes*, i.e., codes in which no proper nonempty prefix (suffix) which is not a code-word, appears followed (preceded) by two different letters. The family of biprefix, overlap-free, and normal codes appears to be deeply connected with ϑ-characteristic morphisms, and especially useful for the proof of our main result.

In Section 6, we prove that every standard ϑ-episturmian word is a morphic image of a standard episturmian word under a suitable injective ϑ-characteristic morphism. This solves another question asked in [4].

A short version of this work was presented at the Developments in Language Theory conference, held in Kyoto in September 2008 [3].

1 Preliminaries

Let A be a nonempty finite set, or *alphabet*. In the following, A^* (resp. A^+) will denote the *free monoid* (resp. *semigroup*) generated by A. The elements of A are called *letters* and those of A^* *words*. The identity element of A^* is called *empty word* and it is denoted by ε. A word $w \in A^+$ can be written uniquely as a product of letters $w = a_1 a_2 \cdots a_n$, with $a_i \in A$, $i = 1, \ldots, n$. The integer n is called the *length* of w and is denoted by $|w|$. The length of ε is conventionally 0. For any $a \in A$, $|w|_a$ denotes the number of occurrences of a in the word w. For any nonempty word w, we will denote by w^f and w^ℓ respectively the first and the last letter of w.

A word u is a *factor* of $w \in A^*$ if $w = rus$ for some words r and s. In the special case $r = \varepsilon$ (resp. $s = \varepsilon$), u is called a *prefix* (resp. *suffix*) of w. A factor u of w is *proper* if $u \neq w$. We denote respectively by $\operatorname{Fact} w$, $\operatorname{Pref} w$, and $\operatorname{Suff} w$ the sets of all factors, prefixes, and suffixes of the word w. For $Y \subseteq A^*$, $\operatorname{Pref} Y$, $\operatorname{Suff} Y$, and $\operatorname{Fact} Y$ will denote respectively the sets of prefixes, suffixes, and factors of all the words of Y.

A factor of w is called a *border* of w if it is both a prefix and a suffix of w. A word is called *unbordered* if its only proper border is ε. A positive integer p is a *period* of $w = a_1 \cdots a_n$ if whenever $1 \leq i, j \leq |w|$ one has that

$$i \equiv j \pmod{p} \Longrightarrow a_i = a_j .$$

As is well known [13], a word w has a period $p \leq |w|$ if and only if it has a border of length $|w| - p$. Thus a nonempty word w is unbordered if and only if its minimal period is $|w|$. We recall the famous *theorem of Fine and Wilf*, stating that if a word w has two periods p and q, and $|w| \geq p + q - \gcd(p, q)$, then w has also the period $\gcd(p, q)$ (cf. [13]).

A word $w \in A^+$ is *primitive* if it cannot be written as a power u^k with $k > 1$. As is well known (cf. [13]), any nonempty word w is a power of a unique primitive word, also called the *primitive root* of w.

A right-infinite word over the alphabet A, called infinite word for short, is a mapping $x : \mathbb{N}_+ \longrightarrow A$, where $\mathbb{N}_+$ is the set of positive integers. One can represent x as

$$x = x_1 x_2 \cdots x_n \cdots ,$$

where for any $i > 0$, $x_i = x(i) \in A$. A (finite) *factor* of x is either the empty word or any sequence $u = x_i \cdots x_j$ with $i \leq j$, i.e., any block of consecutive letters of x. If $i = 1$, then u is a *prefix* of x. We shall denote by $x_{[n]}$ the prefix of x of length n, and by $\operatorname{Fact} x$ and $\operatorname{Pref} x$ the sets of finite factors and prefixes of x respectively. The set of all infinite words over A is denoted by A^ω. We also set $A^\infty = A^* \cup A^\omega$. For any $Y \subseteq A^*$, Y^ω denotes the set of infinite words which can be factorized by the elements of Y. If $w \in A^\infty$, $\operatorname{alph} w$ will denote the set of letters occurring in w.

Let $w \in A^\infty$. An *occurrence* of a factor u in w is any pair $(\lambda, \rho) \in A^* \times A^\infty$ such that $w = \lambda u \rho$. If $v \in A^*$ is a prefix of w, then $v^{-1}w$ denotes the unique word $u \in A^\infty$ such that $vu = w$.

A factor u of w is called *right special* if there exist $a, b \in A$, $a \neq b$, such that ua and ub are both factors of w. Symmetrically, u is said *left special* if $au, bu \in \operatorname{Fact} w$. A word u is called a right (resp. left) special factor of *a set* $Y \subseteq A^*$ if there exist letters $a, b \in A$ such that $a \neq b$ and $ua, ub \in \operatorname{Fact} Y$ (resp. $au, bu \in \operatorname{Fact} Y$). We denote by $RS\,Y$ (resp. $LS\,Y$) the set of right (resp. left) special factors of Y.

The *reversal* of a word $w = a_1 a_2 \cdots a_n$, with $a_i \in A$ for $1 \leq i \leq n$, is the word $\tilde{w} = a_n \cdots a_1$. One sets $\tilde{\varepsilon} = \varepsilon$. A *palindrome* is a word which equals its reversal. We shall denote by $PAL(A)$, or PAL when no confusion arises, the set of all palindromes over A.

A *morphism* (resp. *antimorphism*) from A^* to the free monoid B^* is any map $\varphi : A^* \to B^*$ such that $\varphi(uv) = \varphi(u)\varphi(v)$ (resp. $\varphi(uv) = \varphi(v)\varphi(u)$) for all $u, v \in A^*$. The morphism (resp. antimorphism) φ is *nonerasing* if for any $a \in A$, $\varphi(a) \neq \varepsilon$. A morphism φ can be naturally extended to A^ω by setting for any $x = x_1 x_2 \cdots x_n \cdots \in A^\omega$,

$$\varphi(x) = \varphi(x_1)\varphi(x_2) \cdots \varphi(x_n) \cdots .$$

A *code* over A is a subset Z of A^+ such that every word of Z^+ admits a unique factorization by the elements of Z (cf. [1]). A subset of A^+ with the property that none of its elements is a proper prefix (resp. suffix) of any other is trivially a code, usually called a *prefix* (resp. *suffix*) code. We recall that if Z is a prefix code, then Z^* is *left unitary*, i.e., for all $p \in Z^*$ and $w \in A^*$, $pw \in Z^*$ implies $w \in Z^*$. A *biprefix* code is a code which is both prefix and suffix. We say that a code Z over A is *overlap-free* if no two of its elements overlap properly, i.e., if for all $u, v \in Z$, $\mathrm{Suff}\, u \cap \mathrm{Pref}\, v \subseteq \{\varepsilon, u, v\}$.

For instance, let $Z_1 = \{a, bac, abc\}$ and $Z_2 = \{a, bac, cba\}$. One has that Z_1 is an overlap-free and suffix code, whereas Z_2 is a prefix code which is not overlap-free as bac and cba overlap properly.

A code $Z \subseteq A^+$ will be called *right normal* if it satisfies the following condition:

$$(\mathrm{Pref}\, Z \setminus Z) \cap RS\, Z \subseteq \{\varepsilon\}\ , \tag{1}$$

i.e., any proper and nonempty prefix u of any word of Z such that $u \notin Z$ is not right special in Z. In a symmetric way, a code Z is called *left normal* if it satisfies the condition

$$(\mathrm{Suff}\, Z \setminus Z) \cap LS\, Z \subseteq \{\varepsilon\}\ . \tag{2}$$

A code Z is called *normal* if it is right and left normal.

As an example, the code $Z_1 = \{a, ab, bb\}$ is right normal but not left normal; the code $Z_2 = \{a, aba, aab\}$ is normal. The code $Z_3 = \{a, cad, bacadad\}$ is biprefix, overlap-free, and right normal, and the code $Z_4 = \{a, badc\}$ is biprefix, overlap-free, and normal.

The following proposition and lemma will be useful in the sequel.

Proposition 1.1. *Let Z be a biprefix, overlap-free, and right normal (resp. left normal) code. Then:*

1. *if $z \in Z$ is such that $z = \lambda v \rho$, with $\lambda, \rho \in A^*$ and v a nonempty prefix (resp. suffix) of $z' \in Z$, then $\lambda z'$ (resp. $z'\rho$) is a prefix (resp. suffix) of z, proper if $z \neq z'$.*
2. *for $z_1, z_2 \in Z$, if $z_1^f = z_2^f$ (resp. $z_1^\ell = z_2^\ell$), then $z_1 = z_2$.*

Proof. Let $z = \lambda v \rho$ with $v \in \mathrm{Pref}\, z'$ and $v \neq \varepsilon$. If $v = z'$, there is nothing to prove. Suppose then that v is a proper prefix of z'. Since Z is a prefix code, any proper nonempty prefix of z', such as v, is not an element of Z; moreover, it is not right special in Z, since Z is right normal. Therefore, to prove the first statement it is sufficient to show that $|v\rho| \geq |z'|$, where the inequality is strict if $z \neq z'$. Indeed, if $|v\rho| < |z'|$, then a proper prefix of z' would be a suffix of z, which is impossible as Z is an overlap-free code. If $|v\rho| = |z'|$, then $z' \in \mathrm{Suff}\, z$, so that $z' = z$ as Z is a suffix code.

Let us now prove the second statement. Let $z_1, z_2 \in Z$ with $z_1^f = z_2^f$. By contradiction, suppose $z_1 \neq z_2$. By the preceding statement, we derive that z_1 is a proper prefix of z_2 and z_2 is a proper prefix of z_1, which is clearly absurd. The symmetrical claims can be analogously proved. □

From the preceding proposition, a biprefix, overlap-free, and normal code satisfies both properties 1 and 2 and their symmetrical statements. Some further properties of such codes will be given in Section 4.

Lemma 1.2. *Let $g : B^* \to A^*$ be an injective morphism such that $g(B) = Z$ is a prefix code. Then for all $p \in B^*$ and $q \in B^\infty$ one has that p is a prefix of q if and only if $g(p)$ is a prefix of $g(q)$.*

Proof. The 'only if' part is trivial. Therefore, let us prove the 'if' part. Let us first suppose $q \in B^*$, so that $g(q) = g(p)\zeta$ for some $\zeta \in A^*$. Since $g(p), g(q) \in Z^*$ and Z^* is left unitary, it follows that $\zeta \in Z^*$. Therefore, there exists, and is unique, $r \in B^*$ such that $g(r) = \zeta$. Hence $g(q) = g(p)g(r) = g(pr)$. Since g is injective one has $q = pr$ which proves the assertion in this case. If $q \in B^\omega$, there exists a prefix $q_{[n]}$ of q such that $g(p) \in \operatorname{Pref} g(q_{[n]})$. By the previous argument, it follows that p is a prefix of $q_{[n]}$ and then of q. □

1.1 Standard episturmian words and morphisms

We recall (cf. [9, 12]) that an infinite word $t \in A^\omega$ is *standard episturmian* if it is *closed under reversal* (that is, if $w \in \operatorname{Fact} t$ then $\tilde{w} \in \operatorname{Fact} t$) and each of its left special factors is a prefix of t. We denote by $SEpi(A)$, or by $SEpi$ when there is no ambiguity, the set of all standard episturmian words over the alphabet A.

Given a word $w \in A^*$, we denote by $w^{(+)}$ its *right palindrome closure*, i.e., the shortest palindrome having w as a prefix (cf. [7]). If Q is the longest palindromic suffix of w and $w = sQ$, then $w^{(+)} = sQ\tilde{s}$. For instance, if $w = abacbca$, then $w^{(+)} = abacbcaba$.

We define the *iterated palindrome closure* operator[1] $\psi : A^* \to A^*$ by setting $\psi(\varepsilon) = \varepsilon$ and $\psi(va) = (\psi(v)a)^{(+)}$ for any $a \in A$ and $v \in A^*$. From the definition, one easily obtains that the map ψ is injective. Moreover, for any $u, v \in A^*$, one has $\psi(uv) \in \psi(u)A^* \cap A^*\psi(u)$. The operator ψ can then be naturally extended to A^ω by setting, for any infinite word x,

$$\psi(x) = \lim_{n\to\infty} \psi(x_{[n]}) \ .$$

The following fundamental result was proved in [9]:

Theorem 1.3. *An infinite word t is standard episturmian over A if and only if there exists $\Delta \in A^\omega$ such that $t = \psi(\Delta)$.*

For any $t \in SEpi$, there exists a *unique* Δ such that $t = \psi(\Delta)$. This Δ is called the *directive word* of t. If every letter of A occurs infinitely often in Δ, the word t is called a (standard) *Arnoux-Rauzy word*. In the case of a binary alphabet, an Arnoux-Rauzy word is usually called a *standard Sturmian word* (cf. [2]).

Example 1.4. Let $A = \{a, b\}$ and $\Delta = (ab)^\omega$. The word $\psi(\Delta)$ is the famous *Fibonacci word*

$$\mathrm{f} = abaababaabaababaababa \cdots \ .$$

If $A = \{a, b, c\}$ and $\Delta = (abc)^\omega$, then $\psi(\Delta)$ is the so-called *Tribonacci word*

$$\tau = abacabaabacababacabaabacabaca \cdots \ .$$

A letter $a \in A$ is said to be *separating* for $w \in A^\infty$ if it occurs in each factor of w of length 2. We recall the following well known result from [9]:

Proposition 1.5. *Let t be a standard episturmian word and a be its first letter. Then a is separating for t.*

[1] This operator is denoted by *Pal* in [11] and other papers.

For instance, the letter a is separating for f and τ.

We report here some properties of the operator ψ which will be useful in the sequel. The first one is known (see for instance [7, 9]); we give a proof for the sake of completeness.

Proposition 1.6. *For all $u, v \in A^*$, u is a prefix of v if and only if $\psi(u)$ is a prefix of $\psi(v)$.*

Proof. If u is a prefix of v, from the definition of the operator ψ, one has that $\psi(v) \in \psi(u)A^* \cap A^*\psi(u)$, so that $\psi(u)$ is a prefix (and a suffix) of $\psi(v)$. Let us now suppose that $\psi(u)$ is a prefix of $\psi(v)$. If $\psi(u) = \psi(v)$, then, since ψ is injective, one has $u = v$. Hence, suppose that $\psi(u)$ is a proper prefix of $\psi(v)$. If $u = \varepsilon$, the result is trivial. Hence we can suppose that $u, v \in A^+$. Let $v = a_1 \cdots a_n$ and i be the integer such that $1 \leq i \leq n-1$ and

$$|\psi(a_1 \cdots a_i)| \leq |\psi(u)| < |\psi(a_1 \cdots a_{i+1})|.$$

If $|\psi(a_1 \cdots a_i)| < |\psi(u)|$, then $\psi(a_1 \cdots a_i)a_{i+1}$ is a prefix of the palindrome $\psi(u)$, so that one would have:

$$|\psi(a_1 \cdots a_{i+1})| = |(\psi(a_1 \cdots a_i)a_{i+1})^{(+)}| \leq |\psi(u)| < |\psi(a_1 \cdots a_{i+1})|$$

which is a contradiction. Therefore $|\psi(a_1 \cdots a_i)| = |\psi(u)|$, that implies $\psi(a_1 \cdots a_i) = \psi(u)$ and $u = a_1 \cdots a_i$. □

Proposition 1.7. *Let $x \in A \cup \{\varepsilon\}$, $w' \in A^*$, and $w \in w'A^*$. Then $\psi(w'x)$ is a factor of $\psi(wx)$.*

Proof. By the previous proposition, $\psi(w')$ is a prefix of $\psi(w)$. This solves the case $x = \varepsilon$. For $x \in A$, we prove the result by induction on $n = |w| - |w'|$.

The assertion is trivial for $n = 0$. Let then $n \geq 1$ and write $w = ua$ with $a \in A$ and $u \in A^*$. As $w' \in \operatorname{Pref} u$ and $|u| - |w'| = n - 1$, we can assume by induction that $\psi(w'x)$ is a factor of $\psi(ux)$. Hence it suffices to show that $\psi(ux) \in \operatorname{Fact} \psi(wx)$. We can write

$$\psi(w) = (\psi(u)a)^{(+)} = \psi(u)av = \tilde{v}a\psi(u)$$

for some $v \in A^*$, so that $\psi(wx) = (\tilde{v}a\psi(u)x)^{(+)}$. Since $\psi(u)$ is the longest proper palindromic prefix and suffix of $\psi(w)$, if $x \neq a$ it follows that the longest palindromic suffixes of $\psi(u)x$ and $\psi(w)x$ must coincide, so that $\psi(ux) = (\psi(u)x)^{(+)}$ is a factor of $\psi(wx)$, as desired.

If $x = a$, then $\psi(ux) = \psi(w)$ is trivially a factor of $\psi(wx)$. This concludes the proof. □

The following proposition was proved in [9, Theorem 6].

Proposition 1.8. *Let $x \in A$, $u \in A^*$, and $\Delta \in A^\omega$. Then $\psi(u)x$ is a factor of $\psi(u\Delta)$ if and only if x occurs in Δ.*

For each $a \in A$, let $\mu_a : A^* \to A^*$ be the morphism defined by $\mu_a(a) = a$ and $\mu_a(b) = ab$ for all $b \in A \setminus \{a\}$. If $a_1, \ldots, a_n \in A$, we set $\mu_w = \mu_{a_1} \circ \cdots \circ \mu_{a_n}$ (in particular, $\mu_\varepsilon = \mathrm{id}_A$). The next proposition, proved in [11], shows a connection between these morphisms and iterated palindrome closure.

Proposition 1.9. *For any $w, v \in A^*$, $\psi(wv) = \mu_w(\psi(v))\psi(w)$.*

By the preceding proposition, if $v \in A^\omega$ then one has

$$\begin{aligned}\psi(wv) &= \lim_{n\to\infty} \psi(wv_{[n]}) = \lim_{n\to\infty} \mu_w(\psi(v_{[n]}))\psi(w) \\ &= \lim_{n\to\infty} \mu_w(\psi(v_{[n]})) = \mu_w(\psi(v)) \ .\end{aligned}$$

Thus, for any $w \in A^*$ and $v \in A^\omega$ we have

$$\psi(wv) = \mu_w(\psi(v)) \ . \tag{3}$$

Corollary 1.10. *For any $t \in A^\omega$ and $w \in A^*$, $\psi(w)$ is a prefix of $\mu_w(t)$.*

Proof. Let $t = t_1 t_2 \cdots t_n \cdots$, with $t_i \in A$ for $i \geq 1$. We prove that $\psi(w)$ is a prefix of $\mu_w(t_{[n]})$ for all n such that $|\mu_w(t_{[n]})| \geq |\psi(w)|$. Indeed, by Proposition 1.9 we have, for all $i \geq 1$, $\mu_w(t_i)\psi(w) = \psi(wt_i) = \psi(w)\xi_i$ for some $\xi_i \in A^*$. Hence

$$\mu_w\left(t_{[n]}\right)\psi(w) = \mu_w(t_1)\cdots\mu_w(t_n)\psi(w) = \psi(w)\xi_1\cdots\xi_n \ ,$$

and this shows that $\psi(w)$ is a prefix of $\mu_w(t_{[n]})$. □

From the definition of the morphism μ_a, $a \in A$, it is easy to prove the following:

Proposition 1.11. *Let $w \in A^\infty$ and a be its first letter. Then a is separating for w if and only if there exists $\alpha \in A^\infty$ such that $w = \mu_a(\alpha)$.*

For instance, the letter a is separating for the word $w = abacaaacaba$, and one has $w = \mu_a(bcaacba)$.

We recall (cf. [9, 12, 11]) that a *standard episturmian morphism* of A^* is any composition $\mu_w \circ \sigma$, with $w \in A^*$ and $\sigma : A^* \to A^*$ a morphism extending to A^* a permutation on the alphabet A. All these morphisms are injective. The set $\mathcal{E}$ of standard episturmian morphisms is a monoid under map composition. The importance of standard episturmian morphisms, and the reason for their name, lie in the following (see [9, 12]):

Theorem 1.12. *An injective morphism $\varphi : A^* \to A^*$ is standard episturmian if and only if $\varphi(SEpi) \subseteq SEpi$, that is, if and only if it maps every standard episturmian word over A into a standard episturmian word over A.*

A *pure* standard episturmian morphism is just a μ_w for some $w \in A^*$. Trivially, the set of pure standard episturmian morphisms is the submonoid of $\mathcal{E}$ generated by the set $\{\mu_a \mid a \in A\}$. The following was proved in [9]:

Proposition 1.13. *Let $t \in A^\omega$ and $a \in A$. Then $\mu_a(t)$ is a standard episturmian word if and only if so is t.*

1.2 Involutory antimorphisms and pseudopalindromes

An *involutory antimorphism* of A^* is any antimorphism $\vartheta : A^* \to A^*$ such that $\vartheta \circ \vartheta = \mathrm{id}$. The simplest example is the *reversal operator*:

$$\begin{aligned} R : A^* &\longrightarrow A^* \\ w &\longmapsto \tilde{w} \ . \end{aligned}$$

Any involutory antimorphism ϑ satisfies $\vartheta = \tau \circ R = R \circ \tau$ for some morphism $\tau : A^* \to A^*$ extending an involution of A. Conversely, if τ is such a morphism, then $\vartheta = \tau \circ R = R \circ \tau$ is an involutory antimorphism of A^*.

Let ϑ be an involutory antimorphism of A^*. We call *ϑ-palindrome* any fixed point of ϑ, i.e., any word w such that $w = \vartheta(w)$, and denote by PAL_ϑ the set of all ϑ-palindromes. We observe that $\varepsilon \in PAL_\vartheta$ by definition, and that R-palindromes are exactly the usual palindromes. If one makes no reference to the antimorphism ϑ, a ϑ-palindrome is called a *pseudopalindrome*.

Some general properties of pseudopalindromes, mainly related to conjugacy and periodicity, have been studied in [8]. We mention here the following lemma, which will be useful in the sequel:

Lemma 1.14. *Let w be in PAL_ϑ. If p is a period of w, then each factor of w of length p is in PAL_ϑ^2.*

For instance, let $A = \{a, b\}$ and let $\vartheta(a) = b$, $\vartheta(b) = a$. The word $w = babaababbaba$ is a ϑ-palindrome, having the periods 8 and 10. Any factor of w of length 8 or 10 belongs to PAL_ϑ^2; as an example, $abaababb = (ab)(aababb) \in PAL_\vartheta^2$.

For any involutory antimorphism ϑ, one can define the (right) ϑ-palindrome closure operator: for any $w \in A^*$, $w^{\oplus_\vartheta}$ denotes the shortest ϑ-palindrome having w as a prefix.

In the following, we shall fix an involutory antimorphism ϑ of A^*, and use the notation $\bar{w}$ for $\vartheta(w)$. We shall also drop the subscript ϑ from the ϑ-palindrome closure operator $^{\oplus_\vartheta}$ when no confusion arises. As one easily verifies (cf. [8]), if Q is the longest ϑ-palindromic suffix of w and $w = sQ$, then

$$w^\oplus = sQ\bar{s}\ .$$

Example 1.15. Let $A = \{a, b, c\}$ and ϑ be defined as $\bar{a} = b$, $\bar{c} = c$. If $w = abacabc$, then $Q = cabc$ and $w^\oplus = abacabcbab$.

We can naturally define the *iterated ϑ-palindrome closure* operator $\psi_\vartheta : A^* \to PAL_\vartheta$ by $\psi_\vartheta(\varepsilon) = \varepsilon$ and

$$\psi_\vartheta(ua) = (\psi_\vartheta(u)a)^\oplus$$

for $u \in A^*$, $a \in A$. For any $u, v \in A^*$ one has $\psi_\vartheta(uv) \in \psi_\vartheta(u)A^* \cap A^*\psi_\vartheta(u)$, so that ψ_ϑ can be extended to infinite words too. More precisely, if $\Delta = x_1x_2 \cdots x_n \cdots \in A^\omega$ with $x_i \in A$ for $i \geq 1$, then

$$\psi_\vartheta(\Delta) = \lim_{n\to\infty} \psi_\vartheta(\Delta_{[n]})\ .$$

The word Δ is called the *directive word* of $\psi_\vartheta(\Delta)$, and $s = \psi_\vartheta(\Delta)$ the *ϑ-standard word* directed by Δ. The class of ϑ-standard words was introduced in [8]; some interesting results about such words are in [5].

We denote by $\mathcal{P}_\vartheta$ the set of unbordered ϑ-palindromes. We remark that $\mathcal{P}_\vartheta$ is a *biprefix code*. This means that every word of $\mathcal{P}_\vartheta$ is neither a prefix nor a suffix of any other element of $\mathcal{P}_\vartheta$. We observe that $\mathcal{P}_R = A$. The following result was proved in [4]:

Proposition 1.16. $PAL_\vartheta^* = \mathcal{P}_\vartheta^*$.

This can be equivalently stated as follows: every ϑ-palindrome can be uniquely factorized by the elements of $\mathcal{P}_\vartheta$. For instance, the ϑ-palindrome $abacabcbab$ of Example 1.15 is factorizable as $ab \cdot acabcb \cdot ab$, with $acabcb, ab \in \mathcal{P}_\vartheta$.

Since $\mathcal{P}_\vartheta$ is a code, the map

$$\begin{aligned} f : \mathcal{P}_\vartheta &\longrightarrow A \\ \pi &\longmapsto \pi^f \end{aligned} \tag{4}$$

can be extended (uniquely) to a morphism $f : \mathcal{P}_\vartheta^* \to A^*$. Moreover, since $\mathcal{P}_\vartheta$ is a prefix code, any word in $\mathcal{P}_\vartheta^\omega$ can be uniquely factorized by the elements of $\mathcal{P}_\vartheta$, so that f can be naturally extended to $\mathcal{P}_\vartheta^\omega$.

Proposition 1.17. *Let $\varphi : X^* \to A^*$ be an injective morphism such that $\varphi(X) \subseteq \mathcal{P}_\vartheta$. Then, for any $w \in X^*$:*

1. *$\varphi(\tilde{w}) = \overline{\varphi(w)}$,*
2. *$w \in PAL \Longleftrightarrow \varphi(w) \in PAL_\vartheta$,*
3. *$\varphi(w^{(+)}) = \varphi(w)^\oplus$.*

Proof. The first statement is trivially true for $w = \varepsilon$. If $w = x_1 \cdots x_n$ with $x_i \in X$ for $i = 1, \ldots, n$, then since $\varphi(X) \subseteq \mathcal{P}_\vartheta \subseteq PAL_\vartheta$,

$$\varphi(\tilde{w}) = \varphi(x_n) \cdots \varphi(x_1) = \overline{\varphi(x_n)} \cdots \overline{\varphi(x_1)} = \overline{\varphi(w)}\ .$$

As φ is injective, statement *2* easily follows from *1*.

Finally, let $\varphi(w) = vQ$ where $v \in A^*$ and Q is the longest ϑ-palindromic suffix of $\varphi(w)$. Since $\varphi(w), Q \in \mathcal{P}_\vartheta^*$ and $\mathcal{P}_\vartheta$ is a biprefix code, we have $v \in \mathcal{P}_\vartheta^*$. This implies, as φ is injective, that there exist $w_1, w_2 \in X^*$ such that $w = w_1 w_2$, $\varphi(w_1) = v$, and $\varphi(w_2) = Q$. By *2*, w_2 is the longest palindromic suffix of w. Hence, by *1*:

$$\varphi(w^{(+)}) = \varphi(w_1 w_2 \tilde{w}_1) = vQ\bar{v} = \varphi(w)^\oplus\ ,$$

as desired. □

Example 1.18. Let $X = \{a, b, c\}$, $A = \{a, b, c, d, e\}$, and ϑ be defined in A as $\bar{a} = b$, $\bar{c} = c$, and $\bar{d} = e$. Let $\varphi : X^* \to A^*$ be the injective morphism defined by $\varphi(a) = ab$, $\varphi(b) = ba$, $\varphi(c) = dce$. One has $\varphi(X) \subseteq \mathcal{P}_\vartheta$ and

$$\varphi\left((abc)^{(+)}\right) = \varphi(abcba) = abbadcebaab = (\varphi(abc))^\oplus\ .$$

1.3 Standard ϑ-episturmian words

In [4] *standard ϑ-episturmian* words were naturally defined by substituting, in the definition of standard episturmian words, the closure under reversal with the *closure under ϑ*. Thus an infinite word s is standard ϑ-episturmian if it satisfies the following two conditions:

1. for any $w \in \operatorname{Fact} s$, one has $\bar{w} \in \operatorname{Fact} s$,
2. for any left special factor w of s, one has $w \in \operatorname{Pref} s$.

We denote by $SEpi_\vartheta$ the set of all standard ϑ-episturmian words on the alphabet A. The following two propositions, proved in [4], give methods for constructing standard ϑ-episturmian words.

Proposition 1.19. *Let s be a ϑ-standard word over A, and $B = \mathrm{alph}(\Delta(s))$. Then s is standard ϑ-episturmian if and only if*

$$x \in B,\, x \neq \bar{x} \implies \bar{x} \notin B\ .$$

Example 1.20. Let $A = \{a, b, c, d, e\}$, $\Delta = (acd)^\omega$, and ϑ be defined by $\bar{a} = b$, $\bar{c} = c$, and $\bar{d} = e$. The ϑ-standard word $\psi_\vartheta(\Delta) = abcabdeabcaba\cdots$ is standard ϑ-episturmian.

Proposition 1.21. *Let $\varphi : X^* \to A^*$ be a nonerasing morphism such that*

1. *$\varphi(x) \in PAL_\vartheta$ for all $x \in X$,*
2. *$\mathrm{alph}\,\varphi(x) \cap \mathrm{alph}\,\varphi(y) = \emptyset$ if $x, y \in X$ and $x \neq y$,*
3. *$|\varphi(x)|_a \leq 1$ for all $x \in X$ and $a \in A$.*

Then for any standard episturmian word $t \in X^\omega$, $s = \varphi(t)$ is a standard ϑ-episturmian word.

Example 1.22. Let $A = \{a, b, c, d, e\}$, $\bar{a} = b$, $\bar{c} = c$, $\bar{d} = e$, $X = \{x, y\}$, and $s = g(t)$, where $t = xxyxxxyxxxyxxy\cdots \in SEpi(X)$, $\Delta(t) = (xxy)^\omega$, $g(x) = acb$, and $g(y) = de$, so that

$$s = acbacbdeacbacbacbde\cdots\ . \tag{5}$$

By the previous proposition, the word s is standard ϑ-episturmian, but it is not ϑ-standard, as $a^\oplus = ab \notin \mathrm{Pref}\, s$.

It is easy to prove (see [4]) that every standard ϑ-episturmian word has infinitely many ϑ-palindromic prefixes. By Proposition 1.16, they all admit a unique factorization by the elements of $\mathcal{P}_\vartheta$. Since $\mathcal{P}_\vartheta$ is a prefix code, this implies the following:

Proposition 1.23. *Every standard ϑ-episturmian word s admits a (unique) factorization by the elements of $\mathcal{P}_\vartheta$, that is,*

$$s = \pi_1 \pi_2 \cdots \pi_n \cdots\ ,$$

where $\pi_i \in \mathcal{P}_\vartheta$ for $i \geq 1$.

For a given standard ϑ-episturmian word s, such factorization will be called *canonical* in the sequel. For instance, in the case of the standard ϑ-episturmian word of Example 1.22, the canonical factorization is:

$$acb \cdot acb \cdot de \cdot acb \cdot acb \cdot acb \cdot de \cdots\ .$$

The following important lemma was proved in [4]:

Lemma 1.24. *Let s be a standard ϑ-episturmian word, and $s = \pi_1 \cdots \pi_n \cdots$ be its canonical factorization. For all $i \geq 1$, any proper and nonempty prefix of π_i is not right special in s.*

In the following, for a given standard ϑ-episturmian word s we shall denote by

$$\Pi_s = \{\pi_n \mid n \geq 1\} \tag{6}$$

the set of words of $\mathcal{P}_\vartheta$ appearing in its canonical factorization $s = \pi_1\pi_2\cdots$.

Theorem 1.25. *Let $s \in SEpi_\vartheta$. Then Π_s is a normal code.*

Proof. Any nonempty prefix p of a word of Π_s does not belong to Π_s, since Π_s is a biprefix code. Moreover, $p \notin RS\,\Pi_s$ as otherwise it would be a right special factor of s, and this is excluded by Lemma 1.24. Hence Π_s is a right normal code. Since s is closed under ϑ and $\Pi_s \subseteq PAL_\vartheta$, it follows that Π_s is also left normal. □

The following result shows that no two words of Π_s overlap properly.

Theorem 1.26. *Let $s \in SEpi_\vartheta$. Then Π_s is an overlap-free code.*

Proof. If $\operatorname{card}\Pi_s = 1$ the statement is trivial since an element of $\mathcal{P}_\vartheta$ cannot overlap properly with itself as it is unbordered. Let then $\pi, \pi' \in \Pi_s$ be such that $\pi \neq \pi'$. By contradiction, let us suppose that there exists a nonempty $u \in \operatorname{Suff}\pi \cap \operatorname{Pref}\pi'$ (which we can assume without loss of generality, since it occurs if and only if $\bar{u} \in \operatorname{Suff}\pi' \cap \operatorname{Pref}\pi$). We have $|\pi| \geq 2|u|$ and $|\pi'| \geq 2|u|$, for otherwise u would overlap properly with $\bar{u}$ and so it would have a nonempty ϑ-palindromic prefix (or suffix), which is absurd. Then there exist $v, v' \in PAL_\vartheta$ such that $\pi = \bar{u}vu$ and $\pi' = uv'\bar{u}$.

Without loss of generality, we can assume that π occurs before π' in the canonical factorization of s, so that there exists $\lambda \in (\Pi_s \setminus \{\pi'\})^*$ such that $\lambda\pi \in \operatorname{Pref} s$. Since by Lemma 1.24 any proper prefix of π cannot be right special in s, each occurrence of $\bar{u}$ must be followed by vu; the same argument applies to π', so each occurrence of u in s must be followed by $v'\bar{u}$. Therefore we have

$$s = \lambda(\bar{u}vuv')^\omega = \lambda(\pi v')^\omega .$$

As v' is a ϑ-palindromic proper factor of π', it must be in $(\mathcal{P}_\vartheta \setminus \{\pi'\})^*$, as well as $\pi v'$ and, by definition, λ. Thus we have obtained that $s \in (\Pi_s \setminus \{\pi'\})^\omega$, and so $\pi' \notin \Pi_s$, which is clearly a contradiction. Then π and π' cannot overlap properly. □

The following theorem, proved in [4, Theorem 5.5], shows, in particular, that any standard ϑ-episturmian word is a morphic image, by a suitable injective morphism, of a standard episturmian word. We report here a direct proof based on the previous results.

Theorem 1.27. *Let s be a standard ϑ-episturmian word, and f be the map defined in (4). Then $f(s)$ is a standard episturmian word, and the restriction of f to Π_s is injective, i.e., if π_i and π_j occur in the factorization of s over $\mathcal{P}_\vartheta$, and $\pi_i^f = \pi_j^f$, then $\pi_i = \pi_j$.*

Proof. Since $s \in SEpi_\vartheta$, by Theorems 1.25 and 1.26 the code Π_s is biprefix, overlap-free, and normal. By Proposition 1.1, the restriction to Π_s of the map f defined by (4) is injective. Let $B = f(\Pi_s) \subseteq A$ and denote by $g : B^* \to A^*$ the injective morphism defined by $g(\pi^f) = \pi$ for any $\pi^f \in B$. One has $s = g(t)$ for some $t \in B^\omega$. Let us now show that $t \in SEpi(B)$. Indeed, since s has infinitely many ϑ-palindromic prefixes, by Proposition 1.17 it follows that t has infinitely many palindromic prefixes, so that it is closed under reversal. Let now w be a left special factor of t, and let $a, b \in B$, $a \neq b$, be such that $aw, bw \in \operatorname{Fact} t$. Thus $g(a)g(w), g(b)g(w) \in \operatorname{Fact} s$. Since $g(a)^f \neq g(b)^f$, we have $g(a)^\ell \neq g(b)^\ell$, so that $g(w)$ is a left special factor of s, and then a prefix of it. From Lemma 1.2 it follows $w \in \operatorname{Pref} t$. □

2 Characteristic morphisms

Let X be a finite alphabet. A morphism $\varphi : X^* \to A^*$ will be called ϑ-*characteristic* if

$$\varphi(SEpi(X)) \subseteq SEpi_\vartheta ,$$

i.e., φ maps any standard episturmian word over the alphabet X in a standard ϑ-episturmian word on the alphabet A. Following this terminology, Theorem 1.12 can be reformulated by saying that *an injective morphism $\varphi : A^* \to A^*$ is standard episturmian if and only if it is R-characteristic.*

For instance, every morphism $\varphi : X^* \to A^*$ satisfying the conditions of Proposition 1.21 is ϑ-characteristic (and injective). A trivial example of a non-injective ϑ-characteristic morphism is the constant morphism $\varphi : x \in X \mapsto a \in A$, where a is a fixed ϑ-palindromic letter.

Let $X = \{x, y\}$, $A = \{a, b, c\}$, ϑ defined by $\bar{a} = a$, $\bar{b} = c$, and $\varphi : X^* \to A^*$ be the injective morphism such that $\varphi(x) = a$, $\varphi(y) = bac$. If t is any standard episturmian word beginning in y^2x, then $s = \varphi(t)$ begins with $bacbaca$, so that a is a left special factor of s which is not a prefix of s. Thus s is not ϑ-episturmian and therefore φ is not ϑ-characteristic.

In this section we shall prove some results concerning the structure of ϑ-characteristic morphisms.

Proposition 2.1. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. For each x in X, $\varphi(x) \in PAL_\vartheta^2$.*

Proof. It is clear that $|\varphi(x)|$ is a period of each prefix of $\varphi(x^\omega)$. Since $\varphi(x^\omega)$ is in $SEpi_\vartheta$, it has infinitely many ϑ-palindromic prefixes (see [4]). Then, from Lemma 1.14 the statement follows. □

Let $\varphi : X^* \to A^*$ be a morphism such that $\varphi(X) \subseteq \mathcal{P}_\vartheta^*$. For any $x \in X$, let $\varphi(x) = \pi_1^{(x)} \cdots \pi_{r_x}^{(x)}$ be the unique factorization of $\varphi(x)$ by the elements of $\mathcal{P}_\vartheta$. We set

$$\Pi(\varphi) = \{\pi \in \mathcal{P}_\vartheta \mid \exists x \in X, \exists i : \ 1 \leq i \leq r_x \ \text{ and } \pi = \pi_i^{(x)}\} . \qquad (7)$$

If φ is a ϑ-characteristic morphism, then by Propositions 2.1 and 1.16, we have $\varphi(X) \subseteq PAL_\vartheta^2 \subseteq \mathcal{P}_\vartheta^*$, so that $\Pi(\varphi)$ is well defined.

Proposition 2.2. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. Then $\Pi(\varphi)$ is an overlap-free and normal code.*

Proof. Let $t \in SEpi(X)$ be such that $\operatorname{alph} t = X$, and consider $s = \varphi(t) \in SEpi_\vartheta$. Then the set $\Pi(\varphi)$ equals Π_s, as defined in (6). The result follows from Theorems 1.25 and 1.26. □

Proposition 2.3. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. If there exist two letters $x, y \in X$ such that $\varphi(x)^f \neq \varphi(y)^f$, then $\varphi(X) \subseteq PAL_\vartheta$.*

Proof. Set $w = \varphi((x^2y)^\omega)$. Clearly $\varphi(x)$ is a right special factor of w, since it appears followed both by $\varphi(x)$ and $\varphi(y)$. As w is in $SEpi_\vartheta$, being the image of the standard episturmian word $(x^2y)^\omega$, we have that $\overline{\varphi(x)}$ is a left special factor, and thus a prefix, of w. But also $\varphi(x)$ is a prefix of w, then it must be $\varphi(x) = \overline{\varphi(x)}$, i.e., $\varphi(x) \in PAL_\vartheta$. The same argument can be applied to $\varphi(y)$, setting $w' = \varphi((y^2x)^\omega)$.

Now let $z \in X$. Then $\varphi(z)^f$ cannot be equal to both $\varphi(x)^f$ and $\varphi(y)^f$. Therefore, by applying the same argument, we obtain $\varphi(z) \in PAL_\vartheta$. From this the assertion follows. □

Proposition 2.4. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. If for $x, y \in X$, $\operatorname{Suff} \varphi(x) \cap \operatorname{Suff} \varphi(y) \neq \{\varepsilon\}$, then $\varphi(xy) = \varphi(yx)$, that is, both $\varphi(x)$ and $\varphi(y)$ are powers of a word of A^*.*

Proof. If $\varphi(xy) \neq \varphi(yx)$, since $\operatorname{Suff} \varphi(x) \cap \operatorname{Suff} \varphi(y) \neq \{\varepsilon\}$, there exists a common proper suffix h of $\varphi(xy)$ and $\varphi(yx)$, with $h \neq \varepsilon$. Let h be the longest of such suffixes. Then there exist $v, u \in A^+$ such that

$$\varphi(xy) = vh \quad \text{and} \quad \varphi(yx) = uh \ , \tag{8}$$

with $v^\ell \neq u^\ell$. Let s be a standard episturmian word whose directive word can be written as $\Delta = xy^2x\lambda$, with $\lambda \in X^\omega$, so that $s = xyxyxxyxyxt$, with $t \in X^\omega$. Thus

$$\varphi(s) = \varphi(xy)\varphi(xy)\alpha = \varphi(x)\varphi(yx)\varphi(yx)\varphi(xy)\beta$$

for some $\alpha, \beta \in A^\omega$. By (8), it follows

$$\varphi(s) = v\underline{hv}h\alpha = \varphi(x)uhu\underline{hv}h\beta \ .$$

The underlined occurrences of hv are preceded by different letters, namely v^ℓ and u^ℓ. Since $\varphi(s) \in SEpi_\vartheta$, this implies $hv \in \operatorname{Pref} \varphi(s)$ and then

$$hv = vh \ . \tag{9}$$

In a perfectly symmetric way, by considering an episturmian word s' whose directive word Δ' has yx^2y as a prefix, we obtain that $uh = hu$. Hence u and h are powers of a common primitive word w; by (9), the same can be said about v and h. Since the primitive root of a nonempty word is unique, it follows that u and v are both powers of w. As $|u| = |v|$ by definition, we obtain $u = v$ and then $\varphi(xy) = \varphi(yx)$, which is a contradiction. □

Corollary 2.5. *If $\varphi : X^* \to A^*$ is an injective ϑ-characteristic morphism, then $\varphi(X)$ is a suffix code.*

Proof. It is clear that if φ is injective, then for all $x, y \in X, x \neq y$, one has $\varphi(xy) \neq \varphi(yx)$; from Proposition 2.4 it follows $\operatorname{Suff} \varphi(x) \cap \operatorname{Suff} \varphi(y) = \{\varepsilon\}$. Thus, for all $x, y \in X$, if $x \neq y$, then $\varphi(x) \notin \operatorname{Suff} \varphi(y)$, and the statement follows. □

Proposition 2.6. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. Then for each $x, y \in X$, either*

$$\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) = \emptyset$$

or

$$\varphi(x)^f = \varphi(y)^f .$$

Proof. Let $\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) \neq \emptyset$ and $\varphi(x)^f \neq \varphi(y)^f$. We set p as the longest prefix of $\varphi(x)$ such that $\operatorname{alph} p \cap \operatorname{alph}\varphi(y) = \emptyset$ and $c \in A$ such that $pc \in \operatorname{Pref}\varphi(x)$. Let then p' be the longest prefix of $\varphi(y)$ in which c does not appear, i.e., such that $c \notin \operatorname{alph} p'$. Since we have assumed that $\varphi(x)^f \neq \varphi(y)^f$, it cannot be $p = p' = \varepsilon$. Let us suppose that both $p \neq \varepsilon$ and $p' \neq \varepsilon$. In this case we have that c is left special in $(\varphi(xy))^\omega$, since it appears preceded both by p and p' and, from the definition of p, $\operatorname{alph} p \cap \operatorname{alph} p' = \emptyset$. We reach a contradiction, since c should be a prefix of $\varphi(xy)^\omega$ which is in $SEpi_\vartheta$, and thus a prefix of $\varphi(x)$.

We then have that either $p \neq \varepsilon$ and $p' = \varepsilon$ or $p = \varepsilon$ and $p' \neq \varepsilon$. In the first case we set $z = x$ and $z' = y$, otherwise we set $z' = x$ and $z = y$. Thus we can write

$$\varphi(z) = \lambda c \gamma, \quad \varphi(z') = c\gamma' , \tag{10}$$

with $\lambda \in A^+$, $c \notin \operatorname{alph}\lambda$, and $\gamma, \gamma' \in A^*$. For each nonnegative integer n, $(z^n z')^\omega$ and $(z'^n z)^\omega$ are standard episturmian words, so that $(\varphi(z^n z'))^\omega$ and $(\varphi(z'^n z))^\omega$ are in $SEpi_\vartheta$. Moreover, since

$$(\varphi(zz'))^\omega = \varphi(z')^{-1}(\varphi(z'z))^\omega \quad \text{and} \quad (\varphi(z'z))^\omega = \varphi(z)^{-1}(\varphi(zz'))^\omega ,$$

it is clear that $(\varphi(zz'))^\omega$ and $(\varphi(z'z))^\omega$ have the same set of factors, so that each left special factor of $(\varphi(zz'))^\omega$ is a left special factor of $(\varphi(z'z))^\omega$ and *vice versa*.

Let w be a nonempty left special factor of $(\varphi(z'z))^\omega$; then w is also a prefix. As noted above, w has to be a left special factor (and thus a prefix) of $(\varphi(zz'))^\omega$. Thus w is a common prefix of $(\varphi(z'z))^\omega$ and $(\varphi(zz'))^\omega$, which is a contradiction since the first word begins with c whereas the second begins with λ, which does not contain c. Therefore $\varphi(z'z)^\omega$ has no left special factor different from ε; since each right special factor of a word in $SEpi_\vartheta$ is the ϑ-image of a left special factor, it is clear that $(\varphi(z'z))^\omega$ has no special factor different from ε.

Hence each factor of $(\varphi(z'z))^\omega$ can be extended in a unique way both to the left and to the right, so that by (10) we can write

$$(\varphi(z'z))^\omega = c\gamma'\lambda c \cdots$$

and, as stated above, each occurrence of c must be followed by $\gamma'\lambda c$, which yields that

$$(\varphi(z'z))^\omega = (c\gamma'\lambda)^\omega = (\varphi(z')\lambda)^\omega ,$$

so that this infinite word has the two periods $|\varphi(z'z)|$ and $|\varphi(z')\lambda|$. From the theorem of Fine and Wilf, one derives $\varphi(z'z)(\varphi(z')\lambda) = (\varphi(z')\lambda)\varphi(z'z)$, so that

$$\varphi(zz')\lambda = \lambda\varphi(z'z) . \tag{11}$$

The preceding equation tells us that λ is a suffix of $\lambda\varphi(z'z)$ and so, as $|\varphi(z)| > |\lambda|$, it must be a suffix of $\varphi(z)$; since λ does not contain any c, it has to be a suffix of γ, so that we can write

$$\varphi(z) = \lambda c g \lambda \tag{12}$$

for some word g. Substituting in (11), it follows

$$\varphi(zz') = \lambda\varphi(z')\lambda cg \ .$$

From the preceding equation, we have

$$(\varphi(z'^2 z))^\omega = \varphi(z')\varphi(z')\lambda\varphi(z')\lambda cg \cdots \tag{13}$$

From (12), $\varphi(z)^\ell = \lambda^\ell$. Proposition 2.4 ensures that $\lambda^\ell = \varphi(z)^\ell$ must be different from $\varphi(z')^\ell$, otherwise we would obtain $\varphi(zz') = \varphi(z'z)$ which would imply c is a prefix of $\varphi(z)$, which is a contradiction. Thus, from (13), we have that $\varphi(z')\lambda$ is a left special factor of $\varphi(z'^2 z)^\omega$ and this implies that $\varphi(z')\lambda$ is a prefix of $\varphi(z')^2\varphi(z)$, from which we obtain that λ is a prefix of $\varphi(z'z) = c\gamma'\varphi(z)$, that is a contradiction, since λ does not contain any occurrence of c. Thus the initial assumption that $\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) \neq \emptyset$ and $\varphi(x)^f \neq \varphi(y)^f$, leads in any case to a contradiction. □

Proposition 2.7. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. If $x, y \in X$ and $\varphi(x), \varphi(y) \in PAL_\vartheta$, then either $\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) = \emptyset$ or $\varphi(xy) = \varphi(yx)$. In particular, if φ is injective and $\varphi(X) \subseteq PAL_\vartheta$, then for all $x, y \in X$ with $x \neq y$ we have $\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) = \emptyset$.*

Proof. If $\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) \neq \emptyset$, from Proposition 2.6 we obtain, as $\varphi(x), \varphi(y) \in PAL_\vartheta$, that $\overline{\varphi(x)^\ell} = \varphi(x)^f = \varphi(y)^f = \overline{\varphi(y)^\ell}$. Then $\varphi(x)^\ell = \varphi(y)^\ell$ and, from Proposition 2.4, we have that $\varphi(xy) = \varphi(yx)$.

If φ is injective, then for all $x, y \in X$ with $x \neq y$ we have $\varphi(xy) \neq \varphi(yx)$ so that the assertion follows. □

Corollary 2.8. *Let $\varphi : X^* \to A^*$ be an injective ϑ-characteristic morphism such that $\varphi(X) \subseteq PAL_\vartheta$ and $\operatorname{card} X \geq 2$. Then $\varphi(X) \subseteq \mathcal{P}_\vartheta$.*

Proof. Let $x, y \in X$ with $x \neq y$. Since φ is injective, we have from Proposition 2.7 that $\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) = \emptyset$. Let u be a proper border of $\varphi(x)$. Then there exist two nonempty words v and w such that

$$\varphi(x) = uv = wu.$$

Since $\operatorname{alph}\varphi(x) \cap \operatorname{alph}\varphi(y) = \emptyset$, we have $\varphi(y)^\ell \neq w^\ell$; thus

$$\varphi(yx)^\omega = \varphi(y)uv\varphi(y)wu\varphi(y) \cdots$$

shows that u is a left special factor in $\varphi(yx)^\omega$, but this would imply that u is a prefix of $\varphi(yx)$. As $\operatorname{alph} u \cap \operatorname{alph}\varphi(y) = \emptyset$, it follows $u = \varepsilon$, i.e., $\varphi(x) \in \mathcal{P}_\vartheta$. The same argument applies to $\varphi(y)$. □

The following lemma will be useful in the next section.

Lemma 2.9. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. Then for each $x \in X$ and for any $a \in A$,*

$$|\varphi(x)|_a > 1 \Longrightarrow |\varphi(x)|_{\varphi(x)^f} > 1.$$

Proof. Let b be the first letter of $\varphi(x)$ such that $|\varphi(x)|_b > 1$. Then we can write

$$\varphi(x) = vbwbw'$$

with $w, w' \in A^*$, $b \notin (\operatorname{alph} v \cup \operatorname{alph} w)$, and $|\varphi(x)|_c = 1$ for each c in $\operatorname{alph} v$. If $v \neq \varepsilon$, then we have that $v^\ell \neq (bw)^\ell$, but that means that b is left special in $\varphi(x^\omega)$, which is a contradiction, since each left special factor of $\varphi(x^\omega)$ is a prefix and b is not in $\operatorname{alph} v$. Then it must be $v = \varepsilon$ and $b = \varphi(x)^f$. □

3 Main results

The first result of this section is a characterization of injective ϑ-characteristic morphisms such that the image of any letter is an unbordered ϑ-palindrome.

Theorem 3.1. *Let $\varphi : X^* \to A^*$ be an injective morphism such that for any $x \in X$, $\varphi(x) \in \mathcal{P}_\vartheta$. Then φ is ϑ-characteristic if and only if the following two conditions hold:*

1. $\operatorname{alph} \varphi(x) \cap \operatorname{alph} \varphi(y) = \emptyset$*, for any x, y in X such that $x \neq y$.*
2. *for any $x \in X$ and $a \in A$, $|\varphi(x)|_a \leq 1$.*

Proof. Let φ be ϑ-characteristic. Since φ is injective, from Proposition 2.7 we have that if $x \neq y$, then $\operatorname{alph} \varphi(x) \cap \operatorname{alph} \varphi(y) = \emptyset$. Thus condition 1 holds. Let us now prove that condition 2 is satisfied. This is certainly true if $|\varphi(x)| \leq 2$, as $\varphi(x) \in \mathcal{P}_\vartheta$. Let us then suppose $|\varphi(x)| > 2$. We can write

$$\varphi(x) = ax_1 \cdots x_n b\ ,$$

with $x_i \in A$, $i = 1, \ldots, n$, $\bar{a} = b$, and $a \neq b$.

Let us prove that for any $i = 1, \ldots, n$, $x_i \notin \{a, b\}$. By contradiction, suppose that b has an internal occurrence in $\varphi(x)$, and consider its first occurrence. Since $\varphi(x)$ is a ϑ-palindrome, we can write

$$\varphi(x) = ax_1 \cdots x_i b\lambda = \bar{\lambda} a \bar{x}_i \cdots \bar{x}_1 b\ ,$$

with $\lambda \in A^*$, $1 \leq i < n$, and $x_j \neq b$ for $j = 1, \ldots, i$.

We now consider the standard ϑ-episturmian word $s = \varphi(x^\omega)$, whose first letter is a. We have that no letter $\bar{x}_j$, $j = 1, \ldots, i$, is left special in s, as otherwise $\bar{x}_j = a$ that implies $x_j = b$, which is absurd. Also b cannot be left special since otherwise $b = a$. Thus it follows that $x_i = \bar{x}_1$, $x_{i-1} = \bar{x}_2$, $\ldots$, $x_1 = \bar{x}_i$. Hence, $ax_1 \cdots x_i b$ is a proper border of $\varphi(x)$, which is a contradiction. From this, since $\varphi(x)$ is a ϑ-palindrome, one derives that there is no internal occurrence of a in $\varphi(x)$ as well.

Finally, any letter of $\varphi(x)$ cannot occur more than once. This is a consequence of Lemma 2.9, since otherwise the first letter of $\varphi(x)$, namely a, would reoccur in $\varphi(x)$. Thus condition 2 holds.

Conversely, let us now suppose that conditions 1 and 2 hold; Proposition 1.21 ensures then that φ is ϑ-characteristic. □

A different proof of Theorem 3.1 will be given at the end of this section, as a consequence of a full characterization of injective ϑ-characteristic morphisms, given in Theorem 3.13.

Remark. In the "if" part of Theorem 3.1 the requirement $\varphi(X) \subseteq \mathcal{P}_\vartheta$ can be replaced by $\varphi(X) \subseteq PAL_\vartheta$, as condition 2 implies that $\varphi(x)$ is unbordered for any $x \in X$, so that $\varphi(X) \subseteq \mathcal{P}_\vartheta$. In the "only if" part, in view of Corollary 2.8, one can replace $\varphi(X) \subseteq \mathcal{P}_\vartheta$ by $\varphi(X) \subseteq PAL_\vartheta$ under the hypothesis that $\operatorname{card} X \geq 2$.

Example 3.2. Let X, A, ϑ, and g be defined as in Example 1.22. Then the morphism g is ϑ-characteristic.

As an immediate consequence of Theorem 3.1, we obtain:

Corollary 3.3. *Let $\zeta : X^* \to B^*$ be an R-characteristic morphism, $g : B^* \to A^*$ be an injective morphism satisfying $g(B) \subseteq \mathcal{P}_\vartheta$ and the two conditions in the statement of Theorem 3.1. Then $\varphi = g \circ \zeta$ is ϑ-characteristic.*

Example 3.4. Let X, A, ϑ, and g be defined as in Example 1.22, and let ζ be the endomorphism of X^* such that $\zeta(x) = xy$ and $\zeta(y) = xyx$. Since $\zeta = \mu_{xy} \circ \sigma$, where $\sigma(x) = y$ and $\sigma(y) = x$, ζ is a standard episturmian morphism. Hence the morphism $\varphi : X^* \to A^*$ given by

$$\varphi(x) = acbde, \quad \varphi(y) = acbdeacb$$

is ϑ-characteristic, as $\varphi = g \circ \zeta$.

Theorem 3.5. *Let $\varphi : X^* \to A^*$ be a ϑ-characteristic morphism. Then there exist $B \subseteq A$, a morphism $\zeta : X^* \to B^*$, and a morphism $g : B^* \to A^*$ such that:*

1. *ζ is R-characteristic,*
2. *$g(B) = \Pi(\varphi)$, with $g(b) \in bA^*$ for all $b \in B$,*
3. *$\varphi = g \circ \zeta$.*

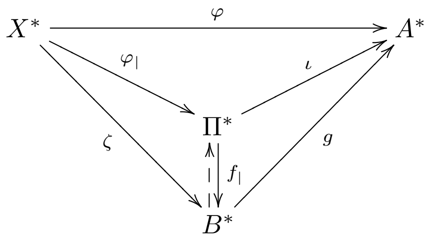

Figure 1: A commutative diagram describing Theorem 3.5

Proof (see Fig. 1). Set $\Pi = \Pi(\varphi)$, as defined in (7), and let $B = f(\Pi) \subseteq A$, where f is the morphism considered in (4). Let $\varphi_| : X^* \to \Pi^*$ and $f_| : \Pi^* \to B^*$ be the restrictions of φ and f, respectively. Setting $\zeta = f_| \circ \varphi_| : X^* \to B^*$, by Theorem 1.27 one derives $\zeta(SEpi(X)) \subseteq SEpi(B)$, i.e., ζ is R-characteristic.

Let $t \in SEpi(X)$ be such that $\operatorname{alph} t = X$, and consider $s = \varphi(t) \in SEpi_\vartheta$. Since Π equals Π_s, as defined in (6), by Theorem 1.27 the morphism f is injective over Π, so that $f_|$ is bijective. Set $g = \iota \circ f_|^{-1}$, where $\iota : \Pi^* \to A^*$ is the inclusion map. Then $g(B) = \Pi$, and $g(b) \in bA^*$ for all $b \in B$. Furthermore, we have

$$\varphi = \iota \circ \varphi_| = \iota \circ (f_|^{-1} \circ f_|) \circ \varphi_| = (\iota \circ f_|^{-1}) \circ (f_| \circ \varphi_|) = g \circ \zeta$$

as desired. □

Example 3.6. Let $X = \{x, y\}$, $A = \{a, b, c\}$, and ϑ be the antimorphism of A^* such that $\bar{a} = a$ and $\bar{b} = c$. The morphism $\varphi : X^* \to A^*$ defined by $\varphi(x) = a$ and $\varphi(y) = abac$ is ϑ-characteristic (this will be clear after Theorem 3.13, see Example 3.14), and it can be decomposed as $\varphi = g \circ \zeta$, where $\zeta : X^* \to B^*$ (with $B = \{a, b\}$) is the morphism such that $\zeta(x) = a$ and $\zeta(y) = ab$, while $g : B^* \to A^*$ is defined by $g(a) = a$ and $g(b) = bac$. We remark that $\zeta(SEpi(X)) \subseteq SEpi(B)$, but $g(SEpi(B)) \nsubseteq SEpi_\vartheta$ as it can be verified using Theorem 3.1. Observe that this example shows that not all ϑ-characteristic morphisms can be constructed as in Corollary 3.3.

Proposition 3.7. *Let $\zeta : X^* \to A^*$ be an injective morphism. Then ζ is R-characteristic if and only if it can be decomposed as $\zeta = \mu_w \circ \eta$, where $w \in A^*$ and $\eta : X^* \to A^*$ is an injective literal morphism.*

Proof. Let $\zeta = \mu_w \circ \eta$, with $w \in A^*$ and η an injective literal morphism. Then η is trivially R-characteristic and μ_w is R-characteristic too, by Theorem 1.12. Therefore also their composition ζ is R-characteristic.

Conversely, let us first suppose that $\zeta(X) \subseteq a_1 A^*$ for some $a_1 \in A$. Then for any $t \in SEpi(X)$, $\zeta(t)$ is a standard episturmian word beginning with a_1, so that by Proposition 1.5 the letter a_1 is separating for $\zeta(t)$. In particular a_1 is separating for each $\zeta(x)$ ($x \in X$); by Proposition 1.11 there exists a morphism $\alpha_1 : X^* \to A^*$ such that $\zeta = \mu_{a_1} \circ \alpha_1$. Since $t \in SEpi(X)$, $\mu_{a_1}(\alpha_1(t))$ is a standard episturmian word over A, so that by Proposition 1.13 the word $\alpha_1(t)$ is also a standard episturmian word over A. Thus α_1 is injective and R-characteristic, and we can iterate the above argument to find new letters $a_i \in A$ and R-characteristic morphisms α_i such that $\zeta = \mu_{a_1} \circ \cdots \circ \mu_{a_i} \circ \alpha_i$, as long as all images of letters under α_i have the same first letter.

If $\operatorname{card} X > 1$, since ζ is injective, we eventually obtain the following decomposition:

$$\zeta = \mu_{a_1} \circ \mu_{a_2} \circ \cdots \circ \mu_{a_n} \circ \eta = \mu_w \circ \eta \ , \tag{14}$$

where $a_1, \ldots, a_n \in A$, $w = a_1 \cdots a_n$, and $\eta = \alpha_n$ is such that $\eta(x)^f \neq \eta(y)^f$ for some $x, y \in X$. If the original requirement $\zeta(X) \subseteq a_1 A^*$ is not met by any a_1, that is, if $\zeta(x)^f \neq \zeta(y)^f$ for some $x, y \in X$, we can still fit in (14) choosing $n = 0$ and $w = \varepsilon$.

Let then $x, y \in X$ be such that $\eta(x)^f \neq \eta(y)^f$. Since η is R-characteristic, by Proposition 2.3 we obtain $\eta(X) \subseteq PAL$. Moreover, since η is injective, by Corollary 2.8 we have $\eta(X) \subseteq \mathcal{P}_R = A$, so that η is an injective literal morphism.

In the case $X = \{x\}$, the lengths of the words $\alpha_i(x)$ for $i \geq 1$ are decreasing. Hence eventually we find an $n \geq 1$ such that $\alpha_n(x) \in A$ and the assertion is proved, for

$$\zeta = \mu_{a_1} \circ \cdots \circ \mu_{a_n} \circ \alpha_n = \mu_w \circ \alpha_n \ ,$$

with $w = a_1 \cdots a_n \in A^*$ and $\alpha_n : X^* \to A^*$ an injective literal morphism. □

Example 3.8. Let $X = \{x, y\}$, $A = \{a, b, c\}$, and $\zeta : X^* \to A^*$ be defined by:

$$\zeta(x) = abacabaabacab = \mu_a(bcbabcb) \quad \text{and} \quad \zeta(y) = abacaba = \mu_a(bcba)\ ,$$

so that $\alpha_1(x) = bcbabcb$ and $\alpha_1(y) = bcba$. Then $\zeta(x)$ can be rewritten also as

$$\zeta(x) = \mu_a(\alpha_1(x)) = (\mu_a \circ \mu_b)(cacb) = (\mu_a \circ \mu_b \circ \mu_c)(ab) = \mu_{abca}(b)\ .$$

In a similar way, one obtains $\zeta(y) = \mu_{abca}(a)$. Hence, setting $\eta(x) = b$ and $\eta(y) = a$, the morphism $\zeta = \mu_{abca} \circ \eta$ is R-characteristic, in view of the preceding proposition.

From Theorem 3.5 and Proposition 3.7 one derives the following:

Corollary 3.9. *Every injective ϑ-characteristic morphism $\varphi : X^* \to A^*$ can be decomposed as*

$$\varphi = g \circ \mu_w \circ \eta\ , \tag{15}$$

where $\eta : X^* \to B^*$ is an injective literal morphism, $\mu_w : B^* \to B^*$ is a pure standard episturmian morphism (with $w \in B^*$), and $g : B^* \to A^*$ is an injective morphism such that $g(B) = \Pi(\varphi)$.

Remarks.

1. From the preceding result, we have in particular that if $\varphi : X^* \to A^*$ is an injective ϑ-characteristic morphism, then $\operatorname{card} X \leq \operatorname{card} A$.

2. Theorem 3.5 and Proposition 3.7 show that a decomposition (15) can always be chosen so that $B = \operatorname{alph} w \cup \eta(X) \subseteq A$ and $g(b) \in bA^* \cap \mathcal{P}_\vartheta$ for each $b \in B$.

3. Corollary 3.9 shows that the code $\varphi(X)$, which is a suffix code by Corollary 2.5, is in fact the *composition* (by means of g) [1] of the code $\mu_w(\eta(X)) \subseteq B^*$ and the biprefix, overlap-free, and normal code $g(B) \subseteq A^*$.

4. From the proof of Proposition 3.7, one easily obtains that if $\operatorname{card} X > 1$, the decomposition (15) is unique.

Proposition 3.10. *Let $\varphi : X^* \to A^*$ be an injective ϑ-characteristic morphism, decomposed as in* (15)*, and ψ be the iterated palindrome closure operator. The word $u = g(\psi(w))$ is a ϑ-palindrome such that for each $x \in X$,*

$$\varphi(x)u = (u\, g(\eta(x)))^{\oplus}\ , \tag{16}$$

and $\varphi(x)$ is either a prefix of u or equal to $ug(\eta(x))$.

Proof. Since $\psi(w)$ is a palindrome and the injective morphism g is such that $g(B) \subseteq \mathcal{P}_\vartheta$, we have $u \in PAL_\vartheta$ in view of Proposition 1.17. Let $x \in X$ and set $b = \eta(x)$. We have

$$\varphi(x)u = g(\mu_w(\eta(x))\psi(w)) = g(\mu_w(b)\psi(w)).$$

By Propositions 1.9 and 1.17 we obtain

$$g(\mu_w(b)\psi(w)) = g(\psi(wb)) = g((\psi(w)b)^{(+)}) = (g(\psi(w)b))^{\oplus} = (ug(b))^{\oplus},$$

and (16) follows. Thus, since $g(b)$ is a ϑ-palindromic suffix of $ug(b)$, we derive $|\varphi(x)| \leq |ug(b)|$. By Proposition 2.1, $\varphi(x) \in \mathcal{P}_\vartheta^*$. Therefore it can be either equal to $ug(b)$ or a prefix of u. Indeed, if $\varphi(x) = ur$ with r a nonempty proper prefix of $g(b) \in \mathcal{P}_\vartheta$, then $r \in \mathcal{P}_\vartheta^*$, as $\mathcal{P}_\vartheta^*$ is left unitary. This gives rise to a contradiction because $\mathcal{P}_\vartheta$ is a biprefix code. □

Corollary 3.11. *Under the same hypotheses and with the same notation as in Proposition 3.10, if $x_1, x_2 \in X$ are such that $|\varphi(x_1)| \leq |\varphi(x_2)|$, then either $\varphi(x_1) \in \operatorname{Pref}\varphi(x_2)$, or $\varphi(x_1)$ and $\varphi(x_2)$ do not overlap, i.e.,*

$$\operatorname{Suff}\varphi(x_1) \cap \operatorname{Pref}\varphi(x_2) = \operatorname{Suff}\varphi(x_2) \cap \operatorname{Pref}\varphi(x_1) = \{\varepsilon\} .$$

Proof. For $i = 1, 2$, let us set $b_i = \eta(x_i)$. By Proposition 3.10, $\varphi(x_i)$ is either a prefix of u or equal to $ug(b_i)$.

If $\varphi(x_1)$ is a prefix of u, then it is a prefix of $\varphi(x_2)$ too, as $|\varphi(x_1)| \leq |\varphi(x_2)|$. Let us then suppose that

$$\varphi(x_i) = ug(b_i) \quad \text{for } i = 1, 2 . \tag{17}$$

Now let v be an element of $\operatorname{Suff}\varphi(x_1) \cap \operatorname{Pref}\varphi(x_2)$. Since $\varphi(x_2) \in \mathcal{P}_\vartheta^*$, we can write $v = v'\lambda$, where v' is the longest word of $\mathcal{P}_\vartheta^* \cap \operatorname{Pref} v$. Then λ is a proper prefix of a word π occurring in the unique factorization of $\varphi(x_2)$ over $\mathcal{P}_\vartheta$. If λ was nonempty, π would overlap with some word π' of the factorization of $\varphi(x_1)$ over $\mathcal{P}_\vartheta$. This is absurd, since for any $t \in SEpi(X)$ such that $x_1, x_2 \in \operatorname{alph} t$, both π and π' would be in $\Pi_{\varphi(t)}$, which is overlap-free by Theorem 1.26. Hence $\lambda = \varepsilon$ and $v \in \mathcal{P}_\vartheta^*$. Therefore by (17) we have $v = g(\xi)$, where ξ is an element of $\operatorname{Suff}(\psi(w)b_1) \cap \operatorname{Pref}(\psi(w)b_2)$.

By Proposition 3.10, (17) is equivalent to $(u\,g(b_i))^{\oplus} = ug(b_i)u$, $i = 1, 2$. Since for $i = 1, 2$ the word $g(b_i)$ is an unbordered ϑ-palindrome, any ϑ-palindromic suffix of $ug(b_i)$ longer than $g(b_i)$ can be written as $g(b_i)\xi_i g(b_i)$, with ξ_i a ϑ-palindromic suffix of u. Hence (17) holds for $i = 1, 2$ if and only if u has no ϑ-palindromic suffixes preceded respectively by $g(b_1)$ or $g(b_2)$. By Proposition 1.17, this implies that for $i = 1, 2$, $\psi(w)$ has no palindromic suffix preceded by b_i, so that $b_i \notin \operatorname{alph} w = \operatorname{alph}\psi(w)$. Therefore, since $b_1 \neq b_2$, the only word in $\operatorname{Suff}(\psi(w)b_1) \cap \operatorname{Pref}(\psi(w)b_2)$ is ε. Hence $v = g(\varepsilon) = \varepsilon$.

The same argument can be used to prove that $\operatorname{Suff}\varphi(x_2) \cap \operatorname{Pref}\varphi(x_1) = \{\varepsilon\}$. □

Example 3.12. Let $X = \{x, y\}$, $A = \{a, b, c, d, e\}$, $B = \{a, d\}$, and ϑ be defined by $\bar{a} = b$, $\bar{c} = c$, and $\bar{d} = e$. As we have seen in Example 3.4, the morphism $\varphi : X^* \to A^*$ defined by $\varphi(x) = acbde$ and $\varphi(y) = acbdeacb$ is ϑ-characteristic. We can decompose φ as $\varphi = g \circ \mu_{ad} \circ \eta$, where $g : B^* \to A^*$ is defined by $g(a) = acb \in \mathcal{P}_\vartheta$, $g(d) = de \in \mathcal{P}_\vartheta$, and η is such that $\eta(x) = d$ and $\eta(y) = a$. We have $u = g(\psi(ad)) = g(ada) = acbdeacb$, and

$$\varphi(x)u = acbdeacbdeacb = (acbdeacbde)^{\oplus} = (u\,g(\eta(x)))^{\oplus} .$$

Similarly, $\varphi(y)u = (u\,g(\eta(y)))^{\oplus}$. In this case, $\varphi(x)$ is a prefix of $\varphi(y)$.

The following basic theorem gives a characterization of all injective ϑ-characteristic morphisms.

Theorem 3.13. *Let $\varphi : X^* \to A^*$ be an injective morphism. Then φ is ϑ-characteristic if and only if it is decomposable as*

$$\varphi = g \circ \mu_w \circ \eta$$

as in (15)*, with $B = \operatorname{alph} w \cup \eta(X)$ and $g(B) = \Pi \subseteq \mathcal{P}_\vartheta$ satisfying the following conditions:*

1. *Π is an overlap-free and normal code,*
2. *$LS\,(\{g(\psi(w))\} \cup \Pi) \subseteq \operatorname{Pref} g(\psi(w))$,*
3. *if $b, c \in A \setminus \operatorname{Suff} \Pi$ and $v \in \Pi^*$ are such that $bv\bar{c} \in \operatorname{Fact} \Pi$, then $v = g(\psi(w'x))$, with $w' \in \operatorname{Pref} w$ and $x \in \{\varepsilon\} \cup (B \setminus \eta(X))$.*

The proof of this theorem, which is rather cumbersome, will be given in Section 5, using some results on biprefix, overlap-free, and normal codes that will be proved in Section 4. We conclude this section by giving some examples and a remark related to Theorem 3.13; moreover, from this theorem we derive a different proof of Theorem 3.1.

Example 3.14. Let $A = \{a, b, c\}$, $X = \{x, y\}$, $B = \{a, b\}$, and let ϑ and $\varphi : X^* \to A^*$ be defined as in Example 3.6, namely $\bar{a} = a$, $\bar{b} = c$, and $\varphi = g \circ \mu_a \circ \eta$, where $\eta(x) = a$, $\eta(y) = b$, and $g : B^* \to A^*$ is defined by $g(a) = a$ and $g(b) = bac$. Then $\Pi = g(B) = \{a, bac\}$ is an overlap-free code and satisfies:

- $(\operatorname{Suff} \Pi \setminus \Pi) \cap LS\,\Pi = \{\varepsilon\}$, so that Π is normal,
- $LS(\{g(\psi(a))\} \cup \Pi) = LS(\{a\} \cup \Pi) = \{\varepsilon\} \subseteq \operatorname{Pref} a$.

The only word verifying the hypotheses of condition 3 is $bac = ba\bar{b} = g(b) \in \Pi$, with $a \in \Pi^*$ and $b \notin \operatorname{Suff} \Pi$. Since $a = g(\psi(a))$ and $B \setminus \eta(X) = \emptyset$, also condition 3 of Theorem 3.13 is satisfied. Hence φ is ϑ-characteristic.

Example 3.15. Let $X = \{x, y\}$, $A = \{a, b, c\}$, ϑ be such that $\bar{a} = a$, $\bar{b} = c$, and the morphism $\varphi : X^* \to A^*$ be defined by $\varphi(x) = a$ and $\varphi(y) = abaac$. In this case we have $\varphi = g \circ \mu_a \circ \eta$, where $B = \{a, b\}$, $g(a) = a$, $g(b) = baac$, $\eta(x) = a$, and $\eta(y) = b$. Then the morphism φ is not ϑ-characteristic. Indeed, if t is any standard episturmian word starting with yxy, then $\varphi(t)$ has the prefix $abaacaabaac$, so that aa is a left special factor of $\varphi(t)$ but not a prefix of it.

In fact, condition 3 of Theorem 3.13 is not satisfied in this case, since $baac = baa\bar{b} = g(b)$, $b \notin \operatorname{Suff} \Pi$, $aa \in \Pi^*$, $B \setminus \eta(X) = \emptyset$, and

$$aa \notin \{g(\psi(w')) \mid w' \in \operatorname{Pref} a\} = \{\varepsilon, a\}\ .$$

If we choose $X' = \{y\}$ with $\eta'(y) = b$, then

$$g(\mu_a(\eta'(y^\omega))) = (abaac)^\omega \in SEpi_\vartheta\ ,$$

so that $\varphi' = g \circ \mu_a \circ \eta'$ is ϑ-characteristic. In this case $B = \operatorname{alph} a \cup \eta'(X')$, $B \setminus \eta'(X') = \{a\}$, and $aa = g(\psi(aa)) = g(aa)$, so that condition 3 is satisfied.

Example 3.16. Let $X = \{x, y\}$, $A = \{a, b, c, d, e, h\}$, and ϑ be the antimorphism over A defined by $\bar{a} = a$, $\bar{b} = c$, $\bar{d} = e$, $\bar{h} = h$. Let also $w = adb \in A^*$, $B = \{a, b, d\} = \operatorname{alph} w$, and $\eta : X^* \to B^*$ be defined by $\eta(x) = a$ and $\eta(y) = b$.

Finally, set $g(a) = a$, $g(d) = dahae$, and $g(b) = badahaeadahaeac$. Then the morphism $\varphi = g \circ \mu_w \circ \eta$ is such that

$$\varphi(y) = adahaeabadahaeadahaeac \quad \text{and} \quad \varphi(x) = \varphi(y)\, adahaea \ ,$$

and it is ϑ-characteristic as the code $\Pi = g(B)$ and the word $u = g(\psi(w)) = g(adabada) = \varphi(x)$ satisfy all three conditions of Theorem 3.13.

Remark. Let us observe that Theorem 3.13 gives an effective procedure to decide whether, for a given ϑ, an injective morphism $\varphi : X^* \to A^*$ is ϑ-characteristic. The procedure runs in the following steps:

1. Check whether $\varphi(X) \subseteq \mathcal{P}_\vartheta^*$.
2. If the previous condition is satisfied, then compute $\Pi = \Pi(\varphi)$.
3. Verify that Π is overlap-free and normal.
4. Compute $B = f(\Pi)$ and then the morphism $g : B^* \to A^*$ given by $g(B) = \Pi$.
5. Since $\varphi = g \circ \zeta$, verify that ζ is R-characteristic, i.e., there exists $w \in B^*$ such that $\zeta = \mu_w \circ \eta$, where η is a literal morphism from X^* to B^*. This can be always simply done, following the argument used in the proof of Proposition 3.7.
6. Compute $g(\psi(w))$ and verify that conditions 2 and 3 of Theorem 3.13 are satisfied. This can also be effectively done.

We now give a new proof of Theorem 3.1, based on Theorem 3.13.

Proof of Theorem 3.1. Let $\varphi : X^* \to A^*$ be an injective morphism such that $\varphi(X) = \Pi \subseteq \mathcal{P}_\vartheta$ and satisfying conditions 1 and 2 of Theorem 3.1. In this case we can assume $w = \varepsilon$, so that $B = \eta(X)$, $u = g(\psi(w)) = \varepsilon$, and $\varphi = g \circ \eta$. Hence $\Pi = g(B) = \varphi(X)$. The code Π is overlap-free by conditions 1 and 2. Since any letter of A occurs at most once in any word of Π, we have $LS(\{\varepsilon\} \cup \Pi) \subseteq \{\varepsilon\} = \operatorname{Pref} u$, whence

$$(\operatorname{Suff} \Pi \setminus \Pi) \cap LS\, \Pi \subseteq \{\varepsilon\} \ ,$$

i.e., Π is a left normal, and therefore normal, code. Let $b, c \in A \setminus \operatorname{Suff} \Pi$, and $v \in \Pi^*$ be such that $bv\bar{c} \in \operatorname{Fact} \pi$ for some $\pi \in \Pi$. This implies $v = \varepsilon = g(\psi(\varepsilon))$, because the equation $v = \pi_1 \cdots \pi_k$ with $\pi_1, \ldots, \pi_k \in \Pi$ would violate condition 1 of Theorem 3.1. Thus all the hypotheses of Theorem 3.13 are satisfied for $w = \varepsilon$, so that $\varphi = g \circ \mu_\varepsilon \circ \eta$ is ϑ-characteristic.

Conversely, let $\varphi : X^* \to A^*$ be an injective ϑ-characteristic morphism such that $\varphi(X) = \Pi \subseteq \mathcal{P}_\vartheta$. We can take $w = \varepsilon$, $B = \eta(X) \subseteq A$ and write $\varphi = g \circ \eta$, so that $g(B) = \varphi(X) = \Pi$. Since $u = \varepsilon$, by Theorem 3.13 we have

$$LS(\{\varepsilon\} \cup \Pi) \subseteq \{\varepsilon\} \ , \tag{18}$$

and, as $B \setminus \eta(X) = \emptyset$, for all $b, c \in A \setminus \operatorname{Suff} \Pi$ and $v \in \Pi^*$,

$$bv\bar{c} \in \operatorname{Fact} \Pi \implies v = g(\psi(\varepsilon)) = \varepsilon \ . \tag{19}$$

Moreover, since $\Pi = \Pi(\varphi)$, we have that Π is normal and overlap-free by Proposition 2.2.

Now let $a \in A$ and suppose $a \in \operatorname{alph} \pi$ for some $\pi \in \Pi$. We will show that *any* two occurrences of a in the words of Π coincide, so that a has exactly one occurrence in Π. Let then $\pi_1, \pi_2 \in \Pi$ be such that

$$\pi_1 = \lambda_1 a \rho_1 \quad \text{and} \quad \pi_2 = \lambda_2 a \rho_2$$

for some $\lambda_1, \lambda_2, \rho_1, \rho_2 \in A^*$, and let us first prove that $\lambda_1 = \lambda_2$.

Let s be the longest common suffix of λ_1 and λ_2, and let $\lambda_i = \lambda_i' s$ for $i = 1, 2$. If both λ_1' and λ_2' were nonempty, their last letters would differ by the definition of s, and therefore sa would be in $LS\,\Pi$, contradicting (18).

Next, we may assume $\lambda_1' = \varepsilon$ and $\lambda_2' \neq \varepsilon$, without loss of generality. Then $sa \in \operatorname{Pref} \pi_1$, so that by Proposition 1.1 we obtain $\lambda_2' \pi_1 \in \operatorname{Pref} \pi_2$; in particular, we have $\pi_1 \neq \pi_2$. Let then r be the longest word of $\Pi^* \cap \operatorname{Suff} \lambda_2'$, and set $\lambda_2' = \xi r$. Since $\lambda_2' \neq \varepsilon$ and Π is a biprefix code, we have $\xi \neq \varepsilon$. Furthermore, ξ^ℓ is not a suffix of any word of Π, for if π' were such a word, by Proposition 1.1 we would derive that $\pi' \in \operatorname{Suff} \xi$, contradicting the definition of r.

Let us now write $\pi_2 = \xi r \pi_1 \delta$. The word δ is nonempty since Π is a biprefix code. Let r' be the longest word in $\Pi^* \cap \operatorname{Pref} \delta$ and set $\delta = r' \zeta$. Since Π is a biprefix code, $\zeta \neq \varepsilon$. By Proposition 1.1, we derive that $\zeta^f \notin \operatorname{Pref} \Pi$. By (19), we obtain that $r \pi_1 r' = \varepsilon$, which is absurd.

Thus $\lambda_1' = \lambda_2' = \varepsilon$, whence $\lambda_1 = \lambda_2$ as desired. From $\lambda_1 a = \lambda_2 a$ it follows $\pi_1^f = \pi_2^f$, so that by Proposition 1.1 we have $\pi_1 = \pi_2$ and hence $\rho_1 = \rho_2$. Therefore, the two (generic) occurrences of a we have considered are the same.

We have thus proved that every letter of A occurs at most once among all the words of $\Pi = \varphi(X)$, so that conditions 1 and 2 of Theorem 3.1 are satisfied. □

4 Some properties of normal codes

In this section, we analyse some properties of left (or right) normal codes, under some additional requirements such as being suffix, prefix, or overlap-free. A first noteworthy result was already given in Section 1 (cf. Proposition 1.1). We stress that all statements of the following propositions can be applied to codes which are biprefix, overlap-free, and normal.

Lemma 4.1. *Let Z be a left normal and suffix code over A. For any $a, b \in A$, $a \neq b$, $\lambda \in A^+$, if $a\lambda, b\lambda \in \operatorname{Fact} Z^*$ and $\lambda \notin \operatorname{Pref} Z^*$, then $a\lambda, b\lambda \in \operatorname{Fact} Z$.*

Proof. By symmetry, it suffices to prove that $a\lambda \in \operatorname{Fact} Z$. By hypothesis there exist words $v, \zeta \in A^*$ such that $va\lambda\zeta = z_1 \cdots z_n$, with $n \geq 1$ and $z_i \in Z$, $i = 1, \ldots, n$. If $n = 1$, then $a\lambda \in \operatorname{Fact} Z$ and we are done. Then suppose $n > 1$, and write:

$$va = z_1 \cdots z_h \delta, \quad \delta\lambda\zeta = z_{h+1} \cdots z_n, \quad z_{h+1} = \delta\xi = z\,, \tag{20}$$

with $\delta \in A^*$, $h \geq 0$, and $\xi \neq \varepsilon$. Let us observe that $\delta \neq \varepsilon$, for otherwise $\lambda \in \operatorname{Pref} Z^*$, contradicting the hypothesis on λ.

If $|\delta\lambda| \leq |z|$, then since $a = \delta^\ell$, we have $a\lambda \in \operatorname{Fact} Z$ and we are done. Therefore, suppose $|\delta\lambda| > |z|$. This implies that ξ is a proper prefix of λ, and by (20), a proper suffix of z. Moreover, as $a = \delta^\ell$, we have $a\xi \in \operatorname{Fact} Z$.

Since $b\lambda \in \operatorname{Fact} Z^*$, in a symmetric way one derives that either $b\lambda \in \operatorname{Fact} Z$, or there exists $\xi' \neq \varepsilon$ which is a proper prefix of λ and a proper suffix of a word

$z' \in Z$. In the first case we have $b\lambda \in \mathrm{Fact}\, Z$, so that $a\xi, b\xi \in \mathrm{Fact}\, Z$, whence $\xi \in \mathrm{Suff}\, Z \cap LS\, Z$, and $\xi \notin Z$ since Z is a suffix code. We reach a contradiction since $\xi \neq \varepsilon$ and Z is left normal.

In the second case, ξ and ξ' are both prefixes of λ. Let $\hat{\xi}$ be in $\{\xi, \xi'\}$ with minimal length. Then $a\hat{\xi}, b\hat{\xi} \in \mathrm{Fact}\, Z$, so that $\hat{\xi} \in \mathrm{Suff}\, Z \cap LS\, Z$. Since $\hat{\xi} \notin Z$, as Z is a suffix code, we reach again a contradiction because $\hat{\xi} \neq \varepsilon$ and Z is left normal. Therefore, the only possibility is that $a\lambda \in \mathrm{Fact}\, Z$. □

Proposition 4.2. *Let Z be a suffix, left normal, and overlap-free code over A, and let $a, b \in A$, $v \in A^*$, $\lambda \in A^+$ be such that $a \neq b$, $va \notin Z^*$, $va\lambda \in \mathrm{Pref}\, Z^*$, and $b\lambda \in \mathrm{Fact}\, Z^*$. Then $a\lambda \in \mathrm{Fact}\, Z$.*

Proof. Since $va\lambda \in \mathrm{Pref}\, Z^*$, there exists $\zeta \in A^*$ such that $va\lambda\zeta = z_1 \cdots z_n$, $n \geq 1$, $z_i \in Z$, $i = 1, \ldots, n$. Then we can assume that (20) holds for suitable $h \geq 0$, $\delta \in A^*$, and $\xi \in A^+$. We have $n > 1$, for otherwise the statement is trivial, and $\delta \neq \varepsilon$ since $va \notin Z^*$. As $\delta^\ell = a$, if $|\delta\lambda| \leq |z|$ we obtain $a\lambda \in \mathrm{Fact}\, Z$ and we are done. Therefore assume $|\delta\lambda| > |z|$. In this case ξ is a proper prefix of λ and a proper suffix of z. If $\lambda \in \mathrm{Pref}\, Z^*$ we reach a contradiction, since $\xi \in \mathrm{Suff}\, Z \cap \mathrm{Pref}\, Z^*$ and this contradicts the hypothesis that Z is a suffix and overlap-free code. Thus $\lambda \notin \mathrm{Pref}\, Z^*$; this implies, by the previous lemma, that $a\lambda \in \mathrm{Fact}\, Z$. □

Proposition 4.3. *Let Z be a biprefix, overlap-free, and right normal code over A. If $\lambda \in \mathrm{Pref}\, Z^* \setminus \{\varepsilon\}$, then there exists a unique word $u = z_1 \cdots z_k$ with $k \geq 1$ and $z_i \in Z$, $i = 1, \ldots, k$, such that*

$$u = z_1 \cdots z_k = \lambda\zeta, \quad z_1 \cdots z_{k-1}\delta = \lambda\,, \tag{21}$$

where $\delta \in A^+$ and $\zeta \in A^*$.

Proof. Let us suppose that there exist $h \geq 1$ and words $z'_1, \ldots, z'_h \in Z$ such that

$$z'_1 \cdots z'_h = \lambda\zeta', \quad z'_1 \cdots z'_{h-1}\delta' = \lambda \tag{22}$$

with $\zeta' \in A^*$ and $\delta' \in A^+$. From (21) and (22) one obtains $u = z_1 \cdots z_k = z'_1 \cdots z'_{h-1}\delta'\zeta$ and $z'_1 \cdots z'_h = z_1 \cdots z_{k-1}\delta\zeta'$, with $z_k = \delta\zeta$ and $z'_h = \delta'\zeta'$. Since Z is a biprefix code, we derive $h = k$ and consequently $z_i = z'_i$ for $i = 1, \ldots, k-1$. Indeed, if $h \neq k$, we would derive by cancellation that $\delta'\zeta = \varepsilon$ or $\delta\zeta' = \varepsilon$, which is absurd as $\delta, \delta' \in A^+$.

Hence we obtain $z_k = \delta'\zeta = \delta\zeta$, whence $\delta = \delta'$. Thus δ is a common nonempty prefix of z_k and z'_k. Since Z is right normal, by Proposition 1.1 we obtain that z_k is a prefix of z'_k and *vice versa*, i.e., $z_k = z'_k$. □

Proposition 4.4. *Let Z be a biprefix, overlap-free, and normal code over A. If $u \in Z^* \setminus \{\varepsilon\}$ is a proper factor of $z \in Z$, then there exist $p, q \in Z^*$, $h, h' \in A^+$ such that $h^\ell \notin \mathrm{Suff}\, Z$, $(h')^f \notin \mathrm{Pref}\, Z$, and*

$$z = hpuqh'\,.$$

Proof. Since u is a proper factor of $z \in Z$, there exist $\xi, \xi' \in A^*$ such that $z = \xi u \xi'$; moreover, ξ and ξ' are both nonempty as Z is a biprefix code. Let p (resp. q) be the longest word in $\mathrm{Suff}\, \xi \cap Z^*$ (resp. $\mathrm{Pref}\, \xi' \cap Z^*$), and write

$$z = \xi u \xi' = hpuqh'$$

for some $h, h' \in A^*$. Since u and hp are nonempty and Z is a biprefix code, one derives that h and h' cannot be empty. Moreover, $h^\ell \notin \operatorname{Suff} Z$ and $(h')^f \notin \operatorname{Pref} Z$, for otherwise the maximality of p and q would be contradicted using Proposition 1.1. $\square$

5 Proof of Theorem 3.13

In order to prove the theorem, we need the following lemma.

Lemma 5.1. *Let $t \in SEpi(B)$ with $\operatorname{alph} t = B$, and let $s = g(t)$ be a standard ϑ-episturmian word over A, with $g : B^* \to A^*$ an injective morphism such that $g(B) \subseteq \mathcal{P}_\vartheta$. Suppose that $b, c \in A \setminus \operatorname{Suff} \Pi_s$ and $v \in \Pi_s^*$ are such that $bv\bar{c} \in \operatorname{Fact} \Pi_s$. Then there exists $\delta \in B^*$ such that $v = g(\psi(\delta))$.*

Proof. Let $\pi \in \Pi_s$ be such that $bv\bar{c} \in \operatorname{Fact} \pi$. By definition, we have $\Pi_s = g(B)$, so that, since $v \in \Pi_s^*$, we can write $v = g(\xi)$ for some $\xi \in B^*$. We have to prove that $\xi = \psi(\delta)$ for some $\delta \in B^*$. This is trivial for $\xi = \varepsilon$. Let then $\psi(\delta')$ be the longest prefix in $\psi(B^*)$ of ξ, and assume by contradiction that $\xi \neq \psi(\delta')$, so that $\psi(\delta')a \in \operatorname{Pref} \xi$ for some $a \in B$. We shall prove that $\psi(\delta' a) = (\psi(\delta')a)^{(+)} \in \operatorname{Pref} \xi$, contradicting the maximality of $\psi(\delta')$.

Since $g(\psi(\delta'))$ is a prefix of v, we have $bg(\psi(\delta')) \in \operatorname{Fact} \pi \subseteq \operatorname{Fact} s$. Moreover $g(\psi(\delta')a) \in \operatorname{Pref} v \subseteq \operatorname{Fact} \pi$. By Proposition 1.17 and since π is a ϑ-palindrome, we have

$$g(a\psi(\delta')) = \overline{g(\psi(\delta')a)} \in \operatorname{Fact} \pi \ .$$

Thus $g(\psi(\delta'))$, being preceded in s both by $b \notin \operatorname{Suff} \Pi_s$ and by $(g(a))^\ell \in \operatorname{Suff} \Pi_s$, is a left special factor of s, and hence a prefix of it.

Suppose first that $a \notin \operatorname{alph} \delta'$, so that $\psi(\delta' a) = \psi(\delta')a\psi(\delta')$. Let λ be the longest prefix of $\psi(\delta')$ such that $\psi(\delta')a\lambda$ is a prefix of ξ. Then $g(\psi(\delta')a\lambda)$ is followed in $v\bar{c}$ by some letter x, i.e.,

$$g(\psi(\delta')a\lambda)x \in \operatorname{Pref}(v\bar{c}) \ . \tag{23}$$

We claim that

$$g(\lambda)x \notin \operatorname{Pref} g(\psi(\delta')) \ . \tag{24}$$

Indeed, assume the contrary. Then x is a prefix of $g(\lambda)^{-1}g(\psi(\delta'))$, which is in Π^* since Π is a biprefix code. Hence $x \in \operatorname{Pref} g(d)$ for some $d \in B$ such that $g(\lambda d) \in \operatorname{Pref} g(\psi(\delta'))$, and then $\lambda d \in \operatorname{Pref} \psi(\delta')$ by Lemma 1.2. As $\bar{c} \notin \operatorname{Pref} \Pi$, we obtain $x \neq \bar{c}$, so that by (23) it follows $g(\psi(\delta')a\lambda)x \in \operatorname{Pref} v$. Therefore $g(\psi(\delta')a\lambda d) \in \operatorname{Pref} v$ by Proposition 1.1, so that $\psi(\delta')a\lambda d \in \operatorname{Pref} \xi$ by Lemma 1.2. This is a contradiction because of our choice of λ.

Let us prove that $\lambda = \psi(\delta')$. Indeed, since $\tilde{\lambda} \in \operatorname{Suff} \psi(\delta')$, by (23) the word $g(\tilde{\lambda}a\lambda)x$ is a factor of π, and so is its image under ϑ, that is $\bar{x}g(\tilde{\lambda}a\lambda)$. By contradiction, suppose $|\lambda| < |\psi(\delta')|$. By (24), $\bar{x}g(\tilde{\lambda}) \notin \operatorname{Suff} g(\psi(\delta'))$, so that the suffix $g(\tilde{\lambda}a\lambda)$ of $g(\psi(\delta')a\lambda)$ is preceded by a letter which is not $\bar{x}$. Thus $g(\tilde{\lambda}a\lambda)$ is a left special factor of $\pi \in \operatorname{Fact} s$, and hence a prefix of s. As we have previously seen, $g(\psi(\delta'))$ is a prefix of s too, so that, as $|\lambda| < |\psi(\delta')|$, it follows by Lemma 1.2 that $\tilde{\lambda}a$ is a prefix of $\psi(\delta')$, contradicting the hypothesis that $a \notin \operatorname{alph} \delta'$. Thus $\lambda = \psi(\delta')$, so that $\psi(\delta' a) \in \operatorname{Pref} \xi$, as we claimed.

Now let us assume $a \in \operatorname{alph} \delta'$ instead, and write $\delta' = \gamma a \gamma'$ with $a \notin \operatorname{alph} \gamma'$, so that $\psi(\delta') = \psi(\gamma) a \rho = \tilde{\rho} a \psi(\gamma)$ and $\psi(\gamma)$ is the longest palindromic prefix (resp. suffix) of $\psi(\delta')$ followed (resp. preceded) by a. Thus

$$\psi(\delta' a) = \tilde{\rho} a \psi(\gamma) a \rho = \psi(\delta') a \rho \ .$$

Let $\lambda \in \operatorname{Pref} \rho$ and $x \in A$ be such that (23) holds and $g(\lambda) x \notin \operatorname{Pref} g(\rho)$. With the same argument as above, one can show that if $|\lambda| < |\rho|$, then $g(\tilde{\lambda} a \psi(\gamma) a \lambda)$ is a left special factor, and then a prefix, of s. Since $g(\psi(\delta'))$ is a prefix of s too, and $|\tilde{\lambda} a \psi(\gamma) a| \leq |\rho a \psi(\gamma)| = |\psi(\delta')|$, by Lemma 1.2 we obtain $\tilde{\lambda} a \psi(\gamma) a \in \operatorname{Pref} \psi(\delta')$. Since $\tilde{\lambda}$ is a suffix of $\tilde{\rho}$, $\tilde{\lambda} a \psi(\gamma)$ is a suffix, and then a border, of $\psi(\delta')$. This is absurd since $\psi(\gamma)$ is the longest border of $\psi(\delta')$ followed by a. Thus $\lambda = \rho$, showing that $\psi(\delta' a)$ is a prefix of ξ also in this case. The proof is complete. □

We can now proceed with the proof of Theorem 3.13.

5.1 Necessity

The decomposition (15) with $B = \operatorname{alph} w \cup \eta(X)$ follows from Corollary 3.9 and subsequent remark.

Since $\Pi = g(B) \subseteq \mathcal{P}_\vartheta$ and φ is ϑ-characteristic, one has by Theorem 3.5 that $\Pi = \Pi(\varphi)$ as defined by (7), so that it is overlap-free and normal by Proposition 2.2.

Let us set $u = g(\psi(w))$, and prove that condition 2 holds. We first suppose that $\operatorname{card} X \geq 2$, and that $a, a' \in \eta(X)$ are distinct letters. Let Δ be an infinite word such that $\operatorname{alph} \Delta = \eta(X)$. Setting $t_a = \psi(wa\Delta)$ and $t_{a'} = \psi(wa'\Delta)$, by (3) we have

$$t_a = \mu_w(\psi(a\Delta)) \quad \text{and} \quad t_{a'} = \mu_w(\psi(a'\Delta)) \ ,$$

so that, setting $s_y = g(t_y)$ for $y \in \{a, a'\}$, we obtain

$$s_y = g(\mu_w(\psi(y\Delta))) \in \mathit{SEpi}_\vartheta$$

as $\psi(y\Delta) \in \eta(\mathit{SEpi}(X)) \subseteq \mathit{SEpi}(B)$ and $\varphi = g \circ \mu_w \circ \eta$ is ϑ-characteristic. By Corollary 1.10 and (3), one obtains that the longest common prefix of t_a and $t_{a'}$ is $\psi(w)$. As $\operatorname{alph} \Delta = \eta(X)$ and $B = \operatorname{alph} w \cup \eta(X)$, we have $\operatorname{alph} t_a = \operatorname{alph} t_{a'} = B$, so that $\Pi_{s_a} = \Pi_{s_{a'}} = \Pi$. Since g is injective, by Theorem 1.27 we have $g(a)^f \neq g(a')^f$, so that the longest common prefix of s_a and $s_{a'}$ is $u = g(\psi(w))$. Any word of $LS(\{u\} \cup \Pi)$, being a left special factor of both s_a and $s_{a'}$, has to be a common prefix of s_a and $s_{a'}$, and hence a prefix of u.

Now let us suppose $X = \{z\}$ and denote $\eta(z)$ by a. In this case we have

$$\varphi(\mathit{SEpi}(X)) = \{g(\mu_w(a^\omega))\} = \{(g(\mu_w(a)))^\omega\} \ .$$

Let us set $s = (g(\mu_w(a)))^\omega \in \mathit{SEpi}_\vartheta$. By Corollary 1.10, $u = g(\psi(w))$ is a prefix of s. Let $\lambda \in LS(\{u\} \cup \Pi)$. Since $\Pi = \Pi_s$, the word λ is a left special factor of the ϑ-episturmian word s, so that we have $\lambda \in \operatorname{Pref} s$.

If $a \in \operatorname{alph} w$, then $B = \{a\} \cup \operatorname{alph} w = \operatorname{alph} w = \operatorname{alph} \psi(w)$, so that $\Pi \subseteq \operatorname{Fact} u$. This implies $|\lambda| \leq |u|$ and then $\lambda \in \operatorname{Pref} u$ as desired.

If $a \notin \operatorname{alph} w$, then by Proposition 3.10 we obtain $\varphi(z) = g(\mu_w(a)) = u\, g(a)$, because $\varphi(z) \notin \operatorname{Pref} u$ otherwise by Lemma 1.2 we would obtain $\mu_w(a) \in \operatorname{Pref} \psi(w)$, that implies $a \in \operatorname{alph} w$. Hence $s = (u\, g(a))^\omega$. Since $\Pi \subseteq \{g(a)\} \cup$

Fact u, we have $|\lambda| \leq |u\, g(a)|$, so that $\lambda \in \operatorname{Pref}(u\, g(a))$. Again, if λ is a proper prefix of u we are done, so let us suppose that $\lambda = u\lambda'$ for some $\lambda' \in \operatorname{Pref} g(a)$, and that λ is a left special factor of $g(a)$. Then the prefix λ' of $g(a)$ is repeated in $g(a)$. The longest repeated prefix p of $g(a)$ is either a right special factor or a border of $g(a)$. Both possibilities imply $p = \varepsilon$, since $g(a)$ is unbordered and Π is a biprefix and normal code. As $\lambda' \in \operatorname{Pref} p$, it follows $\lambda' = \varepsilon$. This proves condition 2.

Finally, let us prove condition 3. Let $b, c \in A \setminus \operatorname{Suff} \Pi$, $v \in \Pi^*$, and $\pi \in \Pi$ be such that $bv\bar{c} \in \operatorname{Fact} \pi$. Let $t' \in \mathit{SEpi}(X)$ with $\operatorname{alph} t' = X$, and set $t = \mu_w(\eta(t'))$, $s_1 = g(t)$. Since φ is ϑ-characteristic, $s_1 = \varphi(t')$ is standard ϑ-episturmian. By Lemma 5.1, we have $v = g(\psi(\delta))$ for some $\delta \in B^*$. If $\delta = \varepsilon$ we are done, as condition 3 is trivially satisfied for $w' = x = \varepsilon$; let us then write $\delta = \delta' a$ for some $a \in B$. The words $bg(\psi(\delta'))$ and $g(a\psi(\delta'))$ are both factors of the ϑ-palindrome π; indeed, $\psi(\delta' a)$ begins with $\psi(\delta')a$ and terminates with $a\psi(\delta')$. Hence $g(\psi(\delta'))$ is left special in π as $b \notin \operatorname{Suff} \Pi$ is different from $(g(a))^{\ell} \in \operatorname{Suff} \Pi$. Therefore $g(\psi(\delta'))$ is a prefix of $g(\psi(w))$, as we have already proved condition 2. Since g is injective and Π is a biprefix code, by Lemma 1.2 it follows $\psi(\delta') \in \operatorname{Pref} \psi(w)$, so that $\delta' \in \operatorname{Pref} w$ by Proposition 1.6. Hence, we can write $\delta = w'x$ with $w' \in \operatorname{Pref} w$ and x either equal to a (if $\delta' a \notin \operatorname{Pref} w$) or to ε. It remains to show that if $w'x \notin \operatorname{Pref} w$, then $x \notin \eta(X)$.

Let us first assume that $\eta(X) = \{x\}$. In this case we have $s_1 = g(\mu_w(\eta(t'))) = g(\psi(wx^{\omega}))$ by (3). Since $bv = bg(\psi(w'x)) \in \operatorname{Fact} \pi$, $g(x)$ is a proper factor of π. Then, as $B = \{x\} \cup \operatorname{alph} w$ and $g(x) \neq \pi$, we must have $\pi \in g(\operatorname{alph} w)$, so that $bv \in \operatorname{Fact} g(\psi(w))$ as $\operatorname{alph} w = \operatorname{alph} \psi(w)$. By Proposition 1.7, $\psi(w'x)$ is a factor of $\psi(wx)$. We can then write $\psi(wx) = \zeta\psi(w'x)\zeta'$ for some $\zeta, \zeta' \in B^*$. If ζ were empty, by Proposition 1.6 we obtain $w'x \in \operatorname{Pref}(wx)$. Since $w'x \notin \operatorname{Pref} w$ we would derive $w = w'$, which is a contradiction since we proved that $bv = bg(\psi(w'x)) \in \operatorname{Fact} g(\psi(w))$. Therefore $\zeta \neq \varepsilon$, and v is left special in s, being preceded both by $(g(\zeta))^{\ell}$ and by $b \notin \operatorname{Suff} \Pi$. This implies that v is a prefix of s and then of $g(\psi(w))$ as $|v| \leq |g(\psi(w))|$. By Lemma 1.2, it follows $\psi(w'x) \in \operatorname{Pref} \psi(w)$ and then $w'x \in \operatorname{Pref} w$ by Proposition 1.6, which is a contradiction.

Suppose now that there exists $y \in \eta(X) \setminus \{x\}$, and let $\Delta \in \eta(X)^{\omega}$ with $\operatorname{alph} \Delta = \eta(X)$. The word $s_2 = g(\psi(wyx\Delta))$ is equal to $g(\mu_w(\psi(yx\Delta)))$ by (3), and is then standard ϑ-episturmian since $\varphi = g \circ \mu_w \circ \eta$ is ϑ-characteristic. By applying Proposition 1.7 to w' and $wy \in w'A^*$, we obtain $\psi(w'x) \in \operatorname{Fact} \psi(wyx)$. We can write $\psi(wyx) = \zeta\psi(w'x)\zeta'$ for some $\zeta, \zeta' \in B^*$. As $w'x \notin \operatorname{Pref} w$ and $x \neq y$, we have by Proposition 1.6 that $\psi(w'x) \notin \operatorname{Pref} \psi(wy)$, so that $\zeta \neq \varepsilon$. Hence $v = g(\psi(w'x))$ is left special in s_2, being preceded both by $(g(\zeta))^{\ell}$ and by $b \notin \operatorname{Suff} \Pi$. This implies that v is a prefix of s_2 and then of $g(\psi(wy))$; by Lemma 1.2, this is absurd since $\psi(w'x) \notin \operatorname{Pref} \psi(wy)$.

5.2 Sufficiency

Let $t' \in \mathit{SEpi}(\eta(X))$ and $t = \mu_w(t') \in \mathit{SEpi}(B)$. Since $g(B) = \Pi \subseteq \mathcal{P}_{\vartheta}$, by Proposition 1.17 it follows that $g(t)$ has infinitely many ϑ-palindromic prefixes, so that it is closed under ϑ.

Thus, in order to prove that $g(t) \in \mathit{SEpi}_{\vartheta}$, it is sufficient to show that any nonempty left special factor λ of $g(t)$ is in $\operatorname{Pref} g(t)$. Since λ is left special, there exist $a, a' \in A$, $a \neq a'$, $v, v' \in A^*$, and $r, r' \in A^\omega$, such that

$$g(t) = va\lambda r = v'a'\lambda r' \; . \tag{25}$$

The word $g(t)$ can be uniquely factorized by the elements of Π. Therefore, $va\lambda$ and $v'a'\lambda$ are in $\mathrm{Pref}\,\Pi^*$. We consider three different cases.

Case 1: $va \notin \Pi^*$, $v'a' \notin \Pi^*$.

Since Π is a biprefix (as it is a subset of $\mathcal{P}_\vartheta$), overlap-free, and normal code, by Proposition 4.2 we have $a\lambda, a'\lambda \in \mathrm{Fact}\,\Pi$. Therefore, by condition 2 of Theorem 3.13, it follows $\lambda \in LS\,\Pi \subseteq \mathrm{Pref}\,g(\psi(w))$, so that it is a prefix of $g(t)$ since by Corollary 1.10, $\psi(w)$ is a prefix of $t = \mu_w(t')$.

Case 2: $va \in \Pi^*$, $v'a' \in \Pi^*$.

From (25), we have $\lambda \in \mathrm{Pref}\,\Pi^*$. By Proposition 4.3, there exists a unique word $\lambda' \in \Pi^*$ such that $\lambda' = \pi_1 \cdots \pi_k = \lambda\zeta$ and $\pi_1 \cdots \pi_{k-1}\delta = \lambda$, with $k \geq 1$, $\pi_i \in \Pi$ for $i = 1, \ldots, k$, $\delta \in A^+$, and $\zeta \in A^*$.

Since g is injective, there exist and are unique the words $\tau, \gamma, \gamma' \in B^*$ such that $g(\tau) = \lambda', g(\gamma) = va, g(\gamma') = v'a'$. Moreover, we have $g(\gamma\tau) = va\lambda' = va\lambda\zeta \in \mathrm{Pref}\,g(t)$ and $g(\gamma'\tau) = v'a'\lambda' = v'a'\lambda\zeta \in \mathrm{Pref}\,g(t)$. By Lemma 1.2, we derive $\gamma\tau, \gamma'\tau \in \mathrm{Pref}\,t$. Setting $\alpha = \gamma^\ell, \alpha' = \gamma'^\ell$, we obtain $\alpha\tau, \alpha'\tau \in \mathrm{Fact}\,t$, and $\alpha \neq \alpha'$ as $a \neq a'$. Hence τ is a left special factor of t; since $t \in SEpi(B)$, we have $\tau \in \mathrm{Pref}\,t$, so that $g(\tau) = \lambda' \in \mathrm{Pref}\,g(t)$. As λ is a prefix of λ', it follows $\lambda \in \mathrm{Pref}\,g(t)$.

Case 3: $va \notin \Pi^*$, $v'a' \in \Pi^*$ (resp. $va \in \Pi^*$, $v'a' \notin \Pi^*$).

We shall consider only the case when $va \notin \Pi^*$ and $v'a' \in \Pi^*$, as the symmetric case can be similarly dealt with.

Since $v'a' \in \Pi^*$, by (25) we have $\lambda \in \mathrm{Pref}\,\Pi^*$. By Proposition 4.3, there exists a unique word $\lambda' \in \Pi^*$ such that $\lambda' = \pi_1 \cdots \pi_k = \lambda\zeta$ and $\pi_1 \cdots \pi_{k-1}\delta = \lambda$, with $k \geq 1$, $\pi_i \in \Pi$ for $i = 1, \ldots, k$, $\delta \in A^+$, and $\zeta \in A^*$. By the uniqueness of λ', $v'a'\lambda'$ is a prefix of $g(t)$.

By (25) we have $va\pi_1 \cdots \pi_{k-1}\delta \in \mathrm{Pref}\,g(t)$. By Proposition 4.2, $a\lambda \in \mathrm{Fact}\,\Pi$, so that there exist $\xi, \xi' \in A^*$, $\pi \in \Pi$, such that

$$\xi a\lambda\xi' = \xi a\pi_1 \cdots \pi_{k-1}\delta\xi' = \pi \in \Pi \; .$$

Since δ is a nonempty prefix of π_k, it follows from Proposition 1.1 that $\pi = \xi a\pi_1 \cdots \pi_k\xi'' = \xi a\lambda'\xi''$, with $\xi'' \in A^*$. By Proposition 4.4, we can write

$$\pi = \xi a\lambda'\xi'' = hp\lambda' qh'$$

with $h, h' \in A^+$, $p, q \in \Pi^*$, $b = h^\ell \notin \mathrm{Suff}\,\Pi$, and $\bar{c} = (h')^f \notin \mathrm{Pref}\,\Pi$.

By condition 3, we have $p\lambda' q = g(\psi(w'x))$ for some $w' \in \mathrm{Pref}\,w$ and $x \in \{\varepsilon\} \cup (B \setminus \eta(X))$. Since $p, \lambda', q \in \Pi^*$ and g is injective, we derive $\lambda' = g(\tau)$ for some $\tau \in \mathrm{Fact}\,\psi(w'x)$. We will show that λ' is a prefix of $g(t)$, which proves the assertion as $\lambda \in \mathrm{Pref}\,\lambda'$.

Suppose first that $p = \varepsilon$, so that $a = b$ and $\tau \in \mathrm{Pref}\,\psi(w'x)$. If $\tau \in \mathrm{Pref}\,\psi(w')$, then $\lambda' \in g(\mathrm{Pref}\,\psi(w')) \subseteq \mathrm{Pref}\,g(\psi(w')) \subseteq \mathrm{Pref}\,g(\psi(w))$, and we are done as $g(\psi(w)) \in \operatorname{Pref} g(t)$. Let us then assume $x \neq \varepsilon$, so that $x \in B \setminus \eta(X)$, and $\psi(w')x \in \operatorname{Pref} \tau$. Moreover, we can assume $w'x \notin \operatorname{Pref} w$, for otherwise we would derive $\lambda' \in \operatorname{Pref} g(\psi(w))$ again. Let $\Delta \in \eta(X)^\omega$ be the directive word of t', so that by (3) we have $t = \psi(w\Delta)$. Since $w' \in \operatorname{Pref} w$, we can write $w\Delta = w'\Delta'$ for some $\Delta' \in B^\omega$, so that $t = \psi(w'\Delta')$.

We have already observed that $v'a'\lambda' \in \operatorname{Pref} g(t)$; as $v'a' \in \Pi^*$, by Lemma 1.2 one derives that τ is a factor of t. Since $\psi(w')x \in \operatorname{Pref} \tau$, it follows $\psi(w')x \in \operatorname{Fact} \psi(w'\Delta')$; by Proposition 1.8, we obtain $x \in \operatorname{alph} \Delta'$. This implies, since $x \notin \eta(X)$, that $w \neq w'$, and we can write $w = w'\sigma x \sigma'$ for some $\sigma, \sigma' \in B^*$. By Proposition 1.7, $\psi(w'x)$ is a factor of $\psi(w'\sigma x)$ and hence of $\psi(w)$, so that, since $\tau \in \operatorname{Pref} \psi(w'x)$, we have $\tau \in \operatorname{Fact} \psi(w)$. Hence we have either $\tau \in \operatorname{Pref} \psi(w)$, so that $\lambda' \in \operatorname{Pref} g(\psi(w))$ and we are done, or there exists a letter y such that $y\tau \in \operatorname{Fact} \psi(w)$, so that $d\lambda' \in \operatorname{Fact} g(\psi(w))$ with $d = (g(y))^\ell \in \operatorname{Suff} \Pi$. In the latter case, since $a = b \notin \operatorname{Suff} \Pi$ and $a\lambda' \in \operatorname{Fact} \Pi$, we have by condition 2 that $\lambda' \in \operatorname{Pref} g(\psi(w))$. Since $g(\psi(w))$ is a prefix of $g(t)$, in the case $p = \varepsilon$ the assertion is proved.

If $p \neq \varepsilon$, we have $a \in \operatorname{Suff} \Pi$. Let then $\alpha, \alpha' \in B$ be such that $(g(\alpha))^\ell = a$ and $(g(\alpha'))^\ell = a'$; as $a \neq a'$, we have $\alpha \neq \alpha'$. Since $p\lambda'$ is a prefix of $g(\psi(w'x))$, $p \in \Pi^*$, and $p^\ell = (g(\alpha))^\ell = a$, by Lemma 1.2 one derives that $\alpha\tau$ is a factor of $\psi(w'x)$. Moreover, as $v'a'\lambda' \in \operatorname{Pref} g(t)$ and $v'a' \in \Pi^*$, we derive that $\alpha'\tau$ is a factor of t.

Let then δ' be any prefix of the directive word Δ of t', such that $\alpha'\tau \in \operatorname{Fact} \psi(w\delta')$. By Proposition 1.7, $\psi(w\delta'x)$ contains $\psi(w'x)$, and hence $\alpha\tau$, as a factor. Thus τ is a left special factor of $\psi(w\delta'x)$ and then of the standard episturmian word $\psi(w\delta'x^\omega)$; as $|\tau| < |\psi(w\delta')|$, it follows $\tau \in \operatorname{Pref} \psi(w\delta')$ and then $\tau \in \operatorname{Pref} t$, so that $\lambda' \in \operatorname{Pref} g(t)$. The proof is now complete.

6 Further results and concluding remarks

Theorem 1.27 shows that every standard ϑ-episturmian word is a morphic image, under a suitable injective morphism, of some standard episturmian word. The following theorem improves upon this, showing that the morphism can always be taken to be ϑ-characteristic.

Theorem 6.1. *Let s be a standard ϑ-episturmian word over A. Then there exists $X \subseteq A$, $t' \in SEpi(X)$ and an injective ϑ-characteristic morphism $\varphi : X^* \to A^*$ such that $s = \varphi(t')$.*

Proof. Set $\Pi = \Pi_s$. By Theorem 1.27, the restriction to Π of the map $f : w \in \mathcal{P}_\vartheta \mapsto w^f \in A$ is injective. Hence, setting $B = f(\Pi) \subseteq A$, we can define an injective morphism g sending any letter $x \in B$ to the only word of Π beginning with x. We have $s = g(t)$, where $t = f(s) \in SEpi(B)$ by Theorem 1.27.

Let now $w \in B^*$ be the longest word such that $\psi(w) \in \operatorname{Pref} t$ and $g(\psi(w)) \in \operatorname{Fact} \Pi$. Such a word certainly exists, as $\varepsilon = \psi(\varepsilon) \in \operatorname{Pref} t$ and $\varepsilon = g(\psi(\varepsilon)) \in \operatorname{Fact} \Pi$. Since $\psi(w) \in \operatorname{Pref} t$, we can write t as $\psi(w\Delta)$ for some $\Delta \in B^\omega$; let us set

$$X = \operatorname{alph} \Delta \subseteq B \quad \text{and} \quad t' = \psi(\Delta) \in SEpi(X) .$$

By (3) we obtain $s = \varphi(t')$, where $\varphi = g \circ \mu_w \circ \eta$ and η is the inclusion map of X in B, i.e., $\eta(X) = X$.

Let us now show that φ is ϑ-characteristic. We have $B = X \cup \operatorname{alph} w$, and $g(B) = \Pi_s \subseteq \mathcal{P}_\vartheta$ is a biprefix code. By Theorems 1.25 and 1.26, Π is also normal and overlap-free, so that condition 1 of Theorem 3.13 is satisfied.

Let us first prove that φ meets condition 3 of that theorem. Indeed, if $v \in \Pi^*$ and $b, c \in A \setminus \operatorname{Suff} \Pi$ are such that $bv\bar{c} \in \operatorname{Fact} \pi$ with $\pi \in \Pi$, then by Lemma 5.1 we have $v = g(\psi(\delta))$ for some $\delta \in B^*$. If $\delta = \varepsilon$ we are done; let us then write $\delta = \delta' a$ for some $a \in B$. The words $bg(\psi(\delta'))$ and $g(a\psi(\delta'))$ are both factors of the ϑ-palindrome π, so that $g(\psi(\delta'))$ is left special in π as $b \notin \operatorname{Suff} \Pi$ is different from $(g(a))^\ell$. Therefore $g(\psi(\delta')) \in \operatorname{Pref} g(t)$, so that by Lemma 1.2 we have $\psi(\delta') \in \operatorname{Pref} t$. Since $g(\psi(\delta')) \in \operatorname{Fact} \Pi$, from the maximality condition on w it follows $|\delta'| \leq |w|$. Moreover, as $\psi(w) \in \operatorname{Pref} t$, by Proposition 1.6 it follows $\delta' \in \operatorname{Pref} w$. Hence, we can write $\delta = w'x$ with $w' \in \operatorname{Pref} w$ and x either equal to a (if $\delta' a \notin \operatorname{Pref} w$) or to ε.

In order to prove condition 3, it remains to show that if $w'x \notin \operatorname{Pref} w$, then $x \notin X$. By contradiction, assume $x \in X = \operatorname{alph} \Delta$ and write $\Delta = \xi x \Delta'$ for some $\xi \in (X \setminus \{x\})^*$ and $\Delta' \in X^\omega$. From (3), it follows $t = \psi(w\xi x\Delta')$. By applying Proposition 1.7 to w' and $w\xi \in w'B^*$, we obtain $\psi(w'x) \in \operatorname{Fact} \psi(w\xi x)$; let us write $\psi(w\xi x) = \zeta\psi(w'x)\zeta'$ for some $\zeta, \zeta' \in B^*$. We claim that $\zeta \neq \varepsilon$, i.e., $\psi(w'x) \notin \operatorname{Pref} \psi(w\xi x)$. Indeed, assume the contrary. Then $w'x \in \operatorname{Pref}(w\xi x)$ by Proposition 1.6, so that $w' = w$ and $\xi = \varepsilon$ since $w'x \notin \operatorname{Pref} w$ and $x \notin \operatorname{alph} \xi$. Thus $g(\psi(wx)) = g(\psi(\delta)) = v \in \operatorname{Fact} \Pi$ and $\psi(wx) \in \operatorname{Pref} t$, but this contradicts the maximality of w. Therefore $\zeta \neq \varepsilon$, so that $g(\psi(w'x))$ is left special in s, being preceded both by $b \notin \operatorname{Suff} \Pi$ and by $(g(\zeta))^\ell \in \operatorname{Suff} \Pi$. Hence $g(\psi(w'x))$ is a prefix of s, and then of $g(\psi(w\xi x))$. By Lemma 1.2, we obtain $\psi(w'x) \in \operatorname{Pref} \psi(w\xi x)$, a contradiction. Thus φ satisfies condition 3 of Theorem 3.13.

Finally, let $u = g(\psi(w)) \in \operatorname{Pref} s$ and let us prove that $LS(\{u\} \cup \Pi) \subseteq \operatorname{Pref} u$. Any word $\lambda \in LS(\{u\} \cup \Pi)$ is left special in s, and hence a prefix of it. If λ is a factor of u, then $|\lambda| \leq |u|$, so that $\lambda \in \operatorname{Pref} u$ as desired.

Let then $\lambda \in LS\,\Pi$, with $\lambda \neq \varepsilon$. Since $\lambda \in \operatorname{Pref} s$, we have $\lambda \in \operatorname{Pref} \Pi^*$, so that by Proposition 4.3 there exists a unique $\lambda' = \pi_1\pi_2 \cdots \pi_k \in \Pi^*$ (with $k \geq 1$ and $\pi_i \in \Pi$ for $i = 1, \ldots, k$) such that $\lambda \in \operatorname{Pref} \lambda'$ and $\pi_1 \cdots \pi_{k-1} \in \operatorname{Pref} \lambda$. Because of its uniqueness, λ' has to be a prefix of s. Moreover, as a consequence of Proposition 1.1, every occurrence of λ as a factor of any $\pi \in \Pi$ can be extended to the right to $\lambda' \in \operatorname{Fact} \pi$, so that $\lambda' \in LS\,\Pi$. As $\lambda' \in \Pi^*$, we can write $\lambda' = g(\tau) \in \operatorname{Pref} g(t)$ for some $\tau \in B^*$. By Lemma 1.2, τ is a prefix of t.

As $\lambda' \in LS\,\Pi$, it is a proper factor of some $\pi \in \Pi$. By Proposition 4.4, we can write $\pi = hp\lambda' qh'$ with $h, h' \in A^+$, $p, q \in \Pi^*$, $b = h^\ell \notin \operatorname{Suff} \Pi$, and $\bar{c} = (h')^f \notin \operatorname{Pref} \Pi$. Therefore, as we have already proved that condition 3 of Theorem 3.13 is satisfied, $p\lambda' q = g(\psi(w'x))$ for suitable $w' \in \operatorname{Pref} w$ and $x \in \{\varepsilon\} \cup (B \setminus X)$. As $p \in \Pi^*$, this implies $\tau \in \operatorname{Fact} \psi(w'x)$.

We claim that $\tau \in \operatorname{Pref} \psi(w)$, so that $\lambda \in \operatorname{Pref} \lambda'$ is a prefix of u. Indeed, suppose this is not the case, so that, since $\tau \in \operatorname{Pref} t$, one has $\psi(w)d \in \operatorname{Pref} \tau$ where d is the first letter of Δ. Then $\psi(w)d \in \operatorname{Fact} \psi(w'x)$. This is absurd if $w'x \in \operatorname{Pref} w$, as $|\psi(w)d| > |\psi(w'x)|$ in that case. If $w'x \notin \operatorname{Pref} w$, since $w' \in \operatorname{Pref} w$ we can write $w = w'yw''$ for some letter $y \neq x$ and $w'' \in B^*$. Then $\psi(w')y$ is a prefix of $\psi(w)d \in \operatorname{Fact} \psi(w'x) \subseteq \operatorname{Fact} \psi(w'x^\omega)$. As $y \notin \operatorname{alph} x^\omega$, we reach a contradiction by Proposition 1.8. Hence all conditions of Theorem 3.13 are met, so that φ is ϑ-characteristic. □

Let us consider the family $SW_\vartheta(N)$, introduced in [4], of all words $w \in A^\omega$

which are closed under ϑ and such that every left special factor of w whose length is at least N is a prefix of w. Moreover, SW_ϑ will denote the class of words which are in $SW_\vartheta(N)$ for some $N \geq 0$. One has that $SW_\vartheta(0) = SEpi_\vartheta$. It has been proved in [4] that the family of ϑ-standard words is included in $SW_\vartheta(3)$, and that SW_ϑ coincides with the family of *ϑ-standard words with seed* introduced in [8, 5].

Proposition 6.2. *Let $\varphi : X^* \to A^*$ be an injective morphism decomposable as $\varphi = g \circ \mu_w \circ \eta$ where $w \in B^*$, $B = \mathrm{alph}\, w \cup \eta(X)$, η a literal morphism, and g is an injective morphism such that $g(B) = \Pi \subseteq \mathcal{P}_\vartheta$. If Π is overlap-free and normal, then $\varphi(SEpi(X)) \subseteq SW_\vartheta(N)$ with $N = \max\{|\pi| \mid \pi \in \Pi\}$.*

Proof. The proof is very similar to the sufficiency of Theorem 3.13 (see Section 5.2). Using the same notation, suppose that λ is a left special factor of $g(t)$ of length $|\lambda| \geq N$ where $t = \mu_w(t') \in SEpi(B)$ and $t' \in SEpi(\eta(X))$. One has that Cases 1 and 3 cannot occur since otherwise one would derive $a\lambda \in \mathrm{Fact}\,\Pi$ that implies $|\lambda| < N$, which is a contradiction. It remains to consider Case 2. By using exactly the same argument one obtains that λ is a prefix of $g(t)$. Finally, since $g(t)$ has infinitely many ϑ-palindromic prefixes one has that $g(t)$ is closed under ϑ. □

In the previous sections we have introduced and studied ϑ-characteristic morphisms and their strict link with normal and overlap-free codes, especially in the biprefix case. Many interesting properties have been proved; in particular, the characterization of injective ϑ-characteristic morphisms given by Theorem 3.13 is a powerful tool for constructing standard ϑ-episturmian words.

Some natural problems could be the subject of further investigation. A first problem is to give a characterization of the endomorphisms of A^* such that $\varphi(SEpi_\vartheta) \subseteq SEpi_\vartheta$. A second, quite general problem is to characterize the injective morphisms $\varphi : X^* \to A^*$ such that $\varphi(X) \subseteq Z^*$, where Z is a biprefix, overlap-free, and normal code, with the condition that if $t \in X^\omega$ is such that any its left special factor is a prefix of t, then $\varphi(t) \in A^\omega$ satisfies the same property. Theorem 3.13 gives a characterization of these morphisms in the special case $Z \subseteq \mathcal{P}_\vartheta$ and t closed under reversal.

Finally, we think that the classes of codes considered here (i.e., normal and overlap-free codes, both in the biprefix and general case) and their combinatorial properties would deserve a deeper analysis.

References

[1] J. Berstel and D. Perrin. *Theory of Codes.* Academic Press, New York, 1985.

[2] J. Berstel and P. Séébold. Sturmian words. In M. Lothaire, editor, *Algebraic Combinatorics on Words.* Cambridge University Press, Cambridge UK, 2002. Chapter 2.

[3] M. Bucci, A. de Luca, and A. De Luca. On a generalization of standard episturmian morphisms. In M. Ito and M. Toyama, editors, *DLT 2008*, volume 5257 of *Lecture Notes in Computer Science*, pages 158–169. Springer, 2008.

[4] M. Bucci, A. de Luca, A. De Luca, and L. Q. Zamboni. On different generalizations of episturmian words. *Theoretical Computer Science*, 393:23–36, 2008.

[5] M. Bucci, A. de Luca, A. De Luca, and L. Q. Zamboni. On some problems related to palindrome closure. *Theoretical Informatics and Applications*, 42:679–700, 2008.

[6] M. Bucci, A. de Luca, A. De Luca, and L. Q. Zamboni. On θ-episturmian words. *European Journal of Combinatorics*, to appear, 2008.

[7] A. de Luca. Sturmian words: structure, combinatorics, and their arithmetics. *Theoretical Computer Science*, 183:45–82, 1997.

[8] A. de Luca and A. De Luca. Pseudopalindrome closure operators in free monoids. *Theoretical Computer Science*, 362:282–300, 2006.

[9] X. Droubay, J. Justin, and G. Pirillo. Episturmian words and some constructions of de Luca and Rauzy. *Theoretical Computer Science*, 255:539–553, 2001.

[10] S. Fischler. Palindromic prefixes and episturmian words. *Journal of Combinatorial Theory, Series A*, 113:1281–1304, 2006.

[11] J. Justin. Episturmian morphisms and a Galois theorem on continued fractions. *Theoretical Informatics and Applications*, 39:207–215, 2005.

[12] J. Justin and G. Pirillo. Episturmian words and episturmian morphisms. *Theoretical Computer Science*, 276:281–313, 2002.

[13] M. Lothaire. *Combinatorics on Words.* Addison-Wesley, Reading MA, 1983.

[14] M. Lothaire. *Applied Combinatorics on Words.* Cambridge University Press, Cambridge UK, 2005.