Journal of Statistical Physics, Vol. 76, Nos. 5/6, 1994

A Dimension Formula for Bernoulli Convolutions

François Ledrappier[1] and Anna Porzio[2]

Received October 25, 1993; final March 25, 1994

We present a "dynamical" approach to the study of the singularity of infinitely convolved Bernoulli measures ν_β, for β the golden section. We introduce ν_β as the transverse measure of the maximum entropy measure μ on the repelling set invariant for contracting maps of the square, the "fat baker's" transformation. Our approach strongly relies on the Markov structure of the underlying dynamical system. Indeed, if $\beta =$ golden mean, the fat baker's transformation has a very simple Markov coding. The "ambiguity" (of order two) of this coding, which appears when projecting on the line, due to passages for the central, overlapping zone, can be expressed by means of products of matrices (of order two). This product has a Markov distribution inherited by the Markov structure of the map. The dimension of the projected measure is therefore associated to the growth of this product; our dimension formula appears in a natural way as a version of the Furstenberg–Guivarch formula. Our technique provides an explicit dimension formula and, most important, provides a formalism well suited for the multifractal analysis of this measure, as we will show in a forthcoming paper.

KEY WORDS: Bernoulli convolutions; Hausdorff dimension; random matrices; Lyapunov exponent.

1. INTRODUCTION

1.1. The Problem

Let $\varepsilon_1, \varepsilon_2, \ldots$ be a sequence of independent random variables each taking the values $+1$ and -1 with equal probability. The probability distribution of the random variable $(1-\beta)\sum_{n=0}^{\infty} \varepsilon_n \beta^n$, $0<\beta<1$, defines a measure ν_β which is called an infinitely convolved Bernoulli measure or simply a Bernoulli convolution.[24, 14, 15, 20, 21, 1]

[1] Laboratoire de Probabilités, Université de Paris VI, 75005 Paris, France.
[2] CPTH Ecole Polytechnique, F 91128 Palaiseau UPR14 CNRS, France.

0022-4715/94/0900-1307$07.00/0

If for $\beta < 1/2$, ν_β has clearly a Cantor distribution, and for $\beta = 1/2$ the uniform (Lebesgue) distribution, for $\beta > 1/2$ it is a difficult, old, and not yet completely solved problem to decide on the nature of ν_β. It is known[24] that ν_β is continuous and always pure, i.e., either absolutely continuous or totally singular, and in 1939 P. Erdös proved the singular continuous nature of ν_β if β^{-1} is a Pisot number (i.e., an algebraic integer whose conjugates lie inside the unit circle). He also proved that for almost all β sufficiently close to 1, ν_β is absolutely continuous.[15]

Twenty years later, A. Garsia considered the entropy H_β^p of the distribution ν_β^p of the discrete random variable $(1-\beta)\sum_{n=0}^p \varepsilon_n \beta^n$. If $\lim_{p\to\infty} H_\beta^p = \inf_p (H_\beta^p/p) \equiv G(\beta)$ takes a value *below* $\log \beta^{-1}$, then ν_β is necessarily singular and this is the case for β^{-1} a Pisot number.[32]

The recent work of Alexander and Yorke[1] relates to dynamics this old arithmetic measure problem. They consider the map $(x, y) \in (-\infty, +\infty) \times [-1, +1] \to T_\beta(x, y)$:

$$T_\beta(x, y) = \begin{cases} \beta x + 1 - \beta, 2y - 1 & \text{if} \quad y \geqslant 0 \\ \beta x - (1 - \beta), 2y + 1 & \text{if} \quad y < 0 \end{cases} \tag{1.1}$$

For $\beta = 1/2$ this is the classical baker's transformation, for $\beta < 1/2$ the dissipative one. For $1/2 < \beta < 1$, T_β is the "fat baker's" transformation: the map is now *not* inversible, the attractor is the whole square $[-1, +1] \times [-1, +1]$, and it possesses a Sinai–Bowen–Ruelle measure whose transverse component is ν_β.

1.2. The Hausdorff and Information Dimension of a Measure

Recall that the Hausdorff dimension (HD) of a Borel probability measure μ on a compact metric space M is the HD of the smallest set of full measure: $HD(\mu) = \inf\{HD(Y), Y: \mu(Y) = 1, Y \subset M\}$. Young[36] proved that if μ is a Borel probability measure on a compact Riemannian manifold, and *if* μ a.e.

$$\lim_{\varepsilon \to 0} \frac{\log \mu(B_\varepsilon(x))}{-\log \varepsilon} = \alpha \tag{1.2}$$

$[B_\varepsilon(x)$ being a ε-ball centered in $x]$, then $HD(\mu) = \alpha$. In the dynamical system context (1.2) the limit exists.[1,36,26,19,17]

Alexander and Yorke prove that if β^{-1} is Pisot number, then $I_D(\nu_\beta) = G(\beta)/(-\log \beta)$ (the Garsia entropy invariant) and find numerically the value $I_D(\nu_\beta)$ for $\beta =$ golden mean.

Alexander and Zagier,[2] as we learned during the preparation of this paper, have now, by a completely different method, a theoretical entropy formula for β = golden mean which agrees with the empirical result of ref. 1. We also learned that Bovier[8] has another proof of the singularity of ν_β in that case.

1.3. Motivations and Method

This paper has a double motivation. First, until now rigorous results in (multi)fractal analysis ["singularity $f(\alpha)$ spectrum;" see, e.g., ref. 11) have been concerned with invariant sets which have essentially a Cantor-set structure. It was tempting to try to extend these ideas and techniques to the more realistic and complicated situations where more than one contracting direction is present and each interferes with the others. Second, the long history and multiple aspects of the problem of the nature of ν_β, and the effort to understand the obscure, fascinating papers of Garsia, led us to concentrate on this example, which offers a beautiful fusion of arithmetic and dynamical aspects. The aim of this paper is to prepare a method to describe in great detail the invariant measure in view of the multifractal analysis of the fat baker's transformation. The singularity of ν_β for β = golden mean is an old result. We give first an explicit (i.e., numerically computable) theoretical formula for the dimension of ν_β in this nice case.

Our approach is a dynamical system approach; we introduce ν_β as the transverse measure of the maximum entropy measure μ on the repelling set invariant for the contracting maps of the square $F_0^{-1} = (\beta x, y/2)$ and $F_1^{-1} = (\beta x + 1 - \beta, (y+1)/2)$.

By refs. 27 and 28 we know that ν_β always satisfies (2), so that all notions of dimension coincide.

Our approach strongly relies on the Markov structure of the two-dimensional system (2.1). (This is the main difference with Garsia and all other works.) Indeed, if β = golden mean, the fat baker's transformation has a very simple Markov coding. The "ambiguity" (of order two) of this coding, which appears when projecting on the line, due to passages for the central, overlapping zone, can be expressed by means of products of matrices (of order two). This product has a Markov distribution inherited by the Markov structure of the underlying dynamical system (2.1). The dimension of the projected measure is therefore associated to the growth of this product; our dimension formula appears in a natural way as a version of the Furstenberg–Guivarch formula. The result of Young (2) ensures that this quantity gives actually the (information) dimension of the measure.

Observe that there are other random products of matrices which might naturally occur in this problem (R. Kenyon, Y. Peres, and S. Lalley, private communications).

It is very likely that our method may be extended to some families of

Pisot numbers. However, for practical purposes, the complexity of the method increases with the size of the matrices. Dimension(s) of measures which are concentrated on attractors is of course of special interest in dynamics. Unfortunately, the singularity of ν_β is more a beautiful arithmetic hazard than a physically relevant property: most of the fat baker's transformations are absolutely continuous. This has to be related to the[18] conjecture.[1,27,28]

2. THE SETTING

We consider the map $(x, y) \to F(x, y)$:

$$F(x, y) = \begin{cases} x/\beta, 2y & \text{if} \quad y \leqslant 1/2, \quad x \leqslant \beta \\ x/\beta - \beta, 2y-1 & \text{if} \quad y \geqslant 1/2, \quad x \geqslant 1-\beta \end{cases} \tag{2.1}$$

with $\beta + \beta^2 = 1$, with inverses

$$F_0^{-1}: \quad [0,1]\times[0,1] \to [0,\beta]\times[0,1/2], \qquad F_0^{-1}(x, y) = (\beta x, y/2)$$

$$F_1^{-1}: \quad [0,1]\times[0,1] \to [1-\beta,1]\times[1/2,1],$$

$$F_1^{-1}(x, y) = (\beta x + 1 - \beta, (y+1)/2)$$

The invariant (repelling) set is

$$X = \left\{ \bigcap_{i=1}^{\infty} \bigcup_{\{k_1, k_2, \dots, k_i\}} F^{-i}_{k_1 k_2 \dots k_i}[0,1]\times[0,1],\ k_i \in \{0,1\} \right\}$$

Let

$$A = [1-\beta, \beta]\times[1/2, 1]$$

$$B = [\beta, 1]\times[1/2, 1]$$

$$C = [0, 1-\beta]\times[0, 1/2]$$

$$D = [1-\beta, \beta]\times[0, 1/2]$$

Remark. If $\beta + \beta^2 = 1$, $\{A, B, C, D\}$ is a Markov partition. The compatibility rules are the following:

$$\begin{aligned} F(A \cap X) &= C \cap X \\ F(B \cap X) &= A \cup B \cup D \cap X \\ F(C \cap X) &= A \cup C \cup D \cap X \\ F(D \cap X) &= B \cap X \end{aligned} \tag{2.2}$$

That is, every point $(x, y) \in X$ is coded by a sequence $\underline{a}(x, y) = a_0 a_1 \dots$ with $a_i \in \{A, B, C, D\}$ such that $(x, y) \in a_0$, $F(x, y) \in a_1, \dots, F^n(x, y) \in a_n, \dots$, and conversely any compatible sequence $a_0 a_1 \ldots$ defines a unique point $(x, y) \in X$.

We describe now the invariant measure we select. $\forall I \in [0, 1] \times [0, 1]$ we set

$$\mu(F_0^{-1} I) = \tfrac{1}{2}\mu(I), \qquad \mu(F_1^{-1} I) = \tfrac{1}{2}\mu(I) \tag{2.3}$$

If we take $I = [0, 1] \times [0, 1]$, we have

$$\mu(A) + \mu(B) = \tfrac{1}{2}, \qquad \mu(C) + \mu(D) = \tfrac{1}{2} \tag{2.4}$$

For $I = A$, B, C, D Markov compatibility rules (2.2) give [where we denote $CA = \{(x, y) \text{ s.t. } (x, y) \in C, F(x, y) \in A\}$]

$$\begin{aligned} \tfrac{1}{2}\mu(A) &= \mu(CA) = \mu(BA) \\ \tfrac{1}{2}\mu(B) &= \mu(DB) = \mu(BB) \\ \tfrac{1}{2}\mu(C) &= \mu(CC) = \mu(AC) \\ \tfrac{1}{2}\mu(D) &= \mu(CD) = \mu(BD) \end{aligned} \tag{2.5}$$

[so that $\mu(AC) + \mu(BD) = 1/4$, etc.]

Observe that Markov rules (2.2) give also

$$\mu(AC) = \mu(A), \qquad \mu(DB) = \mu(D) \tag{2.6}$$

Equations (2.5) and (2.6) give $\frac{1}{2}\mu(C) = \mu(A)$ and $\frac{1}{2}\mu(B) = \mu(D)$, which, combined with (2.4), give also $\frac{1}{2}\mu(C) = \mu(D)$ and $\frac{1}{2}\mu(B) = \mu(A)$, and we conclude that $\mu(B) = \mu(C) = 1/3$, $\mu(A) = \mu(D) = 1/6$.

Observe now that, once the invariance formulas (2.3) and Markov compatibility rules (2.2) are stated, the measure of cylinders of bigger length can be computed, according to the usual rules, via the transition matrix P:

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Any of the 42^N compatible finite sequences $a_0 a_1 \ldots a_N$ of length N has measure $\mu = O(2^{-N})$ (see Section 4) and so this is the maximum entropy ($\log 2$) Markov invariant measure ($\mu P = \mu$).

3. PROJECTION RULES

It is possible, and also easier, to understand the distribution of points $(1 - \beta) \sum \varepsilon_n \beta^n$ on the line by looking at the two-dimensional system (2.1) and not just to its projection. Of course the Markov partition is not necessary for the understanding of the two-dimensional dynamics; it was introduced to set down a "dictionary" for projecting it on the line and vice versa.

Consider the β-adic expansion of $x \in [0, 1]$, $x = \sum_{i>0} \varepsilon_i \beta^i$, $\varepsilon_i \in \{0, 1\}$. If $\beta + \beta^2 = 1$, the expansion of 1 is $\underline{\varepsilon}(1) = 1100\ldots$, so that all admissible β-expansions of $x \in [0, 1]$ are the sequences $\underline{\varepsilon}(x) < \underline{\varepsilon}(1)$, that is, all sequences of 0 and 1 without two adjacent ones. This expansion is also unique up to periodic expansions.[31,5] We have to distinguish three cases.

(a) $0 < x < 1 - \beta$. Then $\underline{\varepsilon}(x) = 00\varepsilon_3\varepsilon_4\ldots$, so that $\underline{\varepsilon}(x/\beta^2) = \varepsilon_3\varepsilon_4\ldots$.

(b) $1 - \beta < x < \beta$. Then $\underline{\varepsilon}(x) = 010\varepsilon_4\varepsilon_5\ldots$, so that $\underline{\varepsilon}((x - \beta^2)/\beta^3) = \varepsilon_3\varepsilon_4\ldots$.

(c) $\beta < x < 1$. Then $\underline{\varepsilon}(x) = 10\varepsilon_3\varepsilon_4\ldots$, so that $\underline{\varepsilon}((x - \beta)/\beta^2) = \varepsilon_3\varepsilon_4\ldots$.

We come now to the Markov coding. We write $C^C_D = \{(x, y)$ s.t. $(x, y) \in C$, $F(x, y) \in C$ or $D\}$).

(a1) $0 < x < 1 - \beta$ and $y < 1/4$. We have $a_0 = C$ and $a_1 = C$ or D. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = C^C_D a_2 a_3 \ldots$. Expand its projection x: $\underline{\varepsilon}(x) = 00\varepsilon_3\varepsilon_4\ldots$, i.e., $\underline{\varepsilon}(x/\beta^2) = \varepsilon_3\varepsilon_4\ldots$. Also, $\underline{a}(F^2(x, y)) = a_2 a_3 \ldots$ and $F^2(x, \cdot) = (x/\beta^2, \cdot)$. This means that $a_2 a_3 \ldots$ projects on $x/\beta^2 = \varepsilon'_1 \beta + \varepsilon'_2 \beta^2 + \ldots$, that is, $\varepsilon'_1 = \varepsilon_3$, $\varepsilon'_1 = \varepsilon_4, \ldots$, etc. In conclusion, $a_2 a_3 \ldots$ projects on $\varepsilon_3\varepsilon_4\ldots$.

(a2) $0 < x < 1 - \beta$ and $1/4 < y < 1/2$. We have $a_0 = C$ and $a_1 = A$. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = CA a_2 a_3 \ldots$. On the other hand its projection x has expansion $\underline{\varepsilon}(x) = 00\varepsilon_3\varepsilon_4\ldots$, i.e., $\underline{\varepsilon}(x/\beta^2) = \varepsilon_3\varepsilon_4\ldots$. Also, $\underline{a}(F^2(x, y)) = a_2 a_3 \ldots$ and $F^2(x, \cdot) = (-\beta + x/\beta^2, \cdot)$. Markov rules force $a_2 = C$, i.e., $F^2(x, y) \in C$, so that $-\beta + x/\beta^2 < \beta^2$. It follows that $-\beta + x/\beta^2 = \varepsilon'_1 \beta + \varepsilon'_2 \beta^2 + \ldots$ with $\varepsilon'_1 = 0$, $\varepsilon'_2 = 0, \ldots$, i.e., $x/\beta^2 = \beta + \varepsilon'_3 \beta^3 + \ldots$. By comparing with $x/\beta^2 = \varepsilon_3 \beta + \varepsilon_4 \beta^2 + \varepsilon_5 \beta^3 + \ldots$, we have $\varepsilon_3 = 1$, $\varepsilon_4 = 0$, $\varepsilon_5 = \varepsilon'_5 \ldots$, etc. In conclusion, $a_2 a_3 \ldots$ projects on $10\varepsilon_5\ldots$.

(b1) $1 - \beta < x < \beta$ and $1/4 < y < 1/4 + 1/8$. We have $a_0 = D$, $a_1 = B$, and $a_2 = D$. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = DBD a_3 \ldots$ and its projection x: $\underline{\varepsilon}(x) = 010\varepsilon_4\ldots$, i.e., $\underline{\varepsilon}((x - \beta^2)/\beta^3) = \varepsilon_4\varepsilon_5\ldots$. One has $\underline{a}(F^3(x, y)) = a_3 a_4 \ldots$ and $F^3(x, \cdot) = ((x - \beta^2)/\beta^3 + \beta, \cdot)$. Markov rules force $a_3 = B$, i.e., $F^3(x, y) \in B$, so that $\beta + (x - \beta^2)/\beta^3 > \beta$. Expand $\beta + (x - \beta^2)/\beta^3 = \varepsilon'_1 \beta + \varepsilon'_2 \beta^2 + \varepsilon'_3 \beta^3 \ldots$ with $\varepsilon'_1 = 1$, $\varepsilon'_2 = 0, \ldots$, i.e., $(x - \beta^2)/\beta^3 = \varepsilon'_3 \beta^3 + \ldots$, and comparing with $(x - \beta^2)/\beta^3 = \varepsilon_4 \beta + \varepsilon_5 \beta^2 + \ldots$, we find $\varepsilon_4 = 0$, $\varepsilon_5 = 0, \ldots$, etc. In conclusion $a_3 a_4 \ldots$ projects on $00\varepsilon_6\ldots$.

(b2) $1 - \beta < x < \beta$ and $1/4 + 1/8 < y < 1/2$. We have $a_0 = D$, $a_1 = B$, and $a_2 = A$ or B. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = DB^A_B a_3 \ldots$ and its projection x: $\underline{\varepsilon}(x) = 010\varepsilon_4\ldots$, i.e., $\underline{\varepsilon}((x - \beta^2)/\beta^3) = \varepsilon_4\varepsilon_5\ldots$. Also,

$\underline{a}(F^3(x, y)) = a_3 a_4 \ldots$. On the other hand, $F^3(x, \cdot) = ((x-\beta^2)/\beta^3, \cdot)$. In conclusion $a_3 a_4 \ldots$ projects on $\varepsilon_4 \varepsilon_5 \ldots$.

(b3) $1-\beta < x < \beta$ and $1/2 < y < 1/2 + 1/8$. We have $a_0 = A$, $a_1 = C$, and $a_2 = C$ or D. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = AC^C_D a_3 \ldots$ and its projection x: $\underline{\varepsilon}(x) = 010\varepsilon_4 \ldots$, i.e., $\underline{\varepsilon}((x-\beta^2)/\beta^3) = \varepsilon_4 \varepsilon_5 \ldots$. Also, $\underline{a}(F^3(x, y)) = a_3 a_4 \ldots$ and $F^3(x, \cdot) = ((x-\beta^2)/\beta^3, \cdot)$. In conclusion $a_3 a_4 \ldots$ projects on $\varepsilon_4 \varepsilon_5 \ldots$.

(b4) $1-\beta < x < \beta$ and $1/2 + 1/8 < y < 1/4$. We have $a_0 = A$, $a_1 = C$, and $a_2 = A$. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = ACAa_3 \ldots$ and its projection x: $\underline{\varepsilon}(x) = 010\varepsilon_4 \ldots$, i.e., $\underline{\varepsilon}((x-\beta^2)/\beta^3) = \varepsilon_4 \varepsilon_5 \ldots$. Also, $\underline{a}(F^3(x, y)) = a_3 a_4 \ldots$ and $F^3(x, \cdot) = (-\beta + (x-\beta^2)/\beta^3, \cdot)$. Markov rules force $a_3 = C$, i.e., $F^3(x, y) \in C$, so that $\beta + (x-\beta^2)/\beta^3 = \varepsilon'_1 \beta + \varepsilon'_2 \beta^2 + \varepsilon'_3 \beta^3 \ldots$ with $\varepsilon'_1 = 0$, $\varepsilon'_2 = 0, \ldots$, and comparing with $(x-\beta^2)/\beta^3 = \varepsilon_4 \beta + \varepsilon_5 \beta^2 + \ldots$, we find $\varepsilon_4 = 1$, $\varepsilon_5 = 0, \ldots$, and in conclusion $a_3 a_4 \ldots$ projects on $10\varepsilon_6 \ldots$.

(c1) $\beta < x < 1$ and $1/2 < y < 1/2 + 1/4$. We have, $a_0 = B$ and $a_1 = D$. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = BDa_2 a_3 \ldots$ and its projection x: $\underline{\varepsilon}(x) = 10\varepsilon_3 \varepsilon_4 \ldots$, i.e., $\underline{\varepsilon}((x-\beta)/\beta^2) = \varepsilon_3 \varepsilon_4 \ldots$. We have $\underline{a}(F^2(x, y)) = a_2 a_3 \ldots$ and $F^2(x, \cdot) = (\beta + (x-\beta)/\beta^2, \cdot)$. Markov rules force $a_2 = B$, i.e., $F^2(x, y) \in B$, which means $\beta + (x-\beta)/\beta^2 = \varepsilon'_1 \beta + \varepsilon'_2 \beta^2 + \ldots$ with $\varepsilon'_1 = 1$, $\varepsilon'_2 = 0, \ldots$, i.e., comparing with $(x-\beta)/\beta^2 = \varepsilon_3 \beta + \varepsilon_4 \beta^2 + \ldots$, we have $\varepsilon_3 = 0$, $\varepsilon_4 = 0, \ldots$, and we conclude that $a_2 a_3 \ldots$ projects on $00\varepsilon_6 \ldots$.

(c2) $\beta < x < 1$ and $1/2 + 1/4 < y < 1$. We have $a_0 = B$ and $a_1 = A$ or B. Expand (x, y): $\underline{a}(x, y) = a_0 a_1 a_2 a_3 \ldots = B^B_A a_2 a_3 \ldots$. Expand its projection x: $\underline{\varepsilon}(x) = 10\varepsilon_3 \varepsilon_4 \ldots$, i.e., $\underline{\varepsilon}((x-\beta)/\beta^2) = \varepsilon_3 \varepsilon_4 \ldots$. We have $\underline{a}(F^2(x, y)) = a_2 a_3 \ldots$ and $F^2(x, \cdot) = ((x-\beta)/\beta^2, \cdot)$. This means that $a_2 a_3 \ldots$ projects on $\varepsilon_3 \varepsilon_4 \ldots$.

We summarize as follows (by * we mean the second or third iterate of F):

$\underline{a}(x, y)$	$\underline{\varepsilon}(x)$	$\underline{a}(F^*(x, y))$	$\underline{\varepsilon}(F^*_x(x, y))$
$C^C_D a_2 \ldots$	$OO\varepsilon_3 \ldots$	$a_2 a_3 \ldots$	$\varepsilon_3 \varepsilon_4 \ldots$
$CAa_2 \ldots$	$OO\varepsilon_3 \ldots$	$Ca_3 a_4 \ldots$	$1O\varepsilon_5 \ldots$
$DB^A_B a_3 \ldots$	$O1O\varepsilon_4 \ldots$	$a_3 a_4 \ldots$	$\varepsilon_4 \varepsilon_5 \ldots$
$DBDa_3 \ldots$	$O1O\varepsilon_4 \ldots$	$Ba_4 a_5 \ldots$	$OO\varepsilon_6 \ldots$
$AC^C_D a_3 \ldots$	$O1O\varepsilon_4 \ldots$	$a_3 a_4 \ldots$	$\varepsilon_4 \varepsilon_5 \ldots$
$ACAa_3 \ldots$	$O1O\varepsilon_4 \ldots$	$Ca_4 a_5 \ldots$	$1O\varepsilon_6 \ldots$
$B^A_B a_2 \ldots$	$1O\varepsilon_3 \ldots$	$a_2 a_3 \ldots$	$\varepsilon_3 \varepsilon_4 \ldots$
$BDa_2 \ldots$	$1O\varepsilon_3 \ldots$	$Ba_3 \ldots$	$O\varepsilon_4 \ldots$

We easily see from the above table that projection rules depend on the splitting of the sequences. In other words, the shift does not commute with the projection. We will be able nevertheless to use these rules to count how many and which Markov sequences have the same given projection. Indeed, even if they do not commute exactly, it is possible to set up a stationary context in which "things go as if they did." We return to this in Section 5.

4. THE MEASURE ν_β

Let $\Sigma(X)$ be the space of admissible sequences $a_1 a_2 \ldots$; $\Sigma_\beta([0, 1])$ the space of admissible sequences $\varepsilon_1 \varepsilon_2 \ldots$; $\Sigma_N(X)$ and $\Sigma_\beta^N([0, 1])$ the spaces of finite sequences of length N. We have constructed a map $\Phi: \Sigma(X) \to \Sigma_\beta([0, 1])$ [and its restriction $\Phi: \Sigma_N(X) \to \Sigma_\beta^N([0, 1])$], $\Phi(a_1 a_2 a_3 \ldots) = \varepsilon_1 \varepsilon_2 \varepsilon_3 \ldots$. Let $P_0 = \{A, B, C, D\}$, $P_N = F_{k_1 k_2 \ldots}^{-N} P_0$. There are 42^N rectangles of length $N+1$ in P_N (Markov rectangles). Let B_N be the partition of $[0, 1]$ obtained by projecting P_N on $[0, 1]$. There are $2F_{N+3}$ intervals in B_N, where $F_0 = 0$, $F_1 = 1$, $F_N + F_{N+1} = F_{N+2}$ are the Fibonacci numbers. They all have length of the order of β^{N+1}: this is a property of the Pisot numbers.

A Markov rectangle has a measure

$$\mu(a_0 a_1 \ldots a_N) = \mu(a_0)\, P(a_1 \mid a_0) \ldots P(a_N \mid a_{N-1})$$

with

$$(1/3)\, 2^{-N} \leqslant \mu(a_0 a_1 \ldots a_N) \leqslant (2/3)\, 2^{-N} \tag{4.1}$$

We define now the projected measure ν_β on $[0, 1]$ as the image of μ via Φ.[9]

Definition 4.1. Let $a_0 a_1 \ldots a_N \in \Sigma_N(X)$ and $\Phi(a_0 a_1 \ldots a_N) = \varepsilon_1 \varepsilon_2 \ldots \varepsilon_N$. Then, $\forall \varepsilon_1 \varepsilon_2 \ldots \varepsilon_N \in \Sigma_\beta^N([0, 1])$,

$$\nu_\beta(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N) = \mu(\Phi^{-1}(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N)) \tag{4.2}$$

Let $\# \{\Phi^{-1}(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N)\} \equiv z(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N)$ be the "ambiguity" of $a_0 a_1 \ldots a_N$. Then, since μ is the maximum entropy measure (4.1),

$$(1/3)\, 2^{-N} z(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N) \leqslant \nu_\beta(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N) \leqslant (2/3)\, 2^{-N} z(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N) \tag{4.3}$$

Finally, observe that ν_β coincides with the probability measure defined in refs. 1, 20, and 21. For, consider $F_0^{-1} = \beta x$, $F_1^{-1} = \beta x + 1 - \beta$, p the probability on the space of functions F; $[0, 1] \to [0, 1]$, which gives weight 1/2 to both F_0^{-1} and F_1^{-1}. Now, p determines a Markov process $\{X_n\}$ in the following way. Choose $X_0 = x \in [0, 1]$ according to a probability ν; then

$$X_{n+1}=\begin{cases}\beta X_n+1-\beta & \text{with probability } 1/2\\ \beta X_n & \text{with probability } 1/2\end{cases}$$

There is only one stationary probability distribution $\nu=\nu_\beta$ for this Markov process, and this is the distribution of the random variable $(1-\beta)\sum_{n=0}^{\infty}\beta^n\varepsilon_n$, where ε_n are the Bernoulli variables $p(\varepsilon_n=0)=p(\varepsilon_n=1)=1/2$.[13]

5. THE AMBIGUITY OF THE PROJECTION

Recall that our aim is "to count ambiguity": the projection rules have been constructed to know which are the Markov rectangles all projecting on the same interval of B_N. We concentrate on the central, overlapping zone. Observe that the β-coding of an interval $I\in B_N$ which lies in "the center," i.e., is contained in $[\beta^2,\beta]$, has the form $\varepsilon(I)=\varepsilon_0\varepsilon_1\ldots\varepsilon_N$, where

$$\varepsilon_0=0,\qquad \varepsilon_1=1,\qquad \varepsilon_2=\cdots=\varepsilon_{n_1}=0,\qquad \varepsilon_{n_1+1}=0,\qquad \varepsilon_{n_1+2}=1$$

$$\varepsilon_{n_1+3}=\cdots=\varepsilon_{n_1+2+n_2}=0,\ldots,\qquad \varepsilon_{n_1+2+n_2+2+\ldots n_{q-1}+1}=0$$

$$\varepsilon_{n_1+2+n_2+2+\ldots n_{q-1}+2}=1,\qquad \varepsilon_{n_1+2+n_2+2+\ldots n_{q-1}+2+n_q}=0$$

with $n_1+2+n_2+2+\ldots+n_q+2=N$, $n_i\geqslant 0$. We denote briefly $\varepsilon(I)=01, n_1, 01, n_2,\ldots, 01, n_q$ for such an interval. There is a repeated structure which allows us to use the projection rules just as if there were commutation between projecting and shift. Observe first that passages through the center (i.e., on 01) are coded by AC or DB when $n+1$ is even, and by AC, DB, or BB when $n+1$ is odd. That is,

$$\{\Phi^{-1}(01,n,01)\}=\left\{AC\ldots\begin{matrix}AC\\DB\end{matrix},\ DB\ldots\begin{matrix}AC\\DB\end{matrix}\right\}$$

or

$$\{\Phi^{-1}(01,n,01)\}=\left\{AC\ldots\begin{matrix}AC\\DB\end{matrix},\ DB\ldots\begin{matrix}AC\\\cdot B\end{matrix}\right\}$$

(where $\cdot B$ means BB or DB and ... stands for any admissible sequence of n symbols), depending on the parity of n. The above BB symbol does not correspond to an initial passage in (01). To overcome this problem, we only have to consider bilateral sequences $\ldots a_{-N}\ldots a_0\ldots a_N\ldots$ and $\ldots\varepsilon_{-N}\ldots\varepsilon_0\ldots\varepsilon_N\ldots$ (see below). Now, the "ambiguity" propagates itself as follows. We have to count how many sequences are produced between two consecutive passages through the center: in the bilateral case this depends *only* on n. It is clear that we can express how ambiguity propagates passage after passage by means of products of matrices of order two. Order two is a simplification allowed by the symmetric behavior of (words beginning with) *BB* or *DB*. These matrices (indexed by n) simply count "how many (words terminating with) *AC* and *DB* are produced by (a word beginning with) *AC* in a passage for the center after the time n, and how many *AC*, *DB*, [*BB*] are produced by *DB* and, which is the same, by *BB*."

So we are led to consider binfinite sequences $\omega = \ldots 01 n_{-1} 01 n_0 01 n_1 \ldots$, the space Ω of these sequences, Ω_N^+ that of finite right ones, and the maps $x_{n_1 \ldots n_q}: \Omega \to \Omega_N^+$ and $x_{n_1 \ldots n_q}(\omega) = 01 n_1 \ldots n_q$.

For notational convenience, we rename $n_i = n_i + 1$, i.e., now $n_i = 1, 2, \ldots$. If $n = 2k + 1$, $k = 0, 1, 2, \ldots$, set

$$B(k) = \begin{pmatrix} 1 & k \\ 1 & k+1 \end{pmatrix} \tag{5.1a}$$

If $n = 2k + 2$, $k = 0, 1, 2, \ldots$, set

$$A(k) = \begin{pmatrix} 1 & k+1 \\ 1 & k+1 \end{pmatrix} \tag{5.1b}$$

We have thus the following result.

Proposition 5.1. The ambiguity

$$\begin{aligned} z(x_{n_1 \ldots n_q}(\omega)) &= z(01, n_1 - 1, 01, n_2 - 1, \ldots, 01, n_q - 1) \\ &\equiv \# \{\Phi^{-1}(01, n_1 - 1, 01, n_2 - 1, \ldots, 01, n_q - 1)\} \end{aligned}$$

is given by

$$|M(n_q)\, M(n_{q-1}) \ldots M(n_1)\, v|$$

where $M(n) = A(k)$ if $n = 2k + 2$ and $M(n) = B(k)$ if $n = 2k + 1$ and $v = \binom{1}{1}$ or $\binom{1}{2}$ ("depending on the parity of n_0"), and

$$\left| \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right| = v_1 + v_2$$

We have this formulation in view of taking a limit. The ambiguity of a finite string is not exactly the same, but this is irrelevant in our context. Also irrelevant is the choice of the initial v.

6. THE MARKOV STRUCTURE UNDERLYING THE PRODUCT OF THE $M(n)$

The $\{\Phi^{-1}(01, n-1, 01)\}$ is a set of Markov rectangles each beginning with *AC*, *DB*, or *BB*, terminating again with *AC*, *DB*, or *BB* and of length

$n+3$. We denote each of them with the triple $(\cdot, n-1, \cdot)$ when we do not specify its beginning or ending; when we do, e.g., we write $(AC, n-1, AC)$; $n-1$ indicates merely that there are $n-1$ unspecified symbols projecting on $n-1$ zeros, $n \geqslant 1$.

If $n = 2k+1$, the description of $\{\Phi^{-1}(01, n-1, 01)\}$ is

$(AC, 2k, AC)$	
$(AC, 2k, DB)$	
$(DB, 2k, AC)$	k strings
$(DB, 2k, DB)$	k strings
$(DB, 2k, BB)$	
$(BB, 2k, AC)$	k strings
$(BB, 2k, DB)$	k strings
$(BB, 2k, BB)$	

[where we mean there are, e.g., k different elements ("strings") denoted $(DB, 2k, AC)$ etc.].

If $n = 2k+2$, the description of $\{\Phi^{-1}(01, n-1, 01)\}$ is

$(AC, 2k+1, AC)$	
$(AC, 2k+1, DB)$	
$(DB, 2k+1, AC)$	$k+1$ strings
$(DB, 2k+1, DB)$	$k+1$ strings
$(BB, 2k+1, AC)$	$k+1$ strings
$(BB, 2k+1, DB)$	$k+1$ strings

Therefore $\{\Phi^{-1}(01, n_1-1, 01, n_2-1, 01, n_3-1, \ldots\}$ consists of Markov rectangles which are built by connecting the above elementary Markov rectangles following the rule that we can connect two of them if and only if the beginning of the following one is equal to the ending of the preceding one.

This means considering the Markov system of the space X of the elementary strings $(\cdot, n-1, \cdot)$, $\forall n \geqslant 1$, the Markov measure defined by $\mu((\cdot, n-1, \cdot)) = (1/12)\, 2^{-n}$, and the transition matrix $\Pi(j \mid i)$ given below $(i, j \in X)$. We have to list separately transitions for "odd" and "even" j [we say, e.g., that i is odd if $i = (\cdot, n-1, \cdot)$ with n odd].

For j odd, we have

$$\begin{pmatrix} & (AC,2k,AC) & (AC,2k,DB) & (DB,2k,AC) & (DB,2k,DB) & (BB,2k,AC) & (BB,2k,DB) & (DB,2k,BB) \\ (\cdot,\cdot,AC) & 2^{-(2k+2)} & 2^{-(2k+2)} & 0 & 0 & 0 & 0 & 0 \\ (\cdot,\cdot,DB) & 0 & 0 & 2^{-(2k+2)} & 2^{-(2k+2)} & 0 & 0 & 2^{-(2k+2)} \\ (\cdot,\cdot,BB) & 0 & 0 & 0 & 0 & 2^{-(2k+2)} & 2^{-(2k+2)} & 0 \end{pmatrix}$$

And for j even

$$\begin{pmatrix} & (AC,2k+1,AC) & (AC,2k+1,DB) & (DB,2k+1,AC) & (DB,2k+1,DB) & (BB,2k+1,AC) & (BB,2k+1,DB) \\ (\cdot,\cdot,AC) & 2^{-(2k+3)} & 2^{-(2k+3)} & 0 & 0 & 0 & 0 \\ (\cdot,\cdot,DB) & 0 & 0 & 2^{-(2k+3)} & 2^{-(2k+3)} & 0 & 0 \\ (\cdot,\cdot,BB) & 0 & 0 & 0 & 0 & 2^{-(2k+3)} & 2^{-(2k+3)} \end{pmatrix}$$

Recall that there are actually k columns occupied by strings denoted $(DB, 2k, AC)$, k by $(DB, 2k, DB)$, k by $(BB, 2k, AC)$, and k by $(BB, 2k, DB)$ for j odd and there are $k+1$ columns occupied by $(DB, 2k+1, AC)$, $k+1$ by $(DB, 2k+1, DB)$, $k+1$ by $(BB, 2k+1, AC)$, and $k+1$ by $(BB, 2k+1, DB)$ for j even.

Here we have a very simple Markov system, very near to Bernoulli, because the transition probabilities $\Pi(j \mid i)$ depend only on the length of j (and of course of the interdictions beginning–end). We are then led to study the growth of the product of matrices

$$|M(x_n)\, M(x_{n-1}) \ldots M(x_1)\, v|$$

where $v = \binom{1}{1}$ or $\binom{1}{2}$, x_i is the sequence of random variables $\{\Phi^{-1}(01, n-1, 01)\}$, $n = 1, 2,\ldots$, which are distributed according to the Markov stationary distribution given by μ and Π; $M(x_i) = M(n)$ if and only if $x_i \in \{\Phi^{-1}(01, n-1, 01)\}$, where $M(n) = A(k)$ if $n = 2k+2$ and $B(k)$ if $n = 2k+1$.

7. THE LYAPUNOV EXPONENT

7.1. Ergodic Theorem

Proposition 7.1 ("Ergodic Theorem"[19,23,25]). Let (X, μ) be a (discrete) probability space, $\Pi(x, y)$ a Markov transition matrix such that $\mu\Pi = \mu$, and $\Pi^n(i, j) > 0$. Let $M: X \to$ nonnegative matrices of order two, such that $\int \log |M(x)|\, d\mu(x) < \infty$. Consider the transition kernel

$$Q(x, \phi, y, \xi) = \Pi(x, y)\, \delta_{M(x)\phi}(\xi) \qquad \text{on} \quad X \times S^1$$

(S^1 is the circle); there is on $X \times S^1$ a measure N left invariant by $Q: NQ = N$; it has the form $N = \mu(x)\, \nu_x(d\phi)$ (see below). Consider the ergodic system $(X^N \times S^1, \hat{\theta}, P_\mu \times \nu_{x_0})$, where $\hat{\theta}$: $(x, \phi) \to (\theta \underline{x}, M(x_0)\, \phi)$, where $\{\theta \underline{x}\}_n = x_{n+1}$ is the shift on the space X^N of the trajectories $\{x_n\}$ of the Markov process; P_μ is the measure on X^N such that if $x_n(\underline{x}) = x_n$, $P_\mu(x_n(\underline{x}) = i) = \mu(i)$ and $P_\mu(x_{n+1}(\underline{x}) = j \mid x_n(\underline{x}) = i) = \Pi(j \mid i)$. Let $F(\underline{x}, \phi) = \log[|M(x_0)\,\phi| / |\phi|]$. Then

$$\frac{1}{n} \log \frac{|M(x_n)\, M(x_{n-1}) \ldots M(x_0)\, \phi|}{|\phi|} = \frac{1}{n} \sum_{i=0}^{n-1} F(\hat{\theta}^i(\underline{x}, \phi))$$

converges $P_\mu(\underline{x}) \times \nu_{x_0}(d\phi)$ almost everywhere to

$$\lambda = \sum_{x_0} \int_{S^1} \log \frac{|M(x_0)\,\phi|}{|\phi|}\, \mu(x_0)\, \nu_{x_0}(d\phi) \tag{7.1}$$

7.2. Ergodicity

We consider actually the ergodic system $(X^N \times S^1_{[\pi/4, \pi/2]}, \hat{\theta}, P_\mu \times \nu_{x_0})$. The measure $P_\mu \times \nu_{x_0}$ is $\hat{\theta}$ invariant (see Sections 7.3 and 7.4), M are the matrices $\{A(k), B(k)\}$, μ is the Markov measure on the space X of the strings $(\cdot, n-1, \cdot)$ (defined in Section 6), and the circle sector $\pi/4 \leqslant \theta \leqslant \pi/2$ is a closed invariant set under the action of M which contains the support of all ν_x.

7.3. The Invariance Equation for the Measure: $NQ = N$

Consider the kernel $Q(x, \phi, y, \xi) = \Pi(x,\ )\,\delta_{M(x)\phi}(\xi)$ and $f \in C^0(X \times S^1)$. We have

$$(Qf)(x, \phi) = \sum_y \Pi(x, y)\, f(y, M(x)\phi)$$

$$N(f) = \sum_x \int_S f(x, \phi)\, \nu_x(d\phi)\, \mu(x)$$

Therefore

$$N(Qf) = \sum_y \sum_x \mu(x) \int_S \Pi(x, y)\, f(y, M(x)\phi)\, \nu_x(d\phi)$$

and the invariance equation $N(f) = N(Qf)$ means that this last expression is equal to $\sum_y \int_S f(x, \phi)\, \nu_y(d\phi)\, \mu(y)$, so we rewrite it as

$$\sum_y \mu(y) \sum_x \frac{\mu(x)}{\mu(y)} \int_S \Pi(x, y)\, f(y, M(x)\phi)\, \nu_x(d\phi)$$

to find that $\forall f \in C^0(X \times S^1)$

$$\sum_x \mu(x) \int_S \Pi(x, y)\, f(y, M(x)\phi)\, \nu_x(d\phi) = \int_S f(y, \phi)\, \nu_y(d\phi)\, \mu(y)$$

Hence, the invariance equation for ν_y reads

$$\mu(y)\, \nu_y(d\phi) = \sum_x \mu(x)\, \Pi(x, y)\, M(x)\, \nu_x(d\phi) \tag{7.2}$$

7.4. The support of ν

Let $\dot{v} = \{w \text{ s.t. } w = \lambda v, \lambda > 0\}$ and $v \in R^2$. Observe that $A(k)\,\dot{v} = \dot{\binom{1}{1}}\ \forall v$. The support of ν_x, as we deduce from the recursive formula (7.2), is therefore the closure of the orbit of $\dot{\binom{1}{1}}$ under $B(k)$. The ν_x are discrete measures carried by the points $\binom{h}{h+q}$, $=1, 2,\ldots$ and $q=0, 1, 2,\ldots$. Actually, there are only *three* measures, which we call ν_{AC}, ν_{DB}, and ν_{BB}, due to the special form of the transition matrix $\Pi(j\mid i)$. We also remark that the invariance equation (7.2) for ν_x gives by recurrence the explicit construction of the $\hat{\theta}$-invariant measure $P_\mu \times \nu_x$ with its existence and unicity.

7.5. Almost Sure Convergence

The Ergodic Theorem 7.1 ensures convergence to λ for almost every trajectory of $\{x_n\}$ and on a set of ν_x measure one. But since ν_x is a (countable) sum of deltas weighted on points $\{\binom{h}{h+q}\}$, there is convergence to λ a.e. $\{x_n\}$ and at all points $\{\binom{h}{h+q}\}$. We remark, however, that, repeating the argument of ref. 25, Chapter III, Corollary 1.3, we can to prove the following stronger statement (although we will not use it):

Proposition 7.2. For all $v \in S_{[\pi/4,\, \pi/2]}$

$$\frac{1}{n} \log \frac{|M(x_n)\, M(x_{n-1}) \ldots M(x_0)\, \phi|}{|\phi|}$$

converges $P_\mu(\underline{x})$ almost everywhere to

$$\lambda = \sum_x \sum_{\substack{\dot{\phi} = \binom{h}{h+q} \\ h \geqslant 1,\, q \geqslant 0}} \log \frac{|M(x)\, \phi|}{|\phi|}\, \mu(x)\, \nu_x(\dot{\phi})$$

8. THE RECURSIVE FORMULA FOR THE COMPUTATION OF THE LYAPUNOV EXPONENT

The recurrence given by (7) can be started at $\dot{\phi} = \binom{1}{1}$. We sometimes write the strings $(\cdot, n-1, \cdot) \in X$ simply as AC, DB, or BB to mean, respectively, $(AC, n-1, \cdot)$, $(DB, n-1, \cdot)$, or $(BB, n-1, \cdot)$, $n \geqslant 1$, and by i to mean any unspecified element of X. We have

$$\nu_{AC}\begin{pmatrix} \dot{1} \\ 1 \end{pmatrix} = \frac{1}{\mu(x)} \sum_{M(i) = A(k)} \mu(i)\, \Pi(x \mid i)$$

where the sum runs on $\{i\colon M(i) = A(k)\} = \{(\cdot, n-1, \cdot)\, \forall n = 2k+2\}$ because $A(k)\,\theta = \binom{\dot{1}}{1}$ for all θ, and $\nu_{AC}(S^1) = 1$ $[\nu_{DB}(S^1) = 1 = \nu_{BB}(S^1)]$. This is equal to

$$\begin{aligned} &= 12 \cdot 2^n \sum_k \big(\mu((AC, 2k+1, AC))\, \Pi(AC \mid (AC, 2k+1, AC)) \\ &\quad + (k+1)\, \mu((DB, 2k+1, AC))\, \Pi(AC \mid (DB, 2k+1, AC)) \\ &\quad + (k+1)\, \mu((BB, 2k+1, AC))\, \Pi(AC \mid (BB, 2k+1, AC))\big) \end{aligned}$$

so that

$$\nu_{AC}\begin{pmatrix}\dot{1}\\1\end{pmatrix}=11/18 \tag{8.1}$$

Similarly

$$\nu_{DB}\begin{pmatrix}\dot{1}\\1\end{pmatrix}=11/18 \tag{8.2}$$

As there are no even strings terminating by BB, we have

$$\nu_{BB}\begin{pmatrix}\dot{1}\\1\end{pmatrix}=0 \tag{8.3}$$

Then the recurrence continues as follows:

1. If $h=qh'$, $h'\geqslant 1$, we have

$$\nu_{AC}\begin{pmatrix}\dot{h}\\h+q\end{pmatrix}=\frac{1}{\mu(x)}\sum_{M(i)=B(h'-1)}\mu(i)\,\Pi(x\mid i)\,\nu_i\begin{pmatrix}\dot{1}\\1\end{pmatrix}$$

where the sum runs on

$$\left\{i\colon M(i)=\begin{pmatrix}1 & h'-1\\1 & h'\end{pmatrix}\right\}$$

because if $h=qh'$, $M(i)^{-1}\binom{h}{h+q}=\dot{\phi}$ if and only if $M(i)=B(h'-1)$ and $\dot{\phi}=\binom{\dot{1}}{1}$. This is equal to

$$=12\cdot 2^{n}\Bigg(\mu((AC,2(h'-1),AC))\,\Pi(AC\mid(AC,2(h'-1),AC))\,\nu_{AC}\begin{pmatrix}\dot{1}\\1\end{pmatrix}$$

$$+\mu((DB,2(h'-1),AC))(h'-1)\,\Pi(AC\mid(DB,2(h'-1),AC))\,\nu_{DB}\begin{pmatrix}\dot{1}\\1\end{pmatrix}\Bigg)$$

so that

$$\nu_{AC}\left(\frac{\dot{h}}{h+q}\right)=\frac{\nu_{AC}\begin{pmatrix}\dot{1}\\1\end{pmatrix}}{2^{2h'}}+(h'-1)\frac{\nu_{DB}\begin{pmatrix}\dot{1}\\1\end{pmatrix}}{2^{2h'}} \tag{8.4}$$

Similarly

$$\nu_{DB}\left(\frac{\dot{h}}{h+q}\right)=\frac{\nu_{AC}\begin{pmatrix}\dot{1}\\1\end{pmatrix}}{2^{2h'}}+(h'-1)\frac{\nu_{DB}\begin{pmatrix}\dot{1}\\1\end{pmatrix}}{2^{2h'}} \tag{8.5}$$

Also,

$$\nu_{BB}\begin{pmatrix}\dot{h}\\ h+q\end{pmatrix}=12\cdot 2^{n}\left(\mu((DB,2(h'-1),BB))\right.$$

$$\left.\times\,\Pi(BB\,|\,(DB,2(h'-1),BB))\,\nu_{DB}\begin{pmatrix}\dot{1}\\ 1\end{pmatrix}\right)$$

$$=\frac{1}{2^{2h'}}\,\nu_{DB}\left(\frac{\dot{1}}{1}\right) \tag{8.6}$$

2. If $h=qh'+r$, $h'\geqslant 0$, $1\leqslant r<q$, we have

$$\nu_{AC}\begin{pmatrix}\dot{h}\\ h+q\end{pmatrix}=\frac{1}{\mu(x)}\sum_{M(i)=B(h')}\mu(i)\,\Pi(x\,|\,i)\,\nu_i\begin{pmatrix}\dot{r}\\ q\end{pmatrix}$$

where the sum runs on

$$\left\{i\colon M(i)=\begin{pmatrix}1 & h'\\ 1 & h'+1\end{pmatrix}\right\}$$

because if $h=qh'+r$, $h'\geqslant 0$, $1\leqslant r<q$, $M(i)^{-1}\,(\begin{smallmatrix}\dot{h}\\ h+q\end{smallmatrix})=\dot{\phi}$ if and only if $M(i)=B(h')$ and $\dot{\phi}=(\begin{smallmatrix}\dot{r}\\ q\end{smallmatrix})$. This is equal to

$$=12\cdot 2^{n}\left(\mu((AC,2h',AC))\,\Pi(AC\,|\,(AC,2h',AC))\,\nu_{AC}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}\right.$$

$$+\,\mu((DB,2h',AC))\,h'\,\Pi(AC\,|\,(DB,2h',AC))\,\nu_{DB}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}$$

$$\left.+\,\mu((BB,2h',AC))\,h'\,\Pi(AC\,|\,(BB,2h',AC))\,\nu_{BB}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}\right)$$

so that

$$\nu_{AC}\left(\frac{\dot{h}}{h+q}\right)=\frac{\nu_{AC}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}}{2^{2h'+2}}+h'\,\frac{\nu_{DB}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}}{2^{2h'+2}}+h'\,\frac{\nu_{BB}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}}{2^{2h'+2}} \tag{8.7}$$

Similarly,

$$\nu_{DB}\left(\frac{\dot{h}}{h+q}\right)=\frac{\nu_{AC}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}}{2^{2h'+2}}+h'\,\frac{\nu_{DB}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}}{2^{2h'+2}}+h'\,\frac{\nu_{BB}\begin{pmatrix}\dot{r}\\ q\end{pmatrix}}{2^{2h'+2}} \tag{8.8}$$

and

$$\nu_{BB}\left(\frac{\dot{h}}{h+q}\right)=\frac{\nu_{DB}\left(\dot{r \atop q}\right)}{2^{2h'+2}}+\frac{\nu_{BB}\left(\dot{r \atop q}\right)}{2^{2h'+2}} \tag{8.9}$$

We can now write the explicit formula for the exponent:

$$\lambda=\frac{1}{6}\sum_{k\geqslant 0}\ \sum_{\substack{\dot{\theta}=(\frac{h}{h+q})\\ h\geqslant 1,\, q\geqslant 0}}\left[\log\frac{|A(k)\,\theta|}{|\theta|}\left(\frac{1}{2^{2k+2}}\nu_{AC}(\dot{\theta})\right.\right.$$
$$\left.+(k+1)\frac{1}{2^{2k+2}}[\nu_{DB}(\dot{\theta})+\nu_{BB}(\dot{\theta})]\right)+\log\frac{|B(k)\,\theta|}{|\theta|}$$
$$\left.\times\left(\frac{1}{2^{2k+1}}\nu_{AC}(\dot{\theta})+(2k+1)\frac{1}{2\cdot 2^{2k+1}}[\nu_{DB}(\dot{\theta})+\nu_{BB}(\dot{\theta})]\right)\right] \tag{8.10}$$

where the measures ν_{AC}, ν_{DB}, and ν_{BB} are given in (8.1)–(8.9).

9. THE DIMENSION OF ν_β

We know that $(1/q)\log|M(x_q)\,M(x_{q-1})\ldots M(x_1)\,v|\to\lambda$ for P_μ a.e. trajectory of the Markov process $\{x_q\}$. Similarly, the law of the large numbers ensures that[12]

$$\frac{g(x_1)+\ldots+g(x_q)}{q}\to E_\mu(g(x_1))\qquad \text{if}\quad E_\mu g<\infty$$

P_μ a.e. trajectory of the Markov process $\{x_q\}$. Take $g(x_i)=\log 2^{n_i+1}$ to have

$$\frac{1}{q}\log 2^{n_1+1\ldots+n_q+1}\to E_\mu(n+1)\log 2$$
$$=\frac{\log 2}{6}\sum_{k\geqslant 0}\frac{(2k+3)^2}{2^{2k+2}}+\frac{(2k+2)^2}{2^{2k+1}}$$
$$=\frac{\log 2}{6}\sum_{n\geqslant 1}\frac{(n+1)^2}{2^n}=E\log 2$$

Similarly

$$\frac{1}{q}\log\beta^{n_1+1\ldots+n_q+1}\to E\log\beta\qquad P_\mu\ \text{a.e.}$$

We have indeed proven convergence P_μ a.e., but now we would like to say something about convergence P_{ν_β} a.e., where $\Phi\mu = \nu_\beta$. Let

$$S_q(\Phi(\underline{x})) = M(\Phi(x_q))\dots M(\Phi(x_1)), \qquad \Phi(\underline{x}) = \underline{n}$$
$$S_q(\underline{n}) = M(n_q)\dots M(n_1)$$

We know that $(1/q)\log|S_q(\Phi(\underline{x}))\,v| \to \lambda$ P_μ a.e., i.e., there exists a set A of P_μ measure 0 such that if $\underline{x} \notin A$, then $(1/q)\log|S_q(\Phi(\underline{x}))\,v| \to \lambda$. But any P_μ null set A has the form $\Phi^{-1}B$, because we know that $S_q(\underline{x}) \equiv M(x_q)\dots M(x_1) = M(n_q)\dots M(n_1)$ if and only if

$$x_{x_1 x_2 \dots x_q}(\underline{x}) = x_1 x_2 \dots x_q \in \{\Phi^{-1}(01n_1 - 101n_2 - 1\dots 01n_q - 1)\}$$

As B has P_{ν_β} measure 0 if and only if $A = \Phi^{-1}B$ has P_μ measure 0, it follows that $(1/q)\log|S_q(\underline{n})\,v| \to \lambda$ for $\underline{n} \notin B$ if and only if $(1/q)\log|S_q(\Phi(\underline{x}))\,v| \to \lambda$ for $\underline{x} \notin A$.[9]

We can conclude that

$$\frac{1}{q}\log|S_q(\underline{n})\,v| \to \lambda \qquad P_{\nu_\beta}\text{ a.e.}$$

$$\frac{g(n_1) + \dots + g(n_q)}{q} \to E\log 2 \qquad P_{\nu_\beta}\text{ a.e.} \qquad \text{for} \quad g(n_i) = \log 2^{n_i+1}$$

$$\frac{f(n_1) + \dots + f(n_q)}{q} \to E\log\beta \qquad P_{\nu_\beta}\text{ a.e.} \qquad \text{for} \quad f(n_i) = \log\beta^{n_i+1}$$

and therefore

$$\dim(\nu_\beta) = \frac{\lambda - E\log 2}{E\log\beta}$$

where λ is given in (8.10).

ACKNOWLEDGMENTS

We wish to thank Philippe Bougerol and Pierre Collet for enlightening discussions.

REFERENCES

1. J. Alexander and J. Yorke, The fat baker's transformation, *Ergod. Theory Dynam. Syst.* 4:1–23 (1984).
2. J. Alexander and D. Zagier, The entropy of a certain infinitely convolved Bernoulli measure, preprint, University of Maryland (1991).

3. T. Bedford, The box dimension of self-affine graphs and repellers, *Nonlinearity* **2**:53–71 (1989).
4. T. Bedford and M. Urbanski, The box and Hausdorff dimension of self-affine sets, *Ergod. Theory Dynam. Syst.* **10**:627–644 (1990).
5. F. Blanchard, Développement des nombres réels en base non entière, *Rev. Mat. Apl.* **9**:9–19 (1988).
6. Ph. Bougerol, Théorèmes limites pour les systèmes linéaires à coefficients markoviens, *Prob. Theory Related Fields* **78** (1988).
7. Ph. Bougerol and J. Lacroix, Products of random matrices with applications to Schrödinger operators, in *Progress in Probability and Statistics*, Vol. 8 (Birkhäuser, Basel, 1985).
8. A. Bovier, Bernoulli convolutions as an invariant measure problem, preprint, University of Bonn (1991).
9. Bourbaki, *Intégration* (Hermann, Paris, 1967), Chapter V.
10. E. Cogen, H. Kesten, and C. Newman, Random matrices and their applications, *Contemp. Math.* **50** (1984).
11. P. Collet, J. Lebowitz, and A. Porzio, The dimension spectrum of some dynamical systems, *J. Stat. Phys.* **47**(5/6) (1987).
12. J. Doob, *Stochastic Process* (Wiley, New York, 1953).
13. L. Dubins and D. Freedman, Invariant probabilities for certain Markov processes, *Ann. Math. Stat.* **37**:837–848 (1966).
14. P. Erdös, On the smoothness properties of a family of Bernoulli convolutions, *Am. J. Math.* **62**:180–186 (1940).
15. P. Erdös, On a family of symmetric Bernoulli convolutions, *Am. J. Math.* **61**:974–976 (1939).
16. J. Farmer, Information dimension and the probabilistic structure of chaos, *Z. Naturforsch.* **37a**:1304–1325 (1982).
17. J. Farmer, E. Ott, and J. Yorke, The dimension of chaotic attractors, in Proceedings Low Alamos Conference "Order in Chaos," *Phys. D* (1982).
18. P. Frederickson, J. Kaplan, E. Yorke, and J. Yorke, The Lyapunov dimension of strange attractors, *J. Diff. Equat.* **49**:185–207 (1983).
19. H. Furstenberg, Non-commuting random products, *Trans. Am. Math. Soc.* **108**:377–428 (1963).
20. A. Garsia, Arithmetic properties of Bernoulli convolutions, *Trans. Am. Math. Soc.* **162**: 409–432 (1962).
21. A. Garsia, Entropy and singularity of infinite convolutions, *Pac. J. Math.* **13**:1159–1169 (1963).
22. Y. Guivarc'h, Quelques propriétés asymptotiques des produits de matrices aléatoires, in *Lecture Notes in Mathematics*, Vol. 774 (Springer-Verlag, Berlin, 1980).
23. Y. Guivarc'h, Exposants caractéristiques des produits de matrices aléatoires en dépendence markovienne, in *Probability Measures on Groups* (Springer-Verlag, Berlin, 1984).
24. B. Jessen and A. Wintner, Distribution functions and the Riemann zeta function, *Trans. Am. Math. Soc.* **38**:48–88 (1935).
25. F. Ledrappier, Quelques propriétés des exposants caractéristiques, in *Lecture Notes in Mathematics*, Vol. 1097 (Springer-Verlag, Berlin, 1984).
26. F. Ledrappier, Some relations between dimension and Lyapunov exposants, *Commun. Math. Phys.* **81**:229–238 (1981).
27. F. Ledrappier and L. S. Young, The metric entropy of diffeomorphisms, *Ann. Math.* **122**:540–574 (1985).

28. F. Ledrappier and L. S. Young, Dimension formula for random transformations, *Commun. Math. Phys.* **117**:529–548 (1988).
29. C. McMullen, The Hausdorff dimension of general Sierpinski carpets, *Nagoya Math. J.* **96**:1–9 (1984).
30. V. I. Oseledec, A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems, *Trudy Moskov. Mat. Obsch.* **19**:179–210 (1968).
31. W. Parry, On the β-expansions of real numbers, *Acta Math. Acad. Sci. Hung.* **11**:401–416 (1960).
32. F. Przytycki and M. Urbanski, On the Hausdorff dimension of some fractal sets, *Studia Math.* **93**:155–186 (1989).
33. D. Ruelle, *Thermodynamic Formalism* (Addison-Wesley, Reading, Massachusetts, 1978).
34. R. Salem, A remarkable class of algebraic integers, *Duke Math. J.* **11**:103–108 (1984).
35. M. Urbanski, The probability distribution and the Hausdorff dimension of self-affine functions, *Prob. Theory Related Fields* **84**:377–391 (1990).
36. L. S. Young, Dimension, entropy, and Lyapunov exponents, *Ergod. Theory Dynam. Syst.* **2**:109–124 (1982).