The Pascal Adic Transformation is Loosely Bernoulli

Élise Janvresse, Thierry de la Rue
CNRS - UMR 6085

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.

Key-words : adic transformation, loosely Bernoulli systems.
Classification AMS 2000: 28D05

La transformation Pascal adique est lâchement 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.

Mots-clefs : transformation adique, systèmes lâchement Bernoulli.

1 Introduction

The notion of *adic transformation* has been introduced by Vershik (see e.g. [5], [4]), as a model in which the transformation acts on infinite paths in some graphs, called *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 *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 σ-algebra $\mathscr{A}$ and the Lebesgue measure μ.

Let $0 < p < 1$ be a fixed parameter. We start by dividing X into two subintervals $P_0 \stackrel{\text{def}}{=} [0,p[$ and $P_1 \stackrel{\text{def}}{=} [p,1[$. Let $\mathscr{P} \stackrel{\text{def}}{=} \{P_0, P_1\}$ 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 τ_0^1 and τ_1^1.

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 τ_0^2, τ_1^2, τ_2^2, with respective heights 1, 2, 1 (see figure 1.1).

After step n, we get $(n+1)$ towers $\tau_0^n, \ldots, \tau_n^n$, with respective heights $\binom{n}{0}, \ldots, \binom{n}{n}$, the width of τ_k^n being $p^{n-k}(1-p)^k$. Denote bu F_k^n the base of τ_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 τ_k^n by sending it linearly onto the left piece of the base F_{k+1}^n of τ_{k+1}^n (both have the same length $p^{n-k}(1-p)^{k+1}$).

Repeting recursively this construction, T is finally defined almost everywhere, and clearly preserves the measure μ.

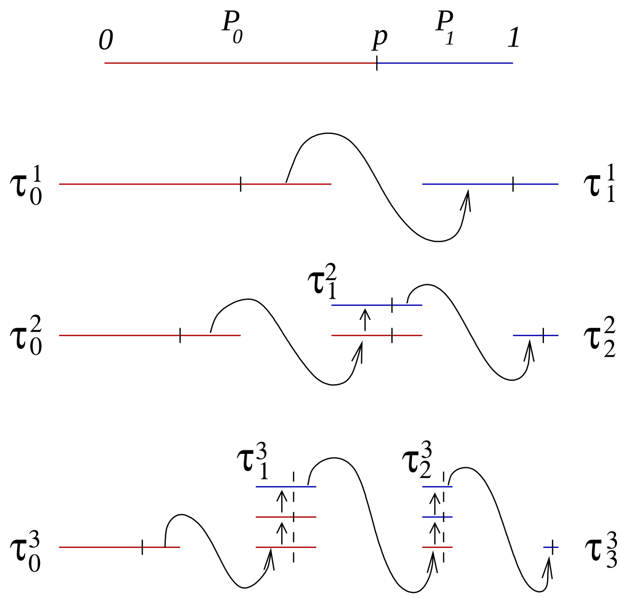

Figure 1: Cutting and stacking construction of the Pascal adic transformation

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 2.1, we consider a general dynamical system $(X, \mathscr{A}, \mu, T)$, where T is an invertible measure-preserving transformation of the Lebesgue probability space $(X, \mathscr{A}, \mu)$. 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 say 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 "$\mathscr{P}$-name" of a point x.

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

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

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

To define the property of being loosely Bernoulli, Feldman introduced the $\overline{f}$ 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 $\overline{f}$ distance between v and w is defined by

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

where s is the greatest integer in $\{0, \dots, l\}$ 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 $v_{i_r} = w_{j_r}$ $(r = 1, \dots, s)$.

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

$$\forall x, y \in A, \quad \overline{f}\left(\mathscr{P}|_0^l(x), \mathscr{P}|_0^l(y)\right) \ < \ \varepsilon.$$

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

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

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

$$\forall (x, y) \in B, \quad \overline{f}\left(\mathscr{P}|_0^l(x), \mathscr{P}|_0^l(y)\right) \ < \ \varepsilon.$$

then the process $(\mathscr{P}, T)$ is LB.

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

$$\forall (x,y) \in B, \quad \overline{f}\left(\mathscr{P}|_0^l(x), \mathscr{P}|_0^l(y)\right) \ < \ \varepsilon/2.$$

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

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

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

$$\overline{f}\left(\mathscr{P}|_0^l(y), \mathscr{P}|_0^l(y')\right) \ < \ \varepsilon.$$

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

Lemma 2.2 *Suppose that for all $\varepsilon > 0$ 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*

$$\overline{f}\left(\mathscr{P}|_0^{l(x,y)}(x), \mathscr{P}|_0^{l(x,y)}(y)\right) \ < \ \varepsilon.$$

then the process $(\mathscr{P}, T)$ is LB.

Proof — Let us fix $\varepsilon > 0$. For $\mu \otimes \mu$-almost every $(x,y) \in X \times X$, we define $l(x,y)$ as the smallest integer $k \geq 1$ such that $\overline{f}\left(\mathscr{P}|_0^k(x), \mathscr{P}|_0^k(y)\right) < \varepsilon/3$. 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

$$M_l \stackrel{\text{def}}{=} \frac{1}{l} \sum_{k=0}^{l-1} \mathbb{1}_{\{l(T^k x, T^k y) \geq n\}}.$$

Using Markov's inequality and the fact that T preserves the measure μ, 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 $B \stackrel{\text{def}}{=} \{M_l < \varepsilon/3\} \subset X \times X$ is such that $\mu \otimes \mu(B) > 1 - \varepsilon$. Let us fix $(x,y) \in B$. We want to show that $\overline{f}\left(\mathscr{P}|_0^l(x), \mathscr{P}|_0^l(y)\right) \ < \ \varepsilon$.
We say that $k \in \{0, \cdots, l-1\}$ 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 $j_0 = r_0 \stackrel{\text{def}}{=} \inf\{r \geq 0 \mid r \text{ is not bad}\}$, and for $i \geq 1$ such that $j_{i-1} \leq l - n$,

$$\begin{aligned} r_i &= \inf\left\{r \geq 0 \mid j_{i-1} + l(T^{j_{i-1}}x, T^{j_{i-1}}y) + r \text{ is not bad}\right\} \\ j_i &= j_{i-1} + l(T^{j_{i-1}}x, T^{j_{i-1}}y) + r_i \end{aligned}$$

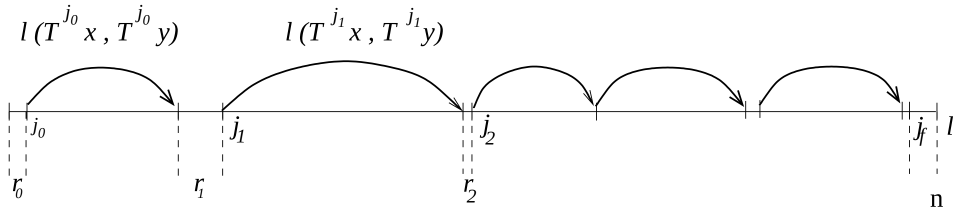

Figure 2: Covering of $\{0, \cdots, l\}$ with good intervals and bad points.

We denote by f the greatest index i such that j_i is defined: $l - j_f < n$.
Recall the definition of $\overline{f}$.

$$
\begin{aligned}
&(l+1)\ \overline{f}\left(\mathscr{P}|_0^l(x), \mathscr{P}|_0^l(y)\right) \\
&\leq \sum_{i=0}^{f-1} (j_{i+1} - j_i)\ \overline{f}\left(\mathscr{P}|_{j_i}^{j_{i+1}}(x), \mathscr{P}|_{j_i}^{j_{i+1}}(y)\right) + (l - j_f) \\
&\leq \sum_{i=0}^{f-1} l(T^{j_i}x, T^{j_i}y)\ \overline{f}\left(\mathscr{P}|_{j_i}^{j_i + l(T^{j_i}x, T^{j_i}y)}(x), \mathscr{P}|_{j_i}^{j_i + l(T^{j_i}x, T^{j_i}y)}(y)\right) \\
&\qquad\qquad + \sum_{i=0}^{f-1} r_i \ + \ (l - j_f) \\
&= \sum_{i=0}^{f-1} l(T^{j_i}x, T^{j_i}y)\ \overline{f}\left(\mathscr{P}|_0^{l(T^{j_i}x, T^{j_i}y)}(T^{j_i}x), \mathscr{P}|_0^{l(T^{j_i}x, T^{j_i}y)}(T^{j_i}y)\right) \\
&\qquad\qquad + \sum_{i=0}^{f-1} r_i \ + \ (l - j_f) \\
&\leq \frac{\varepsilon}{3} \sum_{i=0}^{f-1} l(T^{j_i}x, T^{j_i}y) + \frac{l\varepsilon}{3} + n \ < \ (l+1)\varepsilon.
\end{aligned}
$$

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 $\forall (x, y) \in B$, $\overline{f}\left(\mathscr{P}|_0^l(x), \mathscr{P}|_0^l(y)\right) < \varepsilon$. We conclude with Lemma 2.1. ❏

2.2 Some lemmas on the Pascal adic transformation

From now on, T is the Pascal adic transformation described in section 1.1, and $\mathscr{P}$ 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 *$\mathscr{P}$ is a generating partition for the system $(X, \mathscr{A}, \mu, T)$, i.e.*

$$\bigvee_{k=-\infty}^{+\infty} T^k \mathscr{P} \ = \ \mathscr{A}.$$

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 : $B_{n,0} \stackrel{\text{def}}{=} 0$, $B_{n,n} \stackrel{\text{def}}{=} 1$, and for $1 \leq k \leq n-1$,

$$B_{n,k} \stackrel{\text{def}}{=} B_{n-1,k-1}B_{n-1,k}.$$

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

$$\text{if } \mathscr{P}|_0^{\binom{n}{k}-1}(x) = B_{n,k}, \text{ then } x \in F_k^n. \tag{1}$$

Indeed, (1) is clearly satisfied for $n = 2$. Next, suppose that (1) is satisfied for $n-1$, and pick an x such that $\mathscr{P}|_0^{\binom{n}{k}-1}(x) = B_{n,k}$ $(1 \leq k \leq n-1)$. First, assume that $2 \leq k \leq n-2$. We have then

$$\mathscr{P}|_0^{\binom{n-1}{k-1}-1}(x) = B_{n-1,k-1}, \tag{2}$$

so that $x \in F_{k-1}^{n-1}$, and

$$\mathscr{P}|_0^{\binom{n-1}{k}-1}(T^{\binom{n-1}{k-1}}x) = B_{n-1,k}, \tag{3}$$

which implies $T^{\binom{n-1}{k-1}}x \in F_k^{n-1}$. Thus, after climbing the tower τ_{k-1}^{n-1}, the image of x goes to the next tower τ_k^{n-1}, which is possible only if $x \in F_n^k$ (otherwise, the image of x would go back to F_{k-1}^{n-1}). For the case $k = 1$, we first have to notice that

$$\forall m \geq 2,\ \forall 1 \leq j \leq m-1, \quad B_{m,j} \text{ begins with "0" and ends with "1".} \tag{4}$$

(We leave to the reader the verification of (4) by induction on m.) Now, if $\mathscr{P}|_0^{n-1}(x) = B_{n,1} = 0B_{n-1,1}$, we know that $Tx \in F_1^{n-1}$ because (1) is true for $n-1$, and then we can tell that $x \in F_0^{n-1}$: otherwise, the letter preceeding $B_{n-1,1}$ would be "1". This yields $x \in F_1^n$. The case $k = n-1$ is similar.

Now, for a fixed $n \geq 1$ we observe that the entire $\mathscr{P}$-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 $\mathscr{P}$-name of x gives for any n the value of $k_n(x)$ and tells us in which rung of $\tau_{k_n(x)}^n$ x lies. But the partition $\mathscr{Q}_n$ of X into rungs of towers τ_k^n, $0 \leq k \leq n$ is constituted of intervals whose maximal width is $\max(p, 1-p)^n$; moreover $\mathscr{Q}_{n+1}$ refines $\mathscr{Q}_n$. Therefore $\bigvee_{n \geq 1} \mathscr{Q}_n = \mathscr{A}$. ❑

Lemma 2.4 *For μ-almost every $x \in X$, we have*

$$\frac{k_n(x)}{n} \xrightarrow[n \to +\infty]{} 1-p. \tag{5}$$

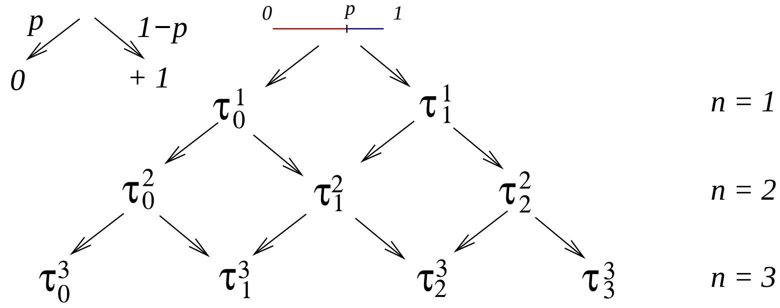

Figure 3: Representing $k_n(x)$ as the sum of n independent Bernoulli random variables.

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 $(X_m)_{\{1 \le m \le n\}}$ with $P(X_m = 0) = p = 1 - P(X_m = 1)$.
By the law of large numbers, we obtain that for μ-almost every $x \in X$, $\frac{k_n(x)}{n} \xrightarrow[n \to +\infty]{} E[X_m] = 1 - p$. ❑

Let $r \ge 1$ be a fixed interger. We consider each tower τ_k^n 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*

$$k_n(x) \;=\; k_n(y), \tag{6}$$

and x and y are both in the first block of level $(n-r)$ in $\tau^n_{k_n(x)}$.

Proof — We have seen in the previous lemma that $k_m(x) = \sum_{i=1}^m X_i$ and $k_m(y) = \sum_{i=1}^m Y_i$, where $(X_i)_{\{1 \le i \le m\}}$ and $(Y_i)_{\{1 \le i \le m\}}$ 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 $X_{m+1}, X_{m+2}, \cdots, X_{m+r}$ and $Y_{m+1}, Y_{m+2}, \cdots, Y_{m+r}$ are equal to 1. One can easily verify that $k_m(x) - k_m(y) = \sum_{i=1}^m (X_i - Y_i)$ 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 $(A_j)_{j \ge 1}$ 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 $j \ge 1$, $P(A_j) = (1-p)^{2r} > 0$;
- $(A_{jr})_{j \ge 1}$ are independent (because $m_{r(j+1)} - m_{rj} \ge r$ for all $j \ge 1$).

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 $(\mathscr{P}, T)$ is LB. For this, we are going to verify that $(\mathscr{P}, T)$ 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 $\tau^n_{k_n(x)}$.

(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 τ^n_k, in which both x and y lie, is very small compared to the height of τ^n_k. Indeed, the height of this $(n-r)$-block is $\binom{n-r}{k-r}$, and we have

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

Thus, if n is chosen large enough, and if we set $l \stackrel{\text{def}}{=} \binom{n}{k}$, both $\mathscr{P}|_0^l(x)$ and $\mathscr{P}|_0^l(y)$ begin with a suffix of $B_{n,k}$ whose length is greater than $(1-\varepsilon/2)l$.

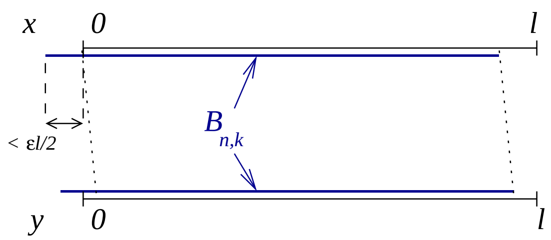

Figure 4: Coupling of large $\mathscr{P}$-names.

It is then easy to find a common subsequence of $\mathscr{P}|_0^l(x)$ and $\mathscr{P}|_0^l(y)$ whose length is greater than $(1-\varepsilon)l$, which gives

$$\bar{f}\left(\mathscr{P}|_0^l(x), \mathscr{P}|_0^l(y)\right) < \varepsilon.$$

❏

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* $\Longrightarrow$ *finite rank* $\Longrightarrow$ *local rank one* $\Longrightarrow$ *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

[1] J. Feldman, *New K-automorphisms and a problem of Kakutani*, Israel Journal of Mathematics **24** (1976), 16–38.

[2] X. Méla, *Dynamical properties of the Pascal adic and related systems*, Ph.D. thesis, Chapel Hill, 2002.

[3] D.S. Ornstein, D.J. Rudolph, and B. Weiss, *Equivalence of measure preserving transformations*, Memoirs of the American Mathematical Society 262, 1982.

[4] A. M. Vershik and A. N. Livshits, *Adic models of ergodic transformations, spectral theory, substitutions, and related topics*, Representation theory and dynamical systems, Adv. Soviet Math., vol. 9, Amer. Math. Soc., Providence, RI, 1992, pp. 185–204.

[5] A.M. Vershik, *A theorem on the Markov approximation in ergodic theory*, Journ. of Soviet Math. **28** (1985), 667–674.

Élise Janvresse, Thierry de la Rue
Université de Rouen, LMRS, UMR 6085 - CNRS
76 821 Mont Saint Aignan, France
Elise.Janvresse@univ-rouen.fr
Thierry.Delarue@univ-rouen.fr