The Pascal adic transformation is loosely Bernoulli

Élise Janvresse, Thierry de la Rue

Université de Rouen, LMRS, UMR 6085, CNRS, 76821 Mont Saint Aignan, France

Received 22 April 2003; accepted 2 September 2003

Abstract

The Pascal adic transformation is one of the simplest examples of adic transformations. We recall its construction by cutting and stacking and prove that it is loosely Bernoulli.

Résumé

La transformation Pascal adique est un des exemples les plus simples de transformations adiques. Nous rappelons sa construction par découpage et empilement et montrons qu’elle est lâchement Bernoulli.

MSC: 28D05

Keywords: Adic transformation; Loosely Bernoulli systems

Mots-clés : Transformation adique ; Systèmes lâchement Bernoulli

1. Introduction

The notion of \textit{adic transformation} has been introduced by Vershik (see e.g. [4,5]), as a model in which the transformation acts on infinite paths in some graphs, called \textit{Bratteli diagrams}. As shown by Vershik, every ergodic automorphism of the Lebesgue space is isomorphic to some adic transformation, with a Bratteli diagram which may be quite complicated. Vershik also proposed to study the ergodic properties of an adic transformation in a given simple graph, such as the Pascal graph which gives rise to the so-called \textit{Pascal adic transformation}.

1.1. The Pascal adic transformation

Here we recall the construction and some basic properties of the Pascal adic transformation with parameter $p$, following the cutting and stacking model exposed in [2]. Our space $X$ is the interval $[0, 1]$, equipped with its Borel $\sigma$-algebra $\mathcal{A}$ and the Lebesgue measure $\mu$. 

E-mail addresses: elise.janvresse@univ-rouen.fr (É. Janvresse), thierry.delarue@univ-rouen.fr (T. de la Rue).
Let $0 < p < 1$ be a fixed parameter. We start by dividing $X$ into two subintervals $P_0 \overset{\text{def}}{=} [0, p]$ and $P_1 \overset{\text{def}}{=} [p, 1]$. Let $P \overset{\text{def}}{=} \{P_0, P_1\}$ be the partition obtained at this first step. We also consider $P_0$ and $P_1$ as “degenerate” Rokhlin towers of height 1, respectively denoted by $\tau_0$ and $\tau_1$.

On second step, $P_0$ and $P_1$ are divided in proportions $(p, 1 - p)$. The transformation $T$ is defined on the right piece of $P_0$ by sending it linearly onto the left piece of $P_1$; note that both intervals have the same length $p(1 - p)$.

This gives a collection of 3 disjoint Rokhlin towers denoted by $\tau_2$, $\tau_3$, $\tau_4$, with respective heights 1, 2, 1 (see Fig. 1).

After step $n$, we get $(n + 1)$ towers $\tau_0^n, \ldots, \tau_n^n$, with respective heights $(0^n), \ldots, (n^n)$, the width of $\tau_k^n$ being $p^{n-k}(1 - p)^k$. Denote by $F_k^n$ the base of $\tau_k^n$. At this step, the transformation $T$ is defined on the whole space except the top of each stack. We then divide each stack in proportions $(p, 1 - p)$, and define $T$ on the right piece of the top of $\tau_k^n$ by sending it linearly onto the left piece of the base $F_k^n$ of $\tau_k^{n+1}$ (both have the same length $p^{n-k}(1 - p)^{k+1}$).

Repeating recursively this construction, $T$ is finally defined almost everywhere, and clearly preserves the measure $\mu$.

It is well-known (see e.g. the proofs given in [2]) that $T$ is ergodic and has zero entropy.

1.2. Loose Bernoullicity

In this section and in Section 2.1, we consider a general dynamical system $(X, A, \mu, T)$, where $T$ is an invertible measure-preserving transformation of the Lebesgue probability space $(X, A, \mu)$. The notion of loose Bernoullicity has been introduced by Feldman in 1976 [1], then used by Ornstein, Rudolph and Weiss [3] to develop the study of Kakutani equivalence for measure preserving transformations. In the zero-entropy case, saying that a transformation $T$ is loosely Bernoulli is equivalent to saying that $T$ is isomorphic to a transformation induced by an irrational rotation. The characterization of loose Bernoullicity given by Feldman makes use of the so-called “$P$-name” of a point $x$.

Let $P = \{P_0, \ldots, P_k\}$ be a finite measurable partition of $(X, A, \mu)$. For $x \in X$, we set $P(x) \overset{\text{def}}{=} j \in \{0, \ldots, k\}$ if $x \in P_j$. For $m < n$ in $\mathbb{Z}$, we define the $P$-name of $x$ (from $m$ to $n$) by

$$P|_m^n(x) \overset{\text{def}}{=} j_m j_{m+1} \cdots j_n,$$

where, for each $m \leq i \leq n$, $j_i \overset{\text{def}}{=} P(T^i x)$. The entire $P$-name of $x$ is the doubly-infinite sequence $P|_{-\infty}^{+\infty}(x)$. 

Fig. 1. Cutting and stacking construction of the Pascal adic transformation.
To define the property of being loosely Bernoulli, Feldman introduced the $\bar{f}$ distance between finite words. Let $V = v_1 \cdots v_l$ and $w = w_1 \cdots w_l$ be two words of length $l$ on the same alphabet. The $\bar{f}$ distance between $v$ and $w$ is defined by

$$\bar{f}(v, w) \overset{\text{def}}{=} \frac{l - s}{l},$$

where $s$ is the greatest integer in $\{0, \ldots, l\}$ such that we can find $1 \leq i_1 < i_2 < \cdots < i_s \leq l$ and $1 \leq j_1 < j_2 < \cdots < j_s \leq l$ with $v_{i_r} = w_{j_r}$ ($r = 1, \ldots, s$).

**Definition 1.1.** Let $T$ be a zero-entropy measure preserving transformation on the probability space $(X, A, \mu)$, and let $P$ be a finite measurable partition of $X$. The process $(P, T)$ is said to be loosely Bernoulli (LB) if for all $\epsilon > 0$ and for all sufficiently large $l$, we can find $A \subset X$ with $\mu(A) > 1 - \epsilon$ such that

$$\forall x, y \in A, \quad \bar{f}(P_{\lfloor l \rfloor}(x), P_{\lfloor l \rfloor}(y)) < \epsilon.$$

The transformation $T$ is said to be LB if for each finite partition $P$ the process $(P, T)$ is LB.

**Remark.** In order to prove that a transformation $T$ is LB, it is enough to verify that $(P, T)$ is LB for some generating partition $P$.

1.3. Main result

**Theorem 1.2.** The Pascal-adic transformation is loosely Bernoulli.

2. Proof of the loose-Bernoullicity

2.1. Equivalence of loose-Bernoullicity with seemingly weaker properties

**Lemma 2.1.** Suppose that for all $\epsilon > 0$ and for all sufficiently large $l$, we can find $B \subset X \times X$ with $\mu \otimes \mu(B) > 1 - \epsilon$ such that

$$\forall (x, y) \in B, \quad \bar{f}(P_{\lfloor l \rfloor}(x), P_{\lfloor l \rfloor}(y)) < \epsilon.$$

then the process $(P, T)$ is LB.

**Proof.** Given $\epsilon > 0$, let $B \subset X \times X$ with $\mu \otimes \mu(B) > 1 - \epsilon$ be such that

$$\forall (x, y) \in B, \quad \bar{f}(P_{\lfloor l \rfloor}(x), P_{\lfloor l \rfloor}(y)) < \epsilon/2.$$

We can find $x \in X$ such that $\mu(B_x) > 1 - \epsilon$, where

$$B_x \overset{\text{def}}{=} \{y \in X \mid (x, y) \in B\}.$$

But, because of the triangular inequality for $\bar{f}$, for all $y$ and $y'$ in $B_x$ we have

$$\bar{f}(P_{\lfloor l \rfloor}(y), P_{\lfloor l \rfloor}(y')) < \epsilon.$$

Thus, the definition of LB is satisfied, with $A \overset{\text{def}}{=} B_x$. $\Box$

**Lemma 2.2.** Suppose that for all $\epsilon > 0$ and for $\mu \otimes \mu$-almost every $(x, y) \in X \times X$, we can find an integer $l(x, y) \geq 1$ such that

$$\bar{f}(P_{\lfloor l(x, y) \rfloor}(x), P_{\lfloor l(x, y) \rfloor}(y)) < \epsilon,$$

then the process $(P, T)$ is LB.
Proof. Let us fix \( \varepsilon > 0 \). For any \( \mu \otimes \mu \)-almost every \((x, y) \in X \times X\), we define \( l(x, y) \) as the smallest integer \( k \geq 1 \) such that \( \hat{f}(\mathcal{P}_{l}^{k}(x), \mathcal{P}_{l}^{k}(y)) < \varepsilon / 3 \). Since \( \mu \otimes \mu(l(x, y) < \infty) = 1 \), there exists \( n \in \mathbb{N}^* \) such that
\[
\mu \otimes \mu(l(x, y) \geq n) < \varepsilon^2 / 3.
\]

For any \( l > 3n / \varepsilon \), we consider
\[
M_l \overset{\text{def}}{=} \frac{1}{l} \sum_{k=0}^{l-1} 1_{\{l(T^k x, T^k y) \geq n\}}.
\]

Using Markov’s inequality and the fact that \( T \) preserves the measure \( \mu \), one can easily check that
\[
\mu \otimes \mu(M_l \geq \varepsilon / 3) \leq \frac{E(M_l)}{\varepsilon / 3} < \frac{\varepsilon^2 / 3}{\varepsilon / 3} = \varepsilon.
\]

Therefore, the set \( B \overset{\text{def}}{=} \{M_l < \varepsilon / 3\} \subset X \times X \) is such that \( \mu \otimes \mu(B) > 1 - \varepsilon \). Let us fix \((x, y) \in B\). We want to show that \( \hat{f}(\mathcal{P}_{l}^{i}(x), \mathcal{P}_{l}^{i}(y)) < \varepsilon \). We say that \( k \in [0, \ldots, l-1] \) is bad if \( l(T^k x, T^k y) > n \). Since \((x, y) \in B\), there are less than \( l \varepsilon / 3 \) such \( k \).

We define \((j_i)_{i \geq 0}\) and \((r_i)_{i \geq 0}\) recursively by \( j_0 = r_0 \overset{\text{def}}{=} \inf\{r \geq 0 \mid r \text{ is not bad}\} \), and for \( i \geq 1 \) such that \( j_{i-1} \leq l - n \),
\[
r_i = \inf\{r > 0 \mid j_{i-1} + l(T^{j_{i-1}} x, T^{j_{i-1}} y) + r \text{ is not bad}\},
\]
\[
j_i = j_{i-1} + l(T^{j_{i-1}} x, T^{j_{i-1}} y) + r_i.
\]

We denote by \( f \) the greatest index \( i \) such that \( j_i \) is defined: \( l - j_f < n \) (see Fig. 2).

Recall the definition of \( \hat{f} \):
\[
(l + 1) \hat{f}(\mathcal{P}_{l}^{i}(x), \mathcal{P}_{l}^{i}(y))
\]
\[
\leq \sum_{i=0}^{f-1} (j_{i+1} - j_i) \hat{f}(\mathcal{P}_{j_i}^{j_{i+1}}(x), \mathcal{P}_{j_i}^{j_{i+1}}(y)) + (l - j_f)
\]
\[
\leq \sum_{i=0}^{f-1} l(T^{j_i} x, T^{j_i} y) \hat{f}(\mathcal{P}_{j_i}^{j_{i+1}}(x), \mathcal{P}_{j_i}^{j_{i+1}}(T^{j_i} x, T^{j_i} y)) + \sum_{i=0}^{f-1} r_i + (l - j_f)
\]
\[
= \sum_{i=0}^{f-1} l(T^{j_i} x, T^{j_i} y) \hat{f}(\mathcal{P}_{0}^{j_{i}}(x), \mathcal{P}_{0}^{j_{i}}(T^{j_i} x, T^{j_i} y)) + \sum_{i=0}^{f-1} r_i + (l - j_f)
\]
\[
\leq \frac{\varepsilon}{3} \sum_{i=0}^{f-1} l(T^{j_i} x, T^{j_i} y) + \frac{l \varepsilon}{3} + n < (l + 1) \varepsilon.
\]

Therefore, we proved that for all sufficiently large \( l \), we can find \( B \subset X \times X \) with \( \mu \otimes \mu(B) > 1 - \varepsilon \) such that \( \Psi(x, y) \in B \), \( \hat{f}(\mathcal{P}_{l}^{i}(x), \mathcal{P}_{l}^{i}(y)) < \varepsilon \). We conclude with Lemma 2.1. \( \square \)
2.2. Some lemmas on the Pascal adic transformation

From now on, \( T \) is the Pascal adic transformation described in Section 1.1, and \( \mathcal{P} \) is the partition \( \{P_0, P_1\} \) given by the first step of the cutting-and-stacking construction. For \( x \in X \) and \( n \geq 1 \), we define \( k_n(x) \) as the element of \( \{0, \ldots, n\} \) telling in which tower of the level \( n \) \( x \) lies: for each \( n \geq 1 \), \( x \in \tau^n_{k_n(x)} \).

**Lemma 2.3.** \( \mathcal{P} \) is a generating partition for the system \((X, A, \mu, T)\), i.e.

\[
\bigvee_{k=+\infty}^{+\infty} T^k \mathcal{P} = \mathcal{A}.
\]

**Proof.** As in [2], for each \( n \geq 1 \), we define the basic blocks of level \( n \) \( B_{n,k} \) \((0 \leq k \leq n)\), which are words on the alphabet \([0, 1]\), by the following induction: \( B_{n,0} \overset{\text{def}}{=} 0 \), \( B_{n,n} \overset{\text{def}}{=} 1 \), and for \( 1 \leq k \leq n - 1 \),

\[
B_{n,k} \overset{\text{def}}{=} B_{n-1,k-1} B_{n-1,k}.
\]

It is straightforward to verify that \( B_{n,k} \) is the \( \mathcal{P} \)-name of length \( n \) of any point \( x \) lying in the base \( F^n_k \) of \( \tau^n_k \). We are now going to prove by induction on \( n \) that \( B_{n,k} \) characterizes the base of \( \tau^n_k \). More precisely, for any \( n \geq 2 \) and \( 1 \leq k \leq n - 1 \),

\[
\text{if } \mathcal{P} \big|^{(n-1)} \mathcal{P}^{-1}(x) = B_{n,k}, \text{ then } x \in F^n_k.
\] (1)

Indeed, (1) is clearly satisfied for \( n = 2 \). Next, suppose that (1) is satisfied for \( n - 1 \), and pick an \( x \) such that

\[
\mathcal{P} \big|^{(n-1)} \mathcal{P}^{-1}(x) = B_{n-1,k-1},
\] (2)

so that \( x \in F^{n-1}_{k-1} \), and

\[
\mathcal{P} \big|^{(n-1)} 
(T^{(n-1)} \mathcal{P}^{-1}) \big|^{(n-1)}(x) = B_{n-1,k},
\] (3)

which implies \( T^{(n-1)} \mathcal{P}^{-1} x \in F^{n-1}_k \). Thus, after climbing the tower \( \tau^{n-1}_k \), the image of \( x \) goes to the next tower \( \tau^{n-1}_k \), which is possible only if \( x \in F^n_k \) (otherwise, the image of \( x \) would go back to \( F^{n-1}_{k-1} \)). For the case \( k = 1 \), we first have to notice that

\[
\forall m \geq 2, \forall 1 \leq j \leq m - 1, \quad B_{m,j} \text{ begins with } \text{“0”} \text{ and ends with } \text{“1”}.
\] (4)

(We leave to the reader the verification of (4) by induction on \( m \).) Now, if \( \mathcal{P} \big|^{(n-1)} \mathcal{P}^{-1}(x) = B_{n,1} = 0 B_{n-1,1} \), we know that \( Tx \in F^{n-1}_1 \) because (1) is true for \( n - 1 \), and then we can tell that \( x \in F^{n-1}_0 \); otherwise, the letter preceding \( B_{n-1,1} \) would be “1”. This yields \( x \in F^n_1 \). The case \( k = n - 1 \) is similar.

Now, for a fixed \( n \geq 1 \) we observe that the entire \( \mathcal{P} \)-name of any point \( x \) is a concatenation of basic blocks of level \( n \). Because of (1), this decomposition into basic blocks \( B_{n,k} \) is unique, and knowing the \( \mathcal{P} \)-name of \( x \) gives for any \( n \) the value of \( k_n(x) \) and tells us in which rung of \( \tau^n_{k_n(x)} \) \( x \) lies. But the partition \( \mathcal{Q}_n \) of \( X \) into rungs of towers \( \tau^n_k \), \( 0 \leq k \leq n \) is constituted of intervals whose maximal width is \( \max(p, 1 - p)^n \); moreover \( \mathcal{Q}_{n+1} \) refines \( \mathcal{Q}_n \). Therefore \( \bigvee_{n \geq 1} \mathcal{Q}_n = \mathcal{A} \).

**Lemma 2.4.** For \( \mu \)-almost every \( x \in X \), we have

\[
k_n(x) \rightarrow \frac{1}{n} \text{ as } n \rightarrow +\infty \quad 1 - p.
\] (5)
Proof. Suppose that $x$ lies in tower $k \in \{0, \ldots, m\}$ at level $m$ ($x \in \tau_k^m$). Then, at level $(m+1)$, $x$ lies either in tower $k$ or tower $(k+1)$, with probability $p$, $1-p$ respectively. Therefore, $k_n(x)$ is the sum of $n$ independent and identically distributed Bernoulli random variables $(X_i)_{1 \leq i \leq n}$ with $P(X_i = 0) = p = 1 - P(X_i = 1)$.

By the law of large numbers, we obtain that for $\mu$-almost every $x \in X$, 

\[
\frac{k_n(x)}{n} \xrightarrow{n \to +\infty} E[X_i] = 1 - p.
\]

Let $r \geq 1$ be a fixed integer. We consider each tower $\tau_k^r$ as a stacking of $2^r$ blocks which are pieces of towers of level $n - r$.

Lemma 2.5. For $\mu \otimes \mu$-almost every $(x,y) \in X \times X$, we can find arbitrarily large $n$ such that

\[
k_n(x) = k_n(y),
\]

and $x$ and $y$ are both in the first block of level $(n-r)$ in $\tau_k^n$.

Proof. We have seen in the previous lemma that if $(x,y)$ follows the law $\mu \otimes \mu$, then $k_m(x)$ and $k_m(y)$ can be respectively represented as $\sum_{i=1}^r X_i$ and $\sum_{i=1}^m Y_i$, where $(X_i)_{1 \leq i \leq m}$ and $(Y_i)_{1 \leq i \leq m}$ are independent and identically distributed Bernoulli random variables with parameter $p$. We want to prove that we can find arbitrarily large $m$ such that $k_m(x) = k_m(y)$ and $X_{m+1}, X_{m+2}, \ldots, X_{m+r}$ and $Y_{m+1}, Y_{m+2}, \ldots, Y_{m+r}$ are equal to 1. One can easily verify that $k_m(x) = k_m(y)$ is a symmetric random walk and is thus recurrent. Hence, we can find arbitrarily large $m$ such that $k_m(x) = k_m(y)$. Let us call $m_1(x,y) < m_2(x,y) < \cdots$ such integers $m$ and consider the events $(A_j)_{j \geq 1}$ defined by

\[
A_j = \{X_{m_j+1} = \cdots = X_{m_j+r} = Y_{m_j+1} = \cdots = Y_{m_j+r} = 1\}.
\]

Using the strong Markov property, we can check that

- for any $j \geq 1$, $P(A_j) = (1 - p)^{2r} > 0$;
- $(A_{jr})_{j \geq 1}$ are independent (because $m_{r(j+1)} - m_r \geq r$ for all $j \geq 1$).

Therefore, we can find arbitrarily large $m_j$ such that $k_{m_j}(x) = k_{m_j}(y)$ and $A_j$ happens. □

2.3. Conclusion

Because of Lemma 2.3, to achieve the proof of Theorem 1.2 it is enough to show that the process $(\mathcal{P}, T)$ is LB. For this, we are going to verify that $(\mathcal{P}, T)$ satisfies the hypotheses of Lemma 2.2. Given $\varepsilon > 0$, choose an integer $r$ such that $(1 - p)^r < \varepsilon/2$. Let $(x,y) \in X \times X$ be such that

- $\frac{k_n(x)}{n} \xrightarrow{n \to +\infty} 1 - p$;
there exist arbitrarily large $n$ satisfying $k_n(x) = k_n(y)$, and $x$ and $y$ are both in the first block of level $(n - r)$ in $\tau_n^{kn}$.

(The preceding lemmas tell us that these properties are satisfied for $\mu \otimes \mu$-almost all $(x, y)$.) Let us consider such an $n$, and note $k$ for $k_n(x)$. Observe that if $n$ is large enough, the height of the first $(n - r)$-block of $\tau_n^{kn}$, in which both $x$ and $y$ lie, is very small compared to the height of $\tau_n^{k}$. Indeed, the height of this $(n - r)$-block is $(\frac{n-r}{k-r})$, and we have

$$\frac{(\frac{n-r}{k-r})}{(\frac{n}{k})} = \frac{k(k-1)\cdots(k-r+1)}{n(n-1)\cdots(n-r+1)} \sim (1-p)^r \quad \text{as } n \to +\infty.$$ 

Thus, if $n$ is chosen large enough, and if we set $l \overset{\text{def}}{=} (\frac{n}{k})$, both $\mathcal{P}|_{l_0}^n(x)$ and $\mathcal{P}|_{l_0}^n(y)$ begin with a suffix of $B_{n,k}$ whose length is greater than $(1 - \varepsilon/2)l$.

It is then easy to find a common subsequence of $\mathcal{P}|_{l_0}^n(x)$ and $\mathcal{P}|_{l_0}^n(y)$ whose length is greater than $(1 - \varepsilon)l$, which gives

$$\bar{f}(\mathcal{P}|_{l_0}^n(x), \mathcal{P}|_{l_0}^n(y)) < \varepsilon.$$ 

$\square$

3. Open questions

So far, very few ergodic properties of the Pascal adic transformation are known. Many important questions concerning its spectral properties remain open; in particular it is not known whether it is weakly mixing or not.

More closely related to the present work, we can point out that the class of zero-entropy and loosely Bernoulli transformations contains several interesting subclasses: rank one, finite rank, local rank one (where rank one $\Rightarrow$ finite rank $\Rightarrow$ local rank one $\Rightarrow$ loosely Bernoulli). To which of these subclasses do the Pascal adic transformation belong? Although the cutting and stacking construction suggests that it is not of local rank one, even proving that it is not rank one seems to be a difficult question.

References