

A class of nonstationary adic transformations

Xavier Méla

IML, 163, avenue de Luminy, case 907, 13288 Marseille cedex 09, France

Received 29 January 2004; received in revised form 7 January 2005; accepted 18 February 2005

Available online 23 September 2005

Abstract

We introduce a class of nonstationary adic transformations generalizing the Pascal adic transformation. We identify the invariant ergodic measures, show that there are no rational eigenvalues, and construct a topological model. We give upper and lower bounds for the complexity function and show that the systems are topologically weakly mixing. We also show that these adics are loosely Bernoulli.

© 2005 Elsevier SAS. All rights reserved.

Résumé

On introduit une classe de transformations adiques non-stationnaires généralisant la Pascal-adique. On détermine l'ensemble des mesures invariantes ergodiques et on montre qu'il n'existe pas de valeur propre rationnelle. Ces systèmes admettent des modèles en dynamique topologique pour lesquels nous donnons des bornes inférieures et supérieures pour la fonction de complexité et montrons qu'ils sont topologiquement faiblement mélangeants. On démontre aussi que ces transformations adiques sont lâchement Bernoulli.

© 2005 Elsevier SAS. All rights reserved.

MSC: primary 37A05, 37A35, 37A50; secondary 37B10, 37B40

Keywords: Adic transformations; Ergodicity; Weak mixing; Loosely Bernoulli; Complexity; Binomial coefficients

1. Introduction

In this paper we introduce a class of nonstationary adic transformations T_d defined on a Bratteli diagram identified to $\{0, 1, \dots, d-1\}^{\mathbb{N}}$, with the particularity that the number of finite paths from the root $(0, 0)$ to a vertex (n, k) is given in terms of certain generalized binomial coefficients. When the number of such paths is given by the ordinary binomial coefficients (the case $d = 2$), the system is just the Pascal adic [20,15,16].

E-mail address: mela@iml.univ-mrs.fr (X. Méla).

The ergodicity of the Bernoulli or Markov measures for these adics is related to the questions of exchangeability and triviality of tail sigma-algebras in probability theory. It is a well-known result (de Finetti, Hajian–Ito–Kakutani, Vershik, [7,21]) that the nonatomic invariant ergodic measures for the Pascal adic are the Bernoulli measures $\mu_{\alpha,1-\alpha}$. Here we extend this result to the adics T_d , establishing that the set of nonatomic invariant ergodic measures is formed by a one-parameter family of Bernoulli measures $\{\mu_{p(a)} \mid a \in (0, 1)\}$ for a certain probability vector $p(a)$ (Theorem 3.6). We show that these measures are totally ergodic, equivalently that there are no rational eigenvalues (Proposition 4.5). For both of the previous results we avoid direct calculations on the generalized binomial coefficients, using instead arguments of isotropy or measure-theoretic tricks. Alternate proofs could be obtained if we had a deeper understanding of the arithmetics of these coefficients.

In Section 3.2 we show that the adic T_d is isomorphic to the Pascal adic on a certain subshift, whose study was in fact the original problem. The ergodic Bernoulli measures for T_d are carried to a family of $(d - 1)$ -step Markov measures for the Pascal on the subshift (Theorem 3.11).

The adics T_d are homeomorphisms but not on the whole space, and extensions lead to many discontinuities. Our topological model, which we present in Section 5, is “almost minimal” and metrically isomorphic to the original one. For this homeomorphism on a compact space, which can be viewed as a dynamical system defined by infinitely many substitutions (or infinitely many letters), we estimate the asymptotics of the complexity function and show that the system is topologically weakly mixing. The question of metric weak mixing (already difficult in the case $d = 2$) seems however to require a more complete understanding of the divisibility properties of the generalized binomial coefficients.

Finally, in Section 6, using the result of de la Rue and Janvresse [3], we show that the adics T_d are loosely Bernoulli.

2. Description of the systems

2.1. The graph construction

Let $d \in \mathbb{N}$ be fixed. Consider the infinite graph divided into levels $n = 0, 1, \dots$, with a root vertex labelled $(0, 0)$ and with $(d - 1)n + 1$ vertices labelled (n, k) at each level n for $k = 0, \dots, (d - 1)n$. From each vertex leave d edges labelled 0 through $d - 1$ connecting to the next level in such a way that for all finite paths going from $(0, 0)$ to a given vertex (n, k) the sum of the labels of the edges is equal to k — see Fig. 1. For $d = 2$ this is the Pascal graph — see [15].

Let X_d be the space of infinite paths starting at the root and going down the graph, which we identify to the space $\{0, \dots, d - 1\}^{\mathbb{N}}$ via the labelling of the edges. We denote by $(n, k_n(x))$ the vertex crossed by the path $x \in X_d$ at level n . We define a partial order on X_d , writing $x < y$ for $x, y \in X_d$ whenever x and y coincide below a certain

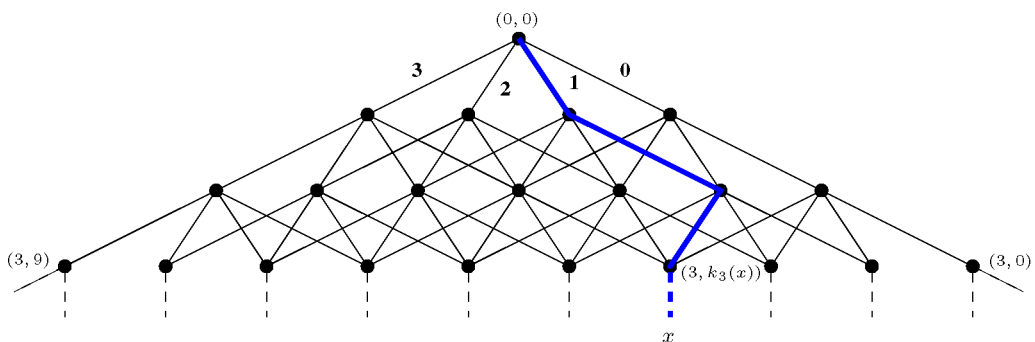


Fig. 1. Bratteli diagram in the case $d = 4$. The picture shows the path $x = 102\dots$ and the label of its vertex at level 3.

level n and $x_n < y_n$. Equivalently, $x < y$ if there is an n such that $x_n < y_n$, $x_i = y_i$ for all $i > n$, and $k_n(x) = k_n(y)$. Let X_d^{\max} and X_d^{\min} be respectively the sets of maximal and minimal paths. The *generalized Pascal adic* is defined by

$$\begin{cases} T_d : X_d \setminus X_d^{\max} \rightarrow X_d \setminus X_d^{\min}, \\ T_d(x) = \text{smallest } y > x. \end{cases}$$

In order to determine the image of x one has to look for the first appearance of a block of the type ij , where $1 \leq i \leq d - 1$ and $0 \leq j < d - 1$. Indeed, if $x = 0^l ij \dots$ and $j \neq d - 1$, we can always find an edge on the right of j and leading to the same vertex as j (giving therefore room for a successor). If there is no such occurrence then we are in the case of a maximal path, and we see a string of 0's followed by $i(d - 1)^\infty$, where $0 \leq i \leq d - 1$. The reverse situation happens for minimal paths; we therefore have:

$$\begin{aligned} X_d^{\max} &= \{0^l i(d - 1)^\infty \mid l \in \mathbb{N}, i = 0, \dots, d - 1\} \cup \{0^\infty, (d - 1)^\infty\}, \\ X_d^{\min} &= \{(d - 1)^l i 0^\infty \mid l \in \mathbb{N}, i = 0, \dots, d - 1\} \cup \{0^\infty, (d - 1)^\infty\}. \end{aligned}$$

There is a way to extend T_d as a bijection to the whole space X_d by sending maximal paths to minimal ones:

$$\begin{aligned} T_d 0^l i(d - 1)^\infty &= (d - 1)^l i 0^\infty, \\ T_d 0^\infty &= 0^\infty, \\ T_d (d - 1)^\infty &= (d - 1)^\infty. \end{aligned}$$

The resulting extended transformation is not a homeomorphism on X_d , but the discontinuities only occur at the minimal and maximal paths. Furthermore, all but the two edge paths (0^∞ and $(d - 1)^\infty$) have infinite orbits.

Remark 1. In [23,21] another generalized version of the Pascal adic is defined called the *n-dimensional Pascal adic*. Its graph lies on the multidimensional lattice $\mathbb{N}^{\mathbb{N}}$ and is (a priori) not related to the one described here.

2.2. Generalized binomial coefficients

A simple combinatorial argument shows that the number of finite paths from the root $(0, 0)$ to a vertex (n, k) is equal to the coefficient of x^k in the polynomial $(1 + x + x^2 + \dots + x^{d-1})^n$:

$$(1 + x + x^2 + \dots + x^{d-1})^n = \sum_{k=0}^{(d-1)n} C_d(n, k) x^k.$$

The coefficients $C_d(n, k)$ can be obtained from the relations

$$C_d(n + 1, k) = C_d(n, k) + C_d(n, k - 1) + \dots + C_d(n, k - d + 1),$$

with the convention that $C_d(n, k) = 0$ if $k > (d - 1)n$ or $k < 0$. In the case $d = 2$ we get the ordinary binomial coefficients satisfying the well-known relation $C_2(n + 1, k) = C_2(n, k) + C_2(n, k - 1)$. The name “generalized binomial coefficients” has various different definitions in the literature, but here it will refer only to the $C_d(n, k)$. The Pascal-type triangle formed by the $C_d(n, k)$ is sometimes called a *generalized Pascal triangle of order d* [2,1]; we will refer to it as the Δ_d -Pascal triangle.

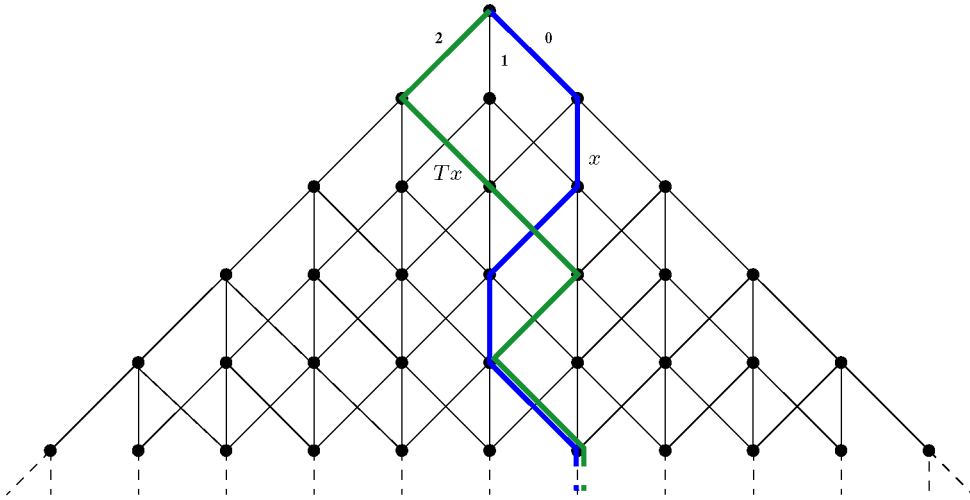
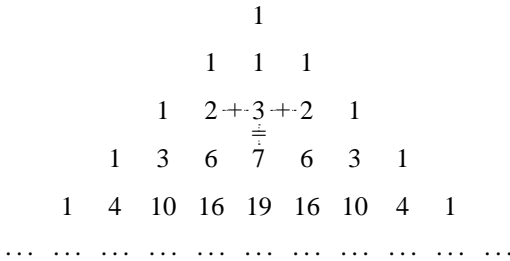


Fig. 2. The picture shows $T_3(01210*) = 20020*$.

Example 2.1. The Δ_3 -Pascal triangle starts as follows:



One can write $C_d(n, k)$ in terms of regular binomial coefficients or as a sum of multinomial coefficients (see [1,2]):

$$C_d(n, k) = \sum_{i=0}^{\lfloor k/d \rfloor} (-1)^i \binom{n}{i} \binom{n-1+k-di}{n-1} = \sum \frac{n!}{(n-k_1)!(k_1-k_2)! \dots (k_{d-2}-k_{d-1})! k_{d-1}!},$$

where the summation is over all (k_1, \dots, k_{d-1}) such that $k_1 + \dots + k_{d-1} = k$ and $k_i \leq k_{i-1}$.

3. A one-parameter family of ergodic measures

As one would suspect, the invariant ergodic measures for T_d are Bernoulli measures. However, not every Bernoulli measure is invariant for the adic T_d ; for each d the set of invariant Bernoulli measures forms a one-parameter family. The fact that each invariant Bernoulli measure is ergodic is a consequence of the Hewitt–Savage 0,1 law. To show the converse, which is that every invariant ergodic measure is in fact a Bernoulli measure, we use an argument of “isotropy”, giving at the same time a new proof for the case $d = 2$ of the Pascal adic.

Proposition 3.1. *The set of Bernoulli measures invariant for the adic T_d on $\{0, \dots, d - 1\}^{\mathbb{N}}$ forms a one-parameter family*

$$\{ \mu_{p(a)} \mid a \in [0, 1] \},$$

where the vector $p(a) \in [0, 1]^{d-1}$ is either $p(0) = (0, \dots, 0, 1)$, or if $a \neq 0$

$$p(a) = \left(a, t_a, \frac{t_a^2}{a}, \frac{t_a^3}{a^2}, \dots, \frac{t_a^{d-1}}{a^{d-2}} \right),$$

where t_a is the unique solution in $[0, 1]$ to the equation

$$a^{d-1} - a^{d-2} + a^{d-3}t + a^{d-4}t^2 + \dots + at^{d-2} + t^{d-1} = 0.$$

Proof. Let's start by examining the case $d = 3$. Denote by $[a_1 a_2 \dots a_r]$ the cylinder of all paths x whose first n edges are labelled $a_1 a_2 \dots a_r$. Let $p = (p_0, p_1, p_2)$ with $p_0 + p_1 + p_2 = 1$, and let μ_p be a Bernoulli measure on $\{0, 1, 2\}^{\mathbb{N}}$ such that $\mu_p[i] = p_i > 0$ for $i = 0, 1, 2$. First note that $\mu_p[20] = \mu_p[11]$ is a necessary and sufficient condition for μ to be invariant. Indeed, assume that $\mu_p[20] = \mu_p[11]$ holds and let C be a cylinder set. Since the set of maximal paths has measure zero (it is countable), it follows that C is a disjoint union up to measure zero of (non-maximal) cylinders of the form

$$C_1 = [0^i 11], \quad C_2 = [0^i 10], \quad C_3 = [0^i 2^j 21], \quad C_4 = [0^i 2^j 20], \quad C_5 = [0^i 12^j 1], \quad C_6 = [0^i 12^j 0],$$

where $i \geq 0$ and $j > 0$. Then respectively

$$TC_1 = [0^i 02], \quad TC_2 = [0^i 01], \quad TC_3 = [2^j 10^i 2], \quad TC_4 = [2^j 10^i 1],$$

$$TC_5 = [2^j 0^i 02], \quad TC_6 = [2^j 0^i 01].$$

Note that $\mu_p(TC_l) = \mu_p(C_l)$ is automatically satisfied for $l = 2, 3, 6$, and that for $l = 1, 4, 5$ the equality follows from our hypothesis that $\mu_p[20] = \mu_p[11]$. Hence if $\mu_p[20] = \mu_p[11]$ then $\mu_p(TC) = \mu_p(C)$ for any cylinder set C . Thus μ_p is invariant under the action of the adic T_3 if and only if the following relation holds:

$$p_0 p_2 = p_1^2.$$

Since in addition $p_0 + p_1 + p_2 = 1$, we can solve for p_1 and p_2 in terms of p_0 ; letting $p_0 = a$, we get

$$p_1 = \left[-a + \sqrt{a(4 - 3a)} \right] / 2,$$

$$p_2 = 1 - a + \left[a - \sqrt{a(4 - 3a)} \right] / 2.$$

For the general case, let $p = (p_0, p_1, \dots, p_{d-1})$ with $p_0 + p_1 + \dots + p_{d-1} = 1$, and let μ_p be a Bernoulli measure on $\{0, 1, \dots, d - 1\}^{\mathbb{N}}$ such that $\mu_p[i] = p_i > 0$ for $i = 0, 1, \dots, d - 1$. Arguing as previously we see that μ_p is invariant if and only if $p_i p_j = \mu_p[ij] = \mu_p[kl] = p_k p_l$ for all $i, j, k, l \in \{0, 1, \dots, d - 1\}$ such that $i + j = k + l$. The previous condition implies $d(d - 1)/2$ nontrivial relations on p_0, p_1, \dots, p_{d-1} . Out of those $d(d - 1)/2$ nontrivial relations only $d - 2$ of them are independent; these are $p_0 p_j = p_1 p_{j-1}$ for $j = 2, \dots, d - 1$ (the other ones follow from them). For simplicity we let $p_0 = a$ and $p_1 = t$. For all $j = 2, \dots, d - 1$ we have

$$p_j = \frac{p_1}{p_0} p_{j-1} = \frac{t}{a} p_{j-1}.$$

We can therefore express each p_j as a function of a and t :

$$p_2 = \frac{t}{a} p_1 = \frac{t^2}{a},$$

$$p_3 = \frac{t}{a} p_2 = \frac{t^3}{a^2},$$

$$\vdots$$

$$p_{d-2} = \frac{t}{a} p_{d-3} = \frac{t^{d-2}}{a^{d-3}},$$

$$p_{d-1} = \frac{t}{a} p_{d-2} = \frac{t^{d-1}}{a^{d-2}}.$$

Now, using that $p_0 + p_1 + \dots + p_{d-1} - 1 = 0$ we get

$$a + t + \frac{t^2}{a} + \frac{t^3}{a^2} + \dots + \frac{t^{d-1}}{a^{d-2}} - 1 = 0.$$

Multiply by a^{d-2} on both sides to obtain

$$a^{d-1} - a^{d-2} + a^{d-2}t + a^{d-3}t^2 + \dots + at^{d-2} + t^{d-1} = 0.$$

To conclude that p_1, p_2, \dots, p_{d-1} can all be expressed as a function of $p_0 = a$ it remains to show that the previous polynomial has a unique root in $[0, 1]$ for each $a \in (0, 1]$. This can be seen by using the intermediate value theorem: let $f(t) = a^{d-1} - a^{d-2} + a^{d-2}t + a^{d-3}t^2 + \dots + at^{d-2} + t^{d-1}$. Then $f(0) = a^{d-1} - a^{d-2} \leq 0$ and $f(1) = a^{d-1} + a^{d-3} + a^{d-4} + \dots + a + 1 > 0$, showing that f has a root in $[0, 1]$. This root is unique since f is strictly increasing on $[0, 1]$: $f'(t) = (d-1)t^{d-2} + a(d-2)t^{d-3} + \dots + 2a^{d-3}t + a^{d-2} > 0$ for all $t \in [0, 1]$.

Finally observe that if $p_0 = a = 0$, the only “trivial” Bernoulli measure which is invariant is $\mathcal{B}(0, \dots, 0, 1)$. \square

Proposition 3.2. *Let μ_p be a Bernoulli measure on $\{0, \dots, d-1\}^{\mathbb{N}}$, where $p = (p_0, \dots, p_{d-1})$. If μ_p is T_d -invariant, then it is ergodic.*

Proof. Simply note that the σ -algebra of invariant sets for the adic T_d is trivial by the Hewitt–Savage 0,1 Law. \square

Definition 3.3. Let (n, k) be a vertex in the graph. We define $\Lambda(n, k)$ to be the number of finite paths from $(0, 0)$ to (n, k) . For every path $x \in X$ let $\Lambda(x_n) = \Lambda(n, k_n(x))$. For every cylinder set $C \subset X$ and every path $x \in X$ define $\Lambda(C, x_n)$ to be the number of paths in C which coincide with x after the n th level.

Lemma 3.4. *Let μ be an invariant nonatomic ergodic Borel probability measure for the adic T_d . Then for every cylinder set C and μ -almost every $x \in X$*

$$\mu(C) = \lim_{n \rightarrow \infty} \frac{\Lambda(C, x_n)}{\Lambda(x_n)}. \tag{3.1}$$

Proof. Note that since μ is nonatomic, minimal and maximal paths form a set of measure zero. For any path $x \in X_d \setminus X_d^{\min} \cup X_d^{\max}$, let $i_n(x)$ and $j_n(x)$ be the integers such that $T^{-i_n(x)}x$ is the smallest path agreeing with x after the n th level, and $T^{j_n(x)}x$ is the greatest path which agrees with x after the n th level — see Fig. 3. Note that we have the relations

$$\Lambda(x_n) = i_n(x) + j_n(x)$$

and

$$\Lambda(C, x_n) = \sum_{i=-i_n(x)}^{j_n(x)} \mathbb{1}_C(T_d^i x).$$

The result follows from the Ergodic Theorem. \square

Remark 2. If μ is invariant but not necessarily ergodic, then the limit in (3.1) is just the conditional expectation $E(\mathbb{1}_C | \mathcal{I})(x)$, where \mathcal{I} is the σ -algebra of T_d -invariant sets.

The following lemma shows in particular that the invariant ergodic measures for the adic T_d are invariant for the (one-sided) shift σ on $\{0, \dots, d-1\}^{\mathbb{N}}$.

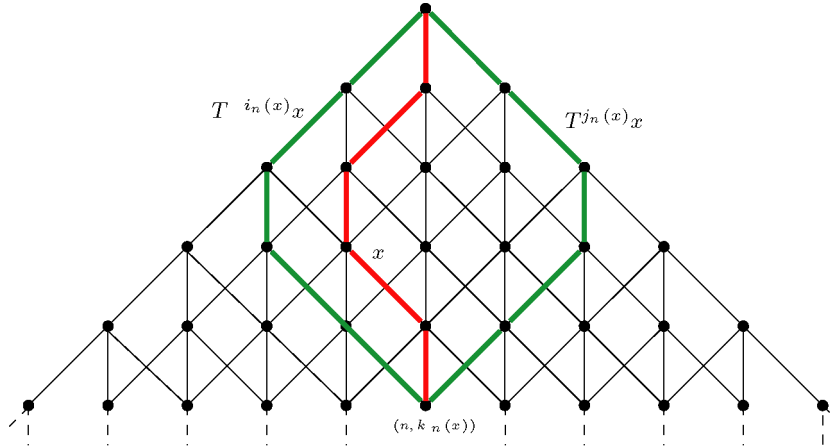


Fig. 3.

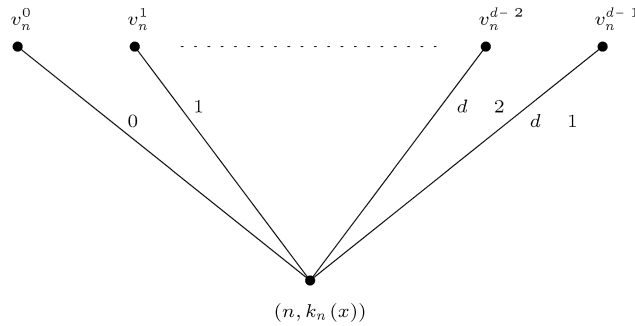


Fig. 4.

Lemma 3.5. Let $\sigma_j x = jx_0x_1 \dots$, for $j = 0, \dots, d - 1$. If μ is invariant and ergodic for the adic, then for any Borel set B

$$\mu(B) = \mu(\sigma_0 B) + \dots + \mu(\sigma_{d-1} B) = \mu(\sigma^{-1} B).$$

Proof. Let C be a cylinder set. By Lemma 3.4, for μ -almost every x we simultaneously have

$$\begin{aligned} \mu(C) &= \lim_{n \rightarrow \infty} \frac{\Lambda(C, x_n)}{\Lambda(x_n)}, \\ \mu(\sigma_j C) &= \lim_{n \rightarrow \infty} \frac{\Lambda(\sigma_j C, x_n)}{\Lambda(x_n)} \quad \text{for all } j = 0, \dots, d - 1. \end{aligned}$$

For every $j = 0, \dots, d - 1$ let x^j be the path whose vertex v_n^j at level $n - 1$ connects to $(n, k_n(x))$ along the edge labelled j , i.e. $v_n^j = (n - 1, k_n(x) - j)$ — see Fig. 4. Every path in C which coincides with x after the n th level must go through one of the vertices v_n^j ; therefore

$$\Lambda(C, x_n) = \Lambda(C, x_n^0) + \dots + \Lambda(C, x_n^{d-1}). \tag{3.2}$$

The following relation can be seen from the “isotropic” structure of the graph and is illustrated in Fig. 5:

$$\Lambda(\sigma_j C, x_n) = \Lambda(C, x_n^j), \quad \text{for all } j = 0, \dots, d - 1. \tag{3.3}$$

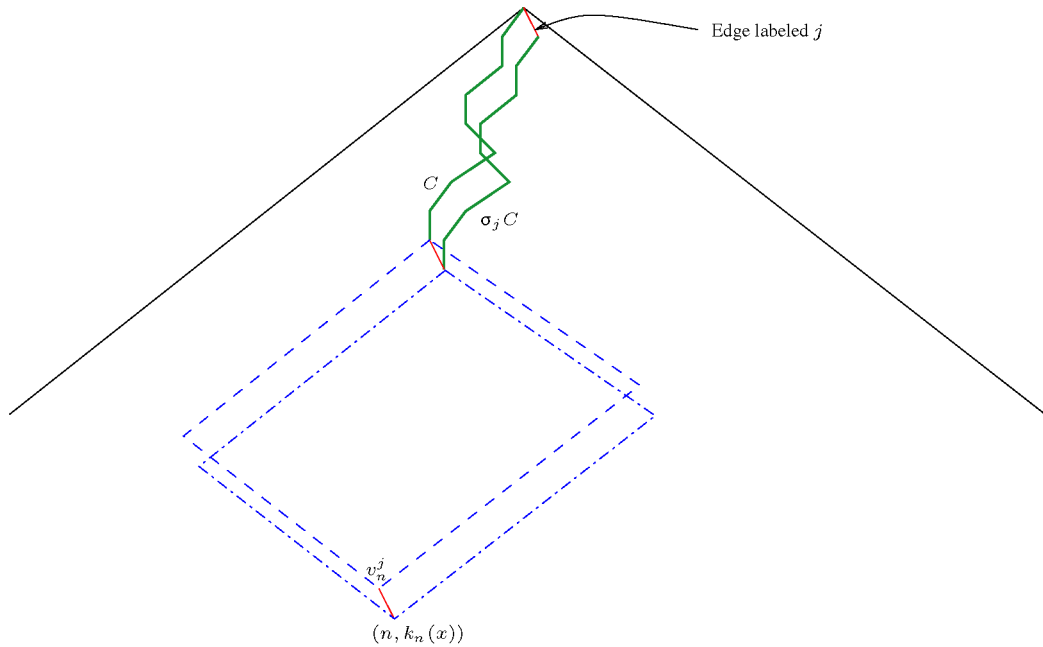


Fig. 5. The number of paths in C which coincide with x after the n th level is exactly the number of finite paths inside the dashed rectangle. The number of paths in $\sigma_j C$ which coincide with x^j after the $(n + 1)$ 'th level is exactly the number of finite paths inside the dot-dashed rectangle. Since these two rectangles have the same dimensions these numbers are equal.

Dividing (3.2) by $\Lambda(x_n)$ and then using (3.3) we get

$$\frac{\Lambda(C, x_n)}{\Lambda(x_n)} = \frac{\Lambda(C, x_n^0)}{\Lambda(x_n)} + \dots + \frac{\Lambda(C, x_n^{d-1})}{\Lambda(x_n)} = \frac{\Lambda(\sigma_0 C, x_n)}{\Lambda(x_n)} + \dots + \frac{\Lambda(\sigma_{d-1} C, x_n)}{\Lambda(x_n)}.$$

The result follows for cylinder sets after taking the limit when $n \rightarrow \infty$, and since cylinders generate the σ -algebra \mathcal{B}_d the lemma is proved. \square

Remark 3. The previous result remains true if μ is not ergodic since the argument in the proof would yield $E(\chi_C | \mathcal{I}) = E(\chi_{\sigma_0 C} | \mathcal{I}) + \dots + E(\chi_{\sigma_{d-1} C} | \mathcal{I})$ μ -a.e. (see Remark 2), and integrating we would get the same result.

Theorem 3.6. *The invariant ergodic measures for the adic T_d are the nonatomic Bernoulli measures described in Proposition 3.1, and $\mathcal{B}(1, 0, \dots, 0)$ and $\mathcal{B}(0, \dots, 0, 1)$, which are atomic.*

Proof. Let μ be an invariant nonatomic ergodic probability measure. If $\mu[j] = 0$ for some j , then $\mu[i] = 0$ for all i . Indeed, suppose that for some $i \neq j$ $\mu[i] > 0$. Then, unless μ only gives masses to maximal and minimal paths in $[i]$, there has to exist a $k \in \mathbb{Z}$ such that $\mu(T^k[i] \cap [j]) > 0$, which contradicts $\mu[j] = 0$. In the following we assume that $\mu[j] > 0$ for all $j \in \{0, \dots, d - 1\}$.

We first introduce some notation. For any cylinder set C and $j, k \in \{0, \dots, d - 1\}$, let C^k denote the cylinder C extended by the edge k and let C^{jk} be the cylinder C extended by the edges jk ; more precisely, if $C = [a_1 a_2 \dots a_r r]$ then $C^k = [a_1 a_2 \dots a_r k]$ and $C^{jk} = [a_1 a_2 \dots a_r j k]$. To prove that μ is a Bernoulli measure it is enough to show that there is a number p_k independent of C such that $\mu(C^k) / \mu(C) = p_k$ for all $k \in \{0, \dots, d - 1\}$. To show this we establish the equality

$$\frac{\mu(C^k)}{\mu(C)} = \frac{\mu(C^{jk})}{\mu(C^j)}, \quad j, k \in \{0, \dots, d - 1\}. \tag{3.4}$$

Then for any cylinder set $C = [a_1 a_2 \dots a_r]$ we have

$$\frac{\mu(C^k)}{\mu(C)} = \frac{\mu[a_1 a_2 \dots a_{r-1}]^{a_r k}}{\mu[a_1 a_2 \dots a_{r-1}]^{a_r}} = \frac{\mu[a_1 a_2 \dots a_{r-1}]^k}{\mu[a_1 a_2 \dots a_{r-1}]} = \dots = \frac{\mu[a_1]^k}{\mu[a_1]}. \tag{3.5}$$

On the other hand, the invariance of μ under T_d implies that for all $j, k, l \in \{0, \dots, d - 1\}$

$$\mu[jl] = \mu[lj] \quad \text{and} \quad \mu[jlk] = \mu[ljk],$$

and again using (3.4) we get

$$\frac{\mu[j]^k}{\mu[j]} = \frac{\mu[j]^{lk}}{\mu[j]^l} = \frac{\mu[jlk]}{\mu[jl]} = \frac{\mu[ljk]}{\mu[lj]} = \frac{\mu[l]^{jk}}{\mu[l]^j} = \frac{\mu[l]^k}{\mu[l]}. \tag{3.6}$$

Together (3.5) and (3.6) show that $\mu(C^k)/\mu(C)$ is independent of the cylinder C .

We now give a proof of the identity (3.4). By Lemma 3.4 we can find a set of full measure E such that for all $x \in E$ and all $j, k \in \{0, \dots, d - 1\}$,

$$\frac{\mu(C^{jk})}{\mu(C^k)} = \lim_{n \rightarrow \infty} \frac{\Lambda(C^{jk}, x_n)}{\Lambda(C^k, x_n)} \quad \text{and} \quad \frac{\mu(C^k)}{\mu(C)} = \lim_{n \rightarrow \infty} \frac{\Lambda(C^k, x_n)}{\Lambda(C, x_n)}.$$

Again using the “isotropic” aspect of the graph (see Fig. 6) we have that

$$\Lambda(C^k, x_n) = \Lambda(C^{jk}, (\sigma_j x)_{n+1}) \quad \text{and} \tag{3.7}$$

$$\Lambda(C, x_n) = \Lambda(C^j, (\sigma_j x)_{n+1}). \tag{3.8}$$

If we divide (3.7) by (3.8), we get

$$\frac{\Lambda(C^k, x_n)}{\Lambda(C, x_n)} = \frac{\Lambda(C^{jk}, (\sigma_j x)_{n+1})}{\Lambda(C^j, (\sigma_j x)_{n+1})}.$$

Now choosing $x \in E$ such that $\sigma_j x \in E$ and taking the limit as $n \rightarrow \infty$ yields (3.4). (Since E is a set of full measure, if $\mu(\sigma_j E) = 0$ for some $j \in \{0, \dots, d - 1\}$, then by Lemma 3.5

$$1 = \mu(E) = \sum_{k \neq j} \mu(\sigma_k E) \leq \sum_{k \neq j} \mu[k] \leq 1.$$

This forces $\mu[j] = 0$, which contradicts our assumption made in the beginning. Therefore $\mu(E \cap \sigma_j E) > 0$ for all j . It follows that $E \cap \sigma_j E \neq \emptyset$ and hence there is a $y \in E$ such that $y = \sigma_j x$ for some $x \in E$, proving that some $x \in E$ is such that $\sigma_j x \in E$.) \square

Remark 4. In the previous proof we showed that $\mu(C^k)/\mu(C)$ was independent of C avoiding calculations on the generalized binomial coefficients. On the other hand, letting (r, s) be the vertex at which the cylinder C ends (where $s = a_1 + \dots + a_r$), we could have tried to show instead that $\lim_{n \rightarrow \infty} C_d(n - r - 1, k_n(x) - s - k) / C_d(n, k_n(x))$ ($= \lim_{n \rightarrow \infty} \Lambda(C^k, x_n) / \Lambda(C, x_n) = \mu(C^k) / \mu(C)$) is independent of C . But if such calculations are easy in the case of ordinary binomial coefficients they seem very messy in the general case. Furthermore our geometric argument implies that these calculations are true; in particular, for $C = [0^r]$ we have that $C_d(n - r, k_n(x)) / C_d(n, k_n(x))$ converges to p_0^r for almost every x , which will be useful in Section 6.

3.1. The cutting and stacking equivalent

Start with the unit interval and cut it into d columns (subintervals in this case) $I_1^1, I_2^1, \dots, I_d^1$ of equal width (step 1). At each step n we cut every column I_i^n for $i = 1, \dots, (d - 1)n + 1$ into d equal subcolumns

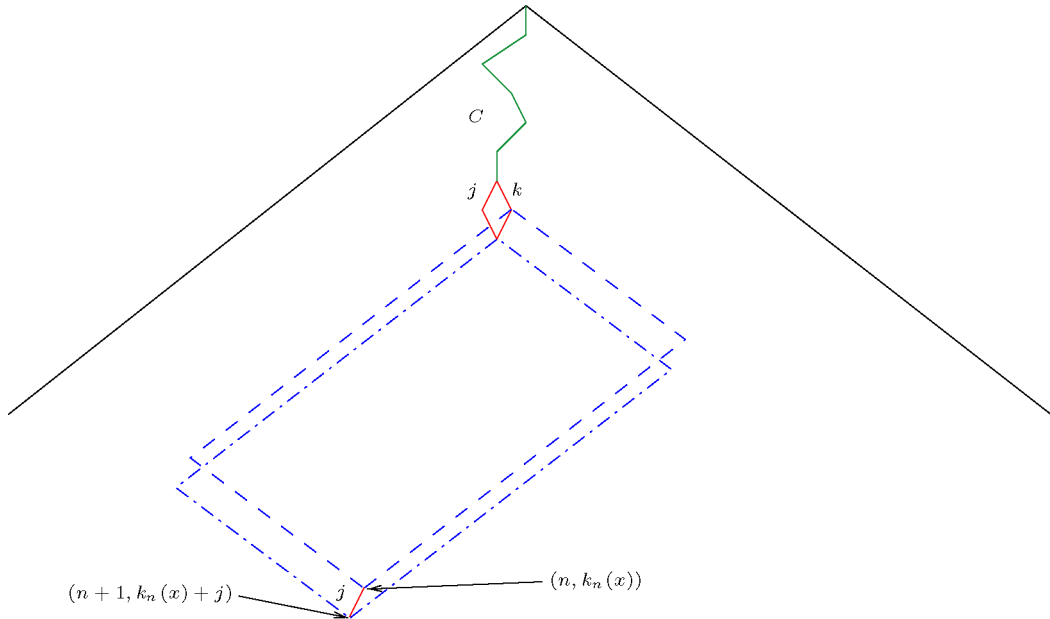


Fig. 6. The number of paths in C^k which coincide with x after the n th level is exactly the number of finite paths inside the dashed rectangle. The number of paths in C^{jk} which coincide with $\sigma_j x$ after the $(n+1)$ 'th level $n+1$ is exactly the number of finite paths inside the dot-dashed rectangle. Since these two rectangles have the same dimensions these numbers are equal.

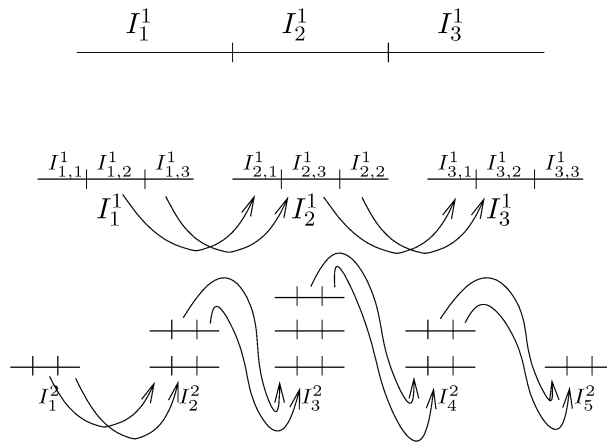


Fig. 7. Cutting and stacking in the case $d = 3$.

$I_{i,1}^n, I_{i,2}^n, \dots, I_{i,d}^n$ of equal width, then stack $I_{i,j}^n$ under $I_{i+1,j-1}^n$ for $i = 1, \dots, dn$ and $j = 2, \dots, d$. This produces $(d-1)(n+1)+1$ new columns $I_1^{n+1}, \dots, I_{(d-1)(n+1)+1}^{n+1}$. If we repeat indefinitely, the resulting map \mathcal{T} (which maps every open interval of each stack to the one above it) is defined everywhere except at the d -adic rationals (which correspond to the paths which are eventually diagonal in the graph construction). Denote by m Lebesgue measure, and let $\mathcal{B}([0, 1])$ be the σ -algebra of Borel sets in $[0, 1]$. $([0, 1], \mathcal{B}([0, 1]), \mathcal{T}, m)$ is a measure-preserving system which we will refer to as the *generalized binomial transformation*. See Fig. 7 for an illustration of the cutting and stacking in the case $d = 3$.

We can show that the generalized Pascal adic transformation equipped with the σ -algebra \mathcal{B}_d of Borel sets and the Bernoulli measure with equal weights is isomorphic to the generalized binomial transformation:

Proposition 3.7. *The systems $([0, 1], \mathcal{B}([0, 1]), \mathcal{T}, m)$ and $(X_d, \mathcal{B}_d, T_d, \mu_p)$ where $p = (1/d, \dots, 1/d)$ are isomorphic.*

Proof. The isomorphism is given by $\psi : [0, 1] \setminus \{d\text{-adic rationals}\} \rightarrow X \setminus \{y : x \in X_d^{\min} \cup X_d^{\max} \text{ and } y \in \mathcal{O}(x)\}$, which at every x with d -adic expansion $\sum_{i=1}^{\infty} x_i d^{-i}$ associates $\psi(x) = x_1 x_2 \dots$. \square

Remark 5. If we cut and stack with proportions p_0, \dots, p_{d-1} where (p_0, \dots, p_{d-1}) is a probability vector as described in Proposition 3.1, which guaranties that the intervals that are being stacked one on top of the other have the same length, then Lebesgue measure carries to the Bernoulli measure $\mathcal{B}(p_0, \dots, p_{d-1})$.

We now give a direct proof that the generalized binomial transformation is ergodic.

Proposition 3.8. *The system $([0, 1], \mathcal{B}([0, 1]), \mathcal{T}, m)$ is ergodic.*

First we need a standard result from probability, which can be found in [19, p. 85]

Lemma 3.9. *Let X_n and Y_n be two independent stationary processes on $([0, 1], \mathcal{B}([0, 1]), m)$ with values in $\{0, 1\}$ such that $\int X_n dm = \int Y_n dm$. Then for almost every (x, y) in $X \times X$ there are infinitely many n for which $\sum_{i=0}^n X_i(x) = \sum_{i=0}^n Y_i(y)$.*

Let $I = [0, 1] \setminus \{d\text{-adic rationals}\}$. For any $x \in I$, let $I_n(x)$ be the d -adic interval of length $1/d^n$ to which x belongs. Let $(x, y) \in I \times I$ have d -adic expansions

$$x = \sum_{i=1}^{\infty} \frac{x_i}{d^i} \quad \text{and} \quad y = \sum_{i=1}^{\infty} \frac{y_i}{d^i}.$$

Then it is not difficult to prove by induction that x and y belong to the same k th stack at step n in the cutting and stacking iff $k = \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. Lemma 3.9 implies that for a.e. $(x, y) \in I \times I$ there are infinitely many n for which $I_n(x)$ and $I_n(y)$ belong to the same stack at level n .

We will also need the following lemma which is a consequence of the Lebesgue density theorem and can be proved using a martingale argument:

Lemma 3.10. *Let $A \subseteq [0, 1]$ be a set of positive measure. Then for almost every $x \in A$*

$$\lim_{n \rightarrow \infty} \frac{m(A \cap I_n(x))}{m(I_n(x))} = 1.$$

Proof. Let \mathcal{F}_n denote the σ -algebra generated by the d -adic intervals $[k/d^n, (k+1)/d^n)$, for $0 \leq k \leq n-1$. Then $\mathcal{F}_n \nearrow \mathcal{B}([0, 1])$, and $X_n = E(\mathbb{1}_A | \mathcal{F}_n)$ is an \mathcal{F}_n -martingale. By the martingale convergence theorem (see for example [18]), X_n converges almost everywhere to $m(A)$. Since for almost every $x \in A$ $E(\mathbb{1}_A | \mathcal{F}_n)(x) = m(A \cap I_n(x))$, the result follows. \square

Proof of the proposition. Let A be a \mathcal{T} -invariant measurable set such that $0 < m(A) < 1$. Then, by Lemma 3.10, for almost every $x \in A$ and almost every $y \in A^c = X \setminus A$

$$\lim_{n \rightarrow \infty} \frac{m(A \cap I_n(x))}{m(I_n(x))} = 1, \tag{3.9}$$

$$\lim_{n \rightarrow \infty} \frac{m(A^c \cap I_n(y))}{m(I_n(y))} = 1. \tag{3.10}$$

On the other hand, for a.e. $(x, y) \in I \times I$ there is a sequence $n_i \nearrow \infty$ such that $I_{n_i}(x)$ and $I_{n_i}(y)$ belong to the same stack at level n_i . Therefore, since A is \mathcal{T} -invariant and $I_{n_i}(x)$ maps onto $I_{n_i}(y)$ (under some iterate of \mathcal{T}), we have

$$m(A \cap I_{n_i}(x)) = m(A \cap I_{n_i}(y)).$$

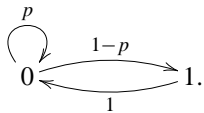
This contradicts (3.9) and (3.10), so $m(A) = 0$ or 1 . \square

3.2. A relationship between T_d and the Pascal adic T_2

Let $X_{\{11\}}$ be the subset of X_2 of all paths in the Pascal graph which do not go along the edge labelled 1 twice in a row. $X_{\{11\}}$ can be identified with the subshift of finite type of all infinite sequences not containing the block 11. Denote by $T_{\{11\}}$ the Pascal adic on $X_{\{11\}}$, that is, the map defined for $x \in X_{\{11\}}$ by $T_{\{11\}}x =$ smallest $y \in X_{\{11\}}$ greater than x . We have $T_{\{11\}}0^r(10)^s100\dots = (10)^s0^r010\dots$. In [21, Theorem 2.11] it was shown that the Pascal adic $T_{\{11\}}$ on the subshift of finite type $X_{\{11\}}$ is isomorphic to the Pascal adic $T = T_2$ on the space $\{0, 1\}^{\mathbb{N}}$. Furthermore, the preimage of a Bernoulli measure μ_p is the Markov measure with transition matrix

$$\begin{bmatrix} p & 1-p \\ 1 & 0 \end{bmatrix}$$

and initial distribution $P(0) = p$ and $P(1) = 1 - p$. The Markov shift has the following graph representation:



Let $X_{\{1^d\}}$ be the subset of X_2 of all paths in the Pascal graph which do not go along the edge labelled 1 d times in a row. $X_{\{1^d\}}$ can be identified with the subshift of finite type consisting of all sequences of 0's and 1's not containing d 1's in a row. Denote by $T_{\{1^d\}}$ the Pascal adic on the subshift $X_{\{1^d\}}$, i.e. the map defined for $x \in X_{\{1^d\}}$ by $T_{\{1^d\}}x =$ smallest $y \in X_{\{1^d\}}$ greater than x . We generalize the previous statement by relating $T_{\{1^d\}}$ and T_d as follows:

Theorem 3.11. *There is a continuous bijective map ψ from the generalized Pascal adic T_d on X_d to the Pascal adic $T_{\{1^d\}}$ on the subshift of finite type $X_{\{1^d\}}$. Furthermore, ψ commutes with the actions of the two adics, and the images of the Bernoulli measures are $(d - 1)$ -step Markov measures.*

Proof. Define a map ψ from X_d to $X_{\{1^d\}}$ by the following block codes:

$$\begin{aligned} 0 &\rightarrow 0 \\ 1 &\rightarrow 10 \\ 2 &\rightarrow 110 \\ &\vdots \\ d-1 &\rightarrow 1^{d-1}0. \end{aligned}$$

It is easy to check that ψ is continuous and bijective. To show that ψ commutes with the actions of the two adics, i.e. $T_d \circ \psi = \psi \circ T_{\{1^d\}}$, it is enough to check that for all $x, y \in X_d$ if $y > x$ then $\psi y > \psi x$. Setting $x = x_1x_2\dots$, $y = y_1y_2\dots$ and assuming $y > x$, there exists a $k \geq 1$ such that

- $y_k > x_k$,
- $x_j = y_j$ for all $j > k$,
- $\sum_{i=1}^k x_i = \sum_{i=1}^k y_i$.

Then

$$\begin{aligned} \psi x &= 1^{x_1} 0 1^{x_2} 0 \dots 0 1^{x_k} 0 \dots = 1^{x_1} 0 \dots \dots \overbrace{0 1 \dots 1}^{x_k} 0 \dots \\ \psi y &= 1^{y_1} 0 1^{y_2} 0 \dots 0 1^{y_k} 0 \dots = 1 1^{y_1} 0 \dots \underbrace{1 \dots 1}_{y_k} \dots 1 0 \dots \end{aligned}$$

Therefore, letting $l = (x_1 + 1) + (x_2 + 1) + \dots + (x_{k-1} + 1)$ be the index of the 0 preceding the string of x_k 1's in ψx , we have

- $(\psi y)_l = 1 > 0 = (\psi x)_l$,
- $(\psi x)_j = (\psi y)_j$ for all $j > l$,
- $\sum_{i=1}^l (\psi x)_i = \sum_{i=1}^k x_i + k = \sum_{i=1}^k y_i + k = \sum_{i=1}^l (\psi y)_i$,

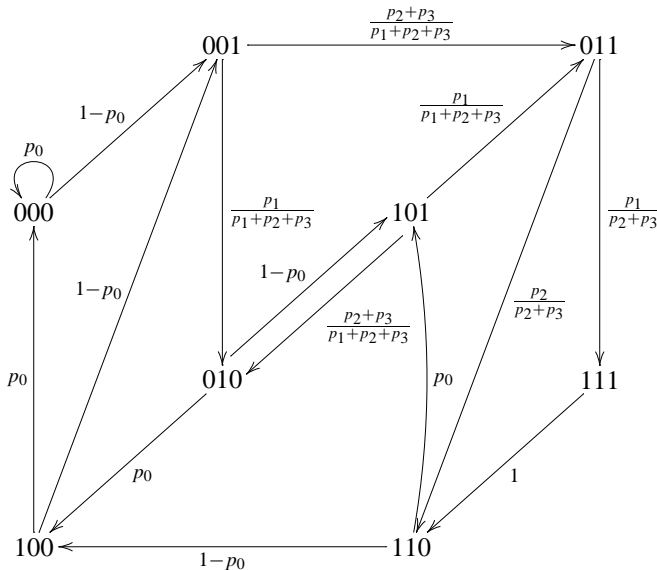
showing that $\psi y > \psi x$.

The fact that the image under ψ of a Bernoulli measure is a $(d - 1)$ -step Markov measure is straightforward, since in order to determine the probability that the n th symbol is 0 or 1 depends only on knowing the $d - 1$ previous ones. \square

Example 3.12. In the case $d = 4$, ψ is defined by

$$X_4 \xrightarrow{\psi} X_{\{1111\}} \begin{cases} 0 & \rightarrow 0, \\ 10 & \rightarrow 1, \\ 110 & \rightarrow 2, \\ 1110 & \rightarrow 3 \end{cases}$$

and the subshift of finite type $X_{\{1111\}}$, with Markov transition probabilities corresponding to $\mathcal{B}(p_0, p_1, p_2, p_3)$ on X_4 , has graph representation:



4. Existence of eigenvalues

In this section we show that the generalized Pascal adics do not have any “rational” eigenvalue, i.e. no complex number $\lambda = e^{2\pi ia/b}$, where a and b are nonzero integers, can be an eigenvalue for T_d . The existence of “irrational” eigenvalues ($\lambda^{2\pi i\alpha}$, $\alpha \notin \mathbb{Q}$) is still an open question.

First note that if a path x is such that $x_{n+1}x_{n+2} = 10$, then x comes back close to itself after $C_d(n, k_n(x))$ steps, more precisely:

Lemma 4.1. *Let $x \in X_d$ be path such that $x_{n+1}x_{n+2} = 10$. Then $T_d^{C_d(n, k_n(x))}x$ and x coincide on the first n edges.*

Proof. Let r (resp. s) be the number of iterations of T_d (resp. T_d^{-1}) necessary to bring the path x to coincide with the n first coordinates of the maximal (resp. minimal) path going through $(n, k_n(x))$. Note that $r + s + 1 = C_d(n, k_n(x))$. Since $T_d(T_d^r x)$ coincides with the n first coordinates of the minimal path going through $(n, k_n(x))$, it follows that $T_d^{r+s+1}x$ and x coincide on the first n edges. \square

This previous observation enables us to formulate a simple eigenvalue condition — cf. [16,8,13,22].

Proposition 4.2. *If λ is an eigenvalue for the generalized Pascal adic T_d , then for almost every path x in X_d (with respect to a Bernoulli measure $\mathcal{B}(p_0, \dots, p_{d-1})$) $\lambda^{C_d(n, k_n(x))}$ converges to 1.*

Proof. Let λ be an eigenvalue for T_d and f be a corresponding measurable eigenfunction ($f \circ T_d = \lambda f$). For n large, by Lusin’s Theorem we may assume that f is continuous and almost constant on the cylinder sets determined by finite paths from $(0, 0)$ to the vertices (n, k) for $k = 0, \dots, (d - 1)n$. For every path x the set of paths x' very close to x such that $x'_{n+1}x'_{n+2} = 10$ has positive measure, i.e. $\mu_p([x_1 \dots x_n 10]) > 0$. By Lemma 4.1, for all $x' \in [x_1 \dots x_n 10]$ $T_d^{C_d(n, k_n(x))}x'$ and x' coincide on the first n edges. Therefore for almost every $x' \in [x_1 \dots x_n 10]$ $f(x') \simeq f(T_d^{C_d(n, k_n(x))}x') = \lambda^{C_d(n, k_n(x))} f(x')$, so that $\lambda^{C_d(n, k_n(x))} \simeq 1$. \square

In the case $d = 2$ it is well known that the Pascal triangle modulo a prime number has a self-similar structure [2]. In particular, a consequence of Lucas’s Theorem [14] is that for any prime number q and any natural number s the binomial coefficient $C(q^s - 1, k)$ is not divisible by q for any $k = 0, \dots, q^s - 1$. Hence, in the Pascal triangle there are infinitely many levels where all binomial coefficients are not divisible by q . If $\lambda = e^{2\pi ia/b}$ ($\gcd(a, b) = 1$) is an eigenvalue, then $\lambda^{C(n, k_n(x))}$ has to converge to 1 a.e., by Proposition 4.2. But if q is a prime divisor of b , this is possible only if $C(n, k_n(x))$ is divisible by q for all n large enough. The existence of entire rows of binomial coefficients not divisible by q , forming “blocking lines”, prevent this from happening.

In the general case things are a little trickier. One has an analogue of Lucas’s Theorem for generalized binomial coefficients, unfortunately it does not imply the existence of “blocking lines”. For example, in the case of the Δ_3 -Pascal triangle modulo 2 it can be shown [1] that there are no such “blocking lines” — see Fig. 8. However, according to Proposition 4.2, in order to prove that a given $\lambda = e^{2\pi ia/b}$ is not an eigenvalue, it is enough to show that the set of paths which hit infinitely many times vertices corresponding to nonzero coefficients modulo q (where q is a prime factor of b) has positive measure. This is what we will do.

First recall the Lucas type theorem for the generalized binomial coefficients $C_d(n, k)$:

Theorem 4.3 [1,2]. *Let q be a prime number. Let $n, k \in \mathbb{N}$ be such that $0 \leq k \leq (d - 1)n$. Then*

$$C_d(n, k) = \sum_{(r_1, r_2, \dots, r_s)} \prod_{i=0}^s C_d(n_i, r_i) \pmod q,$$

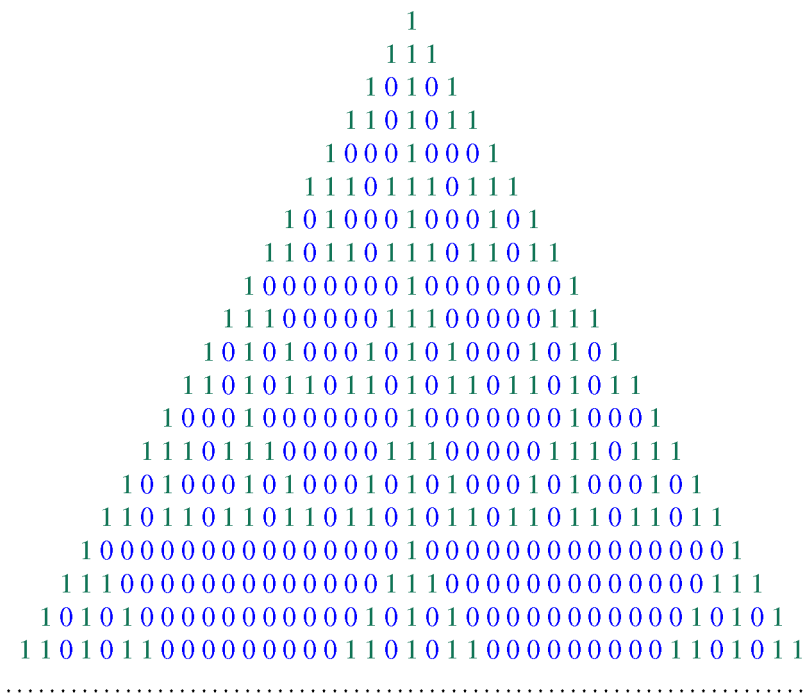


Fig. 8. The Δ_3 -Pascal triangle modulo 2.

where $n = n_0 + n_1q + \dots + n_sq^s$, with $0 \leq n_i < q$, is the q -ary representation of n , and the sum is taken over all (r_1, r_2, \dots, r_s) such that

- $r_1 + r_2q + \dots + r_sq^s = k$,
- $0 \leq r_i \leq (d - 1)n_i$

if such a representation exists; otherwise $C_d(n, k) = 0 \pmod q$.

Note that for $d = 2$ this is just Lucas' Theorem [14]. It has the following immediate consequence:

Corollary 4.4 [2]. *Let q be a prime number. Let $s \in \mathbb{N}$. Then in the row $n = q^s$ of the generalized Pascal triangle of order d , we have*

$$C_d(n, k) = \begin{cases} 1 \pmod q & \text{if } k = 0, q^s, 2q^s, \dots, (d - 1)q^s, \\ 0 \pmod q & \text{otherwise.} \end{cases}$$

This corollary tells us that at each level $n = q^s$ we see, modulo q , a row composed of d 1's each separated by $n - 1$ 0's. The 1's propagate from one level to the next forming diagonal lines (as shown in bold in Fig. 9) and we see $d - 1$ similar adjacent triangles of 0's. For example, in the case $d = 3$, we see two adjacent such triangles below each level $n = q^s$ — see Fig. 9. Recall that in this case the paths have the freedom to go straight, left, or right down the graph. The 1's in bold do not form a “hermetic barrier” playing a similar role as the “blocking lines” in the case $d = 2$. But we will show that with positive probability (independently of n) paths are forced to hit one of those 1's. This argument will be sufficient to show that there are no rational eigenvalues.

Proposition 4.5. *No $\lambda = e^{2\pi ia/b}$, where a and b are nonzero integers, can be an eigenvalue for T_d .*

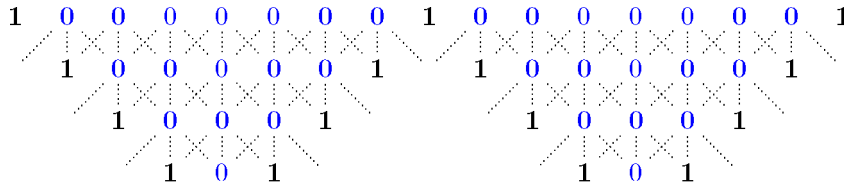


Fig. 9. Two adjacent triangles of 0's with 1's on the outside shown in the case $d = 3$ starting at level $n = 2^3$.

Proof. Assume that $\lambda = e^{2\pi ia/b}$ is an eigenvalue for T_d , where a and b are nonzero integers with $\gcd(a, b) = 1$. Let q be a prime number dividing b . By Proposition 4.2, for almost every path x $C_d(n, k_n(x)) = 0 \pmod q$ when n is large enough. We show that for set of positive measure of paths x this situation cannot happen.

First we introduce some notation. Let s a positive integer, and set $n = q^s$. Let $\{v_{1,j}^n = (n, jn) : j = 0, \dots, d - 1\}$ be the set of vertices corresponding to the d 1's at level n . A part from the two 1's on the edges at $v_{1,0}^n$ and $v_{1,d-1}^n$, each $v_{1,j}^n$ produces at the next level a 1 to its left (connected by the edge labelled $d - 1$) and a 1 to its right (connected by the edge labelled 0); we get $2d - 2$ 1's at that we denote by $v_{2,j}^n$ for $j = 0, \dots, 2d - 1$. More precisely, for $j = 0, \dots, d - 2$ we have $v_{2,2j}^n = (n + 1, jn + (d - 1))$ and $v_{2,2j+1}^n = (n + 1, jn)$. Define in a similar manner $v_{4,j}^n, v_{5,j}^n, \dots, v_{i_n,j}^n$, where the recurrence relation is $v_{i+1,2j}^n = v_{i,2j}^n + (1, d - 1) = (n + 1, jn + i(d - 1))$ and $v_{i+1,2j+1}^n = v_{i,2j+1}^n + (1, 0) = (n + 1, jn)$, and where i_n is the largest positive integer such $n - 1 - i_n(d - 1) > 0$ (corresponding to the smallest possible spacing of 0's between the 1's at the same level). The vertices $v_{i,j}^n$ correspond to the 1's in bold forming the edges of the triangle of 0's in Fig. 9. For $j = 0, \dots, d - 1$ let V_j^n be the set of infinite paths that go through the vertex $v_{1,j}^n$. For $k = 0, \dots, d - 1$ let W_k^n be the set of infinite paths going through some vertex w corresponding to a 0 and connected to a vertex $v_{i,j}^n$ by the edge labelled k . The set W_k^n is the set of all sequences going through one of the 0's in bold in Fig. 9; note that there can be 0's that have connections with several 1's. For any given $n = q^s$, every infinite path in the space must either be in W_k^n for some k , or go through one of the $v_{i,j}^n$ for some i and j . Hence X_d is the (nondisjoint) union of V_j^n and W_k^n over $j, k = 0, \dots, d - 1$, and therefore:

$$\sum_{j=0}^{d-1} \mu_p(V_j^n) + \sum_{k=0}^{d-1} \mu_p(W_k^n) \geq \mu_p(X_d) = 1. \tag{4.1}$$

Denote by C_n the set of paths caught in a 1 just after going through a triangle of zeros below the n th level, i.e. the set of paths which go through one of the $v_{i,j}^n$ for some i and j . Then, in particular, if a path in W_k^n takes the edge labelled k connecting to some $v_{i,j}^n$ it must be in C_n . From this inclusion it follows that

$$\mu_p(C_n) \geq \sum_{k=0}^{d-1} p_k \mu_p(W_k^n). \tag{4.2}$$

Since $\sum_{i=1}^d \mu_p(V_i^n)$ tends to 0 as $n = q^s$ goes to infinity, (4.1) and (4.2) imply that there has to exist a constant $c > 0$ independent of n such that $\mu_p(C_n) \geq \sum_{k=0}^{d-1} p_k \mu_p(W_k^n) \geq c$. It follows that there is a set of paths of positive measure that hit 1 infinitely many times, which concludes the proof. \square

5. Topological model

Recall from Section 2.1 that T_d is not a homeomorphism on the whole space, since it has countably many discontinuities. A natural way to study the adic as a topological dynamical system is to consider only the paths

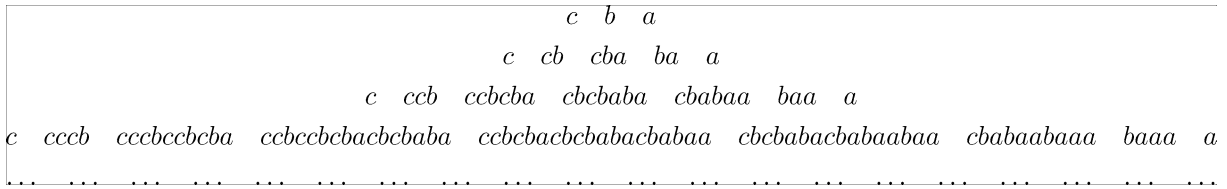


Fig. 10. Triangle of the blocks $B_{n,k}$ ’s up to level $n = 4$ (case $d = 3$).

which are not in the same orbit as a minimal or maximal path, and which therefore have a bi-infinite orbit. If we code all these paths according to the partition determined by the first coordinate we get a system which is “almost minimal”, and where the action of the adic is replaced by the action of the shift on the space of doubly infinite sequences on an alphabet with d letters. This viewpoint can be interpreted as the analogue of associating a substitution to a stationary adic (see [8]).

Let X'_d be the set of paths $x \in X_d$ such that $T_d^i x$ is minimal or maximal for some integer i . Note that X'_d is countable and that every path in $X_d \setminus X'_d$ has a well defined orbit which is bi-infinite under the action of T_d . Let P be the partition of X_d whose atoms are $[j] = \{x \in X_d : x_1 = j\}$ for $j = 0, \dots, d - 1$ and let $\mathcal{A} = \{a_0, a_1, \dots, a_{d-1}\}$ be a finite alphabet. For every $x \in X_d \setminus X'_d$ we define $\phi(x) = \dots \omega_{-2}\omega_{-1}\omega_0\omega_1\omega_2 \dots$, where $\omega_i = a_j$ if $T^i x \in [j]$. Let Σ_d be the closure of $\phi(X_d \setminus X'_d)$. It is not hard to describe all the sequences in Σ_d : they are all doubly infinite sequences in $\{a_0, a_1, \dots, a_{d-1}\}^{\mathbb{Z}}$ for which any finite subblock can be found as a subblock of one of the words $B_{n,k}$ defined as follows:

$$B_{1,0} = a_0, \quad B_{1,1} = a_1, \quad \dots, \quad B_{1,d-1} = a_{d-1},$$

$$B_{n+1,k} = B_{n,k} B_{n,k-1} \cdots B_{n,k-d+1} \quad \text{for } n > 1 \text{ and } 0 \leq k \leq (d - 1)n,$$

(with the convention that $B_{n,k}$ is the empty symbol for each $k < 0$ or $k > (d - 1)n$).

The construction of the blocks $B_{n,k}$ ’s is similar to the construction of the Δ_d -Pascal triangles, the addition is simply replaced by the concatenation of symbols. We will refer to this construction as the *triangles of blocks*. The blocks $B_{n,k}$, for $0 \leq k \leq (d - 1)n$, will be called n -generation blocks. Note that whenever $m < n$ we can always write $B_{n,k}$ as the product of m -generation blocks. We like to think of the block $B_{n,k}$ as the “DNA” of the generalized binomial coefficients $C_d(n, k)$, because not only it contains the number itself but the actual history of creation of that number.

Let $\mu_{p(a)}$ be one of the Bernoulli measures described in Proposition 3.1, set $\nu_a = \mu_{p(a)}\phi^{-1}$, and let σ be the shift transformation on $\{a_0, a_1, \dots, a_{d-1}\}^{\mathbb{Z}}$ ($(\sigma(\omega))_i = \omega_{i+1}$). The system (Σ_d, σ) can be viewed as *countable* substitution dynamical system. The language associated with Σ_d is the set of all words which are subblocks of a $B_{n,k}$ for some integers n and k ; we will denote it by $\mathcal{L}(\Sigma_d)$. (In what follows we use the appellation *block* or *word* indifferently, as well as *subblock* and *subword*.) Then

Proposition 5.1. ϕ is a measurable isomorphism between $(X_d, T_d, \mu_{p(a)})$ and $(\Sigma_d, \sigma, \nu_a)$.

Proof. One can easily check that ϕ separates points, or, equivalently, that P is a generating partition. \square

The system (Σ_d, σ) is a homeomorphism on a compact space. It has two fixed points, a_1^∞ and a_{d-1}^∞ , and only a countable set of sequences do not have dense orbits. For this system we study the complexity of the language it defines, and show that it is topologically weakly mixing ($\Sigma_d \times \Sigma_d$ topologically transitive).

Complexity. Let N be a positive integer. The complexity function $p_d(N)$ is the number of words of length N (or N -words) in $\mathcal{L}(\Sigma_d)$. We are interested in the asymptotic growth of $p_d(N)$. In [16] it was shown that the Pascal

adic transformation (or at least its topological model as presented here) has complexity asymptotic to $N^3/6$. Here we show that

Theorem 5.2.

$$o(N^3) + \frac{N^3}{6} \leq p_d(N) \leq (d-1)\frac{N^3}{6} + o(N^3).$$

Proof. We start counting the N -words of the language from the root of the triangle of blocks. We examine how many new N -words are created from one level to the next, one question being how far down do we need to look until we have seen all the N -blocks. Recall that at each level n there are $(d-1)n+1$ blocks. The $(n+1)$ -generation blocks are made of successive concatenations of adjacent n -generation blocks of the n th level and there are exactly $(d-1)n$ different concatenations performed. Observe that only the first and last letters a_0 and a_{d-1} can have several successive occurrences, but that any other letter a_i is preceded and followed by a different letter. Call *left diagonals* the set of blocks corresponding to the set of vertices $\{(n, i): i = 0, \dots, d-2 \text{ and } n > i\}$; similarly, call *right diagonals* the set of blocks corresponding to the set of vertices $\{(n, (d-1)n-i): i = 0, \dots, d-2 \text{ and } n > i\}$. For our counting we restrict our view to the triangle of blocks with these diagonals removed and explain later that their contribution is little. Note that in the truncated triangle every block starts with a power of a_{d-1} and ends with a power of a_0 , more precisely, at the n th level we see, in order from left to right (and this can be shown by induction), the follow list of blocks:

$$\begin{array}{ll} a_{d-1}^{n-1} \cdots a_0 & \\ a_{d-1}^{n-2} \cdots a_0 & a_{d-1}^{n-2} \cdots a_0 \quad \dots \quad a_{d-1}^{n-2} \cdots a_0 \quad a_{d-1}^{n-2} \cdots a_0^2 \quad (d-1) \text{ terms} \\ a_{d-1}^{n-3} \cdots a_0^2 & a_{d-1}^{n-3} \cdots a_0^2 \quad \dots \quad a_{d-1}^{n-3} \cdots a_0^2 \quad a_{d-1}^{n-3} \cdots a_0^3 \quad (d-1) \text{ terms} \\ \dots & \\ a_{d-1}^2 \cdots a_0^{n-3} & a_{d-1}^2 \cdots a_0^{n-3} \quad \dots \quad a_{d-1}^2 \cdots a_0^{n-3} \quad a_{d-1}^2 \cdots a_0^{n-2} \quad (d-1) \text{ terms} \\ a_{d-1} \cdots a_0^{n-2} & a_{d-1} \cdots a_0^{n-2} \quad \dots \quad a_{d-1} \cdots a_0^{n-2} \quad a_{d-1} \cdots a_0^{n-1} \quad (d-1) \text{ terms} \end{array}$$

(Note that the total number of terms in the list is $1 + (n-2)(d-1)$ which is exactly the total number of terms in the triangle at level n , i.e. $(d-1)n+1$, minus the $2(d-1)$ terms corresponding to the diagonals that were removed.) Note also that when we write $a_{d-1}^k \cdots a_0^{n-k-1}$ we mean that the letter following a_{d-1} is not a_{d-1} and that the letter preceding a_0 is not a_0). To look for new N -words that were not subwords of n -generation blocks we must look at words overlapping two consecutive blocks W and W' in the above list. Write them as the product of $(n-1)$ -generation blocks: $W = A_1 \cdots A_d$ and $W' = A'_1 \cdots A'_d$. Assume that the lengths of $|A_d|$ and $|A'_1|$ are greater than N , which is the case when $n > \sqrt{2N} + 1$ and $(d-1) \leq k \leq (d-1)n-1$ (since $C_d(n, k) \geq C_2(n, 2) > N$). Then observe that any new N -word overlapping W and W' is a subword of $A_d A'_1$ and must contain a central block $a_0^k a_{d-1}^{n-k-1}$ otherwise it would have already appeared at some earlier level. In particular, since the length $|a_0^k a_{d-1}^{n-k-1}|$ of a central block is $n-1$, this shows that all N -words are seen as subblocks at level $N+2$. It follows from this discussion that there are exactly $N-n+2$ new N -word overlapping W and W' , provided $n > \sqrt{2N}$ and $2 \leq k \leq 2n-1$. For the remaining pairs (n, k) , including those in the left and right diagonals, one can easily show using a coarse estimation that the proportion of N -words occurring there is a $o(N^3)$. Now observe that there are $n-2$ types of overlapping central blocks: the blocks $a_0^k a_{d-1}^{n-k-1}$, for $k = 1, \dots, n-2$. Each one is seen $d-1$ times. The present analysis does not enable us to decide whether two different appearances of a given central block $a_0^k a_{d-1}^{n-k-1}$ could produce the same N -word, therefore we estimate that a new N -word can be seen up to $d-1$ times. It follows that

$$\begin{aligned} \frac{1}{d-1} \sum_{n>\sqrt{2N}}^{N+1} (N-n+2)[(d-1)(n-2)] + o(N^3) &\leq p_d(N) \\ &\leq \sum_{n>\sqrt{2N}}^{N+1} (N-n+2)[(d-1)(n-2)] + o(N^3), \end{aligned}$$

which after simplification leads to the result we announced. \square

Remark 6. It would be interesting to find more precise asymptotics or even an exact formula for $P(N)$, but a deeper understanding of the structure of the blocks $B_{n,k}$ is required.

Remark 7. Does the system (Σ_d, σ_d) satisfy a *recognizability* property (see [6] for the definition)? In particular, we conjecture that for $k/n \simeq \alpha$, there is $\delta > 0$, such that for n large enough, if W is an N -word ($N = C_d(n, k)$), then $\bar{d}(B_{n,k}, W) < \delta$ implies that $W = B_{n,k}$. This is connected to the question of rank of the systems.

Topological weak mixing. It is well known that for minimal systems topological weak mixing is equivalent to not having any nonconstant continuous eigenfunctions. Here the system is not minimal, but as in [16] we can show that (Σ_d, σ) is topologically weakly mixing using a criterion from [12]. The proof is a straightforward adaptation of the one given in [16], and we therefore refer the reader to that previous paper for full details.

Keynes and Robertson’s result says that (under certain hypotheses) a system is topologically weakly mixing if one can show that for every bounded measurable eigenfunction f such that $f \circ T = \lambda f$ everywhere, is constant on its set of continuity points.

Theorem 5.3. (Σ_d, σ) is topologically weakly mixing.

Proof. The outline of the proof is as follows:

- (i) We know already from Proposition 4.5 that there are no rational eigenvalues, so let $\lambda = e^{2\pi i\alpha}$ with $\alpha \notin \mathbb{Q}$ be an eigenvalue associated to some bounded measurable function f whose points of continuity is a residual set.
- (ii) Show that every point $\omega \in \Sigma_d$ which has a dense forward orbit and a dense backward orbit is also a point of continuity of f .
- (iii) Use Weyl’s criterion on uniform distribution (in particular implying that $\{C_d(n, k)\alpha : n \geq 1\}$ is dense modulo 1) to produce a point $\omega \in \Sigma_d$ with a dense forward and backward orbit and a sequence of integers $N_k \rightarrow \infty$ such that $\lambda^{N_k} \rightarrow -1$ and $\sigma^{N_k}(\omega) \rightarrow \omega$.
- (iv) Conclude that the equality $f \circ \sigma = \lambda f$ together with the previous facts give a contradiction. Therefore $\lambda = 1$ and by topological ergodicity f is constant on its set of continuity points. \square

6. The loosely Bernoulli property

The notion of loosely Bernoulli was introduced by Feldman [4] in the context of Kakutani equivalence (see also [9–11]). Examples of zero entropy transformations which are loosely Bernoulli include irrational rotations, ergodic interval exchange transformations, and the horocycle flow. For further details see Ornstein, Rudolph and Weiss [17,5]. In this section we use the ideas and results of de la Rue and Janvresse [3] to show that the systems $(X_d, \mathcal{B}_d, \mu_{p(a)}, T_d)$ are loosely Bernoulli.

First we recall the definition of the \bar{f} distance which Feldman (and Katok independently) used to define the notion of loosely Bernoulli. Given two words w and w' in $\{0, 1, \dots, d - 1\}^l$,

$$\bar{f}(w, w') = 1 - \frac{s}{l},$$

where s is the largest integer for which we can find $i_1 < i_2 < \dots < i_s$ and $j_1 < j_2 < \dots < j_s$ such that $w_{i_k} = w'_{j_k}$ for $1 \leq k \leq s$. (For example: $\bar{f}(101010, 010101) = 1/6$.)

In the following assume that (X, \mathcal{A}, μ, T) is an invertible measure-preserving system with zero entropy. Given a partition $P = \{P_0, P_1, \dots, P_{d-1}\}$ of X and $x \in X$, we call P - l -name of x , denoted by $P_l(x)$, the word $\omega_0\omega_1 \dots \omega_{l-1}$, where $\omega_i = j$ if $T^i x \in P_j$.

Definition 6.1. The system (X, \mathcal{A}, μ, T) is loosely Bernoulli if for every finite partition P and every $\epsilon > 0$ there is an integer L such that for all $l > L$ we can find a set $E \in \mathcal{A}$ with $\mu(E) > 1 - \epsilon$ such that for all $x, y \in E$, $\bar{f}(P_l(x), P_l(y)) < \epsilon$.

If the previous statement holds for a generating partition P then it holds for any partition.

The following result, which does not assume ergodicity, provides a simplification of the previous definition:

Lemma 6.2 (de la Rue and Janvresse [3]). *If P is a generating partition, and for all $\epsilon > 0$ and $\mu \otimes \mu$ a.e. $(x, y) \in X \times X$ there is an integer l such that $\bar{f}(P_l(x), P_l(y)) < \epsilon$, then (X, \mathcal{A}, μ, T) is loosely Bernoulli.*

Using the previous lemma, it is not difficult to see that our systems $(X_d, \mathcal{B}_d, \mu_{p(a)}, T_d)$ are loosely Bernoulli. Let P be the partition defined by the first coordinate as in Section 5. Since P is a generating partition (Proposition 5.1), it is enough to check the statement of Lemma 6.2 for P . The idea is the following: If two paths $x, y \in X_d$ cross at some vertex (n, k) it means that each of their P -names begins with a suffix of the block $B_{n,k}$. But say they also crossed at the vertex $(n - 1, k)$ (so that they follow the same edge labelled 0 from level $n - 1$ to level n). Then, since $B_{n,k} = B_{n-1,k}B_{n-1,k-1} \dots B_{n-1,k-d+1}$, it means that the P - l -names (where $l = |B_{n,k}|$) of x and y both actually start somewhere in $B_{n-1,k}$ and hence the \bar{f} distance between them is not too large. Now suppose that x and y actually crossed at some vertex $(n - r, k)$ and coincided up to (n, k) ; for a good choice of r large enough, this would mean that the P_l -names of x and y start in the same $B_{n-r,j}$ in the decomposition of $B_{n,k}$ into $(n - r)$ -generation blocks, which would imply that the $\bar{f}(P_l(x), P_l(y))$ is small. If we can show that this situation is actually generic, then Lemma 6.2 implies loosely Bernoulli. We now turn to the formal proof which begins with a simple lemma which is an exercise in probability:

Lemma 6.3 (see also [3]). *Let X_n and Y_n be two independent i.i.d. (Bernoulli) processes on the probability space (Ω, \mathcal{F}, P) with initial law $P(X_0 = j) = P(Y_0 = j) = p_j$, for $j = 0, \dots, d - 1$ (for all j , $p_j > 0$ and $\sum_{j=0}^{d-1} p_j = 1$). Set $S_n = \sum_{i=1}^n X_i$, $R_n = \sum_{i=1}^n Y_i$, and let $r \in \mathbb{N}$ be given. Then almost surely there are infinitely many n such that $S_n = R_n$ and $X_{n-j} = Y_{n-j} = 0$ for $j = 1, \dots, r$.*

In our context $(\Omega = X_d, X_n(x) = x_n, Y_n(y) = y_n)$ the previous lemma implies that for $\mu \otimes \mu$ a.e. $(x, y) \in X_d \times X_d$, x and y cross infinitely often while joining along the edge labelled 0 r times in row.

Theorem 6.4. $(X_d, \mathcal{B}_d, \mu_{p(a)}, T_d)$ is loosely Bernoulli.

Proof. Let $\epsilon > 0$ be given. Choose r so that $p_0^r < \epsilon/2$. By Lemma 6.3, for $\mu \otimes \mu$ a.e. $(x, y) \in X_d \times X_d$ we can find arbitrarily large n for which $(n, k_n(x)) = (n, k_n(y))$ and $x_{n-j} = y_{n-j} = 0$ for $j = 1, \dots, r$. Let n be an integer such that the previous statement holds and large enough so that $C_d(n - r, k_n(x))/C_d(n, k_n(x)) \simeq p_0^r$ (this is possible since $C_d(n - r, k_n(x))/C_d(n, k_n(x))$ tends to $\mu_{p(a)}[0^r] = p_0^r$ by ergodicity (see Remark 4) — a direct computation

seems difficult). Let $l = C_d(n, k)$ and observe that the P - l -names of x and y share a common word of length $l - 2C_d(n - r, k)$ (see the preliminary discussion). Hence $\tilde{f}(P_l(x), P_l(y)) \leq 2C_d(n - r, k)/C_d(n, k) < \epsilon$. \square

Acknowledgements

The author was supported by the Millennium Nucleus in Information and Randomness, Programa Iniciativa Científica Milenio P01-005. Some of the results come from the author's thesis [15]. I would like to thank Karl Petersen, Xavier Bressaud and Christian Paroissin for helpful discussions. The paper was completed while the author was a postdoctoral researcher with the CNRS at the IML in Luminy, Marseille.

References

- [1] R. Bollinger, C. Burchard, Lucas's theorem and some related results for extended Pascal triangles, *Amer. Math. Monthly* 97 (3) (1990) 198–204.
- [2] B.A. Bondarenko, *Generalized Pascal Triangles and Pyramids, their Fractals, Graphs and Applications*, Fan, Tashkent, 1990.
- [3] T. de la Rue, E. Janvresse, The Pascal adic transformation is loosely Bernoulli, *Ann. Inst. H. Poincaré Probab. Statist.* 40 (2004) 133–139.
- [4] J. Feldman, New K -automorphisms and a problem of Kakutani, *Israel J. Math.* 24 (1976) 16–38.
- [5] S. Ferenczi, Systèmes localement de rang un, *Ann. Inst. H. Poincaré* 20 (1984) 35–51.
- [6] N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics, and Combinatorics*, Lecture Notes in Math., vol. 1794, Springer-Verlag, 2002.
- [7] A. Hajian, Y. Ito, S. Kakutani, Invariant measures and orbits of dissipative transformations, *Adv. in Math* 9 (1972) 52–65.
- [8] B. Host, Substitution subshifts and Bratteli diagrams, in: F. Blanchard, A. Maass, A. Nogueira (Eds.), *Topics in Symbolic Dynamics and Applications*, in: London Math. Soc. Lecture Note Ser., Cambridge University Press, 1999.
- [9] A. Katok, Time change, monotone equivalence, and standard dynamical systems, *Dokl. Akad. Nauk SSSR* 223 (4) (1975) 789–792 (in Russian).
- [10] A. Katok, Monotone equivalence in ergodic theory, *Izv. Akad. Nauk SSSR Ser. Mat.* 41 (1) (1977) 104–157 (in Russian).
- [11] A. Katok, E. Sataev, Standardness of rearrangement automorphisms of segments and flows on surfaces, *Mat. Zametki* 20 (4) (1976) 479–488 (in Russian).
- [12] H. Keynes, J. Robertson, Eigenvalue theorems in topological transformation groups, *Trans. Amer. Math. Soc.* 139 (1969) 359–369.
- [13] A. Livshitz, A sufficient condition for weak mixing of substitutions and stationary adic transformations, *Math. Notes* 44 (1988) 920–925.
- [14] E. Lucas, Théorie des fonctions numériques simplement périodiques, *Amer. J. Math.* 1 (1878) 184–240.
- [15] X. Méla, *Dynamical properties of the Pascal adic and related systems*, Ph.D. thesis, University of North Carolina at Chapel Hill, 2002.
- [16] X. Méla, K. Petersen, Dynamical properties of the Pascal adic transformation, *Ergodic Theory Dynam. Systems* 25 (2005) 227–256.
- [17] D. Ornstein, D. Rudolph, B. Weiss, Equivalence of measure preserving transformations, *Mem. Amer. Math. Soc.* 37 (1982).
- [18] W. Parry, *Topics in Ergodic Theory*, Cambridge University Press, 1981.
- [19] K. Petersen, *Ergodic Theory*, Cambridge University Press, 1989.
- [20] K. Petersen, T. Adams, Binomial coefficient multiples of irrationals, *Monatsh. Math.* 125 (4) (1998) 269–278.
- [21] K. Petersen, K. Schmidt, Symmetric Gibbs measures, *Trans. Amer. Math. Soc.* 349 (7) (1997) 2775–2811.
- [22] B. Solomyak, On the spectral theory of adic transformations, *Adv. Soviet Math.* 9 (1992) 217–230.
- [23] A. Vershik, Uniform algebraic approximation of shift and multiplicative operators, *Dokl. Akad. Nauk SSSR* 218 (24) (1981) 526–529.