INFORMATIQUE THÉORIQUE ET APPLICATIONS

DO LONG VAN
BERTRAND LE SAËC
IGOR LITOVSKY

On coding morphisms for zigzag codes

Informatique théorique et applications, tome 26, nº 6 (1992), p. 565-580
<http://www.numdam.org/item?id=ITA_1992__26_6_565_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 26, n° 6, 1992, p. 565 à 580)

ON CODING MORPHISMS FOR ZIGZAG CODES (*)

by Do LONG VAN (1), Bertrand LE SAËC (1) and Igor LITOVSKY (1)

Communicated by J. BERSTEL

Abstract. – *We are dealing with the zigzag codes in connection with monoid morphisms and free group morphisms. The injectivity of free group morphisms plays here the role of that of free monoid morphisms for ordinary codes. Subsets of free group words describing behaviours of zigzag factorizations allow us to characterize zigzag codes in different ways. Every two-element code is a zigzag code. The free monoid morphisms preserving the property of being a zigzag code are exactly the coding morphisms for the biprefix codes.*

Résumé. – *Les zigzag codes sont étudiés par le biais des morphismes de monoïdes et de groupes libres. L'injectivité des morphismes de groupes libres joue ici le rôle de celle des morphismes de monoïdes pour les codes. Les calculs décrits dans le groupe libre associé aux zigzag factorisations, nous permettent de caractériser de différentes manières les zigzag codes. Les codes à deux mots sont des zigzag codes. Les morphismes de monoïdes libres préservant la propriété d'être un zigzag code sont exactement les morphismes de codages pour les codes bipréfixes.*

1. INTRODUCTION

As well known, a language X over an alphabet A is a code if every word in A^* has at most one factorization on X. The notion of zigzag factorization due to M. Anselmo [1], which consists in allowing backward steps in factorizing a word, is in some sense a two-way version of factorisation. This led naturally to the notion of zigzag operation, denoted by $\uparrow$, on a language and also to the notion of zigzag code. These notions are the main subject of [1, 2, 7, 8, 9].

In this paper, we are studying the zigzag codes in connection with monoid morphisms and free group morphisms induced by the first ones. The central point is an axiomatic definition of zigzag factorizations.

(*) Received June 1991, accepted December 1991.
(1) Laboratoire Bordelais de Recherche en Informatique, U.R.A. au C.N.R.S. n° 1304, Université Bordeaux-I, 351, cours de la Libération, 33405, Talence Cedex, France.

For an ordinary code X, any bijection φ from an alphabet B onto X induces an isomorphism from B^* onto X^*. Furthermore every word in B^* describes a unique factorization on X of some word. For the zigzag codes, the zigzag factorizations on X may not be encoded in B^*. It is necessary to work in the free group generated by B. So letters in $\bar{B}$ encode the backward steps in the zigzag factorizations. This approach of the notion of zigzag factorizations allows us to characterize zigzag codes in different ways.

The zigzag factorizations on X are encoded by zigzag sequences which are words in $(B \cup \bar{B})^*$. The set of zigzag sequences $M_\varphi(B)$ is a very pure submonoid of $(B \cup \bar{B})^*$. This set is used to characterize the zigzag codes. Like for the ordinary codes, X is a zigzag code iff $M_\varphi(B)$ is in an one-to-one fashion with $X^\dagger$. If the injectivity of free monoid morphism is a characteristic property for ordinary codes then the injectivity of the corresponding free group morphism is characteristic for zigzag codes. By using free group words, we obtain different characterizations of zigzag codes. Hence, we prove that any two-element code is a zigzag code and we propose a test for zigzag codes inspired by Sardinas, Patterson's one. Finally, we prove that the class of injective free monoid morphisms preserving the property of being a zigzag code is much smaller than that of those preserving the property of being a code. It is just the coding morphisms for the biprefix codes.

Section 2 contains the preliminaries. In the next section, we define the zigzag decompositions and the zigzag factorizations. In Section 4, we present results concerning the coding morphisms for zigzag codes. In Section 5, we characterize the zigzag codes in different ways. Section 6 is devoted to the injective free monoid morphisms that preserve the zigzag codes.

2. PRELIMINARIES

Let A be an alphabet. As usual A^* is the free monoid of all finite words over A. The empty word is denoted by ε and $A^+ = A^* \setminus \{\varepsilon\}$. Let $(u, v) \in \Sigma^* \times \Sigma^*$, uv denotes the concatenation of the words u and v. We denote by $|u|$ the length of the word u. The notations $u \leqq v$ or $u < v$ mean that u is a prefix or a proper prefix of v, respectively. A language is a subset of A^*. If X is a language then X^* is the submonoid of A^* generated by X. We denote by $\mathrm{Root}(X^*)$ the language $(X^* \setminus \{\varepsilon\}) \setminus (X^* \setminus \{\varepsilon\})^2$.

A language $X \subseteq A^*$ is a code if every word w in A^* has at most one factorization on X. If X is a code then X^* is a free submonoid od A^*. A submonoid M of A^* is very pure if for all $u, v \in A^*$, $uv \in M$ and $vu \in M$ imply

$u \in M$ and $v \in M$. As in [5], we that a submonoid M of A^* satisfies the condition $C(1, 1)$ if for all $u, v, w \in A^*$, $uv, vw \in M$ implies $v \in M$. Any submonoid satisfying the condition $C(1, 1)$ is very pure and any very pure submonoid is free [5].

Let A be an alphabet and $\bar{A}$ be another alphabet in bijection with A. For every $a \in (A \cup \bar{A})$, we denote by $\bar{a}$ the corresponding element of a. If $x = a_1 \ldots a_n$ is a word in $(A \cup \bar{A})^*$ then $\bar{x} = \bar{a}_n \ldots \bar{a}_1$. For every subset X of $(A \cup \bar{A})^*$, we define $\bar{X} = \{ \bar{x} / x \in X \}$. In the sequel, one used to consider X as an alphabet (not necessarily finite) and so one can speak about the free monoid $(X \cup \bar{X})^*$ generated by $X \cup \bar{X}$. Let us note that the notation $*$ is ambiguous. Indeed $(X \cup \bar{X})^*$ may be considered as the free monoid generated by the alphabet $(X \cup \bar{X})$ and also as the submonoid of $(A \cup \bar{A})^*$ generated by the set $(X \cup \bar{X}) \subseteq (A \cup \bar{A})^*$. (the two previous monoids are isomorphic iff $(X \cup \bar{X})$ is a code in $(A \cup \bar{A})^*$). In this paper, we use $(X \cup \bar{X})^*$ in the first sense only. However, for all $(x_1, \ldots, x_n) \in (X \cup \bar{X})^n$, we use ambiguously $x_1 \ldots x_n$ to denote a word on the alphabet $(X \cup \bar{X})$ as well a word on the alphabet $(A \cup \bar{A})$ depending on the context.

We denote by $\approx_X$ the congruence over $(X \cup \bar{X})^*$ generated by: $\forall x \in X$, $x\bar{x} \approx_X \varepsilon$ and $\bar{x}x \approx_X \varepsilon$. The notation $u_{\approx_X}$ is used for the class of the word $u \in (X \cup \bar{X})^*$. We denote by $\overset{X}{\mapsto}$ the relation defined by $\forall (u, v) \in (X \cup \bar{X})^* \times (X \cup \bar{X})^*$, $u \overset{X}{\mapsto} v$ if $u = f\alpha g$, $v = fg$ and $\alpha = x\bar{x}$ or $\alpha = \bar{x}x$ for some $x \in X$. We call X-reduction the reflexive transitive closure of $\overset{X}{\mapsto}$. The X-reduction is confluent [4]. We denote by Red_X the function which associates with each $w \in (X \cup \bar{X})^*$, the unique X-reduced word $\mathrm{Red}_X(w)$.

3. Z-DECOMPOSITIONS AND Z-FACTORIZATIONS

In this section, we present in another form, the notions of zigzag operation and of zigzag factorization introduced by M. Anselmo in [1].

DEFINITION 3.1: Let X be a language in A^*.

A zigzag decomposition (z-decomposition) on X is a word $x_1 \ldots x_n$ in $(X \cup \bar{X})^+$ that satisfies the following two properties:

- P_1: $\forall i \in \{1, \ldots, n\}$, $\mathrm{Red}_A(x_1 \ldots x_i) \in A^+$.
- P_2: $\forall i \in \{1, \ldots, n-1\}$, $\mathrm{Red}_A(x_1 \ldots x_i) < \mathrm{Red}_A(x_1 \ldots x_n)$.

We denote by $Z\text{-Dec}(X)$ the set of all z-decompositions on X together with the empty word.

Remark: Note that the property P_1 implies that $x_1 \in X$ and the property P_2 implies that $x_n \in X$.

DEFINITION 3.2: Let X be a language in A^*.

A zigzag factorization (z-factorization) on X is a z-decomposition $x_1 \ldots x_n$ satisfying the following property:

- P_3: $\forall\, 1 \leqq i < j \leqq n$, $\text{Red}_A(x_1 \ldots x_i) \neq \text{Red}_A(x_1 \ldots x_j)$.

We denote by $Z\text{-Fac}(X)$ the set of z-factorizations on X together with the empty word.

Every such z-factorization f is, exactly, a z-factorization of the word $\text{Red}_A(d)$ in the sense of M. Anselmo where each x_i corresponds to a z-step. In the sequel, we use the notation f_w to mean that f is a z-factorization of w.

DEFINITION 3.3: Let X be a language in A^*.

A zigzag calculus, $c(w_1, u, w_2)$, (z-calculus) on X of a word u with context $(w_1, w_2) \in A^* \times A^*$ is a word $x_1 \ldots x_n$ in $(X \cup \bar{X})^*$ that satisfies:

1. $\text{Red}_A(x_1 \ldots x_n) = u$
2. $\forall\, 1 \leqq i \leqq n$, $\text{Red}_A(w_1 x_1 \ldots x_i) \in A^*$
3. $\forall\, 1 \leqq i \leqq n$, $\text{Red}_A(w_1 x_1 \ldots x_i) \leqq w_1 u w_2$.

FACT 3.4: *We have:*

- *Any z-factorization is X-reduced.*
- $Z\text{-Fac}(X) \subseteq \text{Red}_X(Z\text{-Dec}(X))$.

The previous inclusion is strict as it shown in the following example:

Example 3.5: Let $A = \{a, b\}$ be an alphabet and $X = \{a, ab, ba\}$. The word $w = aabaa$ admits two z-factorizations $f_1 = (a)(a)(ba)(a)$ and the z-decomposition $d = (a)(a)(ba)(\bar{a})(\overline{ab})(a)(ba)(a)$ of w is not a z-factorization, but $\text{Red}_X(d) = d$.

DEFINITION 3.6: We denote by $\approx_l$ the congruence on $(X \cup \bar{X})^*$ generated by the relations: For any z-calculus $x_1 \ldots x_n$ of ε, $x_1 \ldots x_n \approx_l \varepsilon$. We denote by $\overset{l}{\mapsto}$ the relation on $(X \cup \bar{X})^*$ defined by $u \overset{l}{\mapsto} v$ if $u = f \alpha g$, $v = fg$ and $\alpha \approx_l \varepsilon$. We call l-reduction the reflexive transitive closure of $\overset{l}{\mapsto}$. We denote by Red_l the function which associates with each word $w \in (X \cup \bar{X})^*$ the subset of l-reduced words obtained from w.

FACT 3.7: *Let* $d \in Z$-Dec(X). *We have:*

1. $d_{\approx_X} \subseteq d_{\approx_l}$.
2. *If* d *is* l-reduced then d is X-reduced.

Remarks: The converse of 2. is generally false (*see* Corollary 5.4). The l-reduction is notherian but not confluent (*see* Proposition 5.1).

FACT 3.8: $\mathrm{Red}_l(Z\text{-Dec}(X)) = Z\text{-Fac}(X)$.

FACT 3.9: *Let* $(d, d') \in Z\text{-Dec}(X) \times Z\text{-Dec}(X)$ *and* $\alpha \in \{A, X, l\}$.

If $d \approx_\alpha d'$ *then* d *and* d' *are two* z*-decompositions of the same word* $\mathrm{Red}_A(d)$.

DEFINITION 3.10: Let X be a language in A^*.

We call zigzag on X the set $X^\uparrow = \{w \in A^* : \exists f_w \in Z\text{-Fac}(X)\}$. Note that $X^\uparrow$ is a submonoid of A^* which contains X^*. A language L of the form $X^\uparrow$ is called a z-submonoid of A^*.

LEMMA 3.11 [8]: *Let* L *be a language in* A^*. *The following properties are equivalent:*

1. $L = L^\uparrow$.
2. $\varepsilon \in L$, *and* $(uv, v, vw) \in L \times L \times L$ *implies* $uvw \in L$.

Proof: If $L = L^\uparrow$ then the second point is obviously true. Conversely, let m be a shortest word in $L^\uparrow \setminus L$. The word m has a z-factorization $x_1 \ldots x_n$ on L with $n > 1$ otherwise $m = x_1 \in L$. Let i_1 be the index of the longest word u_1 of the set $\{\mathrm{Red}_A(x_1 \ldots x_i), 1 \leq i < n\}$. The word u_1 belongs to $L^\uparrow$. Let i_2 be the index of the shortest word u_2 of the set $\{\mathrm{Red}_A(x_1 \ldots x_i), i_1 \leq i < n\}$. Since $\bar{x}_{i_2} \ldots \bar{x}_{i_1+1}$ is a z-factorization, $u_2^{-1} u_1$ is in $L^\uparrow$. Now, the word $u_3 = u_2^{-1} m$ also belongs to $L^\uparrow \setminus \{\varepsilon\}$ since $x_{i_2+1} \ldots x_n$ is a z-factorization of u_3. Thus by the minimality of m, the words u_1, u_2, u_3 belong to L. Thus, by setting $u = u_2$, $v = u_2^{-1} u_1$ and $w = u_1^{-1} m$, we have $(uv, v, vw) \in L \times L \times L$ whereas $m = uvw \notin L$. So 2. implies 1. □

DEFINITION 3.12 [1]: Let X be a language in A^*.

X is a z-code iff $\forall (f_1, f_2) \in Z\text{-Fac}(X) \times Z\text{-Fac}(X)$: $f_1 \approx_A f_2 \Rightarrow f_1 = f_2$.

DEFINITION 3.13: Let L be a z-submonoid of A^*. We call Z-Root of L the language $Z\text{-Root}(L) = \{w \in L \setminus \{\varepsilon\} : w$ has no z-factorization on L of length $> 1\}$. L is said to be z-free if Z-Root(L) is a z-code.

4. CODING MORPHISMS FOR ZIGZAG CODES

Let A and B be two alphabets and let $\varphi : B^* \to A^*$ be a monoid morphism. In the sequel, φ is extended to: $\varphi : (B \cup \bar{B}) \to (A \cup \bar{A})^*$ by setting $\forall \bar{b} \in \bar{B}, \varphi(\bar{b}) = \overline{\varphi(b)}$ and we denote by $\bar{\varphi}$: $(B \cup \bar{B})^*/_{\approx_B} \to (A \cup \bar{A})^*/_{\approx_A}$ the free group morphism induced by φ. For the sake of simplicity instead of $\bar{\varphi}(x_{\approx_B})$, we write $\bar{\varphi}(x)$. Also, we always understand $\bar{\varphi}(x)$ as the reduced representative of its class, but not the class itself.

DEFINITION 4.1: We call quasi-zigzag sequence with respect to φ, any word $b_1 \ldots b_n$ in $(B \cup \bar{B})^*$ satisfying the following conditions:

- $P'_1 : \forall i \in \{1, \ldots, n\}, \bar{\varphi}(b_1 \ldots b_i) \in A^+$
- $P'_2 : \forall i \in \{1, \ldots, n-1\}, \bar{\varphi}(b_1 \ldots b_i) < \bar{\varphi}(b_1 \ldots b_n)$.

Remark: Note that the property P'_1 implies that b_1 belongs to B and the property P'_2 implies that b_n also belongs to B.

The set of these words together with the empty word is denoted by $Z_\varphi(B)$.

DEFINITION 4.2: We call zigzag sequence with respect to φ any quasi-zigzag sequence satisfying:

- $P'_3 : \forall 1 \leqq i < j \leqq n, \bar{\varphi}(b_1 \ldots b_i) \neq \bar{\varphi}(b_1 \ldots b_j)$.

The set of these words together with the empty word is denoted by $M_\varphi(B)$. Clearly, $M_\varphi(B) \subseteq Z_\varphi(B)$ and $M_\varphi(B) \subseteq \mathrm{Red}_B(Z_\varphi(B))$.

Note: In the sequel, we will use the notations Z_φ *and* M_φ instead of $Z_\varphi(B)$ and $M_\varphi(B)$, if there is no ambiguity on the used alphabet.

LEMMA 4.3: *If $b_1 \ldots b_n$ is a quasi-zigzag sequence then we have* $\forall i \in \{1, \ldots, n\}, \bar{\varphi}(b_i \ldots b_n) \in A^+$.

Proof: We prove that, if $u \in A^+$ and $\mathrm{Red}_A(uv) \in A^+$ with $u < \mathrm{Red}_A(uv)$ and v is an A-reduced word, we have $v \in A^+$. We can write $v = \bar{v}_1 v_2$ with $v_1 \in A^*$ and $v_2 \notin \bar{A}(A \cup \bar{A})^*$. Thus $\mathrm{Red}_A(uv) \in A^+$ *implies* $u = u_1 v_1$ and then $\mathrm{Red}_A(uv) = u_1 v_2$. As $v_2 \notin \bar{A}(A \cup \bar{A})^*$, $u_1 v_2 \in A^+$ implies $v_2 \in A^*$. Since $u < \mathrm{Red}_A(uv)$ that is $u_1 v_1 < u_1 v_2$, we obtain $v_1 < v_2$. Now, since $v = \bar{v}_1 v_2$ is A-reduced, we have $v_1 = \varepsilon$. Hence $v = v_2 \in A^*$ and $v \neq \varepsilon$ since $u < \mathrm{Red}_A(uv)$. Thus, if $b_1 \ldots b_n$ is a quasi-zigzag sequence in Z_φ, we have $\forall 1 \leqq i \leqq n$, $\bar{\varphi}(b_i \ldots b_n) \in A^+$. □

Remark: If φ is injective and $\varphi(B) = X$ then the zigzag sequences and the quasi-zigzag sequences encode in $(B \cup \bar{B})^*$ the z-factorizations and the z-decompositions on X, respectively: Z-Fac(X) is in bijection with M_φ and Z-Dec(X) is in bijection with Z_φ.

Example 4.4:

1. For any non erasing morphism $\varphi : B^* \to A^*$ we have $B^* \subseteq M_\varphi$.

2. Let $A=\{a\}$ and $B=\{x, y\}$. Let $\varphi : B^* \to A^*$ be the morphism given by $\varphi(x)=a$ and $\varphi(y)=aa$. Thus $M_\varphi=\{x, y, y\bar{x}y\}^*$. The word $w=xxx\bar{y}xxx$ satisfies the condition P'_1, P'_2 and so belongs to Z_φ. Since $\bar{\varphi}(x)=\bar{\varphi}(xxx\bar{y})=a$, so w does not satisfy the condition P'_3 and w does not belong to M_φ.

PROPOSITION 4.5: *M_φ and Z_φ are very pure submonoids of $(B \cup \bar{B})^*$ and consequently are free.*

Proof: Clearly from the definitions, M_φ and Z_φ are submonoids. We prove that M_φ satisfies the condition $C(1, 1)$ recalled in Section 2. Let u, v, w be three words in $(B \cup \bar{B})^*$ such that uv and vw belong to M_φ. We can write $u=b_1 \dots b_i$, $v=b_{i+1} \dots b_j$, $w=b_{j+1} \dots b_n$ with all b_k in $(B \cup \bar{B})$. From $uv \in M_\varphi$, it follows that $b_j \in B$ and P'_2 is satisfied for v. Since $vw \in M_\varphi$, we have $b_{i+1} \in B$ and also P'_1,P'_3 are satisfied for v. Hence M_φ satisfies the condition $C(1, 1)$ so it is very pure and free. The same argument can be applied to Z_φ. □

As M_φ and Z_φ are languages in $(B \cup \bar{B})^*$, we can consider $(M_\varphi)^\uparrow$ *and* $(Z_\varphi)^\uparrow$. The following proposition shows that M_φ is not a z-submonoid of $(B \cup \bar{B})^*$ whereas so is Z_φ.

PROPOSITION 4.6: *For any morphism $\varphi : B^* \to A^*$ we have:*

1. $Z_\varphi=(Z_\varphi)^\uparrow$.

2. $M_\varphi \subseteq (M_\varphi)^\uparrow \subseteq Z_\varphi$.

Proof: We prove that $Z_\varphi=(Z_\varphi)^\uparrow$ by using Lemma 3.11.

Let $uv=b_1 \dots b_i \dots b_j$, $v=b_i \dots b_j$ and $vw=b_i \dots b_j \dots b_n$ be three words in Z_φ. If $u=\varepsilon$ or $w=\varepsilon$, then $uvw \in Z_\varphi$. Assume $u \neq \varepsilon$ *and* $w \neq \varepsilon$. The condition P'_1, is trivialy satisfied for uvw. Let $k \in \{1, \dots, n\}$. If $k<i$, $\bar{\varphi}(b_1 \dots b_k) < \bar{\varphi}(uv)$. On the other hand by the Lemma 4.3, $\bar{\varphi}(w) \in A^+$. Thus $\bar{\varphi}(b_1 \dots b_k) < \bar{\varphi}(uvw)$ for $k<i$. Now, if $i \leq k \leq n$, $\bar{\varphi}(b_i \dots b_k) < \bar{\varphi}(vw)$. Since $\bar{\varphi}(u) \in A^+$ we have $\bar{\varphi}(b_1 \dots b_k) < \bar{\varphi}(uvw)$ for $i \leq k \leq n$ and so P'_2 holds for uvw. By Lemma 3.11, $Z_\varphi=(Z_\varphi)^\uparrow$. Now, from $Z_\varphi=(Z_\varphi)^\uparrow$, the inclusions of the second point are immediate. □

The previous inclusions are strict as shown in the following example.

Example 4.7: Let $B=\{a, b, c, d, e, f\}$ and $A=\{x, y, z, t, u\}$. Let $\varphi : B^* \to A^*$ be a morphism given by $\varphi(a)=xyz$, $\varphi(b)=z$, $\varphi(c)=zt$, $\varphi(d)=yzt$, $\varphi(e)=y$, $\varphi(f)=ztu$. It is easy to verify that the words $a\bar{b}\bar{e}d$, d,

$d\bar{c}f$ are in M_φ. Thus the word $w=a\bar{b}\bar{e}d\bar{c}f$ is in $(M_\varphi)^\uparrow$. But w is not in M_φ. because we have $\bar{\varphi}(a\bar{b}\bar{e}d\bar{c})=\bar{\varphi}(a\bar{b})=xyz$. That is the condition P'_3 does not hold for w. Thus $(M_\varphi)^\uparrow \neq M_\varphi$. Now consider $w=a\bar{b}c\bar{d}ef$, one can check that w is in Z_φ and that w does not belong to $(M_\varphi)^\uparrow$. That is $(M_\varphi)^\uparrow \neq Z_\varphi$. □

PROPOSITION 4.8:

1. *If $X \subseteq A^*$ is a z-code and $\varphi : B^* \to A^*$ is a morphism injective on B with $\varphi(B)=X$, then $\bar{\varphi}$ is injective on M_φ.*
2. *If $\varphi : B^* \to A^*$ is a morphism such that $\bar{\varphi}$ is injective on M_φ then $\varphi(B)=X$ is a z-code.*

Proof: 1. Assume that α and β are two elements of M_φ such that $\bar{\varphi}(\alpha)=\bar{\varphi}(\beta)=w$. If one of them is ε then so is the other one. Assume now that $\alpha \neq \varepsilon$ and $\beta \neq \varepsilon$ and $\alpha=b_1 \ldots b_m$ and $\beta=c_1 \ldots c_n$. We have two z-factorizations of $w: f_1=x_1 \ldots x_m$ and $f_2=y_1 \ldots y_n$ such that for all $i \in \{1, \ldots, m\}$, $\varphi(b_i)=x_i$ and for all $i \in \{1, \ldots, n\}$, $\varphi(c_i)=y_i$. Since X is a z-code, these z-factorizations are identical. One has $m=n$ and for all $i \in \{1, \ldots, n\}$, $x_i=y_i$ which implies $b_i=c_i$. Thus φ is injective on M_φ.

2. Assume that a word w has two z-factorizations f and g on X such that $f=x_1 \ldots x_m$ and $g=y_1 \ldots y_n$. There exist $\alpha=b_1 \ldots b_m \in M_\varphi$ and $\beta=c_1 \ldots c_n \in M_\varphi$ such that $\bar{\varphi}(\alpha)=\bar{\varphi}(\beta)=w$ and for all $i \in \{1, \ldots, m\}$, $\varphi(b_i)=x_i$ and for all $i \in \{1, \ldots, n\}$, $\varphi(c_i)=y_i$. By the injectivity of φ, one has $\alpha=\beta$ hence $n=m$ and for all $i \in \{1, \ldots, n\}$, $x_i=y_i$ that is $f=g$ thus X is a z-code. □

So we have the following result which generalizes the notion of coding morphism on codes [5].

COROLLARY 4.9: *Let* $\varphi : B^* \to A^*$ be a morphism injective on B with $\varphi(B)=X$.

Then X is a z-code iff $\bar{\varphi}$ is injective on M_φ.

We will characterize the z-free submonoids by means of M_φ, but before, we need the following fact deduced immediatly from the definitions and Proposition 4.8:

FACT 4.10: *Let $\varphi : B^* \to A^*$ be a morphism with $\varphi(B)=X$. Then we have:*

1. $\bar{\varphi}(Z_\varphi)=\bar{\varphi}(M_\varphi)=X^\uparrow$.
2. *If φ is injective on B and X is a z-code in A^* then M_φ and $X^\uparrow$ are isomorphic by $\bar{\varphi}$.*

The z-free submonoids can be now characterized as follows:

THEOREM 4.11: *Let M be a language of A^*. The following properties are equivalent:*

1. *M is a z-free submonoid of A^*.*

2. *There exist and alphabet B and a morphism $\varphi : B^* \to A^*$ such that M_φ and M are isomorphic by $\bar{\varphi}$.*

Proof:

$1 \Rightarrow 2$. Let M be a z-submonoid of A^*. Then $M = X^\uparrow$ for some z-code X. Choose an alphabet B in bijection φ with X and $\varphi(B) = X$. Then φ induces a morphism from B^* into A^*. By Fact 4.10(1), $\bar{\varphi}(M_\varphi) = X^\uparrow$. By Proposition 4.8 (1), $\bar{\varphi}$ is injective on M_φ. Thus M_φ and M are isomorphic by $\bar{\varphi}$.

$2 \Rightarrow 1$. Let $\varphi : B^* \to A^*$ be a morphism such that M_φ and M are isomorphic by $\bar{\varphi}$. We have $\bar{\varphi}(M_\varphi) = M$. Set now $X = \bar{\varphi}(B)$. By Fact 4.10(1), we have $\bar{\varphi}(M_\varphi) = X^\uparrow$ and therefore M_φ and $X^\uparrow$ are isomorphic by $\bar{\varphi}$. By Proposition 4.8(2), X is a z-code thus $M = X^\uparrow$ is a z-free submonoid.

COROLLARY 4.12: *Every z-free submonoid is free.*

Proof: Let M be a z-free submonoid of A^*. By the previous theorem, there exist an alphabet B and a morphism $\varphi : B^* \to A^*$ such that M_φ and M are isomorphic by $\bar{\varphi}$. By Proposition 4.5, M_φ is free, thus so is M. □

In other words, Corollary 4.12 sounds as follows: for any z-code X there exists a code Y such that $X^\uparrow = Y^*$. The relationship between X and Y is given by:

PROPOSITION 4.13: *Let $X \subseteq A^*$ be a z-code and $\varphi : B^* \to A^*$ be an injective morphism such that $\varphi(B) = X$. Then $Y = \bar{\varphi}(\text{Root}(M_\varphi))$ is a code and $X^\uparrow = Y^*$.*

Proof: By virtue of Proposition 4.5, M_φ is free. Therefore $\text{Root}(M_\varphi)$ is a code in $(B \cup \bar{B})^*$. Let $\psi : C^* \to (B \cup \bar{B})^*$ be a coding morphism for $\text{Root}(M_\varphi)$. Then ψ is injective on C^* and $\psi(C^*) = (\text{Root}(M_\varphi))^* = M_\varphi$. By Proposition 4.8(1), $\bar{\varphi}$ is injective on M_φ. Therefore $\bar{\varphi} \circ \psi$ is an injective morphism from C^* into A^* with $\bar{\varphi} \circ \psi(C) = Y$. Hence Y is a code on A. Now, by Fact 4.10(2), it follows that

$$X^\uparrow = \bar{\varphi}(M_\varphi) = \bar{\varphi}(\text{Root}(M_\varphi)^*) = [\bar{\varphi}(\text{Root}(M_\varphi))]^* = Y^*$$

which completes the proof. □

Examples 4.14: Let $A = \{a, b\}$ and $B = \{x, y\}$, $X = \{a, aba\}$. Let $\varphi : B^* \to A^*$ be the morphism defined by $\varphi(x) = a$ and $\varphi(y) = aba$. It is easy to see that $\mathrm{Root}(M_\varphi)=\{x\}\cup y(\bar{x}y)^*$. Hence $Y=a(ba)^*$. Since X is a z-code, Y is a code and $X^\dagger=Y^*$.

5. CHARACTERIZATIONS OF ZIGZAG CODES

PROPOSITION 5.1: *Let X be a language in A^*.*

X is a Z-code iff Red_l *is confluent on* Z-Dec(X).

Proof: Assume that X is a Z-code and let $d\in Z$-Dec(X) be a z-decomposition of w. Since $\mathrm{Red}_l(d)\subseteq Z$-Fac$(X)$, according to Fact 3.8, we deduce by Fact 3.9 that $\mathrm{Red}_l(d)=\{f\}$ where f is the z-factorization of w, that is Red_l is confluent. Conversely, assume that a word w has two z-factorizations $f=x_1\ldots x_n$ and $f'=x'_1\ldots x'_{n'}$. Let $x\in X$. Then $xx_1\ldots x_n\bar{x}_n\ldots\bar{x}_1x'_1\ldots x'_{n'}x$ is a z-decomposition of the word xwx which can be reduced in two z-factorizations: $f_1=xx_1\ldots x_nx$ and $f_2=xx'_1\ldots x'_{n'}x$. Thus, Red_l is not confluent. □

LEMMA 5.2: *Let X be a language in A^*. Let $x_1\ldots x_n$ be a z-calculus of* ε *such that $x_1\ldots x_{n-1}$ is l-reduced. Let $\bar{u}$ be the longest word in*

$$\{\mathrm{Red}_A(x_1\ldots x_i),\ 1\leqq i\leqq n\}\cap\bar{A}^*,$$

and let v be the longest word in

$$\{\mathrm{Red}_A(x_1\ldots x_i),\ 1\leqq i\leqq n\}\cap A^*.$$

If $n>2$ then the word uv has two distinct z-factorizations.

Proof: First note that if $x_1\ldots x_n$ is a z-calculus of ε such that $x_1\ldots x_{n-1}$ is l-reduced, then for every $i\in\{1,\ldots,n\}$, $x_i\ldots x_nx_1\ldots x_{i-1}$ is a z-calculus of ε such that $x_i\ldots x_nx_1\ldots x_{i-2}$ is l-reduced.

There exist $k\in\{1,\ldots,n\}$ and $m\in\{1,\ldots,n\}$ such that $k\neq m$, $\bar{u}=\mathrm{Red}_A(x_1\ldots x_k)$ and $v=\mathrm{Red}_A(x_1\ldots x_m)$. If $\bar{u}=\varepsilon$ then $x_1\ldots x_m$ and $\bar{x}_n\ldots\bar{x}_{m+1}$ are two distinct z-factorizations of $v=uv$, since $x_m\neq\bar{x}_{m+1}$. Symetrically, if $v=\varepsilon$ then $\bar{x}_k\ldots\bar{x}_1$ and $x_{k+1}\ldots x_n$ *are two distinct* z-factorizations of $u=uv$, since $\bar{x}_k\neq x_{k+1}$.

Now, assume that $u\neq\varepsilon$ and $v\neq\varepsilon$. Let us assume that $m<k$ (the case $k<m$ is similar). Since $\bar{u}=\mathrm{Red}_A(x_1\ldots x_k)\in\bar{A}^+$ and $v=\mathrm{Red}_A(x_1\ldots x_m)\in A^+$, we have $k<n$ and $m<n$. Moreover, $f_1=\bar{x}_k\ldots\bar{x}_{m+1}$ and $f_2=x_{k+1}\ldots x_nx_1\ldots x_m$ are two z-factorizations of uv according to the premiminary remark. Since $\bar{x}_k\neq x_{k+1}$, f_1 and f_2 are distinct. □

Now, we state that a language X is a z-code iff the "loops" in the z-calculus, if exist, are "trivial".

PROPOSITION 5.3: *Let X be a language in A^*.*

X is a z-code iff $\varepsilon_{\approx_l} = \varepsilon_{\approx_X}$.

Proof: Assume that X is not a z-code and let w be a word with two different z-factorizations f_1 and f_2. Since $f_1 \neq f_2$, $f_1 \bar{f}_2$ is not X-congruent with ε. However, $f_1 \bar{f}_2$ is a z-calculus of ε. That is $f_1 \bar{f}_2 \in \varepsilon_{\approx_l} \setminus \varepsilon_{\approx_X}$. Conservely, assume that $x_1 \ldots x_n \in \varepsilon_{\approx_l} \setminus \varepsilon_{\approx_X}$. Without loss of generality, we can assume that $x_2 \ldots x_n$ is l-reduced. Since $x_1 \ldots x_n \notin \varepsilon_{\approx_X}$, we have $n > 2$. Thus in virtue of Lemma 5.2, X is not a z-code. □

From the previous result, we obtain easily two corollaries which show that, if X is a z-code, the l-reductions of the z-decompositions concern only the "trivial" loops and that, in this case, the X-reduced z-decompositions on X are exactly the z-factorizations on X.

COROLLARY 5.4: *Let X be a language in A^*.*

X is a z-code iff $\forall\, d \in Z\text{-Dec}(X)$, $\text{Red}_l(d) = \text{Red}_X(d)$.

COROLLARY 5.5: *Let X be a language in A^*.*

X is a z-code iff $\text{Red}_X(Z\text{-Dec}(X)) = Z\text{-Fac}(X)$.

Now, we characterize the z-codes with some properties of the A-reduction:

PROPOSITION 5.6: *Let X be a language in A^*.*

X is a z-code iff $\forall\, (d_1, d_2) \in Z\text{-Dec}(X) \times Z\text{-Dec}(X)$, $d_1 \approx_X d_2 \Leftrightarrow d_1 \approx_A d_2$.

Proof: Let X be a z-code. Let d_1 and d_2 be two z-decompositions such that $d_1 \approx_A d_2$. Then d_1 and d_2 are z-decompositions of the same word $\text{Red}_A(d_1) = \text{Red}_A(d_2)$. By Fact 3.8, there exist $f_1 \in Z\text{-Fac}(X)$ and $f_2 \in Z\text{-Fac}(X)$ such that $d_1 \approx_l f_1$ and $d_2 \approx_l f_2$. Since X is a z-code $f_1 = f_2 = f$. By Corollary 5.4, we have $d \approx_X d'$. Conversely, if X is not a z-code, there exist two z-factorizations f_1, f_2 such that $f_1 \neq f_2$ and $f_1 \approx_A f_2$. Since f_1 and f_2 are X-reduced, they are not X-congruent. □

LEMMA 5.7 [9]: *Let $X \subseteq A^*$ which is not a z-code. Let w be a shortest word having two distinct z-factorizations $x_1 \ldots x_n$ and $y_1 \ldots y_m$ on X. Then $x_1 \neq y_1$.*

Proof: Assume that $m \leqq n$. Let $k \in \{1, \ldots, m\}$ such that $x_k \neq y_k$ and $\forall\, i \in \{1, \ldots, k\}$, $x_i = y_i$. Let p be the least integer greater than k such that $x_k \ldots x_p \approx_A y_k \ldots y_{p'}$ for some $p' \geqq k$ (such an integer p exists since $x_k \ldots x_n \approx_A y_k \ldots y_m$). Thus $x_k \ldots x_p \bar{y}_{p'} \ldots \bar{y}_k \approx_l \varepsilon$ and $x_k \ldots x_p \bar{y}_{p'} \ldots \bar{y}_{k+1}$ is

l-reduced. Hence, according to the Lemma 5.2, $p=p'=k$. That is $x_k \bar{y}_k \approx_l \varepsilon$, which is a contradiction. □

PROPOSITION 5.8: *Let $X=\{u, v\}$ be a two-element language in A^*. X is a z-code iff X is a code.*

Proof: It is sufficient to prove that if $X=\{u, v\}$ is not a z-code then it is not a code. So, suppose that X is not a z-code. Since any prefix code is a z-code, we may assume that $u<v$. Let n be the greatest integer such that $v=u^n u_1$. Since X is not a z-code, there exists a shortest word $w \in A^*$ that admits two different z-factorizations f_1 and f_2 on X. By Lemma 5.7, we can assume that $f_1 \in u(X \cup \bar{X})^+$ and $f_2 \in v(X \cup \bar{X})^+$. Necessarily, we have $f_1 \in u\{v, uv, u^2 v, \ldots, u^{n-1} v, u^n\}(X \cup \bar{X})^*$, in other words, the longest prefix w' of f_1 in X^+ satisfies $\mathrm{Red}_A(w') \in u^n A^*$. Thus, we have $u_1 < u$. Set $u=u_1 u_2$. If f_1 and f_2 are in X^+ then X is not a code. If f_1, for instance, belongs to $X^+ \bar{X}(X \cup \bar{X})^*$ then f_1 has a factor $v\bar{u}$ or $u\bar{v}$. As $v=u^n u_1$ and $n \geqq 1$, we have $v \in A^+ u_2 u_1$. Thus a factor $v\bar{u}$ or $u\bar{v}$ in f_1 implies that $u_1 u_2 = u_2 u_1$. Hence $uv=vu$ that is X is not a code. □

Remark: There exists three-element codes which are note z-code. For example, $X=\{a, aba, baba\}$ is a code in $\{a, b\}^*$, but it is not a z-code since the word *ababa* admits two different z-factorizations on X.

The next characterization of z-codes is inspired by the Sardinas, Patterson's criterium (*see* [5] for instance) for ordinary codes.

DEFINITION 5.9: We construct by induction a sequence of languages $(U_i)_{i \geqq 0}$ in $A^* \# A^*$, where $\#$ is a new symbol, by setting:

- $U_0 = \{u \# v \mid (u, uv) \in X \times X, v \neq \varepsilon\}$.
- $\forall i \geqq 0$,

$$\begin{aligned} U_{i+1} = &\{ux \# v \mid u \# xv \in U_i, x \in \mathrm{X}\} \\ &\cup \{uv \# y \mid u \# v \in U_i, vy \in X\} \\ &\cup \{u \# xv \mid ux \# v \in U_i, x \in X, u \neq \varepsilon\}. \end{aligned}$$

FACT 5.10: *Let $c(\varepsilon, u, v)$ be a l-reduced z-calculus of u and let $x \in (X \cup \bar{X}) \setminus \{\bar{u}\}$ such that $c(\varepsilon, u, v)x$ has sense. Then the z-calculus $c(\varepsilon, u, v)x$ of* $\mathrm{Red}_A(ux)$ *can be reduced in a z-calculus c' such that* First-Letter $(c') =$ FirstLetter $(c(\varepsilon, u, v))$.

DEFINITION 5.11: A word u is a z-prefix of uv with respect to X if u admits a z-calculus on X with context (ε, v).

LEMMA 5.12: *For any* $i \geqq 0$, *we have the following property* P_i:
$u \# v \in U_i$ *iff there exist a z-factorization* f_{uv} *of* uv *and a l-reduced z-calculus* $c(\varepsilon, u, v)$ *such that*

$$\text{FirstLetter}(\text{Red}_l(c(\varepsilon, u, v))) \neq \text{FirstLetter}(f_{uv}).$$

Proof: We prove the property P_i by induction on i. For $i=0$, the property is clear. Assume that P_i is true for $i \geqq 0$. Let $w \# w' \in U_{i+1}$. Three cases arise:

- If $w \# w' = ux \# v$ where $u \# xv \in U_i$ and $x \in X$. Then u is a z-prefix of uxv and $uxv \in X^\uparrow$. Hence ux is a z-prefix of uxv and according to Fact 5.10, $c(\varepsilon, u, xv)x$ can be l-reduced in a z-calculus c' of ux such that

$$\text{FirstLetter}(\text{Red}_l(c')) \neq \text{FirstLetter}(f_{uv}).$$

- If $w \# w' = uv \# y$ where $u \# v \in U_i$ and $vy \in X$. Then $c(\varepsilon, u, v)vy$ is a z-factorization of uvy and the f_{uv} is a l-reduced z-calculus of uv.
- If $w \# w' = u \# xv$ where $ux \# v \in U_i$ and $x \in X$ and $u \neq \varepsilon$. Then, according to Fact 5.10, $c(\varepsilon, ux, xv)\bar{x}$ can be l-reduced in a z-calculus c' of u such that

$$\text{FirstLetter}(c') = \text{FirstLetter}(c(\varepsilon, u, xv)).$$

Thus

$$\text{FirstLetter}(c') \neq \text{FirstLetter}(f_{uxv}).$$

Thus in all cases P_{i+1} is satisfied. □

PROPOSITION 5.13: *Let* X *be a language in* A^*. X *is a z-code iff* $(\bigcup_{i \geqq 0} U_i) \cap (A^* \#) = \varnothing$.

Proof: If X is not a z-code, according to Lemma 5.7, there exists a word w that has two distinct z-factorizations xf and yg with $(x, y) \in X \times X$ and $x \neq y$. Thus by setting $p = (|xf| + |yg|)$, we have $w \# \in U_{p-2}$. Conversely, if $U_i \cap A^* \# \neq \varnothing$ for some $i > 0$, we have $u \# \in U_i$ that is $u \in X^\uparrow$ and there exists a l-reduced z-calculus $c(\varepsilon, u, \varepsilon)$ of u. That is $c(\varepsilon, u, \varepsilon)$ is a z-factorization of u and f_u is a z-factorization of u. Since $\text{FirstLetter}(f_u) \neq \text{FirstLetter}(c(\varepsilon, u, \varepsilon))$, we have $f_u \neq c(\varepsilon, u, \varepsilon)$. □

Remarks:

- If $X \in \text{Rat}(A^*)$, each $U_i \in \text{Rat}(A^* \# A^*)$.
- The set of languages U_i is not necessarily finite (even if X is finite). However, according to [2], in the rational case, one can compute an upper bound of the length of a shortest word having two different z-factorizations

on X so it is sufficient to compute a finite calculable number of U_i to decide whether a regular language is a z-code.

6. MORPHISMS PRESERVING ZIGZAG CODES

As well know the image and inverse image of a code by an injective morphism is again a code. We shall see below that a similar situation holds for the inverse images of the z-codes, but not for the images. The class of morphisms preserving the property of being a z-code is rather poor. This is nothing but the coding morphisms for the biprefix codes.

Given a morphism $f: A^* \to C^*$ and a language X in A^*. Recall that $Z_f(X)$ encodes in $(X \cup \bar{X})^*$, the z-decompositions on $f(X)$ (*see* Definition 4.1).

PROPOSITION 6.1: *Let $f: A^* \to C^*$ be an injective morphism and X be a subset of A^* such that $Z_f(X) \subseteqq Z\text{-Dec}(X)$. If X is a z-code then so is $f(X)$.*

Proof: Let $\varphi : B^* \to A^*$ be a coding morphism for X and let $\theta = f \circ \varphi$. One has $\theta(B) = f(\varphi(B)) = f(X)$. In order to prove that $f(X)$ is a z-code, we first show that $M_\theta(B) \subseteqq M_\varphi(B)$. Let us have $w = b_1 \ldots b_n \in M_\theta(B)$. For all $i \in \{1, \ldots, n\}$, we denote $x_i = \varphi(b_i)$. Assume $w \in M_\theta(B)$. From the fact that $\bar{\theta}(b_1 \ldots b_i) = \bar{f}(\bar{\varphi}(b_1 \ldots b_i)) = \bar{\varphi}(x_1 \ldots x_i)$, it follows that $\bar{\varphi}(w) \in Z_f(X)$. By the hypothesis, this implies $\varphi(w) \in Z\text{-Dec}(X)$. By $\text{Red}_A(x_1 \ldots x_i) = \bar{\varphi}(b_1 \ldots b_i)$ and the axioms P_1 and P_2, w satisfies P'_1 and P'_2 that is $w \in M_\varphi(B)$. So $M_\theta(B) \subseteqq M_\varphi(B)$.

Let now $\bar{\theta}(u) = \bar{\theta}(v)$ for some $u, v \in M_\theta$. This implies $\bar{f}(\bar{\varphi}(u)) = \bar{f}(\bar{\varphi}(v))$. Since $M_\theta(B) \subseteqq M_\varphi(B)$, $\bar{\varphi}(u)$ and $\bar{\varphi}(v)$ are in A^+. Hence, $f(\bar{\varphi}(u)) = f(\bar{\varphi}(v))$. By the injectivity of f, it gives $\bar{\varphi}(u) = \bar{\varphi}(v)$. Because $\bar{\varphi}$ is injective on $M_\varphi(B)$, we have $u = v$. This means that $\bar{\theta}$ is injective on $M_\theta(B)$. By Corollary 4.5, $\theta(B) = f(X)$ is a z-code. □

The following example shows that the condition $Z_f(X) \subseteqq Z\text{-Dec}(X)$ is not necessary for $f(X)$ to be a z-code.

Example 6.2: Let $A = C = \{a, b\}$, $X = \{a, aba\}$. Let f be given by $f(a) = a$, $f(b) = aba$. The word $w = (aba)(\bar{a})(\bar{a})(aba)$ belongs to $Z_f(X)$, but $\text{Red}_A((aba)) = aba \not< abba = \text{Red}_A(w)$. Moreover $\text{Red}_A((aba)(\bar{a})(\bar{a})) = ab\bar{a} \notin A^+$. However $f(X) = \{a, aabaa\}$ is a z-code.

THEOREM 6.3: *For any injective morphism $f: A^* \to C^*$ the following conditions are equivalent:*

1. *$f(A)$ is a biprefix code.*
2. *For any subset X of A, X is a z-code then so is $f(X)$.*

Proof: 1. $\Rightarrow$ 2. It suffices to prove that if 1. holds then $Z_f(X) \subseteqq Z\text{-Dec}(X)$ for any subset X in A^*. Suppose there exists $x_1 \ldots x_n \in Z_f(X)$ which does not satisfy the axiom P_1 *i.e.* $\text{Red}_A(x_1 \ldots x_m) \notin A^+$ for some $m \in \{2, \ldots, n\}$. Let $\text{Red}_A(x_1 \ldots x_m) = a_1 \ldots a_{i-1} \bar{a}_i \ldots a_k$ where $\bar{a}_i$ is the first occurence of the letters in $\bar{A}$. There exist $x_j \in \{x_2 \ldots x_m\}$ which contains the $\bar{a}_i$: $x_j = x'_j \bar{a}_i x''_j$, x'_j, $x''_j \in \bar{A}^*$, $\text{Red}_A(x_1 \ldots x_{j-1} x'_j \bar{a}_i) = a_1 \ldots a_{i-1} \bar{a}_i$. Because $f(A)$ is biprefix, $\bar{f}(a_1 \ldots a_{i-1} \bar{a}_i)$ and therefore $\bar{f}(x_1 \ldots x_j)$ can not belong to C^+, a contradiction. So every element of $Z_f(X)$ must satisfy the axiom P_1.

We now examine the axiom P_2. Suppose $x_1 \ldots x_n \in Z_f(X)$. Then for all $i \in \{1, \ldots, n\}$, we have $\bar{f}(x_1 \ldots x_i) < \bar{f}(x_1 \ldots x_n)$, or

$$\bar{f}(\text{Red}_A(x_1 \ldots x_i)) < \bar{f}(\text{Red}_A(x_1 \ldots x_n)).$$

By the above, $\text{Red}_A(x_1 \ldots x_i)$ and $\text{Red}_A(x_1 \ldots x_n)$ are in A^+, say $\text{Red}_A(x_1 \ldots x_i) = a_1 \ldots a_p$ and $\text{Red}_A(x_1 \ldots x_n) = a'_1 \ldots a'_q$. We have $\bar{f}(a_1 \ldots a_p) < \bar{f}(a'_1 \ldots a'_q)$ which implies $f(a_1) \ldots f(a_p) < f(a'_1) \ldots f(a'_q)$. Because $f(A)$ is prefix, it follows $q > p$ and $f(a_i) = f(a'_i)$ for all $i \in \{1, \ldots, p\}$. Hence $a_i = a'_i$ for all $i \in \{1, \ldots, p\}$ because of the injectivity of f. So $\text{Red}_A(x_1 \ldots x_i) = a'_1 \ldots a'_p < a'_1 \ldots a'_q = \text{Red}_A(x_1 \ldots x_n)$. Thus the axiom P_2 holds true. We have $Z_f(X) \subseteqq Z\text{-Dec}(X)$.

2 $\Rightarrow$ 1. Suppose that $Z = f(A)$ is not a biprefix code. Then Z is not prefix or not suffix. We treat only the first case, for the other one the argument is similar. There must exist y and z in Z and $t \in C^+$ such that $y = zt$. We have $y = f(a)$ and $z = f(b)$ for some letters $a, b \in A$. The set $X = \{aab, ab, b\}$ is a z-code. But the set $f(X) = \{ztztz, ztz, z\}$ is not a z-code, indeed the word $ztztz$ has two different z-factorizations on $f(X)$. □

Remarks:

- The power of a z-code is not necessarily a z-code. For example, $X = \{a, aba\}$ is a z-code, but $X^2 = \{aa, aaba, abaa, abaaba\}$ is not a z-code because the word *abaaba* has two different z-factorizations on X^2.
- Generally, $f(X^\dagger)$ is strictly contained in $[f(X)]^\dagger$, but if f is a coding morphism for a biprefix code, one has $f(X)^\dagger = [f(X)]^\dagger$ for all language X.

PROPOSITION 6.4: *Let $f: A^* \to C^*$ be an injective morphism.*

Then if Y is a z-code on C, $f^{-1}(Y)$ is a z-code on A.

Proof: Assume that a word w admits two different z-factorizations on $f^{-1}(Y)$. Since f is injective, $f(w)$ has two different z-factorizations on $f(f^{-1}(Y))$. Hence $f(w)$ has two different z-factorizations on Y: a contradiction with the fact that Y is a z-code. □

REFERENCES

1. M. Anselmo, Automates et codes zigzag, *R.A.I.R.O. Inform. Theor. Appl.*, 1991, *25*, 1, pp. 49-66.
2. M. Anselmo, Sur les codes zigzag et leur décidabilité, *Theoret. Comput. Sci.*, 1990, *74*, pp. 341-354.
3. M. Anselmo, Automates bilatères et codes zigzag *Thèse L.I.T.P.*, 1990, n° 90-27.
4. J. Berstel, Transductions and Context Free Languages, *Teubner Studienbucher*, Stuggart, 1979.
5. J. Berstel and D. Perrin, Theory of Codes, *Academic Press*, New York, 1985.
6. J. C. Birget, Two Way Automaton Computations, *R.A.I.R.O. Inform. Theor. Appl.*, 1990, *24*, 1, pp. 47-66.
7. Do Long Van, B. Le Saëc and I. Litovsky, Stability for the Zigzag Submonoids, *Rapp. interne LaBRI* No. 91-37, Bordeaux, 1991, *Theoret. Comput Sci.* (to appear).
8. M. Madonia, S. Salemi and T. Sportelli, On Z-Submonoids and Z-Codes, *R.A.I.R.O. Inform. Theor. Appl.*, 1991, *25*, 4, pp. 305-322.
9. M. Madonia, S. Salemi and T. Sportelli, A Generalization of Sardinas, Patterson's Algorithm to Z-Codes, *Theoret. Comput. Sci.* (to appear).
10. J. P. Pécuchet, Automates Boustrophédons semigroupes de Birget et monoïde inversif libre, *R.A.I.R.O. Inform. Theor. Appl.*, 1985, *19*, 1, pp. 71-100.