Lyndon Factorization of Sturmian Words

GUY MELANÇON[1]
LaBRI, URA 1304 CNRS – Université Bordeaux I

1 Introduction

Infinite sturmian words appear through many chapters of the litterature: number theory, combinatorics of dynamical systems, combinatorics on words, as well as theoretical computer science. These (right) infinite words have both geometrical and combinatorial characterizations (cf [1], [2], [3] or [4]). The present paper proposes to look at characteristic sturmian words, using Lyndon factorization of infinite words.

Lyndon words are minimal representatives of conjugacy classes (w.r.t. the lexicographical order); equivalently they are stricly smaller than their non empty proper right factors. The Lyndon factorization theorem [5] asserts that any word may be expressed as a non increasing product of Lyndon words. A beautiful algorithm by Duval [6], exploiting the combinatorics of Lyndon words, computes this factorization in linear time. Siromoney and al. [7] introduced infinite Lyndon words and gave a generalization of Lyndon's theorem: any right infinite word may be expressed as a non increasing product of Lyndon words (finite or infinite) (cf [7, Th. 2.4].

We give the explicit computation of the factorization of any characteristic sturmian word s as a non increasing product of finite Lyndon words (Th. 3.3):

$$s = \prod_{n\geq 0} [(a\bar{s}_{2n+1})^{c_{2n}-1} a s_{2n} \bar{s}_{2n+1}]^{c_{2n}+1}$$

in terms of exponents $(c_n)_{n\geq 0}$ intimately linked to the word s.

We then show how one can use this information on s and give two applications. First, we prove that the factorization of s gives an ω-division for it. Second we give a short and elegant proof of a recent result by Berstel and de Luca [4]: we compute the set of its factors that are Lyndon words, and show that it is equal to the set of primitive Christoffel words (cf Def. 4.3).

2 Infinite Lyndon words

The basic definitions and notations we use are those usual in theoretical computer science (see [8]). We denote by $A = \{a, b\}$ the two letter alphabet and suppose it is totally ordered by $a < b$. This order is naturally extended to the set of all words A^* lexicographically.

[1]Work partially supported by EC grant CHRX-CT93-0400.

Let us go through basic facts about Lyndon words. All the results concerning Lyndon words (finite or infinite) we state here hold true over an arbitrary alphabet. Recall that a word $w \in A^*$ is a *Lyndon word* if it is strictly smaller than any of its non empty proper right factors (w.r.t. the lexicographical order $<$). An equivalent definition may be given in terms of conjugation of words (cf [8]): any Lyndon word is primitive and minimal in its conjugacy class. Recall that $w \in A^+$ is primitive if it is not a proper power of another word u, that is, $w = u^n$ implies $n = 1$ and $w = u$. For example, with $A = \{a, b\}$, the word $ababb$ is a Lyndon word, with conjugacy class $\{ababb, babba, abbab, bbaba, babab\}$. Remark that, in particular, letters are Lyndon words and that Lyndon words are primitive.

A right *infinite word* s is a sequence $(a_i)_{i \geq 0}$, written as $s = a_0a_1a_2\cdots$ Siromoney and al. [7] introduced *infinite Lyndon words* as inductive limits of sequences of finite Lyndon words. Recall the Lyndon factorization theorem [5]: any non empty word can be expressed as a non increasing product of Lyndon words (cf [8, Th. 5.1.5]). In [7], the authors showed how this theorem may be extended to (right) infinite words. Let us state their result:

Theorem 2.1 ([7, Th. 2.4])
Any right infinite word s may be uniquely expressed as a non increasing product of Lyndon words, finite or infinite, in one of the two following forms:

> *either there exists an infinite non increasing sequence of finite Lyndon words $(\ell_k)_{k \geq 0}$ such that:*
>
> $$s = \ell_0\ell_1\cdots, \tag{1}$$
>
> *or there exist finite Lyndon words $\ell_0, \ldots, \ell_{m-1}$ $(m \geq 0)$ and an infinite Lyndon word ℓ_m such that:*
>
> $$s = \ell_0\cdots\ell_{m-1}\ell_m, \quad \text{with } \ell_0 \geq \cdots \geq \ell_{m-1} > \ell_m. \tag{2}$$

A result of Varricchio [9, Th. 3.7] implicitely shows that certain infinite words admit a factorization of type (1) over Viennot factorizations (for a definition of a Viennot factorization, see [8, Th. 5.4.4]). Fortunately, it is possible to show a complete analog of Th. 2.1 for general Viennot factorizations [10], [11].

In this paper, we compute the explicit Lyndon factorization of characteristic sturmian words. They all have a factorization of type (1). Since infinite Lyndon words do not appear in these factorizations, we do not take time here to define them, and refer the interested reader to [7]. We look at characteristic sturmian words and show how one can get information about them out of their factorization (1). In particular, we give a proof of a recent result by Berstel and de Luca [4].

3 Factorization of characteristic sturmian words

3.1 The Fibonaci infinite word

In this paragraph, we give the explicit computation of the factorization of the Fibonacci word. The computation of the factorization in the more general case of a characteristic sturmian word follows the same line.

Define finite words $f_0 = b$, $f_1 = a$ and for $n \geq 1$, $f_{n+1} = f_n f_{n-1}$. The *Fibonacci* (infinite) *word* is the limit $f = \lim f_n$. Hence, we have $f = abaababaabaababaabab\cdots$

Proposition 3.1 ([10, Prop. 11])
The factorization of the Fibonacci word f is of type (1) and is given by the sequence of words $(\ell_k)_{k\geq 0}$ with $\ell_0 = ab$ and $\ell_{k+1} = \varphi(\ell_k)$, where $\varphi : \{a,b\}^* \to \{a,b\}^*$ is the homomorphism defined by $\varphi(a) = aab$ and $\varphi(b) = ab$. Moreover, we have $|\ell_k| = F_{2k+2}$ (where F_k denotes the kth Fibonacci number, with $F_0 = F_1 = 1$).

Thus the factorization of f is:

$$f = (ab)(aabab)(aabaababaabab)\cdots$$

Before giving the proof, we need to develop a little theory about Lyndon words. Denote by L the set of Lyndon words. Any Lyndon word $w \in L$ of length ≥ 2 may be expressed as a product of two Lyndon words, $w = uv$, with $u, v \in L$ and $u < v$. Let v be the longest right factor of w that qualifies as a Lyndon word. Then $w = uv$, and we have $u \in L$ and $u < uv < v$. This factorization of w is called its *right standard factorization.* Given a Lyndon word $w \in L \setminus A$, we will write $w = w'w''$ to denote the left and right factors of its right standard factorization. Let us gather into a proposition the informations we will need about Lyndon words:

Proposition 3.2 *Let $u, v \in L$ be such that $u < v$ and suppose u has standard factorization $u = u'u''$. Then the factorization uv is standard if and only if $u'' \geq v$.*

Let $u, v \in L$. We have $uv \in L$ iff $u < v$. Consequently, for all $p, q \geq 1$, the word $u^p v^q \in L$ is a Lyndon word. Moreover, suppose both $u = u'u''$ and uv are standard factorizations (i.e. $u'' \geq v$). Then $u^p v^q$ has standard factorization:

$$\begin{cases} (u^p v^q)'' = u^{p-1}v^q \\ (u^p v^q)' = u \end{cases} \text{if } p \geq 2 \qquad \begin{cases} (u^p v^q)'' = v \\ (u^p v^q)' = uv^{q-1} \end{cases} \text{if } p = 1$$

Furthermore, suppose that u and v have a unique factorization into an increasing product of two Lyndon words. Then the same holds true for $u^p v^q$.

The first statement is originally from [5]. The second statement is from Duval [6]. For a proof, the reader is refered to [8].

Proof of Prop. 3.1. Every word f_{2n+1} ends with the letter a; denote by $\bar{w}$ the word obtained from w by deleting the a at its end (if possible). One shows by induction that the words $\ell_n = af_{2n}\bar{f}_{2n+1}$ are Lyndon words. Prop. 3.2 then implies $\ell'_n = a\bar{f}_{2n+1}$ and it follows that the sequence $(\ell_n)_{n\geq 0}$ is strictly decreasing. It is straightforward to compute $f = \prod_{n\geq 0} \ell_n$, after observing that

$$f = f_1 f_0 f_1 f_2 f_3 \cdots \tag{3}$$

The second part of the statement is a consequence of the fact that we have $\ell_0 = ab$ and $\ell_{k+1} = \ell'_k \ell^2_k$ (as shows the preceding induction). The result then follows from the fact that the homomorphism φ respects standard factorization, i.e. $\varphi(\ell'_k\ell''_k) = \varphi(\ell'_k)\varphi(\ell''_k)$. The equality $|\ell_k| = F_{2k+2}$ is easy.

3.2 Characteristic sturmian words

As Berstel and de Luca [4] point out, infinite sturmian words may be defined either geometrically or combinatorially. Combinatorially, they may be defined as infinite words having a minimal number $p(n)$ of factors of length n. Since, $p(n) \geq n+1$, they satisfy $p(n) = n+1$, from which we conclude that they are two letter words. Geometrically they correspond to lines in the planes: read the intersection of this line with the *discrete grid* from the origin (i.e. lines $x = \alpha$ or $y = \beta$, with $\alpha, \beta \in Z$). Writing a letter a for an intersection with a horizontal segment, and a letter b for an intersection with a vertital segment, you get an infinite word with minimal complexity.

Let $(c_n)_{n\geq 0}$ be any sequence of integers satisfying $c_0 \geq 0$ and $c_n \geq 1$ for $n \geq 1$. Define finite words $s_0 = b$ and $s_1 = a$, and $s_{n+1} = s_n^{c_{n-1}} s_{n-1}$ for $n \geq 1$. Then the word $s = \lim s_n$ is infinite sturmian. Infinite sturmian words obtained this way are called *characteristic* sturmian words; they form an important subclass among all sturmian words (they correspond to lines intersecting the origin). For example, if $c_n = 1$ for all $n \geq 0$, we get $s = f$, the Fibonacci word. For more details on infinite sturmian words, the reader may consult [1], [2], [3], [4] and a recent survey by Berstel [12]. The sequence $(c_n)_{n\geq 0}$ is called the *directive sequence* for the word s. Note that s_{2n+1} ends with an a. Our central result is:

Theorem 3.3 *Let s be a characteristic sturmian word with directive sequence* $(c_n)_{n\geq 0}$. *Set* $\ell_n = (a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1}$, *where it is understood* *that $\ell_0 = b$ if $c_0 = 0$.*

Then the words $(\ell_n)_{n\geq 0}$ form a strictly decreasing sequence of Lyndon words and the unique factorization of s as a non increasing product of Lyndon words is:

$$s = \prod_{n\geq 0} \ell_n^{c_{2n+1}}. \tag{4}$$

Lemma 3.4 *The words $a\bar{s}_{2n+1}$, $as_{2n}\bar{s}_{2n+1}$ are Lyndon words. Furthermore, one has $(as_{2n}\bar{s}_{2n+1})' = a\bar{s}_{2n+1}$.*

Consequently, the words $(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1}$ form a strictly decreasing sequence of Lyndon words.

Proceed by induction. For $n = 0$, we have $a\bar{s}_1 = a$ and $as_0\bar{s}_1 = ab$. Now compute:

$$\begin{aligned}
& as_{2n+2}\bar{s}_{2n+3} \\
= \; & a(s_{2n+1}^{c_{2n}}s_{2n})(s_{2n+2}^{c_{2n+1}}\bar{s}_{2n+1}) \\
= \; & a\;(\bar{s}_{2n+1}(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n})(\bar{s}_{2n+1}(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n})^{c_{2n+1}}\bar{s}_{2n+1} \\
= \; & a\;\bar{s}_{2n+1}(a\bar{s}_{2n+1})^{c_{2n}-1}\;as_{2n}\bar{s}_{2n+1}\;[(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1}]^{c_{2n+1}} \\
= \; & (a\bar{s}_{2n+1})\;[(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1}]^{c_{2n+1}+1}
\end{aligned} \tag{5}$$

and

$$\begin{aligned}
& a\bar{s}_{2n+3} \\
= \; & a(s_{2n+2}^{c_{2n+1}}\bar{s}_{2n+1}) \\
= \; & a\;(\bar{s}_{2n+1}(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n})^{c_{2n+1}}\bar{s}_{2n+1} \\
= \; & a\bar{s}_{2n+1}[(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1})^{c_{2n+1}-1}](a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1} \\
= \; & (a\bar{s}_{2n+1})\;[(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1}]^{c_{2n+1}}
\end{aligned} \tag{6}$$

We conclude at once, that $as_{2n+2}\bar{s}_{2n+3}$ and $a\bar{s}_{2n+3}$ are Lyndon words, and that $a\bar{s}_{2n+3} = (as_{2n+2}\bar{s}_{2n+3})'$ by virtue of Prop. 3.2. The sequence of Lyndon words $(\ell_n)_{n\geq 0}$ is strictly decreasing since $(a\bar{s}_{2n+1})^{c_{2n}-1}as_{2n}\bar{s}_{2n+1}$ is a right factor of $as_{2n+2}\bar{s}_{2n+3}$.

Proof of Th. 3.3. First we write an identity, analog to (3):

$$
\begin{aligned}
s &= s_k^{c_{k-2}} s_{k-1} s_{k+1}^{c_{k-1}-1} s_k s_{k+2}^{c_k-1} s_{k+1} \cdots \\
&= s_{k+1} s_{k+1}^{c_{k-1}-1} s_k s_{k+2}^{c_k-1} s_{k+1} \cdots \\
&= s_{k+1}^{c_{k-1}} s_k s_{k+2}^{c_k-1} s_{k+1} \cdots
\end{aligned}
$$

Second, we compute:

$$
\begin{aligned}
& a s_{2n+1}^{c_{2n}-1} s_{2n} s_{2n+2}^{c_{2n+1}-1} s_{2n+1} \\
= \; & [(a\bar{s}_{2n+1})^{c_{2n}-1} a s_{2n} \bar{s}_{2n+1}]^{c_{2n+1}} \, a \\
= \; & \ell_n^{c_{2n+1}} \, a
\end{aligned}
$$

From which it follows, when $c_0 > 0$, that:

$$
\begin{aligned}
s &= s_1^{c_0} s_0 s_2^{c_1-1} s_1 s_3^{c_2-1} s_2 s_4^{c_3-1} s_3 \cdots \\
&= (a s_1^{c_0-1} s_0 s_2^{c_1-1} s_1)(s_3^{c_2-1} s_2 s_4^{c_3-1} s_3) \cdots \\
&= (\ell_0^{c_1} a)(s_3^{c_2-1} s_2 s_4^{c_3-1} s_3) \cdots \quad = \quad \ell_0^{c_1}(a s_3^{c_2-1} s_2 s_4^{c_3-1} s_3) \cdots \\
&= \textstyle\prod_{n\geq 0} \ell_n^{c_{2n+1}}
\end{aligned}
$$

In case $c_0 = 0$, we must be careful. Note that $s_2 = s_1^{c_0} s_0 = s_0 = b$ and compute:

$$
\begin{aligned}
s &= s_1^{c_0} s_0 s_2^{c_1-1} s_1 s_3^{c_2-1} s_2 s_4^{c_3-1} s_3 s_5^{c_4-1} s_4 s_6^{c_5-1} s_5 \cdots \\
&= s_2^{c_1} s_1 s_3^{c_2-1} s_2 s_4^{c_3-1} s_3 s_5^{c_4-1} s_4 s_6^{c_5-1} s_5 \cdots \\
&= s_2^{c_1} (a s_3^{c_2-1} s_2 s_4^{c_3-1} s_3)(s_5^{c_4-1} s_4 s_6^{c_5-1} s_5) \cdots \\
&= b^{c_1} (\ell_0^{c_1} a)(s_5^{c_4-1} s_4 s_6^{c_5-1} s_5) \cdots \\
&= b^{c_1} \ell_1^{c_3} (a s_5^{c_4-1} s_4 s_6^{c_5-1} s_5) \cdots \\
&= \textstyle\prod_{n\geq 0} \ell_n^{c_{2n+1}}
\end{aligned}
$$

Observe that our notation for ℓ_0 is in accordance with the usual notation $a^{-1}v$ corresponding to the deletion of $a \in A$ at the beginning of $v \in A^*$ (if possible). Indeed, we then compute $(a\bar{s}_1)^{c_{2n}-1} a s_0 \bar{s}_1 = a^{-1}(ab) = b$.

4 Applications

We give two applications of Th. 3.3.

4.1 ω-division of infinite words

Recall that a finite word is m-divided if it can be expressed as $w = x_1 \cdots x_m$, such that for all permutation $\sigma \in \Sigma_m$ ($\sigma \neq id$), we have $w > x_{\sigma(1)} \cdots x_{\sigma(m)}$. This definition can be extended to infinite words by asking for a factorisation $s = x_1 x_2 \cdots$ into finite words $x_i \in A^*$, to give rise to m-divided finite words $x_i \cdots x_{i+m-1}$ for all $m \geq 2$.

Corollary 4.1 *Let s be a characteristic sturmian word. Then the factorization of s, $s = \prod_{n \geq 0} x_n$ with $x_n = \ell_n^{c_{2n+1}}$ is an ω-division for s.*

De Luca [13] showed that sturmian words are ω-divided words using a different factorization. Cor. 4.1 is a consequence of Th. 3.3 together with a result by Reutenauer [14] according to which the decreasing factorization of a finite word $w = \ell_1^{n_1} \cdots \ell_m^{n_m}$ into *distinct* Lyndon words is an m-division of that word. In [10, Prop. 15], we give a more general result than Cor. 4.1 according to which any infinite word having a non ultimately periodic factorization of type (1), is ω-divided. Compare [9, Th. 3.7].

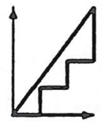
4.2 Lyndon factors of infinite words

Using Th. 3.3 we give a short and elegant proof of a result by Berstel and de Luca [4]. We say that a finite word $v \in A^*$ is a *factor* of an infinite word s if $s = uvt$ (where $u \in A^*$ is finite, and t is infinite). Denote by CSt the set of factors of characteristic sturmian words; thus $CSt = \{v \in A^* : \exists$ a characteristic sturmian word s such that v is a factor of $s\}$. Let $L \cap CSt$ denote the factors of characteristic sturmian words that qualify as Lyndon words.

Corollary 4.2 *The set $L \cap CSt$ of factors of characteristic sturmian words that qualify as Lyndon words is equal to the set CP of* primitive Christoffel *words.*

We recall the definition of primitive Christoffel words. Associate with any word $w \in \{a, b\}^*$ a path in the discrete plane $\mathbf{Z} \times \mathbf{Z}$: to a letter a corresponds a horizontal segment $(i, j) \to (i+1, j)$ and to a letter b corresponds a vertical segment $(i, j) \to (i, j+1)$.

Definition 4.3 *A* primitive Christoffel word *is a word such that its path is below the line segment joining the end points of the path, and such that the region thus formed does not contain points with integer coordinates. By convention, letters are primitive Christoffel words.*



The word $ababbab \in CP$ with associated slope 4/3.

These words play a central role in algorithmic number theory (see [15]). Primitive Christoffel words are easily obtained: draw a line segment with rational slope $m = p/q$ joining two points in $\mathbf{Z} \times \mathbf{Z}$. The unique primitive Christoffel words corresponding to the given line segment is obtained by forming the unique path crossing the points $(i, \lfloor i * p/q \rfloor)$ $(i = 1, \ldots, p + q)$. One can show that primitive Christoffel words are Lyndon words and that the standard factorization of a primitive Christoffel word correspond to its unique factorization into a product of two primitive Christoffel words. Moreover, a primitive Christoffel word is intimately linked to the slope m of the line segment joining the end points of its path: it is maximal amongst all Lyndon words having an associated line segment with slope m (see [15], [16]).

Proposition 4.4 *Let s be an infinite word with unique non increasing factorization (finite or infinite):*

$$s = \ell_0 \ell_1 \ell_2 \cdots$$

A word $u \in L$ is a factor of s if and only if it is a factor of one of the ℓ_i's.

This is a consequence of a general result on factorizations of the free monoid according to which a factor of the form $v\ell_{p+1} \cdots \ell_{q-1} w$ (where v and w are right and left factors of ℓ_p and ℓ_q, respectively) factorizes into a non increasing product of at least two Lyndon words (see [17]). The following proposition is easy.

Proposition 4.5 *Let $u \in L$ have a unique factorization into a product of two Lyndon words, $u = u'u''$. Then any factor v of u qualifying as a Lyndon word either is equal to u itself, or is a factor of u' or u''.*

We shall apply Prop. 4.5 to the words ℓ_n in (4). Indeed, it follows from Prop. 3.2 that they have a unique factorization into an increasing product of two Lyndon words. This, together with Prop. 4.4, implies that any Lyndon factor of the characteristic sturmian word s either is $\ell_n = as_{2n}\bar{s}_{2n+1}$ or $a\bar{s}_{2n+1}$, for some $n \geq 0$.

Proof of Cor. 4.2. Let $\alpha, \beta \geq 1$ be two integers. It is easy to check that $(ab^\alpha)^\beta b$ and $w_1 = a(a^\alpha b)^\beta$ are primitive Christoffel words. Let $u, v \in CP$ be such that $u \in A$ or $u'' \geq v$ (i.e. $uv \in L$ is in standard form). In [16], it is shown that the substitution $a \mapsto u, b \mapsto v$ applied to any primitive Christoffel word gives a primitive Christoffel word. Applying this to w_0 or w_1 (according to the value of c_0), an easy induction shows that $a\bar{s}_{2n+1}$, $as_{2n}\bar{s}_{2n+1}$, and ℓ_n are primitive Christoffel words. Hence any Lyndon factor of a characteristic sturmian word is a primitive Christoffel word.

Conversely, consider a primitive Christoffel word w and let m denote the slope of its associated line segment. One can give a recursive process to generate w in terms of the continued fraction $[\alpha_0; \alpha_1, \cdots \alpha_k]$ associated with m (see [16]). More precisely, consider the words:

$$u_0 = a, v_0 = b$$
$$u_{i+1} = u_i v_i^{\alpha_{2i}}, v_{i+1} = u_{i+1}^{\alpha_{2i+1}} v_i. \tag{7}$$

Then each of the words u_i, v_j are primitive Christoffel words and the word w is either u_k or v_k (according to the parity of k). Now, suppose $\alpha_0 \neq 0$. Let $(c_n)_{n\geq 0}$ be any directive sequence satisfying $c_0 = 0, c_1 = \alpha_0, \ldots, c_{k+1} = \alpha_k$. One computes $u_i = a\bar{s}_{2i+1}$, and $v_j = \ell_j$. The case $\alpha_0 = 0$ is similar. This shows that the primitive Christoffel word w is a Lyndon factor of a characteristic sturmian word s.

Thus, the set $L \cap CSt$ of Lyndon factors of characteristic sturmian words is exactly the set CP of primitive words.

References

[1] G. RAUZY. Automata in Infinite Words. In M. NIVAT and D. PERRIN, editors, *Mots infinis en arithmétique*, pages 164 – 171. Springer Verlag, 1984. Lecture Notes in Computer Science, 192.

[2] J. BERSTEL. Tracés de droite, fractions continues et morphismes itérées. In *Mots – Mélanges offerts à M. P. Schützenberger*, pages 298–309. Hermès, 1990.

[3] A. De LUCA. Sturmian Words: Structure, Combinatorics, and Their Arithmetics. *Theoretical Computer Science*, Special Issue on Formal Languages, to appear.

[4] J. BERSTEL and A. De LUCA. Sturmian Words, Lyndon Words and Trees. to appear, 1995.

[5] K. T. CHEN, R. H. FOX, and R. C. LYNDON. Free Differential Calculus, IV – The Quotient Groups of the Lower Central Series. *Annals of Mathematics*, 68:81 – 95, 1958.

[6] J. P. DUVAL. Factorizing Words over an Ordered Alphabet. *Journal of Algorithms*, 4:363 – 381, 1983.

[7] R. SIROMONEY, L. MATTHEW, V. R. DARE, and K. G. SUBRAMANIAN. Infinite Lyndon Words. *Information Processing Letters*, 50:101 - 104, 1994.

[8] M. LOTHAIRE. *Combinatorics on Words.* Addison Wesley, 1983.

[9] S. VARRICCHIO. Factorizations of Free Monoids and Unavoidable Regularities. *Theoretical Computer Science*, 73:81 - 89, 1990.

[10] G. MELANÇON. Lyndon Factorization of Infinite Words. In C. PUECH and R. REISCHUK, editors, *STACS '96, 13th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Computer Science, 1046*, pages 147 - 154. Springer Verlag, 1996.

[11] G. MELANÇON. Viennot Factorizations of Infinite Words. In preparation, 1996.

[12] J. BERSTEL. Recent Results in Sturmian Words. Invited paper to DLT'95. To be published by World Scientific, 1996.

[13] A. De LUCA. A Division Property of the Fibonacci Word. Technical Report 95/1, Università degli Studi di Roma "La Sapienza", 1995.

[14] C. REUTENAUER. Mots de Lyndon et un théorème de Shirshov. *Annales des Sciences Mathématiques du Québec*, 10(2):237 - 245, 1986.

[15] J. P. BOREL and F. LAUBIE. Quelques mots sur la droite projective réelle. *Journal de Théorie des Nombres de Bordeaux*, 5:23 - 51, 1993.

[16] E. LAURIER. *Opérations sur les mots de Christoffel.* PhD thesis, Université de Limoges, 1995.

[17] G. MELANÇON. Combinatorics of Hall Trees and Hall Words. *Journal of Combinatorial Theory, Series A*, 59(2):285 - 308, 1992.