Self-similarity of the corrections to the ergodic theorem for the Pascal-adic transformation

Élise Janvresse, Thierry de la Rue, Yvan Velenik



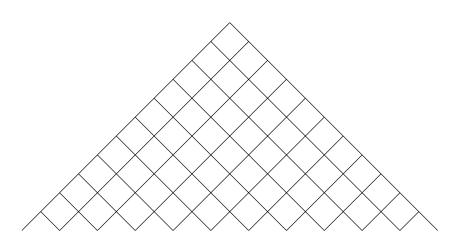
Laboratoire de Mathématiques Raphaël Salem

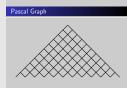




- The Pascal-adic transformation
- Self-similar structure of the basic blocks
- Ergodic interpretation
- 4 Generalizations and related problems

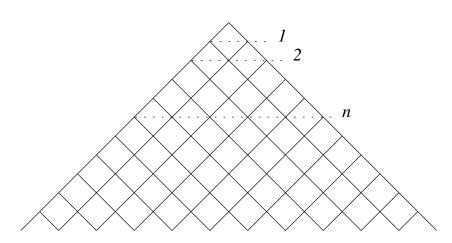
Pascal Graph



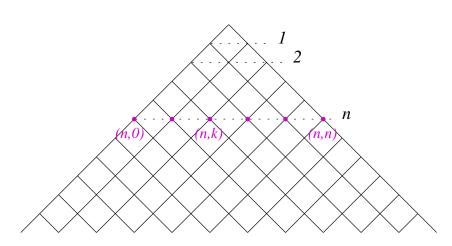


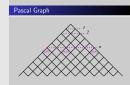
The Pascal graph: it is composed of infinitely many vertices and edges.

Pascal Graph



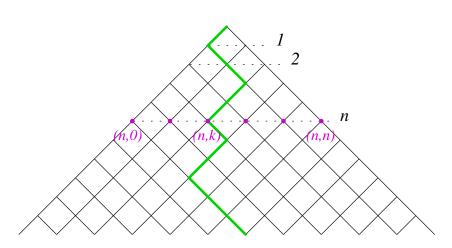
Pascal Graph

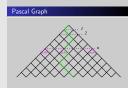




Its vertices are arranged in levels numbered $0,1,2,\ldots,n,\ldots$ Level n contains (n+1) vertices, denoted by $(n,0),(n,1),\ldots,(n,k),\ldots,(n,n)$. Each vertex (n,k) is connected to two vertices at level n+1: (n+1,k) and (n+1,k+1).

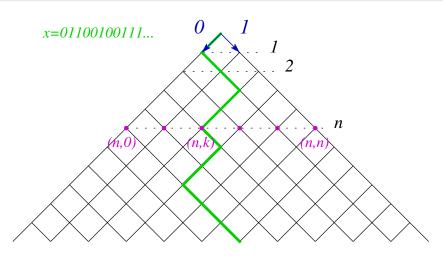
Pascal Graph

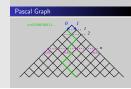




We are interested in trajectories on this graph, starting from the 0-level vertex (the root) and going successively through all its levels.

Pascal Graph

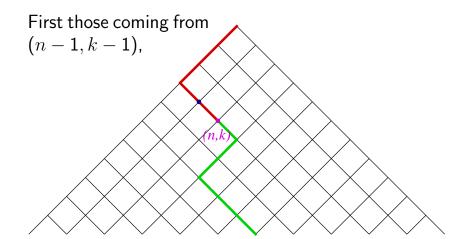




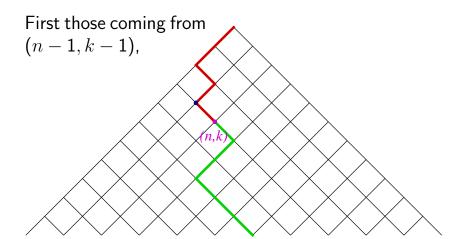
Each step to the left is labelled by a 0, and each step to the right by a 1. We then identify each trajectory with an infinite sequence of 0's and 1's.

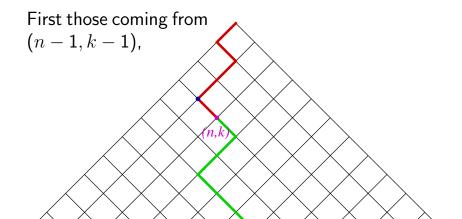
We list all trajectories going through (n, k)and fixed beyond this point.

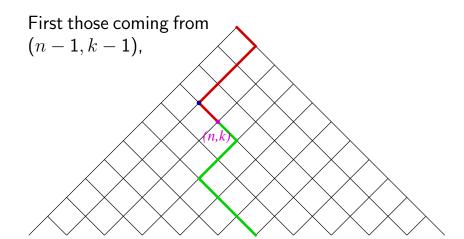
We introduce a partial order on the trajectories. We start by considering all trajectories coinciding from level n on.

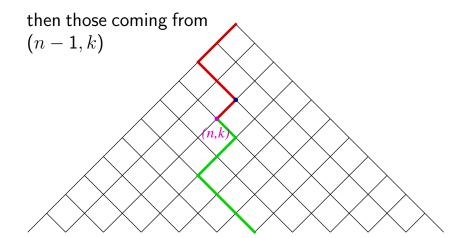


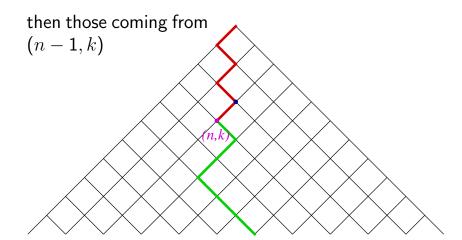
The order is defined in a recursive way: among all trajectories going through (n, k), those going through (n - 1, k - 1) precede those going through (n - 1, k).





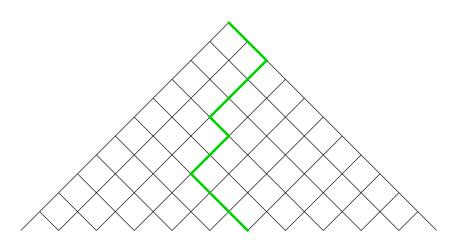




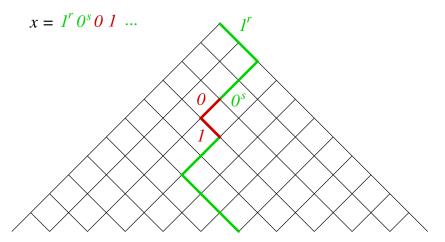


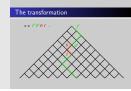
When this is over, we go one level down...

The transformation



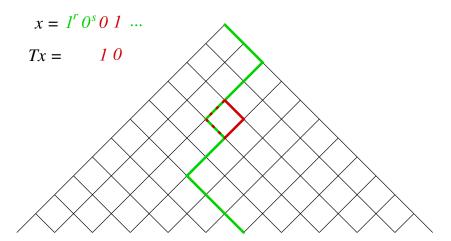
The transformation





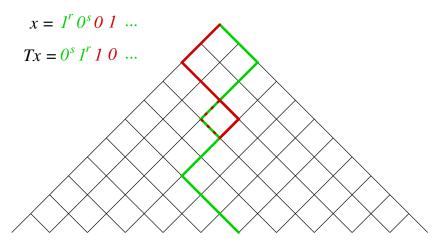
We obtain the successor of a trajectory by looking at the first 01 "left elbow" in the associated word...

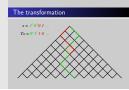
The transformation



... we reverse the "elbow"...

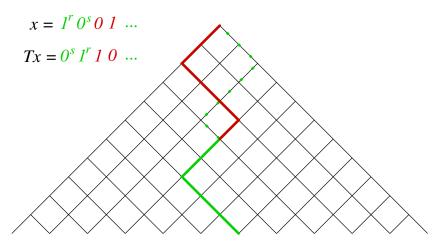
The transformation

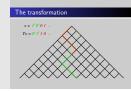




...and we order the preceding letters (i.e. first all 0's, then all 1's).

The transformation



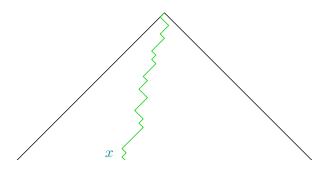


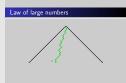
The Pascal-adic transformation, denoted by T in this talk, corresponds simply to the passage from a trajectory to its successor. Note that Tx is defined for all but a countable number of trajectories, namely those of the form $x = 1^r 0000 \dots$ Similarly, $T^{-1}x$ is defined for every x, except x of the form $0^s 1111 \dots$

Ergodic measures

The ergodic measures for T are the Bernoulli measures μ_p , $0 \le p \le 1$, where p is the probability of a step to the right.

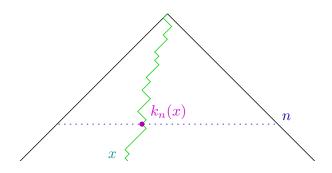
Law of large numbers



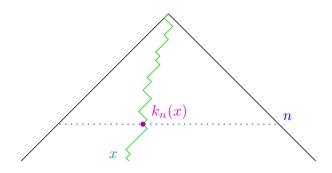


We denote by $k_n(x)$ the index such that at level n, the trajectory x passes through $(n, k_n(x))$. Observe that $k_n(x)$ is the sum of the values associated to the first n steps of x.

Law of large numbers

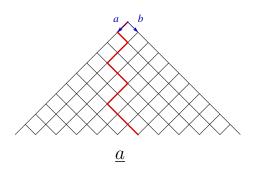


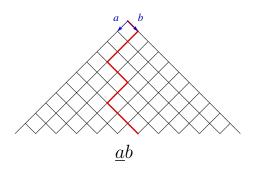
Law of large numbers

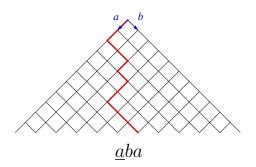


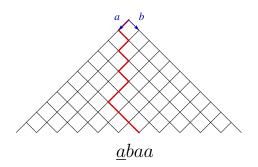
$$\frac{k_n(x)}{n} \underset{n \to \infty}{\longrightarrow} p$$

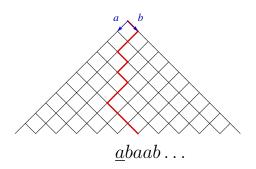
 μ_p -almost surely.

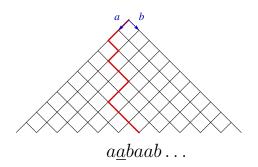


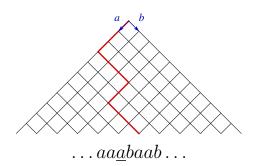


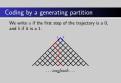






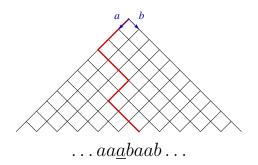






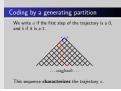
The orbit of a point x under the action of T is thus coded by a doubly-infinite sequence of a's and b's. The underlined letter indicates the origin.

We write a if the first step of the trajectory is a 0, and b if it is a 1.

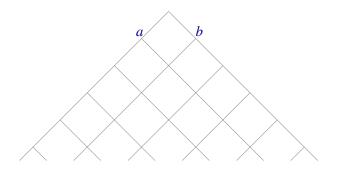


This sequence **characterizes** the trajectory x.

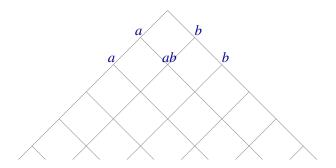




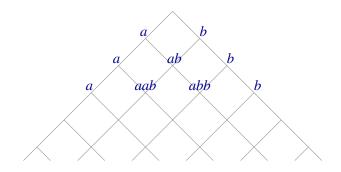
A partition \mathcal{P} is called *generating* if the doubly infinite word it associates (the \mathcal{P} -name) to every point x characterizes the latter.



 $B_{n,k}$: sequence of a's and b's corresponding to the ordered list of trajectories arriving at (n,k).

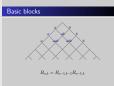


 $B_{n,k}$: sequence of a's and b's corresponding to the ordered list of trajectories arriving at (n,k).

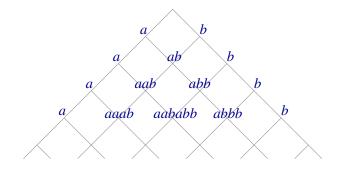


$$B_{n,k} = B_{n-1,k-1}B_{n-1,k}$$



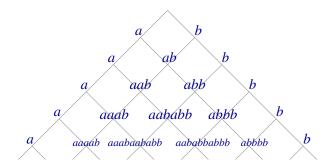


The trajectories arriving at (n, k) are ordered in such a way that those going through (n - 1, k - 1) appear before those going through (n - 1, k). Consequently, $B_{n,k}$ is the concatenation of $B_{n-1,k-1}$ and $B_{n-1,k}$.

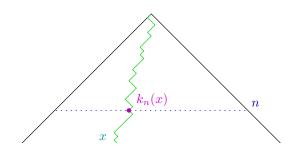


$$B_{n,k} = B_{n-1,k-1}B_{n-1,k}$$

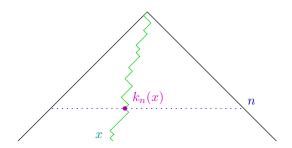


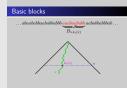


$$B_{n,k} = B_{n-1,k-1}B_{n-1,k}$$



$$B_{n,k_n(x)}$$





The sequence of words $B_{n,k_n(x)}$ observed along x is increasing and converges to the doubly infinite words coding the orbit of x.

- The Pascal-adic transformation
- Self-similar structure of the basic blocks
- Ergodic interpretation
- 4 Generalizations and related problems

These words quickly become complicated:

These words quickly become complicated:

We are interested in the asymptotic structure of the words appearing in the triangle. We first consider the words appearing along the vertical (2k, k). Their length grows very rapidly $(|B_{n,k}| = C_n^k)$; it is therefore useful to represent them differently.

These words quickly become complicated:

ab

These words quickly become complicated:

ab aababb

These words quickly become complicated:

 $ab \\ aababb \\ aaabaababbaababbb$

These words quickly become complicated:

. .

Graph associated to a word

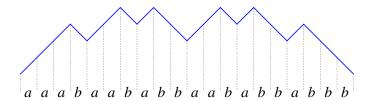
Graphical representation of words: $a \neq a$

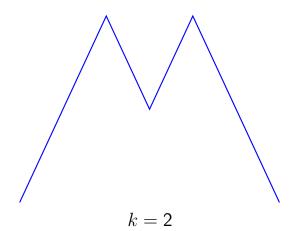


Graph associated to a word

Graphical representation of words: $a \neq b \setminus b$

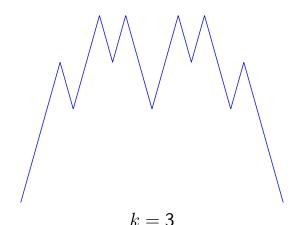
Example: $B_{6,3} = aaabaababbaababbabbb$ becomes

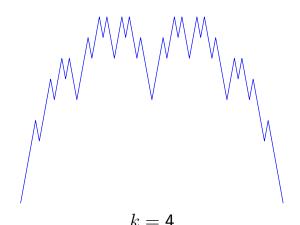


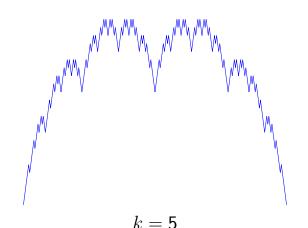




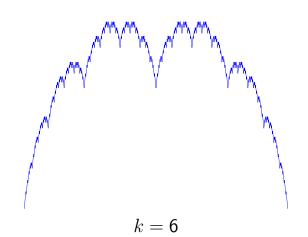
A remarkable phenomenon takes place: after a suitable normalization, the graph associated to the words $B_{2k,k}$ converges to a self-similar curve.



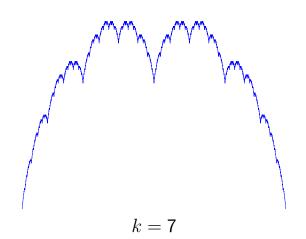




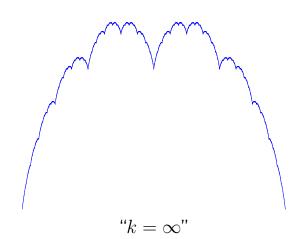
Graph associated to $B_{2k,k}$



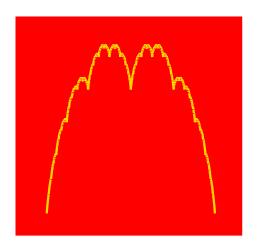
Graph associated to $B_{2k,k}$



Graph associated to $B_{2k,k}$



MacDonald's curve





Some people call this limiting curve the *MacDonald's curve*. We really don't know why!

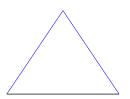
MacDonald's Blancmange curve



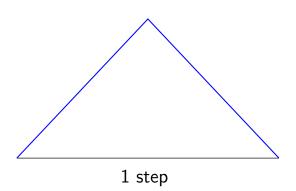
In fact, this curve is known in the litterature under the name *Blancmange curve*, because it looks like the dessert. . .

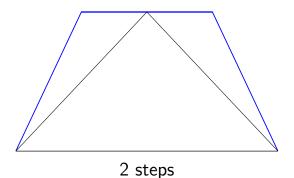
The fractal Blancmange curve (also called Takagi's curve) is the attractor of the family of the two affine contractions

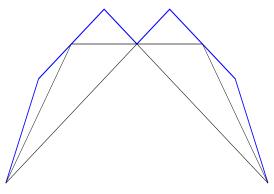
$$(x,y) \mapsto (\frac{1}{2}x, \frac{1}{2}y + x)$$
 $(x,y) \mapsto (\frac{1}{2}x + \frac{1}{2}, \frac{1}{2}y - x + 1)$



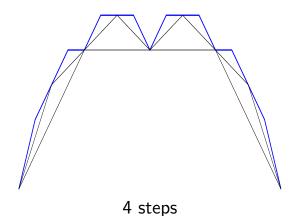
This limiting curve is also called Takagi's curve (who introduced it in 1903 as a simple example of a continuous nowhere-differentiable function). This curve can be constructed in several ways. The relevant construction for this talk is the following one: we introduce two affine transformations, that we apply iteratively starting with the segment [0, 1]. The attractor is the Blancmange curve.

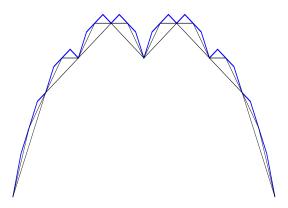






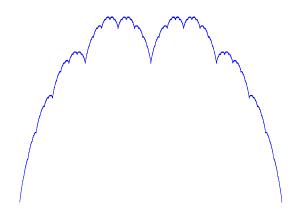
3 steps





5 steps





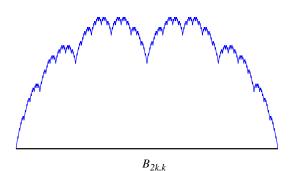
The attractor : $\bigcap_{1/2}$



Result

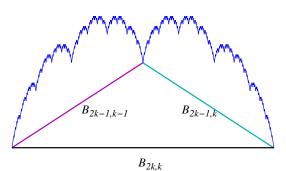
Theorem

After a suitable scaling, the curve associated to the block $B_{2k,k}$ converges in L^{∞} to $\bigcap_{1/2}$.





We are going to compute the coordinates of a dense subset of points of $B_{2k,k}$, when $k \to \infty$.

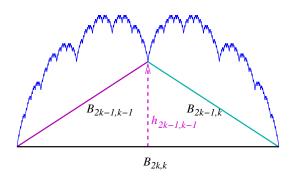






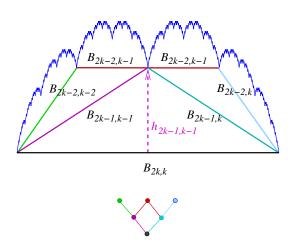


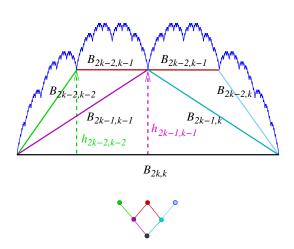
The word $B_{2k,k}$ is the concatenation of the two symmetric words $B_{2k-1,k-1}$ and $B_{2k-1,k}$. The quantities we are interested in are the length and the height of the word $B_{2k-1,k-1}$. This gives us the coordinates of a first point.

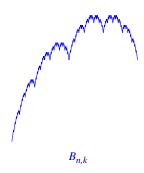


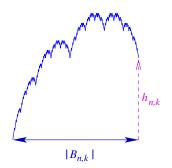


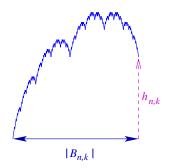










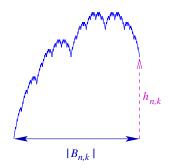


$$|B_{n,k}| = C_n^k$$





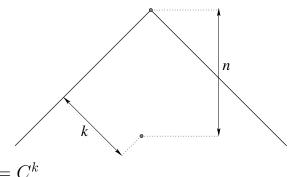
The length of a word $B_{n,k}$ is equal to the number of trajectories connecting (0,0) and (n,k).



$$|B_{n,k}| = C_n^k$$

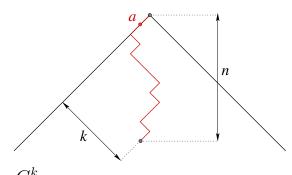
 $h_{n,k} = |B_{n,k}|_a - |B_{n,k}|_b$





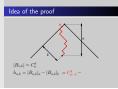
$$|B_{n,k}| = C_n^k$$

 $h_{n,k} = |B_{n,k}|_a - |B_{n,k}|_b$

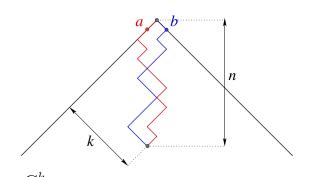


$$|B_{n,k}| = C_n^k$$

 $h_{n,k} = |B_{n,k}|_a - |B_{n,k}|_b = C_{n-1}^k - C_n^k$



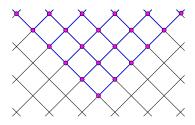
 $|B_{n,k}|_a$, the number of a's in $B_{n,k}$, is equal to the number of trajectories connecting (1,0) and (n,k).



$$|B_{n,k}| = C_n^k$$

 $h_{n,k} = |B_{n,k}|_a - |B_{n,k}|_b = C_{n-1}^k - C_{n-1}^{k-1}$

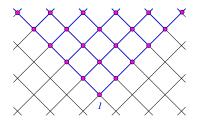




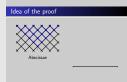
Abscissae



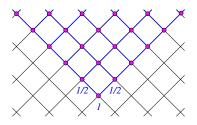
We present now an efficient method to obtain the coordinates of points located at the words boundaries. We start by constructing the "ascending" triangle composed of (2k, k) and all its "ancestors". Consider first the case of abscissae.



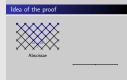
Abscissae



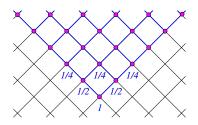
The length of the first word after horizontal normalization is arbitrary. We choose it to be 1.



Abscissae

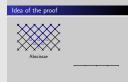


By symmetry, the lengths of its two parents are therefore equal to 1/2.

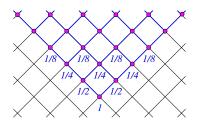


Abscissae



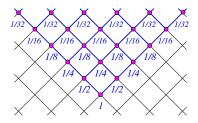


Going on like this, we easily obtain all the abscissae.



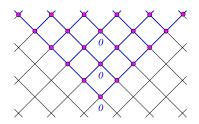
Abscissae





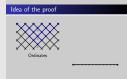
Abscissae



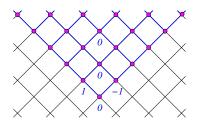


Ordinates

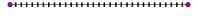


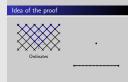


The heights of all the words located along the vertical through (0,0) is equal to 0.

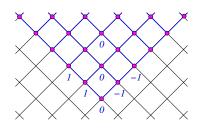


Ordinates





The height, after vertical normalization, of the parents of (2k, k) is arbitrary. We choose it to be 1.



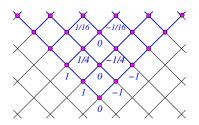
$$h_{n,k} = h_{n-1,k-1} + h_{n-1,k}$$





This allows us to fill in two more values. At this stage, we need an additional information. The following asymptotics easily follows from the formula for $h_{n,k}$:

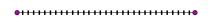
$$\lim_{k \to \infty} \frac{h_{2k-1,k-1}}{h_{2k+1,k+1}} = 1/4.$$

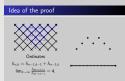


$$h_{n,k} = h_{n-1,k-1} + h_{n-1,k}$$

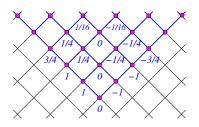
 $\lim_{k \to \infty} \frac{h_{2k+1,k+1}}{h_{2k-1,k-1}} = 4.$





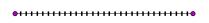


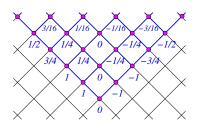
We can now completely fill in the array.



$$h_{n,k} = h_{n-1,k-1} + h_{n-1,k}$$
$$\lim_{k \to \infty} \frac{h_{2k+1,k+1}}{h_{2k-1,k-1}} = 4.$$

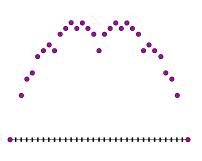


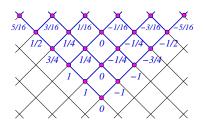




$$h_{n,k} = h_{n-1,k-1} + h_{n-1,k}$$

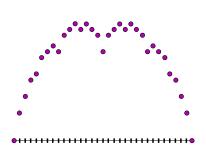
 $\lim_{k \to \infty} \frac{h_{2k+1,k+1}}{h_{2k-1,k-1}} = 4.$

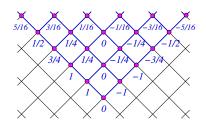




$$h_{n,k} = h_{n-1,k-1} + h_{n-1,k}$$

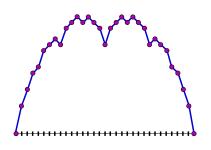
 $\lim_{k \to \infty} \frac{h_{2k+1,k+1}}{h_{2k-1,k-1}} = 4.$

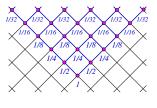




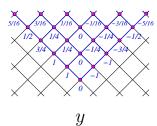
$$h_{n,k} = h_{n-1,k-1} + h_{n-1,k}$$

 $\lim_{k \to \infty} \frac{h_{2k+1,k+1}}{h_{2k-1,k-1}} = 4.$

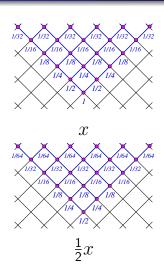


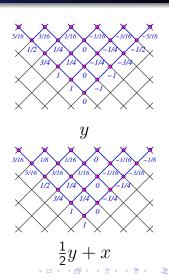


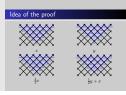
x



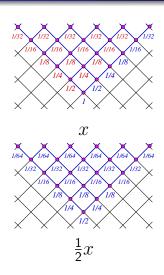
It remains to explain how the link with the Blancmange curve is made. Let us consider the first of the two affinities used in its construction.

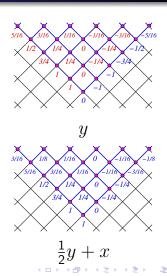


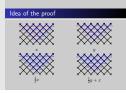




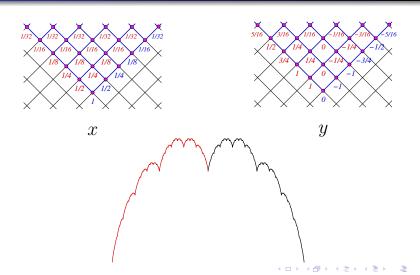
Applying this transformation to the abscissae and ordinates we have just computed, we obtain two new ascending triangles.

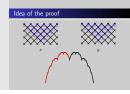




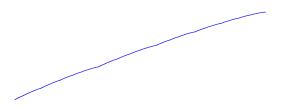


We then observe that these two triangles are precisely those obtained when removing the rightmost border of the original triangles.





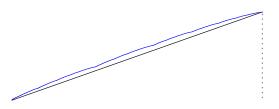
But these two subtriangles correspond to the coordinates of the points of the left half of the curve. The desired self-similarity is therefore seen to hold. The second affinity can be analyzed similarly.



The curve obtained for $B_{33.11}$



Observing, for example, the word associated to $B_{33,11}$, we see that the latter is essentially a straight line. This leads us to the study of what one obtains when renormalizing this graph in order to remove this leading linear growth.

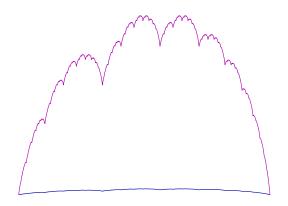


We subtract the straight line. . .



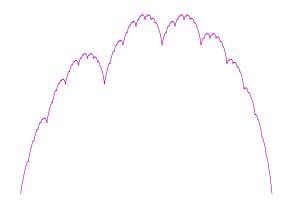
We subtract the straight line. . .





...and we change the vertical scale





The attractor $\cap_{1/3}$





Obviously, this is not $\bigcap_{1/2}$, but a related curve.

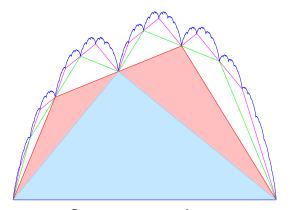
The family of limiting curves

We consider the family of curves \bigcap_p defined as follows: \bigcap_p is the attractor of the family of the two affine contractions

$$(x,y) \mapsto (px, py + x)$$

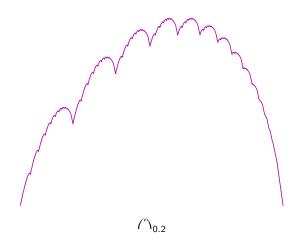
 $(x,y) \mapsto ((1-p)x + p, (1-p)y - x + 1)$

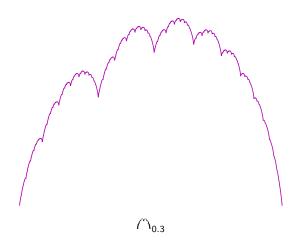
Limiting curve for p = 0.4

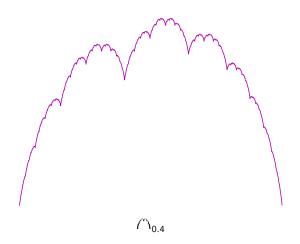


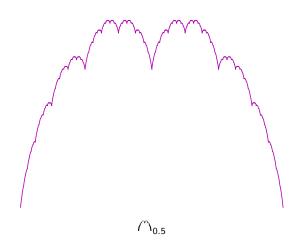
Construction of $\cap_{0.4}$

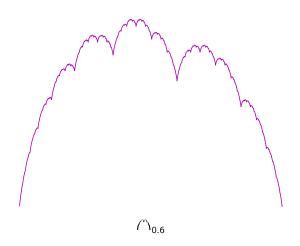
Some examples

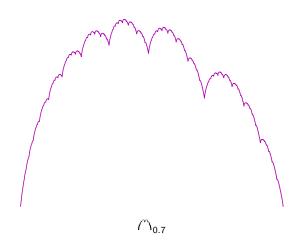


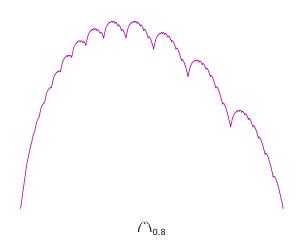












Result

Theorem

Let (k_n) be a sequence such that $\lim_n k_n/n = p \in (0,1)$.

After a suitable normalization, the curve associated to the block B_{n,k_n} converges in L^{∞} to \bigcap_p .

- The Pascal-adic transformation
- Self-similar structure of the basic blocks
- Ergodic interpretation
- 4) Generalizations and related problems

The case of i.i.d. random variables

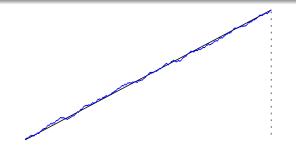


$$t \mapsto \frac{1}{\ell} \sum_{0 \le j < t\ell} X_j$$



We want to interpret the preceding result in term of corrections to the ergodic theorem. For this, we make a parallel with a better known situation: the case of bounded i.i.d. random variables. First consider a large integer ℓ , and the graph representing the partial sums up to time n.

The case of i.i.d. random variables



$$t \mapsto \frac{1}{\ell} \sum_{0 \le j < t\ell} X_j$$

$$t \mapsto \frac{t}{\ell} \sum_{0 \le j < \ell} X_j$$



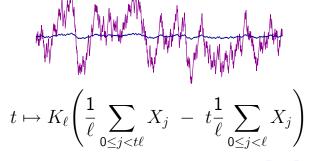
The case of i.i.d. random variables

$$t \mapsto \frac{1}{\ell} \sum_{0 \le j < t\ell} X_j - \frac{t}{\ell} \sum_{0 \le j < \ell} X_j$$

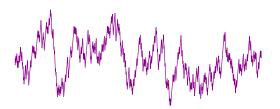
2005-03-31

We substract a linear function so that the graph ends on the line $y=\mathbf{0}$.

The case of i.i.d. random variables



The case of i.i.d. random variables



Brownian bridge



2005-03-31



After a suitable vertical renormalization, we (asymptotically) get a typical trajectory of a Brownian bridge, which describes the corrections to the law of large numbers.

Let
$$g(x) = \begin{cases} 1 & \text{if } x \text{ begins with } 0 \\ -1 & \text{if } x \text{ begins with } 1. \end{cases}$$

Let
$$g(x) = \begin{cases} 1 & \text{if } x \text{ begins with } 0 \\ -1 & \text{if } x \text{ begins with } 1. \end{cases}$$

Since g is integrable, the ergodic theorem yields, for 0 < t < 1

$$\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{0 \leq j < t\ell} g\left(T^j x\right) \; = \; t \lim_{\ell \to \infty} \frac{1}{\ell} \sum_{0 \leq j < \ell} g\left(T^j x\right).$$

Let
$$g(x) = \begin{cases} 1 & \text{if } x \text{ begins with } 0 \\ -1 & \text{if } x \text{ begins with } 1. \end{cases}$$

$$\frac{1}{\ell} \sum_{0 \le j < t\ell} g\left(T^{j}x\right) - t \frac{1}{\ell} \sum_{0 \le j < \ell} g\left(T^{j}x\right)$$

Let
$$g(x) = \begin{cases} 1 & \text{if } x \text{ begins with } 0 \\ -1 & \text{if } x \text{ begins with } 1. \end{cases}$$

$$K_{\ell} \left(\frac{1}{\ell} \sum_{0 \le j < t\ell} g\left(T^{j} x \right) - t \frac{1}{\ell} \sum_{0 \le j < \ell} g\left(T^{j} x \right) \right)$$

Let
$$g(x) = \begin{cases} 1 & \text{if } x \text{ begins with } 0 \\ -1 & \text{if } x \text{ begins with } 1. \end{cases}$$

$$\lim_{\ell \to \infty} K_{\ell} \left(\frac{1}{\ell} \sum_{0 \le j < t\ell} g\left(T^{j} x \right) - t \frac{1}{\ell} \sum_{0 \le j < \ell} g\left(T^{j} x \right) \right) = \bigcap_{p} \left(\prod_{i \le j < t\ell} g\left(T^{j} x \right) \right)$$

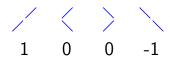
Our theorem is the analog of what happens in the i.i.d. case, when one considers the partial ergodic sums of a special function g depending on the first step of the trajectory during an interval of time corresponding to a basic block. What is remarkable here is that we get a deterministic limit (depending only on the ergodic component).

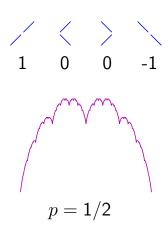
Cylindrical functions

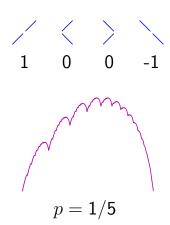
It is natural to extend this study to functions

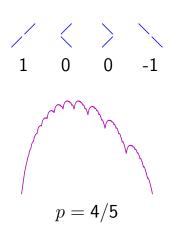
$$g(x_1,\ldots,x_{N_0})$$

depending only on the first N_0 steps of the trajectory.



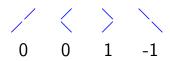


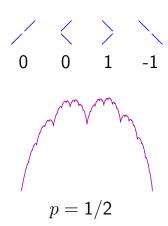


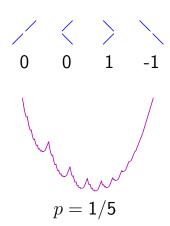


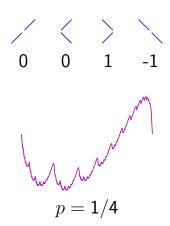


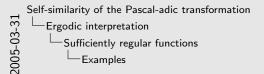
For this particular function g depending on the first 2 steps of the trajectory, the situation seems to be the same as what we observed for the simpler function.













But for this function g, we get an amazing phenomenon: For some values of the parameter p, *i.e.* for some ergodic components, the limiting curve is reversed! And for the special value p=1/4, we get a different kind of limiting curve.

General result

Let g be a cylindrical function depending only on the first N_0 steps, and not cohomologous to a constant.

General result

Let g be a cylindrical function depending only on the first N_0 steps, and not cohomologous to a constant.

There does not exist a function h such that

$$g = h \circ T - h + C.$$

General result

Let g be a cylindrical function depending only on the first N_0 steps, and not cohomologous to a constant.

Theorem

There exists a polynomial P^g of degree N_0+1 such that the behavior of the ergodic sums of the function g is characterized by the sign of $P^g(p)$: if $P^g(p) \neq 0$, the limiting curve is $\operatorname{sign}\left(P^g(p)\right) \cap_p$.

The polynomial P^g

The polynomial P^g is given by the following formula:

$$P^{g}(p) = -\text{cov}_{\mu_{p}}(g, k_{N_{0}}).$$

It has at most $N_0 - 1$ zeros in the interval (0, 1).

The critical case

Question: What happens when $P^g(p) = 0$?

Other classes of functions?

It is easy to construct functions g for which such a result does not hold.

We can either construct them by hand, or use a result of Dalibor Volný stating that one can always find a function satisfying the invariance principle.

Other classes of functions?

It is easy to construct functions g for which such a result does not hold.

Question: If g is such that

$$\lim_{N_{0}\to\infty}\operatorname{cov}_{\mu_{p}}\left(g,k_{N_{0}}\right)$$

exists and is non zero, does one observe the same phenomenon?

- The Pascal-adic transformation
- Self-similar structure of the basic blocks
- Ergodic interpretation
- Generalizations and related problems

In 1988, Conway introduced the following recursive sequence:

$$C(j) = C(C(j-1)) + C(j-C(j-1))$$

with initial conditions C(1) = C(2) = 1.

During a talk at the Bell Laboratories in 1988, John Conway described this recursive sequence, and challenged the audience to find the first n such that, for all $j \geq n$, we have |C(j)/j - 1/2| < 0.05. He thought that it was so difficult to solve that he promised \$10 000 for the solution. But two weeks later, Colin Mallows solved Conway's problem. The two men then agreed that Conway had suffered a "slip of the tongue" in offering \$10 000, when he had meant only \$1 000.

We introduce the infinite word D_{∞} obtained by concatenating all the words $B_{n,k}$:

$$D_{\infty} = B_{1,0}B_{1,1}B_{2,0}B_{2,1}B_{2,2}B_{3,0}\dots$$

Let D_j be the word given by the first j letters of D_{∞} . The following relation holds $(j \ge 3)$

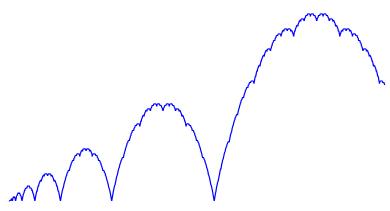
$$C(j) = 1 + |D_{j-2}|_a.$$

 $D_{\infty} = B_{1,0}B_{1,1}B_{2,0}B_{2,1}B_{2,2}B_{3,0}\dots$

Let D_j be the word given by the first j letters of D_{∞} . The following relation holds ($j \ge 3$)

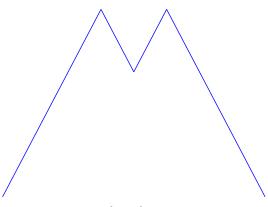
 $C(j) = 1 + |D_{j-2}|_a$

In fact the recursive sequence C(j) is closely related to the basic blocks $B_{n,k}$.

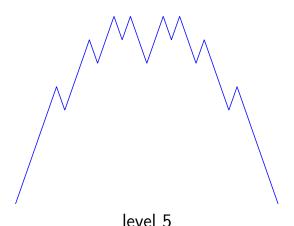


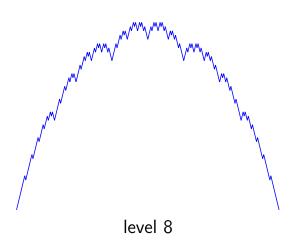
The beginning of the word D_{∞}

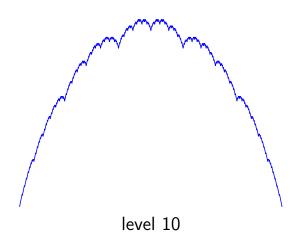


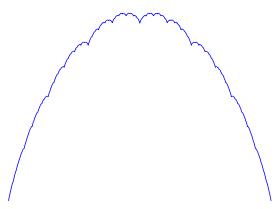


level 3

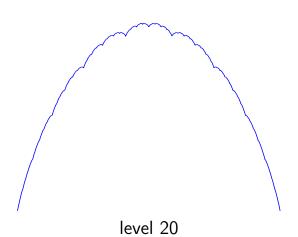


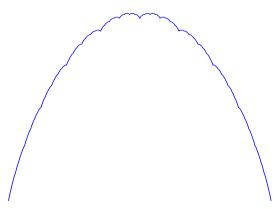




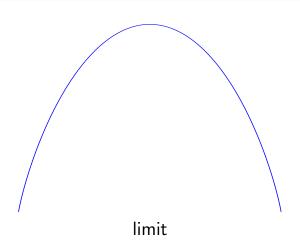


level 15





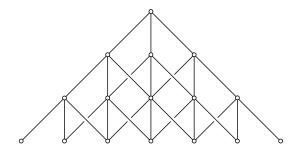
level 27

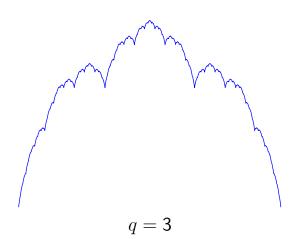


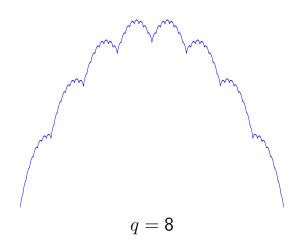
In his analysis of Conway's problem, Mallows observed that the graph associated to D_{∞} , consists in a series of bigger and bigger humps. Each of these humps corresponds to the concatenation of all the basic blocks in a given level n. He proved that, when suitably renormalized, the humps converge to the smooth curve $x=2+2\Phi(u), y=\varphi(u)$, where φ and Φ are respectively the density and the cumulative function of the standard Gaussian distribution. The results we present here can be interpreted as a refinement of this convergence.

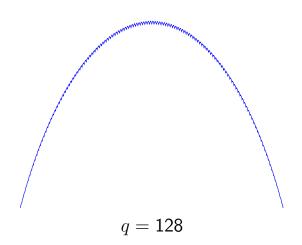
There exists a natural generalization of the Pascal-adic transformation, in which the graph has (q-1)N+1 vertices at level N, but where each vertex has q offsprings.

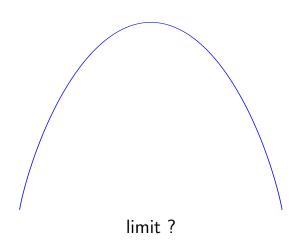
Example: the graph for q = 3







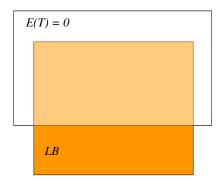


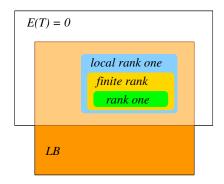


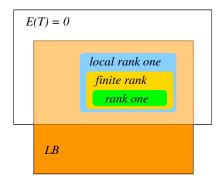


When $q \to \infty$, it seems that the limiting curve converges to the same smooth function as the humps described before.

$$E(T) = 0$$







Is the Pascal-adic transformation of rank one? Of finite rank? Of local rank one?



We know that the Pascal-adic transformation has zero entropy and is loosely Bernoulli. In the class of zero-entropy loosely-Bernoulli transformations, we have the following stronger properties: Rank one \Longrightarrow finite rank \Longrightarrow local rank one. The conjecture is that the Pascal-adic transformation is not of local rank one, but we are not even able to prove that it is not rank one!

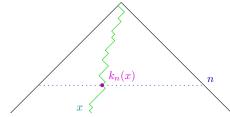
The question of weak mixing

Is the Pascal-adic transformation weakly mixing?

The question of weak mixing

Is the Pascal-adic transformation weakly mixing? If λ is an eigenvalue of T for the ergodic component μ_p , then for μ_p -every x

$$\lambda^{C_n^{k_n(x)}} \underset{n \to \infty}{\longrightarrow} 1.$$



The question of weak mixing

Is the Pascal-adic transformation weakly mixing? If λ is an eigenvalue of T for the ergodic component μ_p , then for μ_p -every x

$$\lambda^{C_n^{k_n(x)}} \xrightarrow[n \to \infty]{} 1.$$

Does this imply that $\lambda = 1$?



To be continued...