ACTA ARITHMETICA
101.2 (2002)

On the number of Arnoux–Rauzy words

by

Filippo Mignosi (Palermo) and Luca Q. Zamboni (Denton, TX)

1. Introduction. Let $\omega = \omega_1\omega_2\ldots$ be a sequence with values in a finite alphabet A. The *complexity function* $p_\omega : \mathbb{N} \to \mathbb{N}$ assigns to each n the number of distinct factors (or subwords) of ω of length n. A fundamental result due to Hedlund and Morse states that a sequence ω is ultimately periodic if and only if for some n the complexity $p_\omega(n) \le n$. (See [11], [25] and [22, Chapter 2].) Sequences of complexity $p(n) = n + 1$ are called *Sturmian sequences* or *Sturmian words* (see [22, Chapter 3]). The best known example is the Fibonacci sequence

$$121121211211212112121121121211211212112121121121211 \ldots$$

fixed by the morphism $1 \mapsto 12$ and $2 \mapsto 1$. It is well known that all Sturmian words can be realized geometrically by an irrational rotation on the circle (see [11, 25]). More precisely, every Sturmian word is obtained by coding the symbolic orbit of a point x on the circle (of circumference one) under a rotation by an irrational angle α where the circle is partitioned into two complementary intervals, one of length α and the other of length $1-\alpha$. And conversely every such coding gives rise to a Sturmian word. The irrational number α is called the *slope*.

Let St_n denote the cardinality of the set of all Sturmian words u of length n, that is, the set of words u of length n in $\{0, 1\}$ which are a factor of some Sturmian sequence in $\{0, 1\}^{\mathbb{N}}$. In [24], Mignosi gave a proof of an explicit formula for St_n in terms of the Euler phi function $\varphi(n)$, conjectured by Dulucq and Gouyou-Beauchamps [13] (see also Corollary 4). A combinatorial proof of this formula was given by de Luca and Mignosi in [23] (see also [22, Chapter 3]). In this paper we describe a multidimensional generalization of the Euler phi function which counts the number of Arnoux–Rauzy words of each length.

2000 *Mathematics Subject Classification*: 11J70, 37B10.

Definition 1. Let $A_k = \{1, \dots, k\}$ with $k \geq 2$. A sequence ω in the alphabet A_k is called an *Arnoux–Rauzy sequence* if it satisfies the following three conditions:

- ω is recurrent,
- the complexity function $p_\omega(m)$ equals $(k-1)m+1$,
- for each m there is exactly one right special and one left special factor of ω of length m.

Recall that a factor u of ω is called *right special* (resp. *left special*) if u is a prefix (resp. suffix) of at least two words of length $|u|+1$ which are factors of ω. A word which is both right and left special is called *bispecial*. Arnoux–Rauzy sequences are a natural generalization of Sturmian words; Sturmian words correspond to taking $k = 2$ in the above definition. For $k = 3$ the combinatorial conditions listed in Definition 1 distinguish them from other sequences of complexity $2n+1$ such as those obtained by coding trajectories of 3-interval exchange transformations [16, 17, 18], or those of *Chacon type*, i.e., topologically isomorphic to the subshift generated by the Chacon sequence [6, 15]. Perhaps the best known example on three letters is the so-called *Tribonacci sequence* defined as the fixed point of the morphism $\tau(1) = 12$, $\tau(2) = 13$ and $\tau(3) = 1$. In [26] Rauzy showed that the subshift generated by τ is isomorphic (in measure) to an exchange of three fractal domains in $\mathbb{R}^2$ which generate a tiling of the plane.

Arnoux and Rauzy [2] showed that each Arnoux–Rauzy sequence may be geometrically realized by an exchange of $2k$ intervals on the circle, and is uniquely ergodic. It was further believed, as in the case of Tribonacci and the Rauzy fractal, that each Arnoux–Rauzy sequence is measure isomorphic to a rotation on the torus, i.e., is obtained by a symbolic coding of the trajectories of points under a rotation on the k-dimensional torus with respect to a natural partition. This was recently disproved by Cassaigne–Ferenczi–Zamboni in [4] where the authors exhibited an Arnoux–Rauzy sequence ω on a 3-letter alphabet $\{0, 1, 2\}$ which is *totally unbalanced* in the following sense: for each $n > 0$ there exist two factors of ω of equal length, with one having at least n more occurrences of the letter 0 than the other. It follows that the cylinder [0] is not a *bounded remainder set* (in the sense of Kesten [21]) and hence via an unpublished result of Rauzy later generalized by Ferenczi [14], either ω is not a natural coding of a rotation in $\mathbb{R}^n$ modulo a lattice, or the A–R sequence $\omega(0)$, obtained by coding ω according to first returns to 0, is not a natural coding of a rotation in $\mathbb{R}^n$ modulo a lattice.

Arnoux–Rauzy sequences have been extensively studied from many different points of view in connection with dynamical systems (see [1, 2, 7, 8, 20]), number theory (see [7, 9, 19, 20, 27, 29, 30]) and combinatorics (see [4, 5, 7, 12, 20, 27]).

Acknowledgements. The second author was supported in part by a grant from the Texas Advanced Research Program, and from the NSF. We wish to thank the referee for useful comments and suggestions.

2. Counting words. We fix $k \geq 2$ and let $\mathcal{AR}_\infty = \mathcal{AR}_\infty(k)$ be the set of all Arnoux–Rauzy sequences on the alphabet A_k. We denote by $\mathcal{AR}$ the set of all *Arnoux–Rauzy words*, that is, the set of all words u (including the empty word) such that u is a factor of some Arnoux–Rauzy sequence $x \in \mathcal{AR}_\infty$. For each $n \geq 0$ we let $\mathcal{AR}_n$ be the set of all $u \in \mathcal{AR}$ of length n.

For each $a \in A_k$ define the morphism τ_a on A_k by $\tau_a(a) = a$ and $\tau_a(b) = ab$ for all $b \in A_k$ different from a. Then it is proved in [2] (see also [27]) that each Arnoux–Rauzy sequence ω is in the shift orbit closure of a unique sequence of the form

$$\omega_* = \lim_{j\to\infty} \tau_{i_1} \circ \ldots \circ \tau_{i_j}(1)$$

where the sequence of indices (i_j) (called the *coding sequence*) takes values in A_k. Moreover each $a \in A_k$ occurs in the coding sequence an infinite number of times. The sequence ω_* is called a *characteristic Arnoux–Rauzy sequence.*

LEMMA 1. *Let $u \in \mathcal{AR}$ and suppose that for some $b, c \in A_k$ distinct, ub and uc are in $\mathcal{AR}$. Then there exists an Arnoux–Rauzy sequence $\omega \in \mathcal{AR}_\infty$ which contains as factors the k words $u1, u2, \ldots, uk$. In other words if u is a right special factor of $\mathcal{AR}$ then u is a right special factor of some Arnoux–Rauzy sequence ω.*

Proof. We proceed by induction on the length of u. The result is clearly true if u is empty, or if $|u| = 1$. Writing $u = av$ with $a \in A_k$ and $|v| \geq 1$, we make the inductive hypothesis that the result of the lemma holds for all words of length smaller than $|u|$. Thus avb and avc are each in $\mathcal{AR}$. Without loss of generality we can assume 1 is the last letter of v.

CASE 1: $a = 1$. If av is of the form $av = 1^n$, then for each Arnoux–Rauzy sequence ω, the Arnoux–Rauzy sequence $\tau_1^n(\omega)$ contains the k words $1^{n+1}, 1^n2, 1^n3, \ldots, 1^nk$. Next suppose that av is not of the form 1^n. If b and c are each different from 1, then we can write $avb = 1vb = \tau_1(v'b)$ and $avc = 1vc = \tau_1(v'c)$ for some v' with $v'b, v'c \in \mathcal{AR}$ and $|v'| < |u|$. By the inductive hypothesis there exists an Arnoux–Rauzy sequence ω which contains both $v'b$ and $v'c$ as factors. It follows that the Arnoux–Rauzy sequence $\tau_1(\omega)$ contains both avb and avc as factors. Hence av is a right special factor of $\tau_1(\omega)$.

Next assume that one of b or c (say b) is equal to 1. Then $avb = 1v1 = 1v'11$ and $avc = 1vc = 1v'1c$ for some v' in $\mathcal{AR}$. Thus we can write $1v'1 = \tau_1(v''1)$ and $1v'1c = \tau_1(v''c)$ for some v'' with $v''1, v''c \in \mathcal{AR}$ and $|v''| < |u|$. By the inductive hypothesis there exists an Arnoux–Rauzy sequence

ω which contains both $v''1$ and $v''c$ as factors. Thus $\tau_1(\omega)$ contains both $1v'1 = \tau_1(v''1)$ and $1v'1c = \tau_1(v''c)$. As τ_1 of each letter begins with 1, it follows that $1v'11$ is also a factor of $\tau_1(\omega)$. Hence $1v'1 = av$ is a right special factor of $\tau_1(\omega)$.

CASE 2: $a \neq 1$. In this case it is easy to see that $1avb$ and $1avc$ are both in $\mathcal{AR}$. Applying the arguments of Case 1 we deduce that $1av$ is a right special factor of some Arnoux–Rauzy sequence ω. Hence so is av. ∎

As an immediate consequence of Lemma 1 we have

COROLLARY 1. *Let $r(n)$ denote the number of right special factors of $\mathcal{AR}$ of length n. Then*

$$\mathrm{Card}(\mathcal{AR}_n) = \mathrm{Card}(\mathcal{AR}_{n-1}) + (k-1)r(n-1).$$

Proof. In fact each right special factor of length $n-1$ is a prefix of k factors of length n. ∎

LEMMA 2. *Suppose $u \in \mathcal{AR}$ is a bispecial factor of $\mathcal{AR}$, that is, there exist letters $a \neq b$ and $c \neq d$ such that au, bu, uc, ud are in $\mathcal{AR}$. Then there exists an Arnoux–Rauzy sequence $\omega \in \mathcal{AR}_\infty$ such that u is a bispecial factor of ω.*

Proof. The proof of Lemma 2 is similar to the proof of Lemma 1: using the τ_i the result follows by induction on the length of the words. ∎

LEMMA 3. *If u is a bispecial factor of an Arnoux–Rauzy sequence, then for each $a \in A_k$ there exists an Arnoux–Rauzy sequence ω such that u is a bispecial factor of ω and au is a right special factor of ω.*

Proof. Let $\nu = \tau_{n_1} \circ \tau_{n_2} \circ \ldots$ be a characteristic Arnoux–Rauzy sequence containing u as its rth bispecial factor, where we order the bispecial factors of ν according to increasing length. Fix $a \in A_k$ and let ω be any characteristic Arnoux–Rauzy sequence whose S-adic expansion begins with $\tau_{n_1} \circ \tau_{n_2} \circ \ldots \tau_{n_r} \circ \tau_a$. Then in [27] it is proved that ω has the same first r bispecial factors of ν (the rth bispecial factor of a characteristic Arnoux–Rauzy sequence is completely determined by the first r terms of its S-adic expansion), and that au is a right special factor of ω (the $r+1$st term of the S-adic expansion of ω determines which of the k factors $1u, 2u, \ldots, ku$ is right special in ω). ∎

COROLLARY 2. *Let $b(n)$ denote the number of bispecial factors of $\mathcal{AR}$ of length n. Then*

$$r(n) = r(n-1) + (k-1)b(n-1).$$

Proof. In fact each bispecial factor of length $n-1$ is a suffix of k right special factors of length n. ∎

Combining Corollaries 1 and 2 we have:

Corollary 3. *Fix k and let $\mathcal{AR}_n$ denote the set of all Arnoux–Rauzy words of length n on the alphabet $A_k = \{1, \ldots, k\}$. Let $b(n)$ denote the number of bispecial words in $\mathcal{AR}$ of length n. Then*

$$\mathrm{Card}(\mathcal{AR}_n) = k + (n-1)k(k-1) + (k-1)^2 \sum_{i=1}^{n-2} (n-i-1)b(i).$$

Proof. By Corollaries 1 and 2 we have

$$\begin{aligned}\mathrm{Card}(\mathcal{AR}_n) &= k + (k-1)\sum_{i=1}^{n-1} r(i)\\ &= k + (k-1)\sum_{i=1}^{n-1}\Big(k + (k-1)\sum_{j=1}^{i-1} b(j)\Big)\\ &= k + (n-1)k(k-1) + (k-1)^2 \sum_{i=1}^{n-1}\sum_{j=1}^{i-1} b(j)\\ &= k + (n-1)k(k-1) + (k-1)^2 \sum_{i=1}^{n-2}(n-i-1)b(i). \ \blacksquare\end{aligned}$$

As a special case of Corollary 3 we recover the formula for the number of Sturmian words of length n.

Corollary 4. *The number St_n of Sturmian words of length n is*

$$1 + \sum_{i=1}^{n} (n-i+1)\varphi(i)$$

where $\varphi(i)$ is the Euler phi function.

Proof. Applying Corollary 3 to the case $k = 2$ and the fact that $b(i) = \varphi(i+2)$ (see [24] or Corollary 5 ahead) gives

$$\begin{aligned}\mathrm{St}_n &= 2n + \sum_{i=1}^{n-2}(n-i-1)b(i) = 2n + \sum_{j=3}^{n}(n-j+1)\varphi(j)\\ &= 1 + \sum_{j=1}^{n}(n-j+1)\varphi(j)\end{aligned}$$

where the last step follows from the equality $\varphi(1) = \varphi(2) = 1$. $\blacksquare$

3. A generalization of the Euler phi function. In the Sturmian case we have $b(n) = \varphi(n+2)$. We now give a general arithmetic interpretation for the quantity $b(n)$ in terms of a multidimensional generalization of the Euclidean algorithm.

Fix k and let

$$E = \{(x_1, \ldots, x_k) : \text{each } x_i \text{ is a nonnegative integer}\}.$$

For $z = (x_1, \ldots, x_k) \in E$ set $|z| = \sum_{i=1}^{k} x_i$. Define a function $f : E \to E$ as follows: For $z = (x_1, \ldots, x_k) \in E$ fix the least $1 \le j \le k$ such that $x_j \le x_i$ for all $1 \le i \le k$ and set

$$f(z) = (x_1 - x_j, \ldots, x_{j-1} - x_j, x_j, x_{j+1} - x_j, \ldots, x_k - x_j).$$

Clearly for each $z \in E$ there exists a (unique) vector $\widetilde{f}(z) \in E$ such that $f^n(z) = \widetilde{f}(z)$ for all n sufficiently large. For $z \in E$ define the *generalized greatest common divisor* of z, denoted $\mathrm{ggcd}(z)$, by

$$\mathrm{ggcd}(z) = |\widetilde{f}(z)|.$$

For instance, $f(4,2,5) = (2,2,3)$ and $f(2,2,3) = (2,0,1)$ so that $\widetilde{f}(4,2,5) = (2,0,1)$ and $\mathrm{ggcd}(4,2,5) = 3$. For $k = 2$ it follows immediately from the definition that $\mathrm{ggcd}(a,b) = \gcd(a,b)$.

To the best of our knowledge, this algorithm was first defined in [5] (in the special case $k = 3$) in connection with a generalization of the Fine–Wilf theorem to three periods.

Set $P = \{z = (x_1, \ldots, x_k) \in E : \mathrm{ggcd}(z) = 1\}$ and $P(n) = \{z \in P : |z| = n\}$. Then we have

Theorem 1. *Fix k and let $\mathcal{AR}$ denote the set of all Arnoux–Rauzy words on the alphabet $A_k = \{1, \ldots, k\}$. Let $b(n)$ denote the number of bispecial words $u \in \mathcal{AR}$ of length n. Then $b(n) = \mathrm{Card}\, P((k-1)n + k)$.*

Proof. Let $B(n)$ denote the set of bispecial words in $\mathcal{AR}$ of length n, so that $b(n) = \mathrm{Card}\, B(n)$. For each $n \ge 1$ we construct a bijection $\psi_n : B(n) \to P((k-1)n + k)$ as follows: Let $u \in B(n)$; according to Lemma 2, the word u is a bispecial factor of some Arnoux–Rauzy sequence $\omega \in \mathcal{AR}_\infty$. For each $1 \le i \le k$ let v (possibly the empty word) denote the longest proper prefix of u so that iv is a right special factor of ω. If such a v exists, set $x_i = ||u| - |v||$. If no such v exists, set $x_i = |u| + 1$. It follows from the so-called "hat algorithm" given in Section III of [27] that for each i the quantity x_i is independent of the choice of ω. Set $\psi_n(u) = (x_1, \ldots, x_k)$.

We now show that $\psi_n : B(n) \to P((k-1)n + k)$ and is a bijection for each n. Taking $n = 1$ we have $B(1) = A_k = \{1, \ldots, k\}$. Fixing $i \in B(1)$ we see by definition of ψ_1 that $\psi_1(i)$ is the vector whose ith coordinate is 1 and all other coordinates are 2, so that $|\psi_1(i)| = (k-1)2 + 1 = (k-1)1 + k$ as required. Moreover $f(\psi_1(i)) = (1, 1, \ldots, 1)$ and $f^2(\psi_1(i)) = (1, 0, 0, \ldots, 0)$ so that $\mathrm{ggcd}(\psi_1(i)) = 1$. Clearly ψ_1 is injective. To see that ψ_1 is also surjective, let $z = (x_1, \ldots, x_k) \in P((k-1)2 + 1)$. Hence $|z| = (k-1)2 + 1$ and $\mathrm{ggcd}(z) = 1$. These conditions clearly imply that each $x_i > 0$. If all $x_i \ge 2$ we would have $|z| \ge 2k$, a contradiction. Hence some $x_i = 1$. We claim that all other coordinates of z are 2. In fact, if $x_s \ne x_t$ for some choice of $s \ne i$ and $t \ne i$, then $\widetilde{f}(z)$ would have two nonzero coordinates, contradicting the fact that $\mathrm{ggcd}(z) = 1$. Hence z is a vector whose ith coordinate is 1 and all other coordinates are equal to one another. As $|z| = (k-1)2+1$, the other coordinates of z must all be 2, whence it follows that $z = \psi_1(i)$ as required.

Now let $n > 1$ and suppose that $\psi_m : B(m) \to P((k-1)m+k)$ is a bijection for all $m < n$. Let u be a bispecial word of length n. We begin by showing that $\psi_n(u) = (x_1, \ldots, x_k)$ as defined above is in $P((k-1)n+k)$. Fix an Arnoux–Rauzy sequence ω in which u is a bispecial factor. Let v denote the longest proper prefix of u which is also bispecial in ω and fix $i \in A_k$ so that iv is a right special factor of ω. Hence $x_i < x_j$ for all $1 \le j \le k$. Then by definition of ψ_n we have $x_i = ||u| - |v||$. Set $\psi_{|v|}(v) = (y_1, \ldots, y_k)$. Thus $y_i = x_i$, in fact $u = vv'$ where v' is a suffix of v of length x_i (see the hat algorithm in Section III of [27]). Moreover, for $j \neq i$ we have $y_j = x_j - x_i$. Hence $\psi_{|v|}(v) = f(\psi_n(u)) = (x_1-x_i, \ldots, x_{i-1}-x_i, x_i, x_{i+1}-x_i, \ldots, x_k-x_i)$. By the inductive hypothesis we have $\mathrm{ggcd}(\psi_{|v|}(v)) = 1$ (and hence $\mathrm{ggcd}(\psi_n(u)) = 1$) and $|\psi_{|v|}(v)| = (k-1)|v| + k = (k-1)(|u| - x_i) + k$, whence $|\psi_n(u)| = \sum_{j=1}^{k} x_j = (k-1)|u| + k$ as required. Hence $\psi_n(B(n)) \subset P((k-1)n+k)$.

If for some $u' \in B(n)$ with $u' \neq u$ we had $\psi_n(u) = \psi_n(u')$ then $\psi_{|v|}$ (where v is as above) would fail to be injective on $B(|v|)$. Hence ψ_n is one-to-one. To see that ψ_n is a surjection, let $z = (x_1, \ldots, x_k) \in P((k-1)n+k)$. Thus $|z| = (k-1)n+k$ and $\mathrm{ggcd}(z) = 1$. As in the case $n = 1$ these conditions imply that each $x_j > 0$. Fix i such that $x_i \le x_j$ for all $1 \le j \le k$. We claim that $x_i < x_j$ for all $j \neq i$. In fact, if for some $j \neq i$ we had $x_j = x_i$, then $f(z)$ would have a coordinate equal to zero. Since $\mathrm{ggcd}(z) = 1$ this would imply that $z = (1, 1, \ldots, 1)$, contradicting $|z| = (k-1)n+k$. Consider $f(z) = (x_1-x_i, \ldots, x_{i-1}-x_i, x_i, x_{i+1}-x_i, \ldots, x_k-x_i)$. Then $\mathrm{ggcd}(f(z)) = 1$ (since $\mathrm{ggcd}(z) = 1$) and hence $f(z) \in P((k-1)(n-x_i)+k)$. By the inductive hypothesis, since $\psi_{n-x_i} : B(n-x_i) \to P((k-1)(n-x_i)+k)$ is onto, we have $f(z) = \psi_{n-x_i}(v)$ for some bispecial word $v \in \mathcal{AR}$ of length $n - x_i$. Let ω be any Arnoux–Rauzy sequence containing iv as a right special factor, and let u be the shortest bispecial factor of ω beginning with vi. Then it follows from the hat algorithm that $u = vv'$ where v' is a suffix of v of length x_i (see Section III of [27]). Hence $|u| = n$ and $\psi_n(u) = (x_1, \ldots, x_k) = z$ as required. ∎

As a special case of Theorem 1 we have:

COROLLARY 5. *The number of bispecial Sturmian words of length n is $\varphi(n+2)$, where φ denotes the Euler phi function.*

Proof. Applying Theorem 1 to the case $k = 2$ gives $b(n) = \mathrm{Card}\, P(n+2)$. But

$$\begin{aligned} P(n+2) &= \{(a,b) \in E \mid a+b = n+2 \text{ and } \mathrm{ggcd}(a,b) = 1\} \\ &= \{(a, n+2-a) \in E \mid \gcd(a, n+2-a) = 1\} \\ &= \{(a, n+2-a) \in E \mid \gcd(a, n+2) = 1\}, \end{aligned}$$

hence $\operatorname{Card} P(n+2) = \operatorname{Card}\{a \mid 1 \le a < n+2 \text{ and } \gcd(a, n+2) = 1\} = \varphi(n+2)$. ■

From Theorem 1 and Corollary 3 we deduce:

COROLLARY 6. *Fix k and let $\mathcal{AR}_n$ denote the set of all Arnoux–Rauzy words of length n on the alphabet $A_k = \{1, \dots, k\}$. Then*

$$\operatorname{Card}(\mathcal{AR}_n) = k + (n-1)k(k-1) + (k-1)^2 \sum_{i=1}^{n-2} (n-i-1) \operatorname{Card} P((k-1)i + k).$$

REMARK 1. In [24], following a suggestion of G. Rauzy, Mignosi establishes a connection between the number of Sturmian words, Farey numbers, and the Riemann hypothesis. Another such connection between the Riemann hypothesis, the Euler phi function and the formula in Corollary 4 was given by Bender, Patashnik, and Rumsey [3] using a result of Codèca [10]. It would be interesting to find similar connections involving the multidimensional generalization of the Euler phi function described in this paper for $k > 2$, two-dimensional Farey numbers in the sense of [2, 29, 30], and deep results and conjectures in analytical number theory. Also, Rychlik points out a possible connection between our multidimensional generalization of the Euler phi function and Gröbner bases [28].

References

[1] P. Arnoux and S. Ito, *Pisot substitutions and Rauzy fractals*, Bull. Belgian Math. Soc. Simon Stevin 8 (2001), 181–207.

[2] P. Arnoux et G. Rauzy, *Représentation géométrique de suites de complexité* $2n+1$, Bull. Soc. Math. France 119 (1991), 199–215.

[3] E. A. Bender, O. Patashnik and H. Rumsey, *Pizza slicing, phi's, and the Riemann hypothesis*, Amer. Math. Monthly 101 (1994), 307–317.

[4] J. Cassaigne, S. Ferenczi and L. Q. Zamboni, *Imbalances in Arnoux–Rauzy sequences*, Ann. Inst. Fourier (Grenoble) 50 (2000), 1265–1276.

[5] M. G. Castelli, F. Mignosi and A. Restivo, *Fine and Wilf's theorem for three periods and a generalization of Sturmian words*, Theoret. Comput. Sci. 218 (1999), 83–94.

[6] R. V. Chacon, *Weakly mixing transformations which are not strongly mixing*, Proc. Amer. Math. Soc. 22 (1969), 559–562.

[7] N. Chekhova, *Propriétés combinatoires des suites d'Arnoux–Rauzy et réponse à une question d'Hedlund et Morse*, preprint, 1999.

[8] —, *Algorithme d'approximation et propriétés ergodiques des suites d'Arnoux–Rauzy*, preprint, 1999.

[9] N. Chekhova, P. Hubert et A. Messaoudi, *Propriétés combinatoires, ergodiques et arithmétiques de la substitution de Tribonacci*, J. Théor. Nombres Bordeaux, to appear.

[10] P. Codèca, *A note on Euler's φ function*, Ark. Mat. 19 (1981), 261–263.

[11] E. Coven and G. A. Hedlund, *Sequences with minimal block growth*, Math. Systems Theory 7 (1973), 138–153.

[12] X. Droubay, J. Justin and G. Pirillo, *Epi-Sturmian words and some constructions of de Luca and Rauzy*, Theoret. Comput. Sci. 255 (2001), 539–553.

[13] S. Dulucq et D. Gouyou-Beauchamps, *Sur les facteurs des suites de Sturm*, ibid. 71 (1990), 381–400.

[14] S. Ferenczi, *Bounded remainder sets*, Acta Arith. 61 (1992), 319–326.

[15] —, *Les transformations de Chacon : combinatoire, structure géométrique, lien avec les systèmes de complexité* $2n+1$, Bull. Soc. Math. France 123 (1995), 271–292.

[16] S. Ferenczi, C. Holton and L. Q. Zamboni, *The structure of 3-interval exchange transformations I: An arithmetic study*, Ann. Inst. Fourier (Grenoble) 51 (2001), 861–901.

[17] —, —, —, *The structure of 3-interval exchange transformations II: Combinatorial properties*, preprint, 2001.

[18] —, —, —, *The structure of 3-interval exchange transformations III: Ergodic and spectral properties*, preprint, 2001.

[19] S. Ferenczi and C. Mauduit, *Transcendence of numbers with a low complexity expansion*, J. Number Theory 67 (1997), 146–161.

[20] T. Kamae and L. Q. Zamboni, *Maximal pattern complexity of discrete dynamical systems*, Ergodic Theory Dynamical Systems, to appear.

[21] H. Kesten, *On a conjecture of Erdös and Szüsz related to uniform distribution mod 1*, Acta Arith. 12 (1966), 193–212.

[22] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge Univ. Press, to appear. Available at http://www-igm.univ-mlv.fr/berstel.

[23] A. de Luca and F. Mignosi, *Some combinatorial properties of Sturmian words*, Theoret. Comput. Sci. 136 (1994), 361–385.

[24] F. Mignosi, *On the number of factors of sturmian words*, Theoret. Comput. Sci. 82 (1991), 71–84.

[25] M. Morse and G. A. Hedlund, *Symbolic dynamics II: Sturmian sequences*, Amer. J. Math. 62 (1940), 1–42.

[26] G. Rauzy, *Nombres algébriques et substitutions*, Bull. Soc. Math. France 110 (1982), 147–178.

[27] R. N. Risley and L. Q. Zamboni, *A generalization of Sturmian sequences: Combinatorial structure and transcendence*, Acta Arith. 95 (2000), 167–184.

[28] M. Rychlik, personal communication (2001).

[29] N. Wozny and L. Q. Zamboni, *Frequencies of factors in Arnoux–Rauzy sequences*, Acta Arith. 96 (2001), 261–278.

[30] L. Q. Zamboni, *Une généralisation du théorème de Lagrange sur le développement en fraction continue*, C. R. Acad. Sci. Paris Sér. I 327 (1998), 527–530.

Dipartimento di Matematica e Applicazioni
Università degli Studi di Palermo
Via Archirafi 34
90123 Palermo, Italia
E-mail: mignosi@dipmat.math.unipa.it

Department of Mathematics
University of North Texas
Denton, TX 76203-5116, U.S.A.
E-mail: luca@unt.edu

Received on 12.10.2000
and in revised form on 15.3.2001 (3896)