

FRANÇOIS BLANCHARD

Non literal transducers and some problems of normality

Journal de Théorie des Nombres de Bordeaux, tome 5, n° 2 (1993),
p. 303-321

http://www.numdam.org/item?id=JTNB_1993__5_2_303_0

© Université Bordeaux 1, 1993, tous droits réservés.

L'accès aux archives de la revue « Journal de Théorie des Nombres de Bordeaux » (<http://jtnb.cedram.org/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

Non literal transducers and some problems of normality

par FRANÇOIS BLANCHARD

ABSTRACT – A new proof of Maxfield's theorem is given, using automata and results from Symbolic Dynamics. These techniques permit to prove that points that are near normality to base p^k (resp. p) are also near normality to base p (resp. p^k), and to study genericity preservation for non Lebesgue measures when going from one base to the other. Finally, similar results are proved to bases the golden mean and its square.

0. Introduction

Consider the expansion of a real number $r \in [0, 1) = \mathbb{T}^1$ to the base p . Several arithmetical manipulations of this expansion have a transductional character: multiplication or division by an integer; addition of a rational number; changing from base p to base p^k , or the other way round. What if these manipulations are performed on a normal number? Do they preserve normality?

Positive answers to these questions were first given before transducers or automata were defined, by using a result from Harmonic Analysis, namely Weyl's criterion: for instance Maxfield's theorem [M] states that normality to base p (i.e. genericity for the uniform measure) is equivalent to normality to base p^k . Recently, proofs using transducers and ergodic techniques were given in [BIDT] concerning addition and multiplication; they allowed to prove that near normality is also preserved, and to say something about genericity for measures different from Lebesgue measure. One purpose of this paper is to prove Maxfield's theorem with the same tools, thus obtaining the same refinements; another, more philosophical, aim is to underline the topological character of this result.

Suppose one has the expansion of r to base p^k , i.e. an infinite sequence of cells, each of them filled with one symbol in the alphabet

$$A = \{0, \dots, p^k - 1\}.$$

Manuscrit reçu le 20 septembre 1991, version définitive le 2 juin 1993.

Une partie des résultats contenus dans cet article ont été exposés au colloque "Thémate", (CIRM, Luminy, Mai 1991).

Any symbol $a \in A$ has expansion $\gamma(a)$ to base p ; for convenience $\gamma(a)$ is supposed to have length uniformly equal to k , which is achieved by adding the suitable number of 0's in the beginning. The rewriting into base p is done in three steps:

- 1- create $k - 1$ cells between any two of the sequence, in order to provide room for writing down the expansion $\gamma(a)$ instead of a ;
- 2- whenever one has $a \in A$ and then $k - 1$ empty cells, write $\gamma(a)$; one thus obtains a sequence on B , together with its scansion into words of $\gamma(A)$;
- 3- drop the scansions, keeping only the expansion to base p (scansion seems not to matter much in this case; in fact it is important in the setting of doubly infinite sequences, where the question is solved, and whenever one makes use of a code with variable length).

This construction was described in [BIP], but without consideration of topological aspects, which are essential in this article. Our job consists of checking that each step transforms some kind of genericity (first to the uniform measure) to another, suitable, one. It is almost the same the other way round, except for an extra difficulty: given the sequence on B , one must first find all its possible scansions in order to invert step 3; if one wants genericity to be preserved, $\gamma(A)$ must be a code, which means any finite concatenation of words of $\gamma(A)$ has unique decomposition into such words. Fortunately this is the case.

This question may be put in the more general context of transductional manipulations of number expansions. A transducer is an automaton with two labels on each arc, namely an oriented graph on a finite set of states C , with arcs having one input label in A and output label in B^* , where A and B are finite alphabets. The expansion of r to the base p can be obtained from its expansion to the base p^k by applying a 1-state transducer with arcs having input label a and output label $\gamma(a)$, $a \in A$. Transducers of different types may be used in the theory of normality: for instance multiplication by k in base p is performed by a transducer with 1-letter output [BIDT]; we also introduce one in Section 4, when using the same techniques for expansions to base the golden mean. Transducers with output labels equal either to the input or to the empty word can extract subsequences; this is done in [K, BrL] with the use of slightly different automata.

Another, more general, question is the action of such a base change on sequences which are generic for some non uniform invariant measure μ . Sometimes the new sequence is also generic for some other measure μ' , an image of μ under some string of operations. This fact has not as strong a

significance as normality preservation; nevertheless our tools provide some insight into this field.

What about using the same tricks when the expansion is to the base θ , with θ greater than 1 but not an integer? This is not obvious. In [Be], A. Bertrand-Mathis proved that a number is generic for the Parry measure to base θ if and only if it is generic for the Parry measure to base θ^k , with k an integer, in case θ is a Pisot number. Here it is only done for $\theta = (1 + \sqrt{5})/2$ and $k = 2$; the proof makes use of the same set of results, but the main difficulty is that even when θ is Pisot, any number has several expansions to the base θ , only one of which is canonical [F1]. Because of this, several transducers, some of them not too easy to build up, have to be used in order to obtain the canonical θ -expansion of a number, starting from its canonical θ^2 -expansion. This is devised as an example that questions like preservation of near genericity for the Parry measure, or preservation of genericity starting from non canonical measures, can sometimes be tackled with the use of transducers; but the purely technical difficulty encountered in this simple case with the size of one transducer suggests that new ideas would be welcome.

The results obtained also extend far beyond changes from base ξ to base ξ^k , since they may be applied to codes with variable length; of course in this case they have hardly any number-theoretic meaning, still another drawback being that measures with maximal entropy do not generally correspond to each other under the tower construction.

After some preliminaries, Section 2 contains all general statements on preservation of normality or genericity that are required for applications. In Section 3 generalizations of Maxfield's theorem for integer bases are dealt with; in Section 4 the same is done for bases $\theta = (1 + \sqrt{5})/2$ and θ^2 .

This research was partly done during a stay at Universidad de Chile, Santiago, in December 1990. I am indebted to Pierre Liardet and especially Jean-Marie Dumont for several stimulating remarks and questions; also, the work in Section 4 parallels some unfinished common research with him and Alain Thomas on multiplication by integers to base the golden mean (the transducers proved even bigger than the ones in Section 4!).

1. Definitions

Let A be a finite set of symbols endowed with the discrete topology; A^* is the set of all finite words on A , the length of the word u is denoted by $|u|$; a *language* on A is any subset of A^* . The sets $A^{\mathbb{Z}}$ and $A^{\mathbb{N}}$ of all doubly and simply infinite sequences on A , endowed with the product topology,

are compact metric spaces. The *shift* $\sigma : A^{\mathbb{Z}} \rightarrow A^{\mathbb{Z}}$ (or $A^{\mathbb{N}} \rightarrow A^{\mathbb{N}}$) defined by $\sigma((x_n)_{n \in \mathbb{Z}, \mathbb{N}}) = ((x_{n+1})_{n \in \mathbb{Z}, \mathbb{N}})$ is a homeomorphism of $A^{\mathbb{Z}}$ (a continuous transformation of $A^{\mathbb{N}}$). A *subshift* on A is any closed σ -invariant subset of $A^{\mathbb{Z}}$ (or $A^{\mathbb{N}}$). A subshift X is unambiguously determined by the language $L(X) = \{u \in A^*, \exists m, n \in \mathbb{Z} \text{ or } \mathbb{N}, \exists x \in X, x(m, n) = u\}$; thus a language having some obvious suitable properties defines a subshift of $A^{\mathbb{Z}}$ and a subshift of $A^{\mathbb{N}}$. For $u \in L(X)$, denote by $[u]$ the set $\{x; x(0, |u| - 1) = u\}$.

A *transitive subshift* is one such that for any $u, v \in L(X)$ there exists $w \in A^*$ such that $uwv \in L(X)$. A *subshift of finite type* is defined by forbidding a finite set of words; a subshift is called *sofic* if $L(X)$ is regular, or recognized by a finite automaton.

A *factor map* is a continuous, onto, shift-commuting map $\phi : X \rightarrow X'$; in this case X' is a *factor* of X and X an *extension* of X' . A conjugacy map is a one-to-one factor map: when such a map exists, X and X' are said to be *conjugate*. A *bounded-to-one* factor map is such that the number of preimages, or *lifts* of $x' \in X'$ is bounded by some k .

Automata and transducers

An automaton \mathcal{A} consists of:

- a finite alphabet A ,
- a finite set of *states* C ,
- a directed graph on C , the arcs each having a label in A .

To a path in the graph, one associates its *label*, i.e. the word spelled by concatenating the labels of its arcs (usually from left to right, but the natural automata for multiplication by an integer work from right to left); the set of labels is called *the language recognized by the automaton*. When using the automaton to recognize this language, all states are initial and final. There is another word one may associate to a path: to any arc, associate the couple (a, c) of its label a and the origin vertex c . Then do the same for the path by concatenating all corresponding couples. The language thus recognized may be called simply the language of \mathcal{A} .

Each of these languages defines a subshift: they are the *factor subshift* $X \subset A^{\mathbb{Z}}$ and the *lift subshift* $Y \subset (A \times C)^{\mathbb{Z}}$ associated to \mathcal{A} . Elements of these subshifts correspond to infinite paths in the graph: an element of X is the label of such a one, and an element of Y is the pair consisting of the sequence of vertices and the label of an infinite path. These two subshifts

have particular properties (Y is a subshift of finite type; X must be at least sofic). \mathcal{A} defines a mapping ϕ from Y to X by projection on A of coordinates of $y \in Y$. It is of course a factor map. It is convenient, though not strictly correct, to call also ϕ the projection map from $(A \times C)^*$ to A^* . The same definitions are fitting for simply infinite sequences.

An automaton \mathcal{A} is said to be *deterministic* if, given $a \in A$ and $c \in C$, there is at most one arc starting from state c with given label a ; *non-ambiguous* if there is at most one path with given label from one state to another. Deterministic automata are nonambiguous, but the converse is not true.

An automaton is said to be *irreducible* if its graph is strongly connected, i.e. if for any two states c, c' there exists a path joining c to c' in the graph. If \mathcal{A} is irreducible, subshifts Y and X are transitive.

An automaton \mathcal{A} is said to be *bounded-to-one* (or a transducer \mathcal{T} *bounded to-one for input*) if the factor map ϕ is bounded-to-one. The corresponding definition applies to output. When the automaton \mathcal{A} is irreducible, the map ϕ is bounded-to-one if and only if \mathcal{A} is nonambiguous [B].

A transducer is an automaton in which arcs have two labels, one input label in alphabet A , and one output label in B^* , where B is another alphabet: thus output labels may be single letters (in this case the transducer is said to be *literal*), or words with length greater than 1, or the empty word. A transducer may be thought of as performing three tasks simultaneously: it recognizes input words, and output words which decompose into concatenations of output labels; it rewrites on B any input word on A . Transducers are said to be deterministic or nonambiguous for input, or output when literal, or irreducible, according to the same definitions as for automata.

Measures

Here we state some relevant facts about invariant measures on a compact metric space X endowed with a continuous transformation T ; examples are subshifts, or the 1-torus \mathbb{T}^1 endowed with multiplication by some integer p . For proofs see [DGS]. Recall the set $\mathcal{M}(X)$ of probability measures on X is a compact metric space for the topology of weak convergence, and the set $\mathcal{I}(X)$ of T -invariant probability measures on X is always nonempty and compact. An invariant measure μ on (X, T) is such that $\mu(T^{-1}(E)) = \mu(E)$ for any Borel set E . The (topological) *support* of an invariant measure μ on X is the intersection of all closed invariant subsets of X having measure

1; its *entropy* is the nonnegative number

$$h_\mu = \lim_{n \rightarrow \infty} -\frac{1}{n} \sum_{c \in C_n} \mu(c) \log \mu(c)$$

where $C_n = \{[u], u \in L(X), |u| = n\}$ (the limit is known to exist). A subshift for which there is exactly one invariant measure with maximal entropy is said to be *intrinsically ergodic*.

For any subshift X , $x \in X$ and $f \in C(X)$, define the measure $S_n(x)$ by the formula

$$S_n(x, f) = \frac{1}{n} \sum_{i=1}^{n-1} \delta_{\sigma^i x}(f).$$

A point $x \in X$ is said to be *generic with regard to the invariant measure μ on X* , or simply *μ -generic*, if $S_n(x)$ converges weakly to μ as n goes to infinity. As genericity is an asymptotic property, depending only on nonnegative coordinates of $x \in X$, it is perfectly defined for sequences in $A^{\mathbb{N}}$. When $X = A^{\mathbb{N}}$ (resp. \mathbb{T}^1) and μ is the Bernoulli measure λ with probability $1/\#A$ for each symbol (resp. Lebesgue), a generic x is called *normal to the base p* .

Define $\mathcal{M}(x)$ as the set of measures *associated to x* , or limits, for the weak convergence of measures, of $(S_n(x))_{n \in \mathbb{N}}$. Compactness implies it is always nonempty. Any measure in $\mathcal{M}(x)$ is invariant; x is generic for μ if and only if $\mathcal{M}(x) = \{\mu\}$. Given some metric generating the topology of weak convergence, and $\delta > 0$, a point x is said to be *(δ, μ) -generic* if $\mathcal{M}(x) \subset B(\mu, \delta)$, (the open ball with centre μ and radius δ); when μ is the uniform measure this is equivalent to (k, ε) -normality for some k and ε depending only on δ (see [BIDT], Section 3]).

Assuming $\phi : Y \rightarrow X$ to be a factor map, denote by $\Phi : \mathcal{M}(Y) \rightarrow \mathcal{M}(X)$ the corresponding map for measures: $\Phi(\nu) = \nu \circ \phi^{-1}$ (generally, for any factor map from Y to X , denoted by a small Greek letter, let the corresponding continuous shift-commuting map from $\mathcal{M}(Y)$ to $\mathcal{M}(X)$ be denoted by the corresponding capital Greek letter). Map Φ is weakly continuous and shift-commuting. As a consequence, $\mathcal{M}(\phi x) = \Phi(\mathcal{M}(x))$ and if $y \in Y$ is ν -generic, $\phi(y)$ is $\Phi(\nu)$ -generic. A *lift* of measure μ on X is a measure ν on Y such that $\Phi(\nu) = \mu$. When ϕ is bounded-to-one, Φ preserves entropy.

There is no essential difference between the theory of invariant measures on subshifts of $A^{\mathbb{Z}}$ and subshifts of $A^{\mathbb{N}}$. Definitions are identical, so there

is a one-to-one correspondence between invariant, or ergodic, measures on a simply infinite subshift and its doubly infinite version. Properties are the same. For instance, genericity may be considered as a property of a doubly infinite sequence, but it depends only on its restriction to positive coordinates. The only (slight) difficulty arises when trying to find preimages of some simply infinite sequence under the tower construction, because in most cases there are several solutions. This will be dealt with in time.

2. Some general statements on genericity preservation

Going from base p^k to base p is done in three steps, as explained in the introduction. So is any more general coding. After describing each step, the corresponding theorems on genericity preservation are stated (when previously known) or proved (when not).

Step 1: the tower construction. In this subsection we deal only with the formal construction. For instance the actual meaning of the function f as regards coding is only explained further on.

Put $Z = A^{\mathbb{Z}}$ for some finite A . Let $f : Z \rightarrow \mathbb{N}^*$ be a continuous map (i.e. for any integer n the set $\{z \in Z / f(z) = n\}$ is closed, which implies of course that f is bounded by some integer k). Let us construct the *tower system* (Z_f, τ) over (Z, σ) under function f , as in [PS]: define

$$Z_f = \{(z, i) / z \in Z, 0 < i \leq f(z)\};$$

put

$$\begin{aligned} \tau(z, i) &= (z, i + 1) \text{ if } i + 1 \leq f(z), \\ \tau(z, i) &= (\sigma z, 1) \text{ otherwise.} \end{aligned}$$

Z_f , endowed with the restriction of the product topology, is a compact metric set and τ is a homeomorphism.

Suppose μ is an invariant measure on (Z, σ) . It is well-known that the measure $\rho_f(\mu)$ defined by

$$(1) \quad \rho_f(\mu)(B) = (\mu(f))^{-1} \cdot \sum_{i \leq k} \mu\{z \in Z / (z, i) \in B\}$$

is τ -invariant on Z_f , and if μ is ergodic, so is $\rho_f(\mu)$; the map ρ_f is 1-to-1 and onto from $\mathcal{I}(Z)$ to $\mathcal{I}(Z_f)$, and Abramov's formula

$$h(\rho_f(\mu)) = \frac{h(\mu)}{\mu(f)}$$

links the entropies of the two measures [A]. None of these facts depends on the topology; all are also true in a purely metric situation.

PROPOSITION 1. *The 1-to-1 map ρ_f is continuous (hence bicontinuous) for the topology of weak convergence of measures.*

Proof. Let $\mu_n \in \mathcal{M}(Z)$ tend weakly to μ as $n \rightarrow \infty$: this means for any continuous function g on Z , $\mu_n(g) \rightarrow \mu(g)$. We wish to check that for any continuous h on Z_f , $\rho_f(\mu_n)(h) \rightarrow \rho_f(\mu)(h)$. Write $h = h_1 + h_2 + \dots + h_k$, where h_i is the (continuous) restriction of h to the set $\{(z, i) \in Z_f\}$. Define $h'_i : Z \rightarrow \mathbb{R}$ by $h'_i(z) = h_i(z, i)$ whenever $f(z) \geq i$, $h'_i(z) = 0$ otherwise. By definition of h_i and (1) one can write

$$\rho_f(\mu_n)(h) = \sum_i \rho_f(\mu_n)(h_i) = (\mu_n(f))^{-1} \cdot \sum_i \mu_n(h'_i).$$

Applying the hypothesis of weak convergence to each continuous function h'_i and f , the last expression goes to $(\mu(f))^{-1} \cdot \sum_i \mu(h'_i)$, which, by (1), is equal to $\rho_f(\mu)(h)$. \square

Now here is the basic tool for all further results.

PROPOSITION 2. *Suppose $z \in Z$, $y = (z, i) \in Z_f$. Then one has $\mathcal{M}(y) = \rho_f(\mathcal{M}(z))$.*

Proof.

1) We prove first the implication $(\mu \in \mathcal{M}(z)) \Rightarrow (\rho_f(\mu) \in \mathcal{M}(y))$. As μ is assumed to be associated to z , there is an infinite subset E of \mathbb{N} such that for any continuous g on Z ,

$$S_n(z, g) = \frac{1}{n} \sum_{i=0}^{n-1} g(\sigma^i z)$$

tends to $\mu(g)$ as n tends to infinity along E . We want to compute

$$S_N(y, h) = \frac{1}{N} \sum_{i=0}^{N-1} h(\tau^i y)$$

for continuous h , as N tends to infinity along some subset E' . As $\mathcal{M}(y)$ does not change under action of τ , it is sufficient to compute $S_N(y, h)$ for $y = (z, 1)$.

For $z \in Z$, define $f_z(n) = \sum_{0 \leq i < n} f(\sigma^i z)$, and $E' = \{f_z(n), n \in E\}$; E' is also an infinite subset of \mathbb{N} , and for given z , $f_z(\cdot)$ defines a 1-to-1 onto map from E to E' .

For continuous $h : Z_f \rightarrow \mathbb{R}$, $0 < i \leq k$, define continuous functions $h_i : Z_f \rightarrow \mathbb{R}$ and $h'_i : Z \rightarrow \mathbb{R}$ as in the proof of Proposition 1. For $N \in E'$, one has

$$S_N(y, h) = \frac{1}{N} \sum_{j=0}^{f_z^{-1}(N)-1} \sum_{i=1}^k h_i(\tau^j z).$$

By permuting the two summations, multiplying and dividing by $f_z^{-1}(N)$, one gets

$$S_N(y, h) = \frac{f_z^{-1}(N)}{N} \sum_{0 < i \leq k} S_N(z, h_i).$$

Putting $f_z^{-1}(N) = n$, since f is continuous and bounded,

$$\frac{N}{f_z^{-1}(N)} = \frac{f_z(n)}{n} = \frac{1}{n} \sum_{0 \leq i < n} f(\sigma^i z)$$

tends to $\mu(f)$ as N tends to infinity along E' , so that $f_z^{-1}(N)/N = n/f_z(n)$ tends to $(\mu(f))^{-1}$. Apply the hypothesis to each continuous function h_i , and then (1):

$$\lim_{N \in E'} S_N(y, h) = (\mu(f))^{-1} \sum_{0 \leq i < k} \mu(h'_i) = \sum_{0 \leq i < k} \rho_f(\mu)(h_i) = \mu_f(h).$$

2) The converse is proved in a similar way. Given $\nu \in \mathcal{I}(Z_f)$, there exists a unique invariant measure μ on Z such that $\nu = \rho_f(\mu)$. Suppose ν is associated to some $(z, i) \in Z_f$: for the same reason as above assume $i = 1$. The convergence of $S_N(y, h)$ to ν occurs on some subset E' . Without loss of generality one may assume $E' \subset f_z(\mathbb{N})$: since f is bounded, for any $p \in E'$ there is $q \in f_z(\mathbb{N})$ with $0 \leq p - q < k$. For given h the difference $|S_p(y, h) - S_q(y, h)|$ is bounded by $k \cdot \sup_{z \in Z} \{|h(z)|\}/q$, so that $S_q(y)$ tends weakly to ν as well as $S_p(y)$.

We claim that for any continuous g , $S_n(z, g) \rightarrow \mu(g)$ as n goes to infinity along $E = f_z^{-1}(E')$. Define continuous functions g' and $f' : Z_f \rightarrow \mathbb{R}$ by $g'(z, 1) = g(z)$, $g'(z, i) = 0$ for $i > 1$, $f'(z, 1) = 1$, $f'(z, i) = 0$ for $i > 1$. One has

$$S_n(z, g) = \frac{1}{n} \sum_{0 \leq i < n} g(\sigma z) = \frac{1}{f_z^{-1}(N)} \sum_{0 \leq i < N} g' \circ \tau^i(z, 1),$$

whence

$$\lim_{n \in E} S_n(z, g) = \lim_{N \in E'} \frac{N}{f_z^{-1}(N)} \cdot \lim_{N \in E'} \frac{1}{N} \sum_{0 \leq i < N} g' \circ \tau^i(z, 1);$$

as g' is continuous, by hypothesis the second limit in the product is equal to $\nu(k) = \mu(g)/\mu(f)$; f_z being defined as above, as

$$\frac{f_z^{-1}(N)}{N} = \frac{1}{N} \sum_{0 \leq i < N} f' \circ \tau^i(z, 1)$$

and f' is continuous, by (1)

$$\lim_{n \in E} \frac{f_z^{-1}(N)}{N} = \nu(f') = \frac{\mu(1)}{\mu(f)} = (\mu(f))^{-1},$$

so $\lim_{n \in E} S_n(z, g) = \mu(g)$. Q.E.D. \square

A direct consequence of Propositions 1 and 2 is the following:

COROLLARY 3. *Given $\varepsilon > 0$, there exists $\delta > 0$ such that $\mathcal{M}(z) \in B(\mu, \delta)$ implies $\mathcal{M}(z, i) \in B(\rho_f(\mu), \varepsilon)$ and conversely $\mathcal{M}(z, i) \in B(\rho_f(\mu), \delta)$ implies $\mathcal{M}(z) \in B(\mu, \varepsilon)$.*

When ε vanishes, this becomes a statement on genericity preservation.

Step 2: writing the letters

This is by far the simplest step.

Suppose now the function f above is generated in the following way. Let α be a map from A to B^* , where B is some finite alphabet; α can be easily extended into a map from A^* to B^* , but not into a shift-commuting map from $A^{\mathbb{Z}}$ to $B^{\mathbb{Z}}$: this is why one has to build the tower. Define f on $Z = A^{\mathbb{Z}}$ by $f(z) = |\alpha(z_0)|$. We want to identify (Z_f, τ) with some convenient symbolic space Y , under the assumption that α is 1-to-1.

Let $X = X_{\alpha(A)}$ be the set of all points $x \in B^{\mathbb{Z}}$ for which there exists a sequence $t \in \{0, 1\}^{\mathbb{Z}}$ such that $t_i = t_j = 1$ implies $x(i, j - 1) \in (\alpha(A))^*$; t is called a *scansion* of x into words of $\alpha(A)$. ϕ and ϕ' being the natural projections from $(B \times \{0, 1\})^{\mathbb{Z}}$ on $B^{\mathbb{Z}}$ and $\{0, 1\}^{\mathbb{Z}}$, define a subshift Y on the alphabet $B \times \{0, 1\}$ by the following condition:

for any $y \in Y$, $\phi(y) \in X$, and $\phi'(y)$ is a scansion of $\phi(y)$ into words of $\alpha(A)$.

For $a \in A$, write $\alpha(a) = \alpha_1(a)\alpha_2(a) \cdots \alpha_{|\alpha(a)|}(a)$, $\alpha_i(a) \in B$. Define the map $\xi : Z_f \rightarrow Y$ by the two conditions:

(2) $(\xi(z, 1))_0 = (\alpha_1(z_0), 1)$; $(\xi(z, i))_0 = (\alpha_i(z_0), 0)$ for $1 < i \leq |\alpha(z_0)|$;

$$(3) \quad \xi \circ \tau = \sigma \circ \xi.$$

This definition consists of replacing each letter z_n in z by the corresponding word $\alpha(z)$, for which Z_f provides convenient room owing to the choice of f , while keeping trace of the beginning of all words $\alpha(z_i)$. ξ is obviously continuous and onto, and by (2) it changes τ to σ ; it is 1-to-1 if and only if α is, which means that in this case ξ is a conjugacy.

Now let us examine the action of ξ (and ξ^{-1}) on points that are generic, or almost generic, for some given measure.

That ξ preserves genericity and almost genericity is a straightforward consequence of its being a factor map.

PROPOSITION 4. *Let π be a factor map from the compact metric space Y to the compact metric space X . For any $\varepsilon > 0$ there exists $\delta > 0$ such that for any invariant measure ν on Y , $y \in Y$, if $\mathcal{M}(y) \in B(\nu, \delta)$, then $\mathcal{M}(\pi(y)) \in B(\Pi(\nu), \varepsilon)$.*

The proof is elementary and left to the reader. As a consequence ξ preserves genericity and, assuming α to be 1-to-1, so does ξ^{-1} .

Step 3: dropping or restoring the scansion

Dropping the scansion only means applying the factor map ϕ ; by Proposition 4 this map preserves genericity. But the converse is far from being always true: some extra conditions are needed. Here is the general ergodic statement we are going to apply: it is an easy consequence of Proposition 3.3 in [BIDT].

PROPOSITION 5. *Suppose $\phi : Y \rightarrow X$ is a factor map, λ is an invariant measure on X such that there is a unique invariant measure ν on Y with $\Phi(\nu) = \lambda$. For any $\varepsilon > 0$, there is $\delta > 0$ such that if $\mathcal{M}(x) \subset B(\lambda, \delta)$ for some $x \in X$, then for any $y \in Y$ with $\phi(y) = x$, $\mathcal{M}(Y) \subset B(\nu, \varepsilon)$.*

Of course one wants sufficient conditions on μ for uniqueness of its lift ν . The next, still general, remark settles the uniform case:

Remark 6. When Y and X are the extension and factor subshift of a nonambiguous irreducible automaton, and μ is the unique measure with maximal entropy on X , then its unique lift on Y is the measure ν with maximal entropy. This is a mere consequence of intrinsic ergodicity of sofic systems and entropy preservation by bounded-to-one factor maps.

A more general sufficient condition for an invariant measure to have a unique invariant preimage under some nonambiguous automaton is given in [BIDT, Propositions 2.7 and 4.9]. And now, how can we use Remark 6 when introducing all possible scansions of $x \in X$? This is done by introducing an important further combinatorial assumption on α : $\alpha(A)$ must be a code.

DEFINITION. $\Gamma \subset B^*$ is said to be a code if

$$x_1x_2 \cdots x_m = y_1y_2 \cdots y_n, \text{ with } x_i, y_j \in \Gamma$$

implies $m = n$ and $x_i = y_i, 1 \leq i \leq n$.

Many finite and infinite codes are described in [BeP]; the ones used in Section 3 and Section 4 are elementary.

When $\Gamma = \alpha(A)$ is a code, one usually wishes α to define a 1-to-1 morphism from A^* into B^* , so as to make it possible to tell what element of A^* any word $u \in \Gamma^*$ is the image of; for this to be true one must also assume α to be 1-to-1.

The *petal automaton* \mathcal{A}_Λ of a language $\Lambda = \{x_1, x_2, \dots, x_s\}$ has state set $\{\varepsilon\} \cup \{(x_i, j) / 1 \leq j < |x_i|, x_i \in \Lambda\}$. There are arcs

- (i) from ε to $(x_i, 1)$ with label $x_i(1), x_i \in \Lambda$;
- (ii) from (x_i, n) to $(x_i, n + 1)$ with label $x_i(n + 1)$ for $n + 1 < |x_i|$;
- (iii) from $(x_i, |x_i| - 1)$ to ε with label $x_i(|x_i|)$.

This automaton is obviously irreducible. When ε is the only initial and final state, \mathcal{A}_Λ recognizes Λ^* ; when all states are initial and final, it recognizes the set $F(\Lambda^*)$ of “pieces” of words of Λ^* .

PROPOSITION 7 [BP]. *The petal automaton of a (finite) language Λ on B is irreducible. It is nonambiguous if and only if Λ is a code.*

In fact, if Λ is not a code, there exists no nonambiguous automaton permitting to obtain the scansions of words belonging to $F(\Lambda^*)$.

Here is now the theorem we have been aiming at; it just collects our previous results.

PROPOSITION 8. *Let μ be an invariant measure on Z , $\alpha : A \rightarrow B^*$, let f be a continuous map from Z to \mathbb{N} , ξ, ϕ the two factor maps defined above, and finally let $\delta > 0$. There exists $\delta' > 0$ such that if z is a (δ', μ) -generic point of Z , then if $x = \phi \circ \xi(z, i)$ for some i , it is a (δ, λ) -generic point*

for $\lambda = \Phi \circ \Xi(\mu_f)$. Conversely assume α is 1-to-1 and $\alpha(A)$ is a code, and $\Phi^{-1}(\lambda)$ is reduced to a unique invariant measure ν . If μ is the unique measure on Z corresponding to $\Phi^{-1}(\lambda)$, then for any (δ, λ) -generic $x \in X$, if $x = \phi \circ \xi(z, i)$, z is (δ', μ) -generic.

Proof. Starting from μ on Z it is enough to apply the classical properties of towers and the direct part of Corollary 3, and Proposition 4, in order to obtain the direct result. Conversely, there is a unique invariant measure ν on Y such that $\Phi(\nu) = \lambda'$; apply Proposition 7 to Y and X , which exactly coincide with the extension and factor subshifts of petal automaton $\mathcal{A}_{\alpha(A)}$ of code $\alpha(A)$; then Proposition 5 to λ , and finally Proposition 4 and the converse part of Corollary 3. \square

Remark 9. For number-theoretic applications, one wants *one-sided* infinite sequences. In order to obtain the same results in this case, just remark that

- (a) given an invariant μ on $A^{\mathbb{N}}$, there is a unique invariant measure μ' on $A^{\mathbb{Z}}$ such that its restriction to the σ -algebra of future events is equal to μ .
- (b) any point x in some subshift of $A^{\mathbb{N}}$ can be extended into at least one point x' in the corresponding subshift of $A^{\mathbb{Z}}$.
- (c) x is μ -generic if and only if x' is μ' -generic.

Thus any question about genericity in $A^{\mathbb{N}}$ can be carried over to $A^{\mathbb{Z}}$ and solved there; all former statements may be applied in $A^{\mathbb{N}}$.

3. Refining and extending Maxfield's theorem

Let us denote a measure on \mathbb{T}^1 , invariant by multiplication by p , by a starred Greek letter, for instance μ^* , and the almost always unique corresponding measure on $\{0, \dots, p - 1\}$ by the same Greek letter without a star: μ .

The next proposition states that δ -normality to the base p is equivalent to δ' -normality to the base p^k for some δ' . The first part, about preservation of normality, is the well-known Maxfield theorem [M]; another ergodic proof can be deduced from the main result in [BrL]. The remainder is mostly original, though J.-M. Dumont [D] has an unpublished combinatorial proof of the converse part, which may also be deduced from [C]. However, the present proof establishes clearly, in the line of [BIDT], that such results are within easy reach of Ergodic Theory, combined with elementary techniques of the Theory of Automata.

PROPOSITION 10. *For any two integers $p \geq 2$, $k \geq 2$, and $r \in (0, 1)$, r is normal to the base p if and only if it is normal to the base p^k . Moreover, for $\delta > 0$, there exists $\delta' > 0$ such that if r is δ' -normal to the base p (resp. p^k), then r is δ -normal to the base p^k (resp. p).*

Proof. Put $A = \{0, \dots, p^k - 1\}$, $B = \{0, \dots, p - 1\}$. Since the Lebesgue measure corresponds to only one measure on $A^{\mathbb{N}}$ or $B^{\mathbb{N}}$, we can work directly on expansions; and according to Remark 9 all propositions on genericity of doubly infinite sequences may be applied to sequences in $A^{\mathbb{N}}$ or $B^{\mathbb{N}}$. To prove the statement we have to perform three tasks:

- first to see how the former construction (tower and factor map) can be applied to deduce the expansion of r in base p from its expansion in base p^k (and conversely);

- then to prove that the measure corresponding to the uniform measure on $A^{\mathbb{N}}$ under this construction is the uniform measure on $B^{\mathbb{N}}$, and conversely;

- finally, to check the hypotheses of Proposition 7 are fulfilled.

1) In order to obtain the expansion of r to base p from its expansion to base p^k , one has to replace each symbol a belonging to $A = \{0, \dots, p^k - 1\}$ by a word $\gamma(a)$ with length k on alphabet $B = \{0, \dots, p - 1\}$: $\gamma(a)$ is the expansion of a in base p , with a suitable number of 0's added in the beginning when necessary. Obviously γ is 1-to-1 onto between A and B^k ; moreover $\gamma(A)$ is evidently a prefix code, therefore ϕ is the canonical factor map of a deterministic automaton.

2) Call μ the uniform measure on $Z = A^{\mathbb{N}}$, λ' the uniform measure on $X = B^{\mathbb{N}}$. We establish that $\lambda' = \Phi \circ \Xi \circ \rho_f(\mu)$, and also $\Xi \circ \rho_f(\mu)$ is the unique element of $\Phi^{-1}(\lambda')$. Starting from μ on $A^{\mathbb{N}}$ one has $f = k$ everywhere, so by Abramov's formula $h(\rho_f(\mu)) = k^{-1}h(\mu) = \log p = h(\lambda')$. As they are associated to bounded-to-one factor maps, Ξ and Φ preserve entropy: since α is 1-to-1, Ξ is a conjugacy; since $\gamma(A)$ is a code, Φ is bounded-to-one. As $\lambda = \Phi \circ \Xi(\rho_f(\mu))$ is the ultimate image of μ on X , $h'(\lambda') = h(\lambda)$; but as λ' is the unique measure on X with entropy $\log p$, one has $\lambda' = \lambda$. Conversely, there is a unique invariant measure ν on Y such that $\Phi(\nu) = \lambda'$: since ϕ is the canonical factor map of a nonambiguous automaton it is bounded-to-one, thus $\Phi(\nu) = \lambda'$ implies $h(\nu) = h(\lambda') = \log p$; as the automaton is irreducible Y is intrinsically ergodic, and ν is unique. As $\Xi^{-1}(\nu)$ also has entropy $\log p$, by Abramov's formula the corresponding measure on Z has entropy $k \cdot \log p$ and so must be equal to μ .

3) The direct part of Proposition 8 only requires that f be continuous, which is evident. For the reverse part, it is sufficient to remark γ is 1-to-1 and also a prefix code, so that, by Proposition 7 and Remark 6, $\phi^{-1}(\mu)$ is a singleton. \square

Now let us drop uniform measures, and suppose $r \in \mathbb{T}^1$ is generic for some $\times p^k$ -invariant measure μ^* (in order to avoid any trouble when passing to the expansions, assume μ is nonatomic). By using Proposition 7 it is easy to prove r is also generic for some corresponding $\times p$ -invariant λ^* . But, supposing one inverts the roles of $\times p$ and $\times p^k$, when is the corresponding statement true ?

An answer to this question can be obtained in particular cases, using some results in [BIDT]. For instance:

PROPOSITION 11. *Suppose the measure λ^* on \mathbb{T}^1 is the image of an ergodic Markov measure λ with topological support $B^{\mathbb{N}}$. Then there is a unique measure ν on Y such that $\Phi(\nu) = \lambda$; thus there is a measure μ^* on \mathbb{T}^1 such that if x is generic for λ^* under $\times p$, it is generic for μ^* under $\times p^k$.*

Proof. The hypothesis implies λ is unique, and by [BIDT, Propositions 2.7 and 4.9] that ν is also unique, thus defining unique measures μ on $A^{\mathbb{N}}$ and μ^* on \mathbb{T}^1 . By Proposition 8 one gets the expected results on preservation of genericity. \square

4. Analogous results for expansions with respect to the golden mean

Given a real number $\beta > 1$, not in \mathbb{N} , a sequence $(s_i, i > 0)$ of symbols in $A = \{0, 1, \dots, [\beta]\}$ is said to be an *expansion* of $r \in \mathbb{T}^1$ to the base β if its *valuation* $V(s)$ checks

$$V(s) := \sum_{i>0} s_i \cdot \beta^{-i} = r.$$

When β is not an integer there are many possible expansions of any real number in \mathbb{T}^1 to the base β . One of these expansions, the greatest for lexicographical order, plays a special role and may be called the canonical one. The matter of β -expansions has been given some attention in the literature, and an analogue to Maxfield's theorem was proved by A. Bertrand-Mathis when β is a Pisot number [Be]. We want to show on a very particular

example how to use transducers and ergodic techniques to get the same result, by the way obtaining the same refinements as for expansions to integer bases in Section 3.

Consider the case $\beta = \theta = (1 + \sqrt{5})/2$: then the canonical expansion of r is the only one in which the word 11 never occurs; the set of all such expansions is the subshift of finite type S_θ defined by this exclusion rule. The canonical expansions to the base $\zeta = \theta^2$ are sequences in which there is at least one occurrence of 0 between two occurrences of 2. They form a sofic system S_ζ . By erasing the output labels in the graph of \mathcal{T}_1 below, one obtains an automaton recognizing S_ζ . On the sets S_θ and S_ζ , there is a unique measure having maximal entropy (equal to $\log \theta$ and $\log 2\theta$ respectively), which is called the Parry measure; its role corresponds to that of the uniform measure with regard to integer bases.

Given $r \in \mathbb{T}^1$, and its canonical expansion s' to the base ζ , we want to compute its canonical expansion $s \in S_\theta$ to the base θ . This has to be done by combining the action of two transducers: on account on their acting in opposite directions, it is impossible to merge them into just one [F1]. The first one, \mathcal{T}_1 , performs two tasks: creating one empty cell between any two symbols of s' , belonging to alphabet $\{0, 1, 2\}$, and then, using the additional room, rewriting these symbols on alphabet $\{0, 1\}$; the sequence s'' thus obtained has valuation r and is on alphabet $A = \{0, 1, \dots, [\beta]\}$ but generally does not satisfy the constraint that 11 must never occur. The second transducer \mathcal{T}_2 is literal; it eliminates all blocks 11 in s' , thus yielding s . A simpler transducer might be used in order to do just this, as in [F1]; but, in order to use the various results of Section 2, we want \mathcal{T}_2 to be nonambiguous, which means its input subshift must be equal to the output subshift of \mathcal{T}_1 ; this makes it more complicated. This also illustrates the kind of difficulties one runs into when using transducers in the field of β -expansions.

Transducer \mathcal{T}_1 :

This transducer acts from left to right; the input words (on alphabet $\{0, 1, 2\}$) are exactly those belonging to S_ζ . Any symbol in $\{0, 1, 2\}$ is replaced by a two-letter word depending on the state the transducer is in. Initially a cell containing a 0 is created between any two symbols in s' ; then any symbol 2 is subtracted and replaced by adding two 1's, one in the first cell to the left and one in the second cell to the right, according to identity $2 = \theta + \theta^{-2}$ (thus $V(s'') = V(s)$). The cell to the left formerly contained a 0; but the second one to the right may have contained a 1, which creates

a new 2, and this must be eliminated in its turn (this is done by the loop with labels 1/10 on state b). Its graph appears in Figure 1.

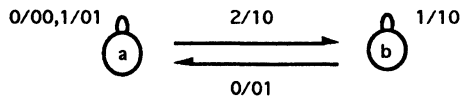


Figure 1

The input subshift X_1 is exactly S_C . The output subshift X'_1 has two relevant features: 1^3 never occurs, and two occurrences of 11 must be separated by at least one occurrence of 0^2 . Figure 2 below pictures an automaton recognizing it.

It is important to notice \mathcal{T}_1 performs the three operations described in Section 2 at the same time. In order to apply previous results it may be thought of as the composition of two transducers. Put $u = 00$, $v = 01$, $w = 10$, and replace output words in \mathcal{T}_1 by these new symbols: one thus obtains a literal, deterministic transducer \mathcal{T}'_1 with the same graph, but with output alphabet $\{u, v, w\}$. The second one, \mathcal{T}''_1 , has only one state and replaces letters of $\{u, v, w\}$ by their value on alphabet $\{0, 1\}$. It is a typical tower transducer, performing steps 1, 2 and 3 of the Introduction.

Remark. it is important to notice $\{u, v, w\}$ is a code.

Transducer \mathcal{T}_2 :

This one acts from right to left. Before constructing it one should build up an automaton recognizing the output subshift of \mathcal{T}_1 , X'_1 , from right to left (this makes it easier to check that the input subshift of \mathcal{T}_2 is equal to the output subshift of \mathcal{T}_1). One such, deterministic, automaton is represented in Figure 2.

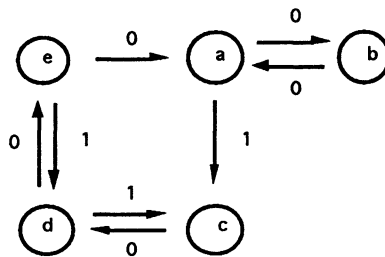


Figure 2

Now this is how \mathcal{T}_2 acts. Everytime the word 11 occurs in the input sequence s'' , the two 1's are replaced by 0's and a 1 is created to the left, according to the identity $1 = \theta^{-1} + \theta^{-2}$ (which ensures $V(s') = V(s'') = V(s)$). As X'_1 , the output subshift of \mathcal{T}_1 , never contains 1^3 or 1^201^2 this never creates 2's, or words 1^n , $n > 2$; when a new word 11 is created to the left of the one just eliminated it must be dealt with in its turn: this is achieved by the loop on states g and h in the graph. Its graph is featured out in Figure 3.

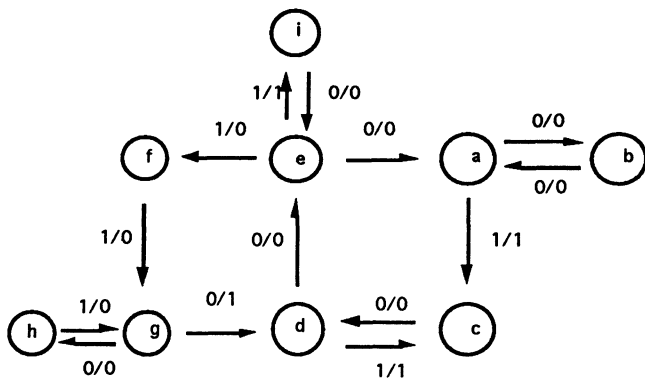


Figure 3

Finally, \mathcal{T}_2 is nonambiguous for input and output.

Let μ_θ and μ_ζ be the Parry measures on \mathbb{T}^1 corresponding to θ and ζ . There only remains to apply results of Section 2 to the transducers \mathcal{T}'_1 , \mathcal{T}''_1 , and \mathcal{T}_2 to prove the following

PROPOSITION 12. *For any $\delta > 0$, there exists $\delta' > 0$ such that the two following properties are equivalent :*

- (1) $r \in \mathbb{T}^1$ is (μ_θ^*, δ) -generic under multiplication by θ .
- (2) r is (μ_ζ^*, δ') -generic under multiplication by ζ .

Proof. It is closely related to the proof of Proposition 10. First one reformulates the problem into its symbolic equivalent on S_θ and S_ζ . Then one has to identify measures μ_θ and μ_ζ as corresponding to each other under action of the transducers \mathcal{T}_1 and \mathcal{T}_2 : this is an easy job with the use of entropy. Finally apply Proposition 8 to \mathcal{T}''_1 (this is possible since $\{u, v, w\}$ is a code) and Propositions 4 and 6 to \mathcal{T}'_1 and \mathcal{T}_2 to reach the conclusion. \square

Remarks.

1- Recent work by C. Frougny [F2] and by C. Frougny and B. Solomyak [F3] suggests this proof may be generalised to all Pisot numbers τ such that 1 has a finite decreasing canonical expansion to base τ [F2].

2- The same set of results suggests there are other measures for which genericity is preserved.

REFERENCES

- [A] L. M. ABRAMOV, *The entropy of a derived automorphism*, Amer. Math. Soc. Transl. **49** (1965), 162–166.
- [BP] J. BERSTEL, D. PERRIN, *Theory of codes*, London, Academic Press, 1985.
- [Be] A. BERTRAND-MATHIS, *Développements en base θ et répartition modulo 1 de la suite $(x\theta^n)$* , Bull. Soc. math. Fr. **114** (1986), 271–324.
- [BIP] F. BLANCHARD, D. PERRIN, *Relèvements d'une mesure ergodique par un codage*, Z. Wahrscheinlichkeist. v. Gebiete **54** (1980), 303–311.
- [BIDT] F. BLANCHARD, J.-M. DUMONT, A. THOMAS, *Generic sequences, transducers and multiplication of normal numbers*, Israel J. Math. **80** (1992), 257–287.
- [BrL] A. BROGLIO, P. LIARDET, *Predictions with automata*, Symbolic Dynamics and its Applications, AMS, providence, RI, P. Walters ed., Contemporary Math **135**, 1992.
- [C] J. W. S. CASSELS, *On a paper of Niven and Zuckerman*, Pacific J. Math. **3** (1953), 555–557.
- [D] J.-M. DUMONT, *Private communication*.
- [F1] C. FROUGNY, *Representations of numbers and finite automata*, Math. Systems Theory **25** (1992), 37–60.
- [F2] C. FROUGNY, *How to write integers in non-integer base*, Lecture Notes in Computer Science (Proceedings of Latin '92) **583**, 154–164.
- [F3] C. FROUGNY, B. SOLOMYAK, *Finite Beta-expansions*, Ergod. Th. & Dynam. Syst. **12** (1992), 713–723.
- [K] T. KAMAE, *Subsequences of normal sequences*, Israel J. Math. **16** (1973), 121–149.
- [M] J. E. MAXFIELD, *Normal k -tuples*, Pacif. J. Math. **3** (1953), 189–196.
- [PS] K. PETERSEN, L. SHAPIRO, *Induced flows*, Trans. Amer. Math. Soc. **177** (1973), 375–390.

François Blanchard
 LMD - Case 930
 163 avenue de Luminy
 F-13288 Marseille Cedex 9
 FRANCE