Complexity of sequences and dynamical systems

Sébastien Ferenczi
Institut de Mathématiques de Luminy
CNRS - UPR 9016
Case 930 - 163 avenue de Luminy
F13288 Marseille Cedex 9
France

In recent years, there has been a number of papers about the combinatorial notion of **symbolic complexity**: this is the function counting the number of factors of length n for a sequence. The complexity is an indication of the degree of randomness of the sequence: a periodic sequence has a bounded complexity, the expansion of a normal number has an exponential complexity. For a given sequence, the complexity function is generally not of easy access, and it is a rich and instructive work to compute it; a survey of this kind of results can be found in [ALL].

We are interested here in further results in the theory of symbolic complexity, somewhat beyond the simple question of computing the complexity of various sequences. These lie mainly in two directions; first, we give a survey of an open question which is still very much in progress, namely: to determine which functions can be the symbolic complexity function of a sequence. Then, we investigate the links between the complexity of a sequence and its associated dynamical system, and insist on the cases where the knowledge of the complexity function allows us to know either the sequence, or at least the system. This leads to another vast open question, the *S-adic conjecture*, and to a conceptual (though still conjectural) link with the notion of Kolmogorov-Chaitin complexity for infinite sequences (see section 6). Also, these links with dynamical systems have been of considerable help to ergodic theory, and this prompted the ergodicians to create their own notions of complexity, mimicking the theory of symbolic complexity; we give in section 5 a brief overview of these promising developments.

I thank V. Berthé and J. Cassaigne for their advices and encouragements.

1 Definitions and examples

We start with a purely combinatorial notion.

Let A be a finite alphabet and $u = (u_n, n \in \mathbb{N}) = u_0u_1u_2...$ be a one-sided sequence on A.

A **word** of **length** k is a finite string $w_1...w_k$ of elements of A; the set of all words on A is called $A^\star$; the **concatenation** of two words $w = w_1...w_r$ and $v = v_1...v_s$ is the word $wv = w_1...w_rv_1...v_s$. A word $w_1...w_k$ is said to **occur** at **place** i in the sequence u if $u_i = w_1, ..., u_{i+k-1} = w_k$.

The **language** $L_n(u)$ is the set of all words of length n occurring in u, while $L(u)$ is the union of all the $L_n(u)$; the **complexity** of u is the function $p_u(n)$ which associates to each integer $n \geq 1$ the cardinality of $L_n(u)$.

As first examples, we can compute the complexity of a **periodic** sequence like 010101...; in this case $p_u(n) = 2$ for every n. On the other hand, if u is the **Champernowne** sequence 011011100101110... made by concatenating the binary expansions of the integers 0, 1, 2, ...n..., then $p_u(n) = 2^n$ for every n.

These sequences provide in fact the two extremal examples of complexity ; we are more interested in "intermediate" sequences, and most of our examples will be provided by **substitutions** - see for example [QUE] for more details.

A substitution is an application σ from an alphabet A into the set $A^\star$ of finite words on A; it extends to a morphism of $A^\star$ for the concatenation by $\sigma(ww') = \sigma(w)\sigma(w')$.

It is called **primitive** if there exists k such that a occurs in $\sigma^k b$ for any $a \in A$, $b \in A$.

If for some letter a, the word σa begins by a, then $\sigma^n a$ begins by $\sigma^{n-1}a$ for every $n \geq 1$, and there is a unique sequence u beginning by $\sigma^n a$ for every n; it is called the **fixed point** of the substitution σ beginning by a.

Among classical examples of substitutions are the **Morse** substitution

$$\begin{array}{rcl} 0 & \to & 01 \\ 1 & \to & 10 \end{array}$$

whose fixed point beginning by 0 is the **Morse sequence**

$$0110100110010110...,$$

the **Fibonacci** substitution

$$\begin{array}{rcl} 0 & \to & 01 \\ 1 & \to & 0 \end{array}$$

whose fixed point beginning by 0 is the **Fibonacci sequence**

$$0100101...,$$

the non-primitive **Chacon** substitution

$$\begin{array}{rcl} 0 & \to & 0010 \\ 1 & \to & 1 \end{array}$$

and the **primitive Chacon** substitution

$$\begin{array}{rcl} 0 & \to & 0012 \\ 1 & \to & 12 \\ 2 & \to & 012. \end{array}$$

The sequence u is **minimal** if every word occurring in u occurs in an infinite number of places with bounded gaps.

A fixed point of a primitive substitution is always minimal; so this is the case for the Morse sequence, the Fibonacci sequence and the fixed points of the primitive Chacon substitution; the fixed point of the non-primitive Chacon substitution beginning by 0 is also a minimal sequence.

2 Dynamical systems

Let $\Omega = A^{\mathbb{N}}$ be the set of all one-sided sequences on A and let S be the one-sided **shift** on A:

$$S(x_0x_1...x_n....) = x_1x_2....$$

(Ω, S), equipped with the product topology of the discrete topology on each copy of A, is a topological dynamical system, called the **full shift**.

Let u be a minimal sequence, and X_u be the set of all sequences $x \in \Omega$ such that $L_n(x) = L_n(u)$ for every n; X_u is a closed shift-invariant subset of Ω, and (X_u, S) is the (topological) **symbolic dynamical system** associated to u.

Hence $p_x(n) = p_u(n)$ for every n and every $x \in X_u$; if we change the system by a **topological conjugacy** (that is: a bicontinuous bijection commuting with the shift), then:

Proposition 1 *([FER2])If the symbolic systems (X_u, S) and (X_v, S) associated to minimal sequences u and v are topologically conjugate, then there exists a constant c such that, for all $n > c$,*

$$p_u(n-c) \leq p_v(n) \leq p_u(n+c).$$

Hence a relation like $p_u(n) \leq an^k + o(n^k)$ when $n \to +\infty$ is preserved by topological conjugacy; and the same is true if we replace $(p_u(n) \leq)$ by $(p_u(n) \geq)$, or $o(n^k)$ by $O(n^k)$. Indeed, the *order of magnitude* of the complexity function (relatively to a fixed scale of functions such as the $n^{\alpha_0}(\log n)^{\alpha_1}(\log\log n)^{\alpha_2}...(\log_{(k)} n)^{\alpha_k}$ for any $\alpha_0 \in \mathbb{R}$, $\alpha_0 > 0$ and any $\alpha_1 \in \mathbb{R}$, ..., $\alpha_k \in \mathbb{R}$) is a topological invariant. The boundedness of first-order or second-order differences of the complexity function is also preserved.

The complexity function itself is not preserved: if u and v are the fixed points beginning by zero of the non-primitive and primitive Chacon substitution, then (X_u, S) and (X_v, S) are topologically conjugate, while $p_u(n) = 2n - 1$ if $n > 1$ and $p_v(n) = 2n + 1$ for all n ([FER1]).

But with this abuse of notation, we also call $p_u(n)$ the **symbolic complexity** of the associated topological system (X_u, S).

In our examples, the complexity of the Fibonacci system, or of the Fibonacci sequence, is $p(n) = n + 1$ for all n, while for the Morse system, the complexity was computed in [BRL] and [deL-VAR]: for $n \geq 3$, if $n = 2^r + p + 1$, $0 < p \leq 2^r$, then $p_u(n) = 3.2^r + 4p$ if $0 < p \leq 2^{r-1}$ and

$p_u(n) = 4.2^r + 2p$ if $2^{r-1} < p \leq 2^r$; note that $p_u(n)$ is smaller than $4n$ and that the differences $p_u(n+1) - p_u(n)$ take two values, 2 and 4. Both for Morse and Fibonacci, and in fact for every primitive substitution, the complexity function is sub-linear ([PAN] using the results in [EHR-LEE-ROZ], or [COB] in the particular case of constant length; see [QUE] for a short proof) and its differences are bounded ([CAS1]); see also [DUR2] for a proof using more recent methods.

Note that proposition 1 fails if we weaken the notion of topological conjugacy to **semi-conjugacy**, by allowing a countable number of discontinuities for the bijection and its inverse: let α be an irrational number, $Ry = y + \alpha$ mod 1; let $u_n = 1$ whenever $R^n x \in [1-\alpha, 1[$, $u_n = 0$ otherwise; let $v_n = 1$ whenever $R^n x \in [\frac{1}{2}, 1[$, $v_n = 0$ otherwise. (X_u, S) and (X_v, S) are trivially semi-conjugate, but $p_u(n) = n + 1$ ([HED-MOR2]) and $p_v(n) = 2n$ ([ALE]).

Complexities have been computed for many other systems (see [ALL] and its bibliography), but what is more interesting is to be able to deduce the system from the complexity; the first example of this situation appears in [HED-MOR1], and we give a proof as an example of the techniques involved.

Proposition 2 *If u is a periodic or an ultimately periodic sequence, $p_u(n)$ is a bounded function. If there exists an n such that $p_u(n) \leq n$, u is an ultimately periodic sequence.*

Proof

The first part is trivial. In the other direction, we have $p_u(1) \geq 2$ otherwise u would be constant, so $p_u(n) \leq n$ implies that $p_u(k+1) = p_u(k)$ for some k. We now write a fundamental equality :

$$p_u(k+1) - p_u(k) = \sum_{w \in L_k(u)} (\#\{a \in A; wa \in L_{k+1}(u)\} - 1).$$

And for every word of length k occurring in u, there exists at least one word of the form wa occurring in u, for some letter a; as $p_u(k+1) = p_u(k)$, there can be only one such word. Hence, if $u_i ... u_{i+k-1} = u_j ... u_{j+k-1}$ then $u_{i+k} = u_{j+k}$. As the set $L_k(u)$ is finite, there exist $j > i$ such that $u_i ... u_{i+k-1} = u_j ... u_{j+k-1}$, and hence $u_{i+p} = u_{j+p}$ for every $p \geq 0$. QED

3 Complexity functions

In view of the previous proposition, a natural question arises: which functions from $\mathbb{N}$to $\mathbb{N}$may be the complexity function of a sequence u? This question is still open, and we state here all the necessary conditions we know at this moment; as for sufficient conditions, these take the form of a list of examples: we try here to describe the present state of this quickly moving question; the reader looking for additional (historical) references should consult [ALL].

Necessary conditions

- p_u is non-decreasing.
- $p_u(m+n) \leq p_u(n)p_u(m)$; this implies that $\frac{\log p_u(n)}{n}$ has a limit, which is the **topological entropy** of the sequence.
- Whenever $p_u(n) \leq n$ for some n, or $p_u(n+1) = p_u(n)$ for some n, $p_u(n)$ is bounded (see proposition 2 above).
- If the alphabet has k letters, $p_u(n) \leq k^n$; if $p_u(n) < k^n$ for some n, then there exists a real number $\kappa < k$ such that $p_u(n) \leq \kappa^n$ for all n.
- If $p_u(n) \leq an$ for all n, then $p_u(n+1) - p_u(n) \leq Ca^3$ for all n, for a universal constant C ([CAS1], [CAS2]).
- If $s_u(n) = p_u(n+1) - p_u(n)$ is bounded, the set of n such that $s_u(n+1) > s_u(n)$ has density zero; in particular, if $s_u(n)$ is ultimately periodic, it is ultimately constant ([ALE]).

Sufficient conditions

- For any pair of integers (a, b), with $a \geq 0$, and $b > 0$ if $a \leq 1$, there exists a sequence u with $p_u(n) = an + b$ ultimately ([ALE], [CAS2]).
- For $(a, b) \in \mathbb{N} \times \mathbb{Z}$, there exists a sequence u with $p_u(n) = an + b$ for all $n \geq 1$ if and only if $a + b \geq 1$ and $2a + b \leq (a + b)^2$ ([CAS2]).

- For any increasing unbounded function ϕ from $\mathbb{N}$ to $\mathbb{N}$ such that $\phi(r^{i+1}) \leq h\phi(r^i)$ for some $r > 1$, some real number h and any integer $i \geq 1$, there exist two constants c and d and a (minimal) sequence u such that ([GOY])

$$cn\phi(n) \leq p_u(n) \leq dn\phi(n);$$

  in particular, we can take $\phi(n) = n^{\alpha_0}(\log n)^{\alpha_1}(\log\log n)^{\alpha_2}...(\log_{(k)} n)^{\alpha_k}$ for any $\alpha_0 \in \mathbb{R}$, $\alpha_0 > 0$ and any $\alpha_1 \in \mathbb{R}$, ..., $\alpha_k \in \mathbb{R}$.
  The sequences used are in the class of *Toeplitz sequences*; partial results in this direction are published in [KOS] and [CAS-KAR]; in [PAN], orders of magnitude of $n \log n$, $n \log\log n$ and n^2 are obtained using non-primitive substitutions.

- For any $1 < \alpha < \beta$, there exists a sequence with $\liminf_{n\to+\infty} \frac{p_u(n)}{n^\alpha} = 0$ and $\limsup_{n\to+\infty} \frac{p_u(n)}{n^\beta} = +\infty$ ([GOY]).

- For any increasing unbounded twice derivable function ϕ from $\mathbb{R}$ to $\mathbb{R}$ such that $0 \leq \phi"(x) \leq 1$ for x large enough and $\phi'(x) \to +\infty$ when $x \to +\infty$, there exists a sequence with $\lim_{n\to+\infty} \frac{p_u(n)}{\phi(n)} = 1$; in particular, we can take $\phi(n) = n^\alpha$ for any real $1 < \alpha < 2$ ([CAS2]).

- There exists a complexity function which is smaller than an^2 but has unbounded second-order differences.
  There exists a sequence with $\liminf_{n\to+\infty} \frac{p_u(n)}{n} = 2$ and $\limsup_{n\to+\infty} \frac{p_u(n)}{n^\beta} = +\infty$ for any $\beta > 1$ ([FER2]).

- Examples of sequences with $p_u(n) = (a_1n+b_1)\dots(a_sn+b_s)$, for integers a_i and b_i, can be buit by generalizing the examples of [ALE], produced by one-dimensional rotations, to dimension s.

- Other explicit examples of (exactly) polynomial complexities can be found by coding the trajectories of billiards in hypercubes of dimension s by the sequences of numbers of the faces which are met; for totally irrational directions, we get

$$p(n, s) = \sum_{i=0}^{\inf(n,s)} \frac{n!s!}{(n-i)!i!(s-i)!}\ .$$

This was conjectured by Rauzy for $s = 2$, $p(2, s) = n^2+n+1$, proved for $s = 2$ and conjectured in the general case in [ARN-MAU-SHI-TAM], and proved in the general case in [BAR].

- Among sequences of positive topological entropy, there are obvious Champernowne-type examples where $p_u(n) = k^n$ for integer k.
  Explicit systems with $\lim_{n\to+\infty} \frac{\log p_u(n)}{n} = a$ for any given a are built in [GRI].
  The only example whose behaviour is more precisely described is due to [BLA-BLE], and satisfies $C_1 n\sqrt{2}^n \leq p_u(n) \leq C_2 n\sqrt{2}^n \leq p_u(n)$ for constants C_1 and C_2.

4 Systems of low complexity

If a system does not satisfy proposition 2, then its complexity is at least $n + 1$. Sequences with complexity $n + 1$ are called **Sturmian** sequences: note that for a Sturmian sequence $p(1) = 2$, hence we can suppose A has two letters, and take $A = \{0, 1\}$. The Sturmian sequences have been completely characterized in [HED-MOR2]; if we know that a sequence is Sturmian, the associated dynamical system has a simple geometric expression (it is semi-conjugate to an irrational rotation); and, better still, we can produce explicitely this sequence, using two arithmetic parameters; every sequence thus produced is Sturmian. Namely

Proposition 3 *The sequence u is Sturmian if and only if there exist an irrational number α and a real number x such that, if $Ry = y + \alpha \bmod 1$, then $u_n = 1$ whenever $R^n x \in [1 - \alpha, 1[$, $u_n = 0$ otherwise; or else $u_n = 1$ whenever $R^n x \in]1 - \alpha, 1]$, $u_n = 0$ otherwise.*

The Fibonacci sequence falls in this category: in this case α is the golden ratio number $\theta = \frac{\sqrt{5}-1}{2}$ and $x = \theta$. Similar results also hold for sequences of complexity $n + k$ ([ALE], [FER-MAU]).

A full geometric characterization of a similar type also exists ([ARN-RAU]) for a subclass of the sequences of complexity $2n + 1$, the **Arnoux-Rauzy** sequences. However, in the most general case ([FER2]) we can only characterize the language, or, equivalently, the associated dynamical system.

Proposition 4 *Let u be a minimal sequence on a finite alphabet A such that $p_u(n) \leq an$ for all n; then there exist a finite number of substitutions σ_i, $1 \leq i \leq c$, on an alphabet $D = \{0, ..., d-1\}$, an application α from D to A, and an infinite sequence $(1 \leq i_n \leq c, n \geq 1)$ such that*

$$\inf_{0 \leq r \leq d-1} |\sigma_{i_1}\sigma_{i_2}...\sigma_{i_n} r| \rightarrow +\infty$$

when $n \rightarrow +\infty$, and any word of the language of the system occurs in $\alpha\sigma_{i_1}\sigma_{i_2}...\sigma_{i_n}0$ for some n.

This proposition can be read as: minimal systems with sub-linear complexity are generated by a finite number of substitutions. Using a terminology initiated by Vershik, we propose to call such systems **S-adic systems**. Among S-adic systems are of course the Sturmian systems; another important class of S-adic systems is the class of *interval exchanges*, coded by the partition into intervals of continuity, see [RAU] for example.

The converse of proposition 4 is false: it can be made true under additional hypotheses but we have not yet been able to find an equivalent condition; see section 6 below for further discussion.

For Sturmian and Arnoux-Rauzy sequences, the substitutions are known and their number is respectively two and three, see [ARN-RAU] or [DUR1] for further details; in the Sturmian case, the system is a coding of an irrational rotation (because of proposition 3 above), and the sequence (i_n) is indeed closely related to the continued fraction approximation of the argument. For general sequences of complexity $2n+1$, or any sequence such that $p_u(n+1) - p_u(n)$ is identically 2, we have still a bound on the number of substitutions, but this disappears in the general case $p_u(n) \leq an$.

Minimal (or, more generally, aperiodic) symbolic systems of sub-linear complexity are known to have only a finite number of ergodic invariant probability measures ([BOS1], [BOS2]). Also, in a different direction, given a sequence u over $A = \{0, ..., k-1\}$, we can define the number $\theta = \sum_{n=0}^{+\infty} \frac{u_n}{k^n}$ whose expansion in base k is u; a recent result ([FER-MAU]) says that this number is **transcendental** if u is a generalized Sturmian sequence ($p_u(n) = n+k-1$ for all n) or an Arnoux-Rauzy sequence.

5 Other complexities

The notion of complexity of dynamical systems has seen recent developments, which we just want to present informally.

Given a (non-necessarily symbolic) topological dynamical system (X, T) or a measure-theoretic dynamical systems (X, T, μ), we can associate to it a symbolic system, by taking a suitable finite partition Q of X, and by associating to each point x the sequence (x_n) where $x_n = i$ whenever $T^n x$ falls into the i-th atom of Q (this sequence is called the Q-**name** of x). Under mild conditions, we can speak of the complexity function $p_Q(n)$ of this system, and would like to use it as an invariant of the initial system. Unfortunately, this fails in both cases.

If we are interested in a topological invariant, the partition Q has to be made with open sets; this is possible for symbolic systems, but not on systems on the interval, for example. Hence we have to replace partitions by **coverings** made of open sets; for a covering U, the significant quantity $q_U(n)$ is the minimum number of sets $\cap_{i=1}^{n} T^{-i} U_{i_j}$ necessary to cover X. Starting from $q_U(n)$, it is possible to define a **topological complexity** for every system ([BLA-HOS-MAA]) - which coincides with the symbolic complexity for symbolic systems.

For measure-theoretic systems, the isomorphism notion is much weaker; it involves measurable mappings, one-to-one on sets of full measure. An invariant would be some $\sup p_Q(n)$ on all measurable partitions, but this is impossible to compute in general; hence we need some quantity which behaves continuously when we change the partition. Hence, instead of just counting the number of possible Q-names of length n, we have to compute $K_{Q,\epsilon}(n)$, which is the number of ϵ-balls (for the usual Hamming distance) of Q-names of length n necessary to cover a subset of X of measure at least $1 - \epsilon$. Then, the rate of growth of $K_{Q,\epsilon}(n)$, after taking some limit in ϵ and some supremum in Q, provides a notion of **measure-theoretic complexity** ([FER3]).

Both for the topological and the measure-theoretic complexity, one fundamental result is true, and it parallels the symbolic theory: the complexity is smallest (namely, essentially bounded) if and only if the system is the simplest possible (equicontinuous in the topological case, isomorphic to a translation of a compact group in the measure-theoretic one).

6 Open problems

The question of finding which functions can be symbolic complexity functions is still far from solved; particularly, in the domain of positive topological entropy (or, equivalently, exponential complexity), there are not many known conditions, either necessary or sufficient.

We have still to find a stronger form of S-adicity which would be equivalent to sub-linear complexity. This is the *S-adic conjecture*, due to Host, stating that minimal systems are of sub-linear complexity if and only if they are S-adic.

The *Kolmogorov-Chaitin complexity* of an infinite sequence associates to n the length of the shortest program (for example on a universal Turing machine) able to produce the first n terms of the sequence (see [LI-VIT]). Now, our proposition 4 states that, if the symbolic complexity of a minimal sequence is small enough, there is an explicit algorithm which produces the system, with bounds on its length in some sub-cases. Can this result be improved to get an equivalence between a reasonably low" symbolic complexity and a reasonably low" Kolmogorov-Chaitin complexity? A result in the same direction has been proved recently in [BLA-KUR]: a Sturmian sequence for which the α in proposition 3 is an algebraic number of degree two or the number e can be recognized by a Turing machine in linear time.

Both the topological and measure-theoretic theory are just nascent, and raise two kind of questions inspired by the symbolic theory: we would like to build examples showing what kind of behaviour the complexity function may have, and to find more systems which are determined by their complexity; the bounded case is already solved, but is only an encourageing beginning.

References

[ALE] P. ALESSANDRI: Codages de rotations et basses complexités, *Ph. D. Thesis*, University Aix-Marseille 2 (1996).

[ALL] J.-P. ALLOUCHE: Sur la complexité des suites infinies, *Bull. Belg. Math. Soc.* 1, 2 (1994), p. 133-143.

[ARN-MAU-SHI-TAM] P. ARNOUX, C. MAUDUIT, I. SHIOKAWA, J.-I. TAMURA: Complexity of sequences defined by billiards in the cube, *Bull. Soc. Math. France* 122 (1994), p. 1-12.

[ARN-RAU] P. ARNOUX, G. RAUZY: Représentation géométrique de suites de complexité $2n+1$, *Bull. Soc. Math. France* 119 (1991), p. 199-215.

[BAR] Yu. BARYSHNIKOV: Complexity of trajectories in rectangular billiards, *Commun. Math. Phys.* 174 (1995), p. 43-56.

[BLA-BLE] G. BLANC, N. BLEUZEN - GUERNALEC, : Production en temps réel et complexité de structures de suites infinies, *Informatique théorique et applications, RAIRO* 23 (1989), p. 195-216.

[BLA-HOS-MAA] F. BLANCHARD, B. HOST, A. MAASS: Topological complexity, *preprint*, Institut de Mathématiques de Luminy (1998).

[BLA-KUR] F. BLANCHARD, P. KŮRKA: Language complexity of rotations and Sturmian sequences, to appear in *Theoret. Comput. Sci.*

[BOS1] M. BOSHERNITZAN: Unique ergodicity of minimal symbolic flows with linear block growth, *J. Analyse Math.* 44 (1985), p. 77-96.

[BOS2] M. BOSHERNITZAN: A condition for unique ergodicity of minimal symbolic flows, *Ergodic Th. Dyn. Syst.* 12 (1992), p. 425-428.

[BRL] S. BRLEK: Enumeration of factors in the Thue-Morse word, *Discr. Appl. Math.* 24 (1989), p. 83-96.

[CAS1] J. CASSAIGNE: Special factors of sequences with linear subword complexity, in *Developments in Language Theory*, World Scientific (1996), p. 25-34.

[CAS2] J. CASSAIGNE: Complexité et facteurs spéciaux, *Bull. Belg. Math. Soc.* 4 (1997), p. 67-88.

[CAS-KAR] J. CASSAIGNE, J. KARHUMÄKI: Toeplitz words, generalized periodicity and periodically iterated morphisms, to appear in *Europ. J. Combinatorics* 18 (1997).

[COB] A. COBHAM : Uniform tag sequences, *Math. Systems Theory* 6 (1972), p. 164-192.

[deL-VAR] A. de LUCA, S. VARRICHIO: Some combinatorial properties of the Thue-Morse sequence, *Theor. Comput. Sci.* 63 (1989), p. 333-348.

[DUR1] F. DURAND: Contribution à l'étude des suites et systèmes dynamiques substitutifs, *Ph. D. Thesis*, University Aix-Marseille 2 (1996).

[DUR2] F. DURAND: A characterization of substitutive sequences using return words, *Discrete Math.* 179 (1998), p. 89-101.

[EHR-LEE-ROZ] A. EHRENFEUCHT, K.P. LEE, G. ROZENBERG: Subword complexities of various classes of deterministic developmental languages without interactions, *Theor. Comput. Sci.* 1 (1975), p. 59-75.

[FER1] S. FERENCZI: Les transformations de Chacon: combinatoire, structure géométrique, liens avec les systèmes de complexité $2n+1$, *Bull. Soc. Math. Fr.* 123 (1995), p. 271-292.

[FER2] S. FERENCZI: Rank and symbolic complexity, *Ergodic Th. Dyn. Syst.* 16 (1996), p. 663-682.

[FER3] S. FERENCZI: Measure-theoretic complexity of ergodic systems, *Israël J. of Math.* 100 (1997), p. 189-207.

[FER-MAU] S. FERENCZI, C. MAUDUIT: Transcendence of numbers with a low complexity expansion, *J. Number Theory* 67 (1997), p. 146-161.

[GOY] J. GOYON: Construction de suites dont la complexité a un ordre de grandeur donné, *preprint*, Institut de Mathématiques de Luminy (1997).

[GRI] C. GRILLENBERGER: Construction of strictly ergodic systems I. Given entropy, *Z. Wahrscheinlichkeitstheorie verw. Geb.* 25 (1973), p. 323-334.

[HED-MOR1] G.A HEDLUND, M. MORSE: Symbolic dynamics, *Amer. J. Math.* 60 (1938), p. 815-866.

[HED-MOR2] G.A HEDLUND, M. MORSE: Symbolic dynamics II. Sturmian trajectories, *Amer. J. Math.* 62 (1940), p. 1-42.

[KOS] M. KOSKAS: Complexity functions for some Toeplitz sequences, to appear in *Theoret. Comput. Sci.* (1997).

[LI-VIT] M. LI, P. VITÁNYI: Kolmogorov complexity and its applications, in: *Handbook of Theoretical Computer Science, Volume A: Algorithms and Complexity*, MIT Press (1990), p. 187-254.

[PAN] J.-J. PANSIOT: Complexité des facteurs des mots infinis enegendrés par morphismes itérés, in *ICALP '84, Lecture Notes in Computer Science* 172, Springer-Verlag (1984), p. 380-389.

[QUE] M. QUEFFELEC: Substitution dynamical systems - Spectral analysis, Lecture Notes in Math. vol. 1294 (1987), Springer-Verlag.

[RAU] G. RAUZY: Echanges d'intervalles et transformations induites, *Acta Arith.* 34 (1979), p. 315-328.