On barycentric subdivision

Persi Diaconis, Laurent Miclo

On barycentric subdivision

Persi Diaconis[†] and Laurent Miclo[‡]

[†]Departments of Statistics and Mathematics
Stanford University, USA
and Université de Nice-Sophia Antipolis, France

[‡]Laboratoire d'Analyse, Topologie, Probabilités, UMR 6632
Université de Provence and CNRS
France

Abstract

Consider the barycentric subdivision which cuts a given triangle along its medians to produce six new triangles. Uniformly choosing one of them and iterating this procedure gives rise to a Markov chain. We show that almost surely, the triangles forming this chain become flatter and flatter in the sense that their isoperimetric values goes to infinity with time. Nevertheless, if the triangles are renormalized through a similitude to have their largest edge equals to $[0,1]\subset\mathbb{C}$ (with 0 also adjacent to the shortest edge), their aspect does not converge and we identify the limit set of the opposite vertex with the segment [0,1/2]. In addition we prove that the largest angle converges to π in probability. Our approach is probabilistic and these results are deduced from the investigation of a limit iterated random function Markov chain living on the segment [0,1/2] and in particular of its invariant probability measure. Some estimates will need the help of the computer.

Keywords: barycentric subdivision, triangle-valued Markov chain, isoperimetric functional, flat triangles, iterated random functions, invariant probability measure, Wasserstein distance.

MSC2000: first: 60J05, secondary: 60D05, 60F15, 26A18, 28C10, 37A30.

1 Introduction

Let $\triangle$ be a given triangle on the plane (to avoid triviality the vertices will always be assumed not to be all the same). Using the medians of $\triangle$ (all intersecting at the barycenter), we can cut it into six small triangles, say $\triangle_1$, $\triangle_2$, $\triangle_3$, $\triangle_4$, $\triangle_5$, $\triangle_6$. Next, each $\triangle_i$, for $i \in [\![1,6]\!]$, can itself be subdivised in the same way into six triangles, $(\triangle_{i,j})_{j\in[\![1,6]\!]}$. Iterating this barycentric subdivision procedure, we get 6^n triangles $(\triangle_I)_{I\in[\![1,6]\!]^n}$ at stage $n \in \mathbb{N}$. It is known, numerically (see for instance [?]) and rigorously (cf. Diaconis and McMullen [5]) that as the barycentric subdivision goes on, the triangles have a tendency to become flat. The original motivation for this kind of results was to show that the barycentric subdivision is not a good procedure to construct nice triangularizations of surfaces. The goal of this paper is to propose a new probabilistic approach to this phenomenon.

First, we adopt a Markovian point of view: Let $\triangle(0) := \triangle$ and throw a fair dice to choose $\triangle(1)$ among the six triangles $\triangle_i$, $i \in [\![1,6]\!]$. By pursuing in the same way, we get a Markov chain $(\triangle(n))_{n\in\mathbb{N}}$: if the n^{th} first triangles have been constructed, the next one is obtained by choosing uniformly (and independently from what have been done before) one of the six triangles of the barycentric subdivision of the last obtained triangle. Of course, at any time $n \in \mathbb{N}^*$, the law of $\triangle(n)$ is the uniform distribution on the set of triangles $\{\triangle_I : I \in [\![1,6]\!]^n\}$. So to deduce generic properties under this distribution it is sufficient to study the chain $(\triangle(n))_{n\in\mathbb{N}}$.
In order to describe more analytically our results, let us renormalize the triangles. For any non-trivial triangle $\triangle$ on the plane, there is a similitude of the plane transforming $\triangle$ into a triangle whose vertices are $(0,0)$, $(0,1)$ and $(x,y) \in [0,1/2]\times[0,\sqrt{3}/2]$, such that the largest (respectively the smallest) edge of $\triangle$ is sent to $[(0,0),(0,1)]$ (resp. $[(0,0),(x,y)]$). The point (x,y) is uniquely determined and characterizes the aspect of $\triangle$ (as long as orientation is not concerned, otherwise we would have to consider positive similitudes and x would have to belong to $[0,1]$). So any time we will be interested in quantities which are invariant by similitude, we will identify triangles with their characterizing points. In particular, this identification will endow the set of triangles with the topology (not separating triangles with the same aspect) inherited from the usual topology of the plane. This convention will implicitely be enforced in the whole paper. The triangle $\triangle$ will be said to be flat if $y = 0$. So up to similitude the set of flat triangles can be identified with $[0,1/2]$. For $n \in \mathbb{N}$, let (X_n, Y_n) be the characterizing point of $\triangle(n)$. Next result gives a justification to the assertion that as the barycentric subdivision goes on, the triangles become flat.

Theorem 1 *Almost surely (a.s.) the stochastic sequence $(Y_n)_{n\in\mathbb{N}}$ converges to zero exponentially fast: there exists a constant $\chi > 0$ such that a.s.:*

$$\limsup_{n\to\infty} \frac{1}{n}\ln(Y_n) \quad \leq \quad -\chi$$

We will show that we can take $\chi = 0.035$ (but this is not the best constant, indeed we will present a numerical experiment suggesting that the above bound should hold with $\chi \approx 0.07$), nevertheless the previous result remains asymptotical. But contrary to Blackwell [2] (see also the remark at the end of section 6), we have not been able to deduce a more quantitative bound in probability on Y_n for any given $n \in \mathbb{N}$.

In particular, we recover the convergence in probability toward the set of flat triangles which was already proven by Diaconis and McMullen [5], through abstract dynamical system arguments.

There is a stronger notion of convergence to flatness asking for the triangles to have an angle which is almost equal to π. With the preceding notations, for $n \in \mathbb{N}$, let A_n be the angle between $[(0,0),(X_n,Y_n)]$ and $[(X_n,Y_n),(0,1)]$, this is the largest angle of $\triangle(n)$.

Theorem 2 *The sequence $(\triangle(n))_{n\in\mathbb{N}}$ is strongly becoming flat in probability:*

$$\forall\ \epsilon > 0, \qquad \lim_{n\to\infty} \mathbb{P}[A_n < \pi - \epsilon] \quad = \quad 0$$

Of course this result implies that $(Y_n)_{n\in\mathbb{N}}$ converges to zero in probability. Note that the converse is not true in general: one can consider isosceles triangles becoming flatter an flatter, then their maximum angle is converging to $\pi/2$. Indeed, Theorem 2 will be more difficult to obtain than Theorem 1 because $(X_n)_{n\in\mathbb{N}}$ is not converging:

Theorem 3 *Almost surely, the limit set of $(X_n)_{n\in\mathbb{N}}$ is $[0,1/2]$, so it follows from Theorem 1 that a.s. the limit set of a trajectory of the triangle Markov chain $(\triangle(n))_{n\in\mathbb{N}}$ is the whole set of flat triangles.*

A crucial tool behind these results is a limit flat Markov chain Z. Strictly speaking the stochastic chain $(X_n)_{n\in\mathbb{N}}$ is not Markovian, but eventually its evolution becomes almost Markovian. Indeed, we note that the above barycentric subdivision procedure can formally also be applied to flat triangles and their set is stable by this operation. This means that if $Y_0 = 0$, then for any $n \in \mathbb{N}$, $Y_n = 0$ a.s. and in this particular situation $(X_n)_{n\in\mathbb{N}}$ is Markovian. Let M be its transition kernel, from $[0,1/2]$ to itself. In what follows, $Z \coloneqq (Z_n)_{n\in\mathbb{N}}$ will always designate a Markov chain on $[0,1/2]$ whose transition kernel is M. An important part of this paper will be devoted to the investigation of the Markov chain Z since it is the key to the above asymptotical behaviors. We will see that Z is ergodic in the sense that it admits an attracting (and thus unique) invariant measure μ on $[0,1/2]$. We will also show that μ is continuous and that its support is $[0,1/2]$ (but we don't know if μ is absolutely continuous).

The plan of the paper is the following: next section contains some global preliminaries, in particular we will show, by studying the evolution of the isoperimetric value, that the triangle Markov chain always return as close as we want to the set of flat triangles. This is a first step in the direction of Theorem 1. In section 3, we begin our investigation of the limiting Markov chain Z, to obtain some informations valid in a neighborhood of the set of flat triangles. Then in section 4 we put together the previous global and local results to prove Theorem 1. Ergodicity and the attracting invariant measure μ of the Markov chain Z are studied in section 5, using results of Dubins and Freedman [6], Barnsley and Elton [1] and Diaconis and Freedman [4] on iterated random functions. This will lead to the proofs of Theorem 2 and Theorem 3 in section 6. In the last section we present some numerical experiments which helped us in some technical proofs. They will also illustrate the obtained results, in particular to evaluate the constant χ of Theorem 1 and to see the profil of μ.

2 A weak result on attraction to flatness

The purpose of this section is to give some preliminary informations and bounds on the triangle Markov chain obtained by barycentric subdivisions. By themselves, these results are not sufficient to conclude to the a.s. convergence toward the set of flat triangles, but at least they give an heuristic reason for this behavior: the barycentric subdivision divides the area by six, but in the mean, the perimeter is divided by a constant smaller than $\sqrt{6}$. So the isoperimetric value of the triangles has a tendency to diverge to infinity, which is a hint that they are becoming flat.

To measure the separation between a given triangle $\triangle$ and the set of flat triangles we use the inverse of the isoperimetric value $I(\triangle)$ of $\triangle$: it is the perimeter of $\triangle$ divided by the square root of the area of $\triangle$ (this quantity is well-defined in $\mathbb{R}_+^* \sqcup \{+\infty\}$, since remember that the vertices are assumed not to be all the same). Note that $I(\triangle) = +\infty$ if and only if $\triangle$ is flat. Furthermore the isoperimetric functional is invariant by similitude, so it depends only on the characteristic point (x,y) of $\triangle$ and we have

$$I(\triangle) \quad = \quad \sqrt{2}\frac{1+\sqrt{x^2+y^2}+\sqrt{(1-x)^2+y^2}}{\sqrt{y}}$$

in particular we get

$$\sqrt{y}/3 \ \le\ (I(\triangle))^{-1} \ \le\ \sqrt{y} \tag{1}$$

so that the convergence of y to zero is equivalent to the divergence of $I(\triangle)$ to $+\infty$.
With the notations of the introduction, we define for $n \in \mathbb{N}$, $I_n := I(\triangle(n))$. Our first goal is to show the

Proposition 4 *Almost surely, we have* $\limsup_{n\to\infty} I_n = +\infty$.

The proof will be based on elementary considerations on one step of the barycentric subdivision. Let consider $\triangle$ a triangle under the normalized form given in the introduction. For simplicity, we denote A, B and C the vertices $(0,0)$, (x,y) and $(1,0)$ of $\triangle$. Let also D, E, F and G be respectively the middle points of $[A,B]$, $[B,C]$ and $[A,C]$ and the barycenter of $\triangle$ (namely the intersection of the lines (A,E), (B,F) and (D,C)). We indice the small triangles obtained by the barycentric subdivision as

$$\begin{array}{lll} \triangle_1 \ := \ \{A,D,G\}, & \triangle_2 \ := \ \{D,B,G\}, & \triangle_3 \ := \ \{B,E,G\} \\ \triangle_4 \ := \ \{E,C,G\}, & \triangle_5 \ := \ \{C,F,G\}, & \triangle_6 \ := \ \{F,A,G\} \end{array} \tag{2}$$

The next result is classical, but we will give a proof for completeness:

Lemma 5 *All the triangles* $\triangle_i$*, for* $i \in [\![1,6]\!]$*, have the same area.*

Proof

First note that the triangles $\{A,B,F\}$ and $\{B,C,F\}$ have the same area. Indeed, let H be the orthogonal projection of B on the line (A,B). Denoting by $\mathcal{A}$ the area and $|\cdot|$ the length, we have

$$\begin{aligned} \mathcal{A}(\{A,B,F\}) &= \frac{1}{2}\,|[A,F]|\,|[B,H]| \\ &= \frac{1}{2}\,|[F,C]|\,|[B,H]| \\ &= \mathcal{A}(\{B,C,F\}) \end{aligned}$$

This argument shows that

$$\begin{aligned} \mathcal{A}(\triangle_1) &= \mathcal{A}(\triangle_2) \\ \mathcal{A}(\triangle_3) &= \mathcal{A}(\triangle_4) \\ \mathcal{A}(\triangle_5) &= \mathcal{A}(\triangle_6) \end{aligned}$$

But since we have on one hand,

$$\begin{aligned} \mathcal{A}(\{A,B,F\}) &= \mathcal{A}(\triangle_1) + \mathcal{A}(\triangle_2) + \mathcal{A}(\triangle_6) \\ &= 2\mathcal{A}(\triangle_2) + \mathcal{A}(\triangle_6) \end{aligned}$$

and on the other hand

$$\begin{aligned} \mathcal{A}(\{B,C,F\}) &= \mathcal{A}(\triangle_3) + \mathcal{A}(\triangle_4) + \mathcal{A}(\triangle_5) \\ &= 2\mathcal{A}(\triangle_3) + \mathcal{A}(\triangle_6) \end{aligned}$$

it follows that $\mathcal{A}(\triangle_2) = \mathcal{A}(\triangle_3)$. Similar reasoning leads to $\mathcal{A}(\triangle_4) = \mathcal{A}(\triangle_5)$ and $\mathcal{A}(\triangle_6) = \mathcal{A}(\triangle_1)$ and this concludes that all the triangles $\triangle_i$, for $i \in [\![1,6]\!]$, have the same area.

■

Next let define

$$\begin{array}{lll} L_1 \ := \ |[A,B]|\,, & L_2 \ := \ |[B,C]|\,, & L_3 \ := \ |[C,A]| \\ l_1 \ := \ |[D,C]|\,, & l_2 \ := \ |[E,A]|\,, & l_3 \ := \ |[F,B]| \end{array}$$

We will need another technical point.

Lemma 6 *The quotient* $(l_1+l_2+l_3)/(L_1+L_2+L_3)$ *is minimal when* $\{A,B,C\}$ *is equilateral, so we always have*

$$l_1+l_2+l_3 \quad \geq \quad \frac{\sqrt{3}}{2}(L_1+L_2+L_3)$$

Proof

Define the perimeter of $\{A,B,C\}$ in terms of the coordinates (x,y) of B:

$$\begin{aligned} p &\coloneqq L_1+L_2+L_3 \\ &= \frac{3}{2}+2\left(x-\frac{1}{2}\right)^2+2y^2 \end{aligned}$$

We first work with p equal to a fixed value. Note that the set of the corresponding points B is included into a circle centered in $(1/2,0)$ and by considering its intersection with the line parallel to the ordinate axis and passing through $(1/2,0)$, it appears that p can vary between the values 2 (isosceles flat triangle) and 3 (equilateral triangle), due to our renormalization asking for $L_1 \leq L_2 \leq L_3 = 1$.

So let $p\in[2,3]$ be fixed. We can express the length of the medians in terms of x and p:

$$\left\{\begin{aligned} l_1 &= \sqrt{\tfrac{p}{2}-\tfrac{3}{4}} \\ l_2 &= \sqrt{\tfrac{3}{4}x+\tfrac{p}{8}} \\ l_3 &= \sqrt{-\tfrac{3}{4}x+\tfrac{p}{8}+\tfrac{3}{4}} \end{aligned}\right. \tag{3}$$

and thus to find where the quotient $(l_1+l_2+l_3)/(L_1+L_2+L_3)$ is minimal, we have to minimize in x (belonging to a subinterval of $[0,1/2]$ depending on p) the quantity

$$\varphi(x) \quad \coloneqq \quad \sqrt{\frac{3}{4}x+\frac{p}{8}}+\sqrt{-\frac{3}{4}x+\frac{p}{8}+\frac{3}{4}}$$

But this expression is clearly a concave function of x and where it is defined we can write

$$\begin{aligned} \varphi(x) &\geq 2\sqrt{\frac{1}{2}\left(\frac{3}{4}x+\frac{p}{8}-\frac{3}{4}x+\frac{p}{8}+\frac{3}{4}\right)} \\ &= 2\sqrt{\frac{p}{8}+\frac{3}{8}} \\ &= \varphi(1/2) \end{aligned}$$

Thus now we are led to minimize over $p\in[2,3]$ the quantity

$$\begin{aligned} \frac{1}{p}\left(\sqrt{\frac{p}{2}-\frac{3}{4}}+2\sqrt{\frac{p}{8}+\frac{3}{8}}\right) &= \frac{1}{p}\left(\sqrt{\frac{p}{2}-\frac{3}{4}}+\sqrt{\frac{p}{2}+\frac{3}{2}}\right) \\ &= \frac{9}{4\psi(p)} \end{aligned}$$

with

$$\psi(p) \quad \coloneqq \quad p\left(\sqrt{\frac{p}{2}+\frac{3}{2}}-\sqrt{\frac{p}{2}-\frac{3}{4}}\right)$$

We compute its derivative

$$\psi'(p) \quad = \quad \left(\frac{3}{2}p+3\right)\sqrt{p-\frac{3}{2}}-\left(\frac{3}{2}p-\frac{3}{2}\right)\sqrt{p+3}$$

For $p \in [2,3]$, we have $\psi'(p) \geq 0$ if and only if

$$\left(\frac{p+2}{p-1}\right)^2 \quad \geq \quad \frac{p+3}{p-\frac{3}{2}} \tag{4}$$

and this leads to a second order equation (the third order terms cancel out) where one can directly check that (4) is indeed true for $p \in [2,3]$. So we have obtained that

$$\max_{p\in[2,3]} \psi(p) \quad = \quad \psi(3)$$

and it follows that the quotient $(l_1+l_2+l_3)/(L_1+L_2+L_3)$ is minimal when $\{A, B, C\}$ is equilateral. In this situation we compute directly that $l_1 = l_2 = l_3 = \sqrt{3}/2$ while $L_1 = L_2 = L_3 = 1$, so the announced bound follows.

■

We can transform these ingredients into a probabilistic statement:

Lemma 7 *Let $\alpha := \frac{\sqrt{3}+1-\sqrt{6}}{\sqrt{6}} > 0$, then for any $n \in \mathbb{N}$ we have the following bound on conditional expectation*

$$\mathbb{E}[I_{n+1}|\mathcal{T}_n] \quad \geq \quad (1+\alpha)I_n$$

where $\mathcal{T}_n$ is the σ-algebra generated by $\triangle(n)$, $\triangle(n-1)$, ..., $\triangle(0)$.

Proof

By the Markov property, the above bound is equivalent to the fact that for any $n \in \mathbb{N}$,

$$\mathbb{E}[I_{n+1}|\triangle(n)] \quad \geq \quad (1+\alpha)I_n$$

Since the Markov chain $(\triangle(n))_{n\in\mathbb{N}}$ is time-homogeneous, it is sufficient to deal with the case $n = 0$. We come back to the notations introduced above. Because the small triangles have the same area, we get that

$$\begin{aligned}
\mathbb{E}[I(\triangle(1))|\triangle(0) = \triangle] \quad &= \quad \frac{1}{6}\sum_{i\in[\![1,6]\!]} I(\triangle_i) \\
&= \quad \frac{1}{6}\frac{L_1+L_2+L_3+2(l_1+l_2+l_3)}{\sqrt{\mathcal{A}(\triangle)/6}} \\
&\geq \quad \frac{1}{\sqrt{6}}\frac{(1+\sqrt{3})(L_1+L_2+L_3)}{\sqrt{\mathcal{A}(\triangle)}} \\
&= \quad (1+\alpha)I(\triangle)
\end{aligned}$$

which is what we looked for.

■

In general the previous submartingale information is not enough to deduce a.s. convergence. Taking expectations, we get that for any $n \in \mathbb{N}$, $\mathbb{E}[I_{n+1}] \geq (1+\alpha)\mathbb{E}[I_n]$, thus $\mathbb{E}[I_n] \geq (1+\alpha)^n I(\triangle)$, so we can just deduce $\mathbb{L}^1$-divergence of I_n for large $n \in \mathbb{N}$, but this is not a very useful result.

To prove Proposition 4, recall that the numbers I_n, $n \in \mathbb{N}$, are uniformly bounded below by a positive constant. Indeed the isoperimetric functional can be defined for relatively general bounded subsets of the plane and its minimal value is obtained for discs (cf. for instance Osserman [7]). Thus we have

$$\forall\, n \in \mathbb{N}, \qquad I_n \quad \geq \quad 2\sqrt{\pi}$$

But from Lemma 7, we see that

$$\forall\, n \in \mathbb{N}, \qquad \mathbb{P}[I_{n+1} \geq (1+\alpha) I_n | \mathcal{T}_n] \;\; \geq \;\; \frac{1}{6}$$

and by consequence

$$\forall\, n, m \in \mathbb{N}, \qquad \mathbb{P}[I_{n+m} \geq (1+\alpha)^m 2\sqrt{\pi} | \mathcal{T}_n] \;\; \geq \;\; \frac{1}{6^m} \tag{5}$$

Let $M > 1$ be an arbitrary large number and consider $m \in \mathbb{N}^*$ such that $(1+\alpha)^m 2\sqrt{\pi} \geq M$. The sequence $(\mathbb{1}_{I_{m(n+1)} \geq M})_{n \in \mathbb{N}}$ stochastically dominates a sequence of independent Bernoulli variables of parameter $1/6^m$. It follows that a.s. we have

$$\limsup_{n\to\infty} I_n \;\; \geq \;\; M$$

and since M can be chosen arbitrarily large, Proposition 4 is proven.

To finish this section, we will prove another simple preliminary result.

Lemma 8 *There exists two constants* $0 < a < b < +\infty$ *such that*

$$\forall\, n \in \mathbb{N}, \qquad aI_n \;\leq\; I_{n+1} \;\leq\; bI_n$$

Proof

Again it is sufficient to consider the first barycentric subdivision and to prove that we can find two constants $0 < a < b < +\infty$ such that

$$\forall\, i \in [\![1, 6]\!], \qquad aI(\triangle) \;\leq\; I(\triangle_i) \;\leq\; bI(\triangle)$$

where we came back to the previous notations for the normalized representation of $\triangle$ and its barycentric subdivision. Such inequalities are obvious for flat triangles, so we can assume that $\triangle$ is not flat. Since the areas are easy to compare, we just need to consider the perimeters and we will denote by $\mathcal{P}(\triangle)$ the perimeter of $\triangle$.
Let us begin with the lower bounds. The existence of a constant $a' > 0$ such that

$$\forall\, i \in [\![1, 6]\!], \qquad a'\mathcal{P}(\triangle) \;\;\leq\;\; \mathcal{P}(\triangle_i)$$

comes from the following observations.
- By the normalized representation, $\max(L_1, L_2) \leq L_3 = 1$, so we have $\mathcal{P}(\triangle) \leq 3$.
- Due to the fact that $L_3 = 1$, we have $\mathcal{P}(\triangle_6) \geq 2\,|[A,F]| = 1$ and $\mathcal{P}(\triangle_5) \geq 2\,|[F,C]| = 1$.
- In quite the same way, it appears that $\mathcal{P}(\triangle_3) \geq 1/2$ and $\mathcal{P}(\triangle_4) \geq 1/2$, because we have $|[B,C]| \geq 1/2$ (since $x \leq 1/2$).
- For the same reason, we have $|[G,C]| \geq 1/2$, we deduce that $|[D,G]| = |[G,C]|/2 \geq 1/4$ and it follows that $\min(\mathcal{P}(\triangle_1), \mathcal{P}(\triangle_2)) \geq 2\,|[D,G]| \geq 1/2$.

For the upper bound, it is clear that

$$\forall\, i \in [\![1, 6]\!], \qquad \mathcal{P}(\triangle_i) \;\;\leq\;\; b'\mathcal{P}(\triangle)$$

should be satisfied with $b' = 1$. But since we are not interested here in precise constants, let us just give a coarse argument for the existence of a worse constant b'. It follows, on one hand, from $\mathcal{P}(\triangle) \geq 2$ and on the other hand, from the fact that all the medians l_1, l_2 and l_3 are smaller than $\sqrt{1 + 3/4}$, due to the fact that the triangle $\{A, B, C\}$ is included into $[0,1] \times [0, \sqrt{3}/2]$.

■

3 Near the limit flat Markov chain

Our goal here is two-fold. First we show that the kernel of the triangle Markov chain nicely converges to the kernel of the flat triangle Markov chain as the triangle becomes more and more flat. Second we study the evolution of a perimeter related functional for the flat triangle Markov chain, to get a bound on the evolution of the isoperimetric functional for the triangle Markov chain, valid at least in a neighborhood of the set of flat triangles.

Let Q be the transition kernel of the Markov chain $(X_n, Y_n)_{n\in\mathbb{N}}$ considered in the introduction. For any $(x, y) \in \mathcal{D}$, the set of characterizing points of triangles, we can write

$$Q((x,y),\,\cdot\,) \;=\; \frac{1}{6}\sum_{i\in[\![1,6]\!]} \delta_{(x_i,y_i)}$$

where δ stands for the Dirac mass and where for any $i \in [\![1, 6]\!]$, (x_i, y_i) is the characterizing point of the triangle $\triangle_i$ described in (2). Of course, the x_i and y_i, for $i \in [\![1, 6]\!]$, have to be seen as functions of (x, y). For $i \in [\![1, 6]\!]$, let us define

$$\forall\ x \in [0, 1/2], \qquad z_i(x) \;\coloneqq\; x_i(x, 0) \tag{6}$$

The transition kernel M on $[0, 1/2]$ of the flat triangle Markov chain alluded to in the introduction can expressed as

$$\forall\ x \in [0, 1/2], \qquad M(x,\,\cdot\,) \;=\; \frac{1}{6}\sum_{i\in[\![1,6]\!]} \delta_{z_i(x)} \tag{7}$$

The next result enables to measure the discrepancy between Q and M as the triangles become more and more flat.

Lemma 9 *There exists a constant $K > 0$ such that*

$$\forall\ i \in [\![1, 6]\!],\ \forall\ (x, y) \in \mathcal{D}, \qquad \begin{cases} |x_i(x, y) - z_i(x)| \;\le\; Ky \\ |y_i(x, y)| \;\le\; K\sqrt{y} \end{cases}$$

Before coming to the proof of these bounds, we make a change of coordinate for renormalized triangles $\triangle$: as in the proof of Lemma 6, it is more convenient to work with the modified coordinates $(x, p) \coloneqq (x, \mathcal{P}(\triangle))$. A useful relation between y and p is that

$$\forall\ (x, y) \in \mathcal{D}, \qquad y^2/3 \;\le\; p - 2 \;\le\; 2y \tag{8}$$

Indeed, for the upper bound we have

$$\begin{aligned} p &= 1 + \sqrt{x^2 + y^2} + \sqrt{(1-x)^2 + y^2} \\ &\le 1 + x + y + (1-x) + y \\ &= 2 + 2y \end{aligned}$$

For the lower bound, taking into account that for $(x, y) \in \mathcal{D}$, we have $y^2/(1-x)^2 \le (3/4)/(1/4) = 3$, we get by concavity of the mapping $[0, 3] \ni z \mapsto \sqrt{1+z}$,

$$\begin{aligned} p &\ge 1 + x + (1-x)\sqrt{1 + \frac{y^2}{(1-x)^2}} \\ &\ge 1 + x + (1-x)\left(1 + \frac{\sqrt{1+3} - 1}{3}\right)\frac{y^2}{(1-x)^2} \\ &= 2 + \frac{y^2}{3(1-x)} \\ &\ge 2 + \frac{y^2}{3} \end{aligned}$$

The lower bound is of the right order when x is not too close to zero: for any $0 < \epsilon < 1/2$, there exists $K_\epsilon > 0$ such that for all $(x, y) \in \mathcal{D} \cap [\epsilon, 1/2] \times \mathbb{R}_+$, $p - 2 \leq K_\epsilon y^2$.

Next let $\triangle$ be any (non-trivial) triangle and denote by $L_1 \leq L_2 \leq L_3$ the lengths of its edges. Let also (x, p) be the modified coordinates of its normalized form.

Lemma 10 *Let $\epsilon > 0$ be given. On the domain $\{(L_1, L_2, L_3) : 0 \leq L_1 \leq L_2 \leq L_3 \in [\epsilon, \epsilon^{-1}]\}$, the mapping $(L_1, L_2, L_3) \mapsto (x, p)$ is Lipschitz. Furthermore on normalized triangles (namely satisfying $L_3 = 1$), the mapping $(x, p) \mapsto (L_1, L_2)$ is also Lipschitz.*

Proof

Indeed, this is clear on the explicit formulas of these mappings. First, we have

$$p = \frac{L_1 + L_2 + L_3}{L_3}$$

and next x satisfies the equations

$$\begin{aligned} L_1^2 &= h^2 + (xL_3)^2 \\ L_2^2 &= h^2 + ((1-x)L_3)^2 \end{aligned}$$

where h is the height orthogonal to the edge of length L_3. It follows that

$$x = \frac{L_3^2 - L_2^2 + L_1^2}{2L_3^2}$$

If $L_3 = 1$, by replacing L_2 by $p - 1 - L_1$, we compute that

$$\begin{aligned} L_1 &= \frac{2x - 1 + (p-1)^2}{2(p-1)} \\ L_2 &= \frac{1 - 2x + (p-1)^2}{2(p-1)} \end{aligned}$$

■

We can now come to the

Proof of Lemma 9

Let $\triangle$ be a normalized triangle, we use the notation of last section. For $i \in [\![1, 6]\!]$, let (x_i, p_i) (respectively (x, p)) be the modified coordinates of $\triangle_i$ (resp. $\triangle$). Taking into account that $2 \leq p \leq 3$, formulas (3) and Lemma 10, it appears without difficulties that for $i \in [\![1, 6]\!]$, the mapping $(x, p) \mapsto (x_i, p_i)$ is uniformly Lipschitz on its domain of definition. Then (8) enables to conclude to the validity of Lemma 9.

■

The second goal of this section is to study the sign of quantities like $\mathbb{E}[\ln(I_{n+1}/I_n)|\triangle(n) = \triangle]$, at least when $\triangle$ is close to a flat triangle, where we recall that I_n is the isoperimetric value of $\triangle(n)$. By the Markov property, it amounts to evaluate the sign of $1/6 \sum_{i \in [\![1,6]\!]} \ln(I(\triangle_i)/I(\triangle))$. Of course the previous quotients are not rigorously defined if the triangle $\triangle$ was to be flat, Nevertheless, let (x, y) be the characterizing point of $\triangle$. When y goes to zero 0_+, $I(\triangle_i)/I(\triangle) = \sqrt{6}\mathcal{P}(\triangle_i)/\mathcal{P}(\triangle)$ converges to $G(i, x)$, which is just the latter quotient for the flat triangle $\triangle$ whose characterizing point is $(x, 0)$. We have, for any $x \in [0, 1/2]$ (see the computations of section 5 for more details)

$$\begin{aligned} G(1,x) &= \sqrt{\tfrac{2}{3}}(1+x), & G(2,x) &= \sqrt{\tfrac{1}{6}}(2-x) \\ G(3,x) &= \sqrt{\tfrac{3}{2}}(1-x), & G(4,x) &= \sqrt{\tfrac{2}{3}}(2-x) \\ G(5,x) &= \sqrt{\tfrac{2}{3}}(2-x), & G(6,x) &= \sqrt{\tfrac{3}{2}} \end{aligned} \tag{9}$$

From the previous considerations, we easily get that this convergence is uniform over x, in the sense that for any $i \in [\![1,6]\!]$,

$$\lim_{y\to 0_+} \sup_{(x,y)\in\mathcal{D}} \left| \frac{I(\triangle_i)}{I(\triangle)} - G(i,x) \right| \quad = \quad 0$$

So to prove that $\mathbb{E}[\ln(I_{n+1}/I_n)|\triangle(n) = \triangle] > 0$ for nearly flat triangles $\triangle$, it is sufficient to show that the mapping $[0,1/2] \ni x \mapsto \sum_{i\in[\![1,6]\!]} \ln(G(i,x))$ only takes positive values. Unfortunately, this is not true, since it takes negative values in a neighborhood of $1/2$ (see section 7). To get around this problem, we iterate the barycentric subdivision one more step.

Proposition 11 *There exist a constant $\gamma > 0$ and a neighborhood $\mathcal{N}$ of the set of the flat triangles, such that*

$$\forall\, n \in \mathbb{N}, \forall\, \triangle \in \mathcal{N}, \qquad \mathbb{E}[\ln(I_{n+2}/I_n)|\triangle(n) = \triangle] \quad \geq \quad \gamma$$

(for flat triangles $\triangle$, the quotient is defined as a limit as above, or equivalently, as a quotient of perimeters, before renormalisation, up to the factor 6).

Proof

Coming back to the notations of the beginning of the introduction, we want to find $\mathcal{N}$ and γ as above and satisfying

$$\forall\, \triangle \in \mathcal{N}, \qquad \frac{1}{36} \sum_{i,j\in[\![1,6]\!]} \ln\left(\frac{6\mathcal{P}(\triangle_{i,j})}{\mathcal{P}(\triangle)}\right) \quad \geq \quad \gamma \tag{10}$$

Let (x,y) be the characterizing point of $\triangle$, as y goes to 0_+, the left hand side converges (uniformly over x) to

$$F(x) \quad \coloneqq \quad \frac{1}{36} \sum_{i,j\in[\![1,6]\!]} \ln(G(j, z_i(x))G(i,x)) \tag{11}$$

where the $z_i(x)$, for $i \in [\![1,6]\!]$, were defined in (6). More explicitely, we will compute in section 5 (see Lemma 15) that on each of the segments $[0,1/5]$, $[1/5,2/7]$ and $[2/7,1/2]$, the z_i, for $i \in [\![1,6]\!]$, are homographical mappings. So it seems more convenient to consider the rational fraction (by pieces)

$$\begin{aligned} R(x) \quad &\coloneqq \quad \exp(36F(x)) \\ &= \quad \prod_{i,j\in[\![1,6]\!]} G(i, z_j(x))G(j,x) \end{aligned} \tag{12}$$

After computations (see section 7), it appears that it is indeed a polynomial function by pieces. By numerically studying the zeroes of $R-1$ of the three underlying polynomial functions, we show that F does not vanishes on $[0,1/2]$. So by continuity, we get that $\gamma \coloneqq \min_{[0,1/2]} F/2 > 0$. Then using the above uniform convergence, we can find a neighborhood $\mathcal{N}$ of the set of flat triangles so that (10) is fulfilled.

■

We will see more precisely in section 7 that F is decreasing, so we can take $\gamma = F(1/2)/2 \approx 0.035$.

4 Almost sure convergence to flatness

We are now in position to prove Theorem 1. The principe behind the proof is that there is a neighborhood $\mathcal{N}'$ of the set of flat triangles such that if the triangle Markov chain is inside $\mathcal{N}'$, then it has a positive probability to always stay in this neighborhood and then to converge exponentially fast to the set of flat triangles. This event will eventually occur, since triangle Markov chain always returns to $\mathcal{N}'$.

In order to see that the triangle Markov chain has a positive chance to remain trapped in a neighborhood of the set of flat triangles, we will use a general martingale argument. To do so, we introduce some notations. On some underlying probability space, let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration, namely a non-decreasing sequence of σ-algebras. Let $\gamma > 0$ and $A > 0$ be two given constants. We assume that for any R large enough, say $R \geq R_0 > 0$, we are given a chain $(V_n^{(R)})_{n\in\mathbb{N}}$ and a martingale $(N_n^{(R)})_{n\in\mathbb{N}}$, adapted to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, satisfying $V_0^{(R)} = R$, $N_0^{(R)} = 0$ and such that for any time $n \in \mathbb{N}$,

$$\begin{aligned} \left|N_{n+1}^{(R)} - N_n^{(R)}\right| &\leq A &\quad (13)\\ V_{n+1}^{(R)} - V_n^{(R)} &\geq \gamma + N_{n+1}^{(R)} - N_n^{(R)} &\quad (14) \end{aligned}$$

The next result shows that if R is large enough, with high probability $V^{(R)}$ will never go below $R/2$.

Lemma 12 *There exists a constant $A' > 0$ only depending on A and γ such that*

$$\mathbb{P}[\exists\, n \in \mathbb{N} : V_n^{(R)} < R/2] \leq \exp(-\gamma R/A') \frac{1}{1 - \exp(-\gamma^2/(4A'))}$$

Furthermore, we have a.s.

$$\liminf_{n\to\infty} \frac{V_n^{(R)}}{n} \geq \gamma$$

Proof
The constant $A' > 0$ is defined as

$$A' := \max\left(A^2 \sup_{|t|\leq A,\, t\neq 0} \frac{\exp(t) - 1 - t}{t^2}; \frac{\gamma}{2}\right)$$

Thus we have for any $\lambda \in [-1, 1]$ and any $n \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}[\exp(-\lambda(N_{n+1}^{(R)} - N_n^{(R)}))|\mathcal{F}_n] &\leq \mathbb{E}[1 + \lambda(N_{n+1}^{(R)} - N_n^{(R)}) + A'A^{-2}\lambda^2(N_{n+1}^{(R)} - N_n^{(R)})^2|\mathcal{F}_n]\\ &= 1 + A'A^{-2}\lambda^2\mathbb{E}[(N_{n+1}^{(R)} - N_n^{(R)})^2|\mathcal{F}_n]\\ &\leq 1 + A'\lambda^2\\ &\leq \exp(A'\lambda^2) \end{aligned}$$

By induction, we deduce the following exponential martingale bound:

$$\forall\, n \in \mathbb{N},\, \lambda \in [-1, 1], \qquad \mathbb{E}[\exp(-\lambda N_n^{(R)})] \leq \exp(\lambda^2 A' n)$$

Indeed, it is true for $n = 0$ and if we assume it is satisfied for some integer $n \in \mathbb{N}$, then

$$\begin{aligned} \mathbb{E}[\exp(-\lambda N_{n+1}^{(R)})] &\leq \mathbb{E}[\mathbb{E}[\exp(-\lambda(N_{n+1}^{(R)} - N_n^{(R)}))|\mathcal{F}_n]\exp(-\lambda N_n^{(R)})]\\ &\leq \exp(A'\lambda^2)\mathbb{E}[\exp(-\lambda N_n^{(R)})]\\ &\leq \exp(\lambda^2 A'(n+1)) \end{aligned}$$

Next we note that for any $n \in \mathbb{N}$,

$$V_n^{(R)} \geq R + n\gamma + N_n^{(R)} \tag{15}$$

so we get for any $\lambda \in [0,1]$,

$$\begin{aligned}
\mathbb{P}[V_n^{(R)} < R/2] &\leq \mathbb{P}[-N_n^{(R)} > R/2 + n\gamma] \\
&\leq \exp(-\lambda(R/2+n\gamma))\mathbb{E}[\exp(-\lambda N_n^{(R)})] \\
&\leq \exp(-\lambda(R/2+n\gamma))\exp(\lambda^2 A' n) \\
&= \exp(-\lambda R/2)\exp(n(\lambda^2 A' - \gamma\lambda))
\end{aligned}$$

The quantity is minimal when $\lambda = \gamma/(2A')$ and this number belong to $[0,1]$ by our definition of A'. It follows that

$$\forall\, n \in \mathbb{N}, \qquad \mathbb{P}[V_n^{(R)} < R/2] \leq \exp(-\gamma R/A')\exp\left(-\frac{\gamma^2}{4A'}n\right)$$

The first announced bound now follows by summation over $n \in \mathbb{N}$.

The second bound is classical: due to the fact that the increments of the martingale $N^{(R)}$ are bounded, the iterated logarithm law (see for instance Stout [8]) implies that a.s.

$$\limsup_{n\to\infty} \frac{\left|N_n^{(R)}\right|}{\sqrt{n\ln(\ln(n))}} \leq A$$

thus (15) enables us to conclude.

■

The latter lemma will be applied with $V^{(R)}$ a stochastic chain of logarithm of isoperimetric values, or rather with a sequence of the kind $(\ln(I_{2n}))_{n\in\mathbb{N}}$.
More precisely, consider the neighborhood $\mathcal{N}$ obtained in Proposition 11. By compactness of $[0,1/2]$, there exists a small constant $\epsilon > 0$ such that $\mathcal{N}$ contains $\{(x,y) \in \mathcal{D} : 0 \leq y < \epsilon\}$ and so taking into account (1), there exists $R_1 > 1$ such that $\{\triangle : \ln(I(\triangle)) > R_1\} \subset \mathcal{N}$ (again we are slightly abusing of notations here, identifying triangles with the characterizing points of their normalized forms, this should not lead to confusion). Let T be a finite stopping time for the triangle Markov chain $(\triangle(n))_{n\in\mathbb{N}}$. Assume that $R \coloneqq \ln(I(\triangle(T)))$ satisfies $R \geq 2R_1$. Define a stopping time τ for the shifted chain $(\triangle(T+2n))_{n\in\mathbb{N}}$ by

$$\tau \coloneqq \inf\{n \in \mathbb{N} : \ln(I(\triangle(T+2n))) \leq R_1\}$$

which is infinite if the set of the r.h.s. is empty. Let $\gamma > 0$ the constant appearing in Proposition 11. We construct a stochastic chain $V^{(R)}$ in the following way:

$$\forall\, n \in \mathbb{N}, \qquad V_n^{(R)} \coloneqq \begin{cases} \ln(I(\triangle(T+2n))) & \text{, if } n \leq \tau \\ \ln(I(\triangle(T+2\tau))) + \gamma(n-\tau) & \text{, otherwise} \end{cases}$$

Let us check that the above assumptions are satisfied. Following the traditional Doob-Meyer semi-martingale decomposition (see for instance Dellacherie and Meyer [3]), we define

$$\forall\, n \in \mathbb{N}, \qquad N_n^{(R)} \coloneqq \sum_{m\in[\![1,n]\!]} V_m^{(R)} - \mathbb{E}[V_m^{(R)}|\mathcal{F}_{m-1}]$$

where for any $n \in \mathbb{N}$, $\mathcal{F}_n$ is the σ-algebra generated by the trajectory-valued variable $(\triangle(m \wedge (T + n)))_{m\in\mathbb{N}}$. Using classical stopping time notations, this is the σ-algebra $\mathcal{T}_{T+n}$, where the filtration $(\mathcal{T}_m)_{m\in\mathbb{N}}$ was introduced in Lemma 7. After conditioning by $\mathcal{F}_0$ and taking advantage of the strong

Markov property, we can apply Lemma 8 to see that (13) is satisfied with $A = (b/a)^2$ (we even have $N^{(R)}_{n+1} - N^{(R)}_n = 0$ for $n \geq \tau$). Furthermore, we have for any $n \in \mathbb{N}$,

$$\begin{aligned} V^{(R)}_{n+1} - V^{(R)}_n &= \mathbb{E}[V^{(R)}_{n+1}|\mathcal{F}_n] - V^{(R)}_n + V^{(R)}_{n+1} - \mathbb{E}[V^{(R)}_{n+1}|\mathcal{F}_n] \\ &= \mathbb{E}[V^{(R)}_{n+1} - V^{(R)}_n|\mathcal{F}_n] + N^{(R)}_{n+1} - N^{(R)}_n \\ &= \mathbb{E}[\ln(I_{T+2(n+1)}/I_{T+2n})|\triangle(T+2n)]\mathbb{1}_{n\leq\tau} + \gamma\mathbb{1}_{n>\tau} + N^{(R)}_{n+1} - N^{(R)}_n \\ &\geq \gamma + N^{(R)}_{n+1} - N^{(R)}_n \end{aligned}$$

where the last inequality comes from Proposition 11. Then Lemma 12 implies that

Proposition 13 *Let $\mathcal{N}' \coloneqq \{\triangle : \ln(I(\triangle)) > R_1\}$. There exists a large enough constant $R_2 \geq 2R_1$ such that for any finite stopping time T for the triangle Markov chain $(\triangle(n))_{n\in\mathbb{N}}$ satisfying $\ln(I(\triangle(T))) \geq R_2$, we have*

$$\mathbb{P}[\exists\ n \in \mathbb{N} : \triangle(T+n) \notin \mathcal{N}'|\mathcal{T}_T] \ <\ 1/2$$

Furthermore on the event $\{\forall\ n \in \mathbb{N} : \triangle(T+n) \in \mathcal{N}'\}$, we have a.s.

$$\liminf_{n\to\infty} \frac{\ln(I_n)}{n} \ \geq\ \gamma/2$$

Indeed, Lemma 12 only gives one hand that we can find $R_2 \geq 2R_1$ such that

$$\mathbb{P}[\tau < \infty|\mathcal{T}_T] \ =\ \mathbb{P}[\exists\ n \in \mathbb{N} : \triangle(T+2n) \notin \mathcal{N}'|\mathcal{T}_T] \ <\ 1/2$$

and on the other hand that on the event $\{\forall\ n \in \mathbb{N} : \triangle(T+2n) \in \mathcal{N}'\}$, we have a.s.

$$\liminf_{n\to\infty} \frac{\ln(I_{T+2n})}{n} \ \geq\ \gamma$$

But Lemma 8 permits to extend these results to the above statement (up to replacement of R_2 by bR_2/a.

Now the proof of Theorem 1 is standard. By iteration, we introduce two sequences $(S_n)_{n\in\mathbb{N}}$ and $(T_n)_{n\in\mathbb{N}}$ of stopping times for the triangle Markov chain: we start with $S_0 = 0$ and for any $n \in \mathbb{N}$, if S_n has been defined, we take

$$\begin{aligned} T_n &\coloneqq \inf\{m > S_n : \ln(I(\triangle(m))) > R_2\} \\ S_{n+1} &\coloneqq \inf\{m > T_n : \ln(I(\triangle(m))) < R_1\} \end{aligned}$$

Of course, if for some $n \in \mathbb{N}$, $S_n = \infty$ then for any $m \geq n$, $S_m = T_m = \infty$. Conversely, via Lemma 4, we see that if $S_n < \infty$, then a.s., $T_n < +\infty$, so the events $\{S_n < \infty\}$ and $\{T_n = \infty\}$ are the same, up to a negligible set. For $n \in \mathbb{N}$, let us define the event

$$\begin{aligned} E_n &\coloneqq \{S_n < \infty \text{ and } S_{n+1} = \infty\} \\ &= \{T_n < \infty \text{ and } \forall\ m \in \mathbb{N},\ \triangle(T_n + m) \in \mathcal{N}'\} \end{aligned}$$

Up to condionning with respect to $\{S_n < \infty\}$, Lemma 13 shows that

$$\mathbb{P}[S_{n+1} = \infty|S_n < \infty] \ =\ \mathbb{P}[E_n|S_n < \infty] \ \geq\ 1/2$$

thus it follows easily that $\mathbb{P}[\cup_{n\in\mathbb{N}} E_n] = 1$. As Lemma 13 also shows that on all the E_n, the sequence $(I_m^{-1})_{m\in\mathbb{N}}$ converges exponentially fast to zero with rate at least γ, the bound (1) enables to conclude to the validity of Theorem 1 with $\chi = \gamma/2$.

Remark 14 Let $\gamma_2 \coloneqq F(1/2) = \min_{x\in[0,1/2]} F(x)$. A closer look at the proof of Proposition 11 shows that for any $\gamma < \gamma_2$, we can find a neighborhood $\mathcal{N}$ of the set of flat triangles such that the lower bound of Proposition 11 is satisfied. By the above arguments, it follows that Theorem 1 also holds with $\chi = \gamma_2/2$, so we win a factor $1/2$.
But one can go further. For $N \in \mathbb{N} \setminus \{0,1\}$ and $x \in [0,1/2]$, consider

$$\begin{aligned} F_N(x) &\coloneqq \frac{1}{6^N} \sum_{(i_1,\ldots,i_N)\in[\![1,6]\!]^N} \ln\left(G(i_N, z_{i_{N-1}} \circ \cdots \circ z_{i_1}(x)) \cdots G(i_2, z_{i_1}(x)) G(i_1, x)\right) \\ &= \mathbb{E}_x\left[\sum_{n\in[\![0,N-1]\!]} \ln(G(I_{n+1}, Z_n))\right] \end{aligned}$$

where $(I_n)_{n\in\mathbb{N}^*}$ is a sequence of independent random variables uniformly distributed on $[\![1,6]\!]$ and $(Z_n)_{n\in\mathbb{N}}$ is the Markov chain starting from x ($Z_0 \coloneqq x$) constructed from $(I_n)_{n\in\mathbb{N}^*}$ through the relations

$$\forall\, n \in \mathbb{N}, \qquad Z_{n+1} \coloneqq z_{I_{n+1}}(Z_n) \tag{16}$$

Then define

$$\gamma_N \coloneqq \min_{x\in[0,1/2]} F_N(x)$$

An easy extension of the previous proof shows that Theorem 1 equally holds with $\chi = \gamma_N/N$ and by consequence with $\chi = \lim_{N\to\infty} \gamma_N/N$. The quantity γ_N/N is converging due to the weak convergence of the Markov chain $(Z_n)_{n\in\mathbb{N}}$, uniformly in its initial distribution, as we will show it in next section. Indeed, if μ is the attracting invariant probability associated to $(Z_n)_{n\in\mathbb{N}}$, we will see that for any $x \in [0,1/2]$,

$$\lim_{n\to\infty} \mathbb{E}_x\left[\ln(G(I_{n+1}, Z_n))\right] = L$$

with

$$L \coloneqq \frac{1}{6} \sum_{i\in[\![1,6]\!]} \int \ln(G(i,x))\, \mu(dx) \tag{17}$$

It follows from Cesaro's lemma that

$$\begin{aligned} \lim_{n\to\infty} \frac{F_N(x)}{N} &= \lim_{n\to\infty} \frac{1}{N} \sum_{n\in[\![0,N-1]\!]} \mathbb{E}_x\left[\ln(G(I_{n+1}, Z_n))\right] \\ &= L \end{aligned}$$

Since this convergence holds uniformly in $x \in [0,1/2]$, we get that Theorem 1 is satisfied with $\chi = L$. In section 7, we will numerically evaluate that $L \approx 0.07$.

□

5 Ergodicity of the limit flat Markov chain

We study here more precisely the limit flat Markov chain $Z \coloneqq (Z_n)_{n\in\mathbb{N}}$. First we will see that it admits a unique invariant probability μ and that it converges exponentially fast to it in the Wasserstein distance. Next we will show that μ is continuous and that its support is the whole state space $[0,1/2]$.

We begin by describing more explicitely the kernel of Z given in (7) under an iterated random functions form.

Lemma 15 *With the notations of the previous sections, we have for all $x \in [0, 1/2]$,*

$$\begin{aligned} z_1(x) &= \frac{3x}{2+2x} \\ z_2(x) &= \frac{3x}{2-x}\mathbb{1}_{x<2/7} + \frac{2-4x}{2-x}\mathbb{1}_{x\geq 2/7} \\ z_3(x) &= \frac{1+x}{3-3x}\mathbb{1}_{x<1/5} + \frac{2-4x}{3-3x}\mathbb{1}_{x\geq 1/5} \\ z_4(x) &= \frac{1+x}{4-2x} \\ z_5(x) &= \frac{1-2x}{4-2x} \\ z_6(x) &= \frac{1-2x}{3} \end{aligned}$$

Proof

First, consider a general triangle $\triangle$ under the normalized form. With the notations of section 2, in particular (2), the lengths of the edges of the triangles obtained by the barycentric subdivision are given by

$$\begin{array}{ccccc} \triangle_1 & : & L_1/2, & l_1/3, & 2l_2/3 \\ \triangle_2 & : & L_1/2, & l_1/3, & 2l_3/3 \\ \triangle_3 & : & L_2/2, & l_2/3, & 2l_3/3 \\ \triangle_4 & : & L_2/2, & 2l_1/3, & l_2/3 \\ \triangle_5 & : & 1/2, & 2l_1/3, & l_2/3 \\ \triangle_6 & : & 1/2, & 2l_2/3, & l_3/3 \end{array}$$

This is also valid if $\triangle$ is a flat triangle with $(x, 0)$ as characteristic point and in this situation we can express the above lengths in terms of x since we have

$$\begin{array}{ccc} L_1 \coloneqq x, & L_2 \coloneqq 1-x, & L_3 \coloneqq 1 \\ l_1 \coloneqq 1-\frac{x}{2}, & l_2 \coloneqq \frac{1+x}{2}, & l_3 \coloneqq \frac{1}{2}-x \end{array}$$

Taking into account that for any flat triangle, the abscissa of the characteristic point is the ratio of the smallest edge by the largest edge, the announced formulas follow.

■

To see that the Markov kernel M of Z is ergodic, in the sense that it admits an invariant and attracting probability, we apply a result due to Barnsley and Elton [1]: let S be a compact segment of $\mathbb{R}$ (more generally it can be a complete, separable metric space) on which we are given n Lipschitz functions $f_i : S \to S$, for $i \in [\![1, n]\!]$. Let $p = (p_i)_{i\in[\![1,n]\!]}$ be a probability on $[\![1, n]\!]$ and consider the Markov kernel N from S to S given by

$$\forall\, x \in S, \qquad N(x, \cdot) \;\coloneqq\; \sum_{i\in[\![1,n]\!]} p_i \delta_{f_i(x)} \tag{18}$$

Then under the assumption that there exists a constant $r < 0$ such that

$$\forall\, x \neq y \in S, \qquad \sum_{i\in[\![1,n]\!]} p_i \ln\left(\frac{|f_i(y) - f_i(x)|}{|y - x|}\right) \;\leq\; r \tag{19}$$

the kernel N is ergodic: it admits a unique invariant and attracting probability μ, namely satisfying $\mu N = \mu$ and for any probability ν on S, $\lim_{n\to\infty} \nu N^n = \mu$ (in the weak topology). Furthermore, Barnsley and Elton [1] show that there exists $q \in (0, 1]$ and $\rho \in (0, 1)$ such that

$$\forall\, x, y \in S, \qquad \sum_{i\in[\![1,n]\!]} p_i\, |f_i(y) - f_i(x)|^q \;\leq\; \rho\, |y - x|^q \tag{20}$$

Let us rewrite this bound in a more probabilistic way. Let $(I_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables taking values in $[\![1,n]\!]$ with distribution $(p_i)_{i\in[\![1,n]\!]}$. For any $x \in S$, we denote by $U^x := (U_n^x)_{n\in\mathbb{N}}$ the stochastic chain constructed as follow: $U_0^x = x$ and for any $n \in \mathbb{N}$, $U_{n+1}^x = f_{I_{n+1}}(U_n^x)$. It is a Markov chain with transition kernel N. This construction enables us to couple together all the Markov chains U^x, for $x \in S$. Then the above bound can be written

$$\forall\, x, y \in S, \qquad \mathbb{E}[|U_1^y - U_1^x|^q] \;\;\leq\;\; \rho\,|y-x|^q$$

and admits an immediate extension:

$$\forall\, n \in \mathbb{N}, \forall\, x, y \in S, \qquad \mathbb{E}[|U_n^y - U_n^x|^q] \;\;\leq\;\; \rho^n\,|y-x|^q$$

This leads us to consider the Wasserstein distance D between probability measures on S: if ν_1 and ν_2 are two such measures,

$$D(\nu_1, \nu_2) \;\;:=\;\; \sup_{f\in\mathcal{L}(1)} |\nu_1[f] - \nu_2[f]|$$

where $\mathcal{L}(1)$ is the set of Lipschitz functions on S whose Lipschitz constant is less (or equal) than 1. Then it appears that

Lemma 16 *Under the above assumption (24), we have for any time $n \in \mathbb{N}$ and any $x \in S$,*

$$D(N^n(x,\cdot), \mu) \;\;\leq\;\; \mathrm{diam}(S)\rho^n$$

where ρ and q are as in (20) and $\mathrm{diam}(S)$ *is the diameter of S. It follows that U^x satisfies the law of large numbers: for any continuous function f on S, we have a.s.,*

$$\lim_{N\to\infty} \frac{1}{N+1} \sum_{n\in[\![0,N]\!]} f(U_n^x) \;\;=\;\; \mu[f]$$

Proof

Let $f \in \mathcal{L}(1)$, we compute that

$$\begin{aligned}
|N^n(x,f) - \mu[f]| &= \left|\int \mu(dy)(N^n(x,f) - N^n(y,f))\right| \\
&\leq \sup_{x,y\in S} |N^n(x,f) - N^n(y,f)| \\
&= \sup_{x,y\in S} |\mathbb{E}[f(U_n^x) - f(U_n^y)]| \\
&\leq \sup_{x,y\in S} \mathbb{E}[|U_n^x - U_n^y|] \\
&= \mathrm{diam}(S) \sup_{x,y\in S} \mathbb{E}\left[\frac{|U_n^x - U_n^y|}{\mathrm{diam}(S)}\right] \\
&\leq \mathrm{diam}(S) \sup_{x,y\in S} \mathbb{E}\left[\left|\frac{U_n^x - U_n^y}{\mathrm{diam}(S)}\right|^q\right] \\
&\leq \mathrm{diam}(S)^{1-q} \sup_{x,y\in S} \rho^n\,|y-x|^q \\
&\leq \mathrm{diam}(S)\rho^n
\end{aligned}$$

The announced bound follows by taking the supremum over all functions $f \in \mathcal{L}(1)$. The law of large numbers is deduced from a traditional martingale argument based on the existence of a bounded solution to the Poisson equation. More precisely, for $f \in \mathcal{L}(1)$, we can define

$$\forall\, x \in S, \qquad \varphi(x) \;\;:=\;\; \sum_{n\in\mathbb{N}} \mathbb{E}[f(U_n^x) - \mu[f]]$$

since the r.h.s. converges exponentially fast and uniformly with respect to $x \in S$. Furthermore we easily see that φ is a Lipschitz function and that it is solution to the Poisson equation

$$\left\{ \begin{array}{rcl} \forall\, x \in S, \qquad \varphi(x) - N(x,\varphi) & = & f(x) - \mu[f] \\ \mu[\varphi] & = & 0 \end{array} \right.$$

This enables us to write for any $n \in \mathbb{N}$,

$$f(X_0) + f(X_1) + \cdots + f(X_n) \quad = \quad (n+1)\mu[f] + \varphi(X_0) - \varphi(X_{n+1}) + \mathcal{M}_{n+1}$$

where $(\mathcal{M}_n)_{n\in\mathbb{N}}$ is a martingale whose increments are bounded. The law of large numbers for functions f belonging to $\mathcal{L}(1)$ then follows from the well-known fact that $\mathcal{M}_n/n$ converges a.s. to zero. So it is also true for all Lipschitz functions f and by usual approximations, for all continuous functions f.

■

Let us discuss condition (19). Note that since the functions f_i, for $i \in [\![1,n]\!]$, are Lipschitz, they are absolutely continuous, so let us write f_i' for their respective weak derivatives. By letting y and x become close in criterion (19), we get that almost everywhere in $x \in S$,

$$\sum_{i\in[\![1,n]\!]} p_i \ln(|f_i'(x)|) \quad \leq \quad r \tag{21}$$

Conversely, this condition is not sufficient to insure that the kernel N is ergodic. Consider the following example with $S = [0,1]$, $n = 2$ and the functions f_1 and f_2 defined by

$$\forall\, x \in [0,1], \qquad f_i(x) \quad \coloneqq \quad \left\{ \begin{array}{l} \min(2x,1) \qquad\qquad\quad ,\text{ if } i = 1 \\ \max(0,-1+2x),\text{ if } i = 2 \end{array} \right.$$

In this case (21) is even satisfied with $r = -\infty$ and the set of invariant probability measures is $\{a\delta_0 + (1-a)\delta_1 \,:\, a \in [0,1]\}$, so none of them can be attractive (but the law of a corresponding Markov chain converges exponentially fast to one of the invariant probability measures).

Nevertheless, under some circumstances, the necessary condition (21) is also sufficient. This is in particular the case if for all the functions $|f_i'|$, with $i \in [\![1,n]\!]$, there exist $a_i, b_i \in \mathbb{R}$ such that almost everywhere (a.e.) in $x \in S$,

$$\left|f_i'(x)\right| \quad = \quad (a_i x + b_i)^{-2} \tag{22}$$

(in particular $-b_i/a_i$ cannot belong to S otherwise f_i' would not be integrable over this interval). Indeed in this situation we can write that for any $x < y \in S$,

$$\begin{aligned} \frac{|f_i(y) - f_i(x)|}{|y-x|} &= \frac{1}{|y-x|}\left|\int_x^y f_i'(z)\,dz\right| \\ &\leq \frac{1}{|y-x|}\int_x^y \left|f_i'(z)\right|\,dz \\ &= \frac{1}{|y-x|}\int_x^y \frac{1}{(a_i z + b)^2}\,dz \\ &= \frac{1}{a_i\,|y-x|}\left(\frac{1}{a_i x + b_i} - \frac{1}{a_i y + b_i}\right) \\ &= \frac{1}{a_i\,|y-x|}\frac{a_i(y-x)}{(a_i y + b_i)(a_i x + b_i)} \\ &= \frac{1}{|a_i y + b_i|\,|a_i x + b_i|} \\ &= \sqrt{|f_i'(y)|\,|f_i'(x)|} \end{aligned}$$

where the last equality has to be understood a.e. It follows that, at least for a.e. $x, y \in S$,

$$\ln\left(\frac{|f_i(y)-f_i(x)|}{|y-x|}\right) \leq \frac{\ln(|f_i'(y)|)+\ln(|f_i'(x)|)}{2} \tag{23}$$

and by consequence

$$\sum_{i\in[\![1,n]\!]} p_i \ln\left(\frac{|f_i(y)-f_i(x)|}{|y-x|}\right) \leq \frac{1}{2}\sum_{i\in[\![1,n]\!]} p_i \ln(|f_i'(y)|)+\frac{1}{2}\sum_{i\in[\![1,n]\!]} p_i \ln(|f_i'(x)|)$$

formula which enables to pass from (21) to (19), which has only to be checked for a.e. $x, y \in S$.

It is time now to come back to the flat triangle Markov chain, namely we consider the setting where $N = M$, i.e. $S = [0, 1/2]$, $n = 6$, $f_i = z_i$ for $i \in [\![1, 6]\!]$ and p the uniform distribution on $[\![1, 6]\!]$. A good news is that condition (22) is satisfied. Since $z_2'(2/7-) = -z_2'(2/7+)$ and $z_3'(1/5-) = -z_3'(1/5+)$, we see that, contrary to $z_i'(x)$, $|z_i'|(x)$ can be defined everywhere (convention that we will adopt from now on). Indeed, we compute that for any $x \in [0, 1/2]$,

$$\begin{array}{lcllcl}
|z_1'|(x) & = & \dfrac{3}{2(1+x)^2}, & |z_2'|(x) & = & \dfrac{6}{(2-x)^2} \\
|z_3'|(x) & = & \dfrac{2}{3(1-x)^2}, & |z_4'|(x) & = & \dfrac{3}{2(2-x)^2} \\
|z_5'|(x) & = & \dfrac{3}{2(2-x)^2}, & |z_6'|(x) & = & \dfrac{2}{3}
\end{array}$$

So we are wondering if (21) would not be satisfied in this situation. Unfortunately this is not true and surprisingly it is a computation we have already encountered : comparing with (9), it appears that

$$\forall\, i \in [\![1, 6]\!],\ \forall\, x \in [0, 1/2], \qquad |z_i'|(x) = \frac{1}{G^2(i, x)}$$

thus by the observation before Proposition 11, we know that $\sum_{i\in[\![1,6]\!]} \ln(|z_i'(x)|)$ is positive for x near $1/2$. As in section 3, we get around this difficulty by iterating one more time the kernel M (this iterating trick was also used by Barnsley and Elton in one example of their paper [1]). So we consider $N = M^2$, namely $S = [0, 1/2]$, $n = 36$, $f_{i,j} = z_i \circ z_j$ for $(i, j) \in [\![1, 6]\!]^2$ and p the uniform distribution on $[\![1, 6]\!]^2$. Its advantage is that we have for any $i, j \in [\![1, 6]\!]$ and any $x \in [0, 1/2]$,

$$\begin{aligned}
|f_{i,j}'|(x) &= |z_i'|(z_j(x))\,|z_i'|(x) \\
&= (G(i, z_j(x))G(j, x))^{-2}
\end{aligned}$$

Thus it appears that

$$\forall\, x \in [0, 1/2], \qquad \sum_{i,j\in[\![1,6]\!]} \ln(|f_{i,j}'|(x)) = -2F(x)$$

and in particular the left hand side is negative due to Proposition 11 (it is even increasing as a function of $x \in [0, 1/2]$ according to the observation made at the end of section 3). But (22) is no longer satisfied by the functions $f_{i,j}$. To avoid this problem, we come back directly to the bound (23): for $i, j \in [\![1, 6]\!]$ and $y > x \in [0, 1/2]$, we write

$$\begin{aligned}
\ln\left(\frac{|f_{i,j}(y)-f_{i,j}(x)|}{|y-x|}\right) &= \ln\left(\frac{|z_i(z_j(y))-z_i(z_j(x))|}{|z_j(y)-z_j(x)|}\right)+\ln\left(\frac{|z_j(y)-z_j(x)|}{|y-x|}\right) \\
&\leq \frac{\ln(|z_i'(z_j(y))|)+\ln(|z_i'(z_j(x))|)}{2}+\frac{\ln(|z_j'(y)|)+\ln(|z_j'(x)|)}{2} \\
&= \frac{\ln(|f_{i,j}'(y)|)+\ln(|f_{i,j}'(x)|)}{2}
\end{aligned}$$

Thus in this situation we can also come back from (22) to (21) and the results of Barnsley and Elton [1] insure that the iterated Markov kernel M^2 is ergodic. To come back from M^2 to M is not difficult:

Proposition 17 *The kernel M is ergodic and the Markov chain Z satisfies the strong law of large numbers.*

Proof

Let μ be the attracting and invariant probability for M^2. Then we have $(\mu M)M^2 = (\mu M^2)M = \mu M$, so μM is invariant for M^2 and by uniqueness it follows that $\mu M = \mu$. Next for any probability measure ν on $[0,1/2]$, the (weak) limit set of $(\nu M^n)_{n\in\mathbb{N}}$ is included into $\{\mu, \mu M\} = \{\mu\}$, so μ is also attracting for M and the uniqueness of μ as the invariant probability of M follows. Finally the strong law of large numbers for Z can be deduced from that of the two Markov chains $(Z_{2n})_{n\in\mathbb{N}}$ and $(Z_{1+2n})_{n\in\mathbb{N}}$.

■

There was a cruder way to deduce the ergodicity of M:

Remark 18 Diaconis and Freedman [4] considered a simpler criterion for ergodicity of a random function Markov kernel (18): for $i \in [\![1,n]\!]$, let $K_i \coloneqq \sup_{x\neq y} |f_i(y) - f_i(x)|/|y-x|$ stands for the Lipschitz constant of f_i and assume that there exists a constant $r < 0$ such that

$$\forall\, x,y \in S, \qquad \sum_{i\in[\![1,n]\!]} p_i \ln(K_i) \quad \leq \quad r \tag{24}$$

then the kernel N is ergodic and Diaconis and Freedman [4] showed that the convergence is exponentially fast in the Prokhorov distance (but for us the Wasserstein distance is more convenient because in the end we would like to couple the two Markov chains $(X_n, Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$). Of course condition (24) implies (19). Since for $i \in [\![1,n]\!]$, K_i is the essential supremum of $|f_i'(x)|$, (24) corresponds to the exchange of essential supremum and sum in (22). Let us now come back to our flat triangle Markov chain. From the previous considerations, (24) cannot be satisfied with $N = M$. It does not work either with $N = M^2$, so this is an example were the criterion (19) is fulfilled while (24) is not. But condition (24) is satisfied with $N = M^3$, namely $S = [0,1/2]$, $n = 216$, $f_{i,j,k} = z_i \circ z_j \circ z_k$ for $(i,j,k) \in [\![1,6]\!]^3$ and p the uniform distribution on $[\![1,6]\!]^3$. For the details of the underlying numerical computions, we refer to section 7.

□

To finish we prove two properties of μ which will be needed in next section. The first one is

Lemma 19 *The probability μ contains no atom, in particular $\mu(\{0\}) = 0$.*

Proof

The proof needs a few steps and notations. Let us define

$$\begin{aligned} \mu^* &\coloneqq \sup\{\mu(\{x\}) \,:\, x \in [0,1/2]\} \\ S^* &\coloneqq \{x \in [0,1/2] \,:\, \mu(\{x\}) = \mu^*\} \end{aligned}$$

and for any $x \in [0,1/2]$,

$$\begin{aligned} \bar{S}(x) &\coloneqq \{(i,y) \in [\![1,6]\!] \times [0,1/2] \,:\, z_i(y) = x\} \\ S(x) &= \{y \in [0,1/2] \,:\, \exists\, i \in [\![1,6]\!] \text{ with } z_i(y) = x\} \end{aligned}$$

The first step is:

- We have $\mu(\{0\} = \mu(\{1/2\}) \leq \mu^*/2$.

By invariance of μ we can write that

$$\begin{aligned}\mu(\{0\}) &= \mu(M[\mathbb{1}_{\{0\}}]) \\ &= \frac{1}{6}\sum_{(i,y)\in\bar{S}(0)} \mu(\{y\}) \\ &= \frac{2}{6}\mu(\{0\}) + \frac{4}{6}\mu(\{1/2\})\end{aligned}$$

and this relation implies that $\mu(\{0\}) = \mu(\{1/2\})$. Next considering the point 1/2, we get

$$\begin{aligned}\mu(\{1/2\}) &= \frac{1}{6}\sum_{(i,y)\in\bar{S}(1/2)} \mu(\{y\}) \\ &= \frac{2}{6}\mu(\{1/2\}) + \frac{1}{6}\mu(\{1/5\}) + \frac{1}{6}\mu(\{2/7\}) \\ &\leq \frac{2}{6}\mu(\{1/2\}) + \frac{1}{3}\mu^*\end{aligned}$$

so it follows that $\mu(\{1/2\}) \leq \mu^*/2$.

Next step is:

• For any $x \in S^*$, we have $S(x) \subset S^* \cup \{0\}$.

Looking at the graphs of the functions z_i, for $i \in [\![1,6]\!]$ (see Figure 1 in section 7), we get that

$$\forall\ x \in [0,1/2], \qquad \mathrm{card}(\bar{S}(x)) = \begin{cases} 4 & \text{, if } x = 1/2 \\ 7 & \text{, if } x = 1/4 \text{ or } x = 1/3 \\ 6 & \text{, otherwise} \end{cases}$$

So consider $x \in S^* \setminus \{1/4, 1/3, 1/2\}$, writting that

$$\begin{aligned}\mu(\{x\}) &= \frac{1}{6}\sum_{(i,y)\in\bar{S}(x)} \mu(\{y\}) \\ &\leq \mu^*\end{aligned}$$

it appears that equality is possible only if $\mu(\{y\}) = \mu^*$ for all $y \in S(x)$, namely $S(x) \subset S^*$.
We study now the three particular cases of 1/4, 1/3 and 1/2.
- For 1/2: as seen in the first step, $1/2 \in S^*$ implies that $\mu^* = 0$, so $S^* = [0,1/2]$ and the inclusion $S(1/2) \subset S^*$ is trivial.
- For 1/4: there exist five distinct points $y_1', y_2', y_3', y_4', y_5' \in [0,1/2]$ such that we have

$$S(1/4) = \{(4,0),(5,0),(1,y_1'),(2,y_2'),(2,y_3'),(3,y_4'),(6,y_5')\}$$

so by invariance of μ we get

$$\begin{aligned}\mu(\{1/4\}) &= \frac{1}{6}\left(2\mu(\{0\}) + \sum_{i\in[\![1,5]\!]} \mu(\{y_i'\})\right) \\ &\leq \frac{1}{6}(2\mu(\{0\}) + 5\mu^*)\end{aligned}$$

and since we know that $\mu(\{0\}) \leq \mu^*/2$, the equality $\mu(\{1/4\}) = \mu^*$ is possible only if $\mu(\{y_i'\}) = \mu^*$ for $i \in [\![1,5]\!]$, so we can conclude that $S(1/4) \subset S^* \cup \{0\}$.
- The same argument holds for 1/3 (even if $1/3 \in S(1/3)$). For the last step, let us denote by $\widetilde{z}_2$ the restriction of z_2 to $[0,2/7]$. This mapping is one-to-one from $[0,2/7]$ to $[0,1/2]$ and we denote by $\widetilde{z}_2^{-1}$ its inverse.

- For $x \in [0,1/2] \setminus \{0\}$, the set $\{\widetilde{z}_2^{-n}(x) : n \in \mathbb{N}\}$ is infinite, so S^* is infinite.

The first assertion comes from the fact that for any $x \in (0,1/2]$, $0 < \widetilde{z}_2^{-1}(x) < x$, so $(\widetilde{z}_2^{-n}(x))_{n\in\mathbb{N}}$ is indeed a decreasing sequence (converging to 0). By the first step, S^* cannot be reduced to $\{0\}$, so there exists $x \in S^* \setminus \{0\}$. By the previous step, the sequence $(\widetilde{z}_2^{-n}(x))_{n\in\mathbb{N}}$ is included into S^*, since none of its elements can be equal to 0. It follows that S^* is infinite.

Of course the last statement implies that $\mu^* = 0$, because μ is a probability measure.

■

If all the functions z_i, for $i \in [\![1,6]\!]$, were strictly monotone, the fact that $\mu(\{0\}) = 0$ could have been deduced more directly from the uniqueness of μ and Theorem 2.10 from Dubins and Freedman [6]. The second information we will need on μ is a direct consequence of a result of the latter paper.

Lemma 20 *The support of μ is the whole segment $[0,1/2]$.*

Proof

By Theorem 4.9 of Dubins and Freedman [6], the support of μ is the whole segment $[0,1/2]$ if we can cover it by the images of the functions z_i which are strict contractions. But this is the case here, since z_4 and z_5 are strict contractions and $z_4([0,1/2]) = [1/4,1/2]$ and $z_5([0,1/2]) = [0,1/4]$.

■

6 More on the asymptotic behavior

Our main goal here is to prove Theorems 2 and 3. The underlying tool is to couple the Markov chains $(X_n, Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ to take advantage of the informations we have on the latter one.

Indeed, there is a natural coupling between the above chains based on the construction alluded to in Remark 14. Assume that (X_0, Y_0) and Z_0 are given and let $(I_n)_{n\in\mathbb{N}^*}$ be a sequence of independent random variables uniformly distributed on $[\![1,6]\!]$ and independent from the previous initial conditions. We consider $(Z_n)_{n\in\mathbb{N}}$ constructed as in (16) and similarly we iteratively define $(X_n, Y_n)_{n\in\mathbb{N}}$ via

$$\forall\, n \in \mathbb{N}, \qquad (X_{n+1}, Y_{n+1}) \;:=\; (x_{I_{n+1}}(X_n), y_{I_{n+1}}(Y_n))$$

In these relations, the indices refer to the conventions made in (2) and (6). A first simple property of this coupling is:

Lemma 21 *The random variable $|X_n - Z_n|$ converges in probability to zero for large time $n \in \mathbb{N}$.*

Proof

First we iterate Lemma 9 to show that there exists a constant $K' > 0$ such that

$$\forall\, i,j \in [\![1,6]\!],\ \forall\, (x,y) \in \mathcal{D}, \qquad |x_i(x_j(x,y), y_j(x,y)) - z_i(z_j(x))| \;\le\; K'\sqrt{y}$$

Indeed, using Lemma 9 and the fact that all the functions z_i, $i \in [\![1,6]\!]$ have a Lipschitz constant less (or equal) than 8/3, we deduce that for any $i,j \in [\![1,6]\!]$ and any $(x,y) \in \mathcal{D}$,

$$\begin{aligned}
&|x_i(x_j(x,y), y_j(x,y)) - z_i(z_j(x))| \\
&\quad\le\; |x_i(x_j(x,y), y_j(x,y)) - z_i(x_j(x,y))| + |z_i(x_j(x,y)) - z_i(z_j(x))| \\
&\quad\le\; K\,|y_j(x,y)| + \frac{8}{3}\,|x_j(x,y) - z_j(x)| \\
&\quad\le\; K\sqrt{y} + \frac{8K}{3}\,|y| \\
&\quad\le\; K'\sqrt{y}
\end{aligned}$$

with $K' := K(1 + 2^{\frac{5}{2}}3^{\frac{3}{4}})$, where we have taken into account that the largest possible $\sqrt{y}$ for $(x, y) \in \mathcal{D}$ is attained for the equilateral triangle and that is value is then $(3/4)^{1/4}$.
Let $q \in (0, 1]$ and $\rho \in (0, 1)$ as in (20) but relatively to the kernel $N = M^2$. Then for any $n \in \mathbb{N}$, we can write

$$\begin{aligned}
&\mathbb{E}[|X_{n+2} - Z_{n+2}|^q|X_n, Y_n, Z_n] \\
&\quad = \mathbb{E}[|x_{I_{n+2}}(x_{I_{n+1}}(X_n, Y_n), y_{I_{n+1}}(X_n, Y_n)) - z_{I_{n+2}}(z_{I_{n+1}}(Z_n))^q||X_n, Y_n, Z_n] \\
&\quad \leq \mathbb{E}[|x_{I_{n+2}}(x_{I_{n+1}}(X_n, Y_n), y_{I_{n+1}}(X_n, Y_n)) - z_{I_{n+2}}(z_{I_{n+1}}(X_n))|^q|X_n, Y_n, Z_n] \\
&\qquad +\mathbb{E}[|z_{I_{n+2}}(z_{I_{n+1}}(X_n)) - z_{I_{n+2}}(z_{I_{n+1}}(Z_n))|^q|X_n, Y_n, Z_n] \\
&\quad \leq (K')^q Y_n^{\frac{q}{2}} + \rho|X_n - Z_n|^q
\end{aligned}$$

For $n \in \mathbb{N}$, denote

$$\begin{aligned}
a_n &:= \mathbb{E}[|X_n - Z_n|^q] \\
b_n &:= (K')^q \mathbb{E}[Y_n^{\frac{q}{2}}]
\end{aligned}$$

so that after integration the above bound leads to

$$\forall\, n \in \mathbb{N}, \qquad a_{n+2} \leq \rho a_n + b_n$$

We deduce that

$$\forall\, n \in \mathbb{N}, \qquad a_{2n} \leq a_0\rho^n + \sum_{m\in[\![0,n-1]\!]} b_{2m}\rho^{n-1-m} \tag{25}$$

where it appears that $\lim_{n\to\infty} a_{2n} = 0$ is a consequence of $\lim_{n\to\infty} b_n = 0$. A similar computation shows that this latter condition also implies that $\lim_{n\to\infty} a_{2n+1} = 0$, i.e. at the end we will be insured of

$$\lim_{n\to\infty} \mathbb{E}[|X_n - Z_n|^q] = 0$$

and thus of the announced convergence in probability.
But we already know that $(Y_n)_{n\in\mathbb{N}}$ converges a.s. to zero and since this sequence is uniformly bounded, we see by the dominated convergence theorem that $\lim_{n\to\infty} b_n = 0$.

■

Now Theorem 2 follows quite easily:

Proof of Theorem 2

For $n \in \mathbb{N}$, let H_n be the orthogonal projection of (X_n, Y_n) on the abscissa axis and denote A'_n (respectively A''_n) the angle between $[(0, 0),(X_n, Y_n)]$ and $[(X_n, Y_n), H_n]$ (resp. $[H_n,(X_n, Y_n)]$ and $[(X_n, Y_n),(1, 0)]$), so that $A_n = A'_n + A''_n$. Since the length of $[H_n,(1, 0)]$ is larger than $1/2$ and Y_n converges a.s. to zero for large $n \in \mathbb{N}$, it is clear that A''_n converges a.s. to $\pi/2$. Furthermore, we have $\tan(A'_n) = X_n/Y_n$, so to see that A_n converges in probability toward π, we must see that Y_n/X_n converges in probability toward 0. Let $\epsilon, \eta > 0$ be given, we have for any $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}[Y_n/X_n \geq \epsilon] &\leq \mathbb{P}[Y_n \geq 2\epsilon\eta] + \mathbb{P}[X_n \leq 2\eta] \\
&\leq \mathbb{P}[Y_n \geq 2\epsilon\eta] + \mathbb{P}[|X_n - Z_n| \leq \eta] + \mathbb{P}[Z_n \leq \eta]
\end{aligned}$$

Taking into account that the attractive probability μ of Z is continuous, we get by letting n going to infinity

$$\limsup_{n\to\infty} \mathbb{P}[Y_n/X_n \geq \epsilon] \leq \mu([0, \eta])$$

Using again Lemma 19 and letting η go to zero, we obtain that $\lim_{n\to\infty}\mathbb{P}[Y_n/X_n \geq \epsilon] = 0$ and by consequence the announced convergence in probability.

■

Remark 22 We don't know if $(A_n)_{n\in\mathbb{N}}$ is converging to zero a.s. One way to deduce this result, via Borel-Cantelli lemma, would be to show that for any given $\epsilon > 0$,

$$\sum_{n\in\mathbb{N}} \mathbb{P}[Y_n/X_n \geq \epsilon] \quad < \quad +\infty$$

In view of the above arguments, one of the main problems is that we have no bound on the way $\mu([0,\eta]))$ goes to zero as η goes to zero. We would like to find $\alpha > 0$ such that $\limsup_{\eta\to 0_+} \mu([0,\eta])/\eta^\alpha$ $< +\infty$, but we were not able to prove such an estimate. If we knew that μ is absolutely continuous, Figure 7 in section 7 would suggest that this property holds with $\alpha = 1$ (and $\lim_{\eta\to 0_+}\mu([0,\eta])/\eta \leq 1$).

□

In order to prove Theorem 3, we will need two technical results. In all that follows, let fix some $a \in [0, 1/2]$ and $\epsilon > 0$ and define $\mathcal{O} \coloneqq [a-\epsilon, a+\epsilon] \cap [0, 1/2]$.

Lemma 23 *There exist $\eta > 0$ and $N \in \mathbb{N}^*$ such that*

$$\inf_{z\in[0,1/2]} \mathbb{P}_z[Z_N \in \mathcal{O}] \quad \geq \quad \eta$$

(the index z means that $Z_0 = z$).

Proof

This is a consequence of Lemma 16 applied to M^2: there exists $\rho \in (0, 1)$ such that for any $z \in [0, 1/2]$ and $n \in \mathbb{N}$, $D(M^{2n}, \mu) \leq \rho^n$. Let φ a function from $\mathcal{L}(1)$ vanishing outside $\mathcal{O}$ and whose value is ϵ at a (φ is uniquely determined if $\mathcal{O}$ is strictly included into $[0, 1/2]$). Since the support of μ is [0,1/2], we have $\eta \coloneqq \mu[\varphi]/2 > 0$. So if we choose $N \in \mathbb{N}$ such that $\rho^{\lfloor N/2\rfloor} < \eta$, we get that for any $z \in [0, 1/2]$, $\mathbb{P}_z[X_N \in \mathcal{O}] \geq M^N(z, \varphi) > \eta$.

■

For the second preliminary, we need a convenient notation: for $(x, y) \in \mathcal{D}$ and $(i_1, i_2, \cdots, i_N) \in [\![1, 6]\!]^N$, we denote $x_{i_1,i_2,\cdots,i_N}(x, y)$, $y_{i_1,i_2,\cdots,i_N}(x, y)$ and $z_{i_1,i_2,\cdots,i_N}(x)$ the values of X_n, Y_n and Z_n when $(X_0, Y_0, Z_0) = (x, y, z)$ and $(I_1, I_2, \cdots, I_N) = (i_1, i_2, \cdots, i_N)$.

Lemma 24 *There exists a constant K_N such that*

$$\forall\ (i_1, i_2, \cdots, i_N) \in [\![1, 6]\!]^N,\ \forall\ (x, y) \in \mathcal{D}, \qquad \begin{cases} |x_{i_1,i_2,\cdots,i_N}(x, y) - z_{i_1,i_2,\cdots,i_N}(x)| & \leq & K_N y^{1/2^{N-1}} \\ |y_{i_1,i_2,\cdots,i_N}(x, y)| & \leq & K_N y^{1/2^N} \end{cases}$$

Proof

For $N = 1$ this is just Lemma 9. The general case is proven by an easy iteration, similar to the one already used in the proof of Lemma 21, starting with

$$\begin{aligned} & |x_{i_1,i_2,\cdots,i_N}(x, y) - z_{i_1,i_2,\cdots,i_N}(x)| \\ & \quad \leq \quad \left|x_{i_N}(x_{i_1,i_2,\cdots,i_{N-1}}(x, y), y_{i_1,i_2,\cdots,i_{N-1}}(x, y) - z_{i_N}(x_{i_1,i_2,\cdots,i_{N-1}}(x, y))\right| \\ & \qquad + \left|z_{i_N}(x_{i_1,i_2,\cdots,i_{N-1}}(x, y)) - z_{i_N}(z_{i_1,i_2,\cdots,i_{N-1}}(x))\right| \end{aligned}$$

■

We can now come to the

Proof of Theorem 3

Let $\mathcal{O}' := [a-2\epsilon, a+2\epsilon]\cap[0,1/2]$. We want to show an analogous result to Lemma 23 but for the chain $(X_n,Y_n)_{n\in\mathbb{N}}$, namely to find $\eta'>0$ and $N'\in\mathbb{N}^*$ such that

$$\inf_{(x,y)\in\mathcal{D}} \mathbb{P}_{(x,y)}[X_{N'}\in\mathcal{O}'] \quad\geq\quad \eta' \tag{26}$$

(the index (x,y) recalls that initially $(X_0,Y_0)=(x,y)$). To do so, we first consider η and N as in Lemma 23 and consider $\delta>0$ sufficiently small such that $K_N\delta^{1/2^{N-1}}<\epsilon$. Then according to Lemmas 23 and 24, we have

$$\inf_{(x,y)\in\mathcal{D}\,:\,y<\delta} \mathbb{P}_{(x,y)}[X_N\in\mathcal{O}'] \quad\geq\quad \eta$$

To extend such an estimate to the whole domain $\mathcal{D}$, we come back to (1) and (5) which enable us to find $N''\in\mathbb{N}$ such that

$$\eta'' \quad:=\quad \inf_{(x,y)\in\mathcal{D}} \mathbb{P}_{(x,y)}[Y_{N''}<\delta] \;>\; 0$$

Now the Markov property implies (26) with $\eta=\eta'\eta''$ and $N'=N+N''$.

Since this bound is uniform over $(x,y)\in\mathcal{D}$, the sequence $(X_{nN'})_{n\in\mathbb{N}}$ is stochastically bounded below by an independent familly of Bernoulli variables of parameter η' and we deduce that a.s.

$$\limsup_{n\to\infty} \mathbb{1}_{\mathcal{O}'}(X_n) \quad=\quad 1$$

The announced result follows because $a\in[0,1/2]$ and $\epsilon>0$ are arbitrary.

■

We would have liked to get more quantitative bounds but unfortunately we didn't:

Remark 25 To complete Remark 22, we are also missing another estimate of the kind

$$\exists\; K,p,\chi>0 \;:\; \forall\; n\in\mathbb{N}, \qquad \mathbb{E}[Y_n^p] \quad\leq\quad K\exp(-\chi n) \tag{27}$$

Blackwell [2] succeeded in obtaining such a bound (with $p=1/2$) by exhibiting an appropriate supermartingale with the help of the computer. His result also implies Theorem 1, with $\chi=0.04$. Furthermore it allows for a more direct proof of Theorem 3. Indeed, by the arguments of the proof of Lemma 21, we deduce from (27) that $|X_n-Z_n|$ converges a.s. to zero. Thus the law of large numbers for Z and Lemma 20 imply immediately Theorem 3. We can even go further to justify the assertion made in the introduction that asymptotically $(X_n)_{n\in\mathbb{N}}$ is almost Markovian. Let us introduce the supremum distance d on $[0,1/2]^{\mathbb{N}}$, seen as the set of trajectories from $\mathbb{N}$ to $[0,1/2]$:

$$\forall\; x:=(x_n)_{n\in\mathbb{N}}, y:=(y_n)_{n\in\mathbb{N}}\in[0,1/2]^{\mathbb{N}}, \qquad d(x,y) \quad:=\quad \sup_{n\in\mathbb{N}}|x_n-y_n|$$

For $m\in\mathbb{N}$, let $X_{[\![m,\infty[\![}=(X_{m+n})_{n\in\mathbb{N}}\in[0,1/2]^{\mathbb{N}}$ and consider

$$D_m \quad:=\quad \inf\mathbb{E}[d(X_{[\![m,\infty[\![},Z)]$$

where the infimum is taken over all couplings of $X_{[\![m,\infty[\![}$ with a Markov chain Z whose transition kernel is M. Then we have $\lim_{m\to\infty}D_m=0$. To be convinced of this convergence, consider for fixed $m\in\mathbb{N}$, $(\widetilde{X}_n,\widetilde{Y}_n)_{n\in\mathbb{N}}$ and $(\widetilde{Z}_n)_{n\in\mathbb{N}}$ two chains coupled as in the beginning of this section and starting from the initial conditions $(\widetilde{X}_0,\widetilde{Y}_0)=(X_m,Y_m)$ and $\widetilde{Z}_0=X_m$. Then (27) and (25) imply that the quantity $\sum_{n\in\mathbb{N}}\mathbb{E}[|\widetilde{X}_n-\widetilde{Z}_n|]$ converges exponentially fast to zero as m goes to infinity.

□

7 Numerical computations

We present here the technical facts needed to conclude the proof of Proposition 11 (and by consequence also of Proposition 17). We used the free numerical computational package Scilab 4.1.2. Next we investigate the Lipschitz constants alluded to in Remark 18. Finally we illustrate the results obtained in this paper by giving a numerical approximation of the invariant measure μ of the chain Z and of rate L defined in (17).

But as a preliminary, let us give in Figure 1 the graphs of the six functions z_i, for $i \in [\![1, 6]\!]$, since they are at the heart of this paper.

Figure 1: Graphs of the z_i, for $i \in [\![1, 6]\!]$

7.1 Computations relative to Proposition 11

The goal is to show that the function F defined in (11) does not vanish in $[0, 1/2]$. Since we want to work formally as long as we can, it is better to consider the rational fraction (by pieces) R given in (12). Here are the codes we wrote and the results given by the computer.

- On the interval $(2/7, 1/2)$.

First we compute the matrix $A \coloneqq (\sqrt{6}G(i, z_j(x))\sqrt{6}G(j, x))_{i,j\in[\![1,6]\!]}$, where x is a "free variable". We multiplied the function G by $\sqrt{6}$ to avoid square roots in the computations.

```
x=poly(0,'x');
z1=3/2*x/(1+x);
z2=(2-4*x)/(2-x);
z3=(2-4*x)/(3-3*x);
z4=(1+x)/(4-2*x);z5=(1-2*x)/(4-2*x);z6=(1-2*x)/3;
z=[z1 z2 z3 z4 z5 z6];
g1=2*(1+x);g2=(2-x);g3=3*(1-x);
g4=2*(2-x);g5=g4;g6=3;
g=[g1;g2;g3;g4;g5;g6];
a=horner(g,z);
A=a*diag(g)
```

Running this code, we get the following polynonial matrix:

$$
A \quad := \quad \begin{pmatrix}
4+10x & 8-10x & 10-14x & 10-2x & 10-8x & 8-4x \\
4+x & 2+2x & 4-2x & 7-5x & 7-2x & 5+2x \\
6-3x & 9x & 3+3x & 9-9x & 9 & 6+6x \\
8+2x & 4+4x & 8-4x & 14-10x & 14-4x & 10+4x \\
8+2x & 4+4x & 8-4x & 14-10x & 14-4x & 10+4x \\
6+6x & 6-3x & 9-9x & 12-6x & 12-6x & 9
\end{pmatrix}
$$

By construction, we have

$$
\begin{aligned}
R(x) \quad &= \quad \frac{1}{6^{36}} \prod_{i,j \in [\![1,6]\!]} A_{i,j}(x) \\
&= \quad \frac{2^{29}3^{17}}{6^{36}} \prod_{i,j \in [\![1,6]\!]} B_{i,j}(x)
\end{aligned}
$$

where the matrix B is defined by

$$
B \quad := \quad \begin{pmatrix}
2+5x & 4-5x & 5-7x & 5-x & 5-4x & 2-x \\
4+x & 1+x & 2-x & 7-5x & 7-2x & 5+2x \\
2-x & x & 1+x & 1-x & 1 & 1+x \\
4+x & 1+x & 2-x & 7-5x & 7-2x & 5+2x \\
4+x & 1+x & 2-x & 7-5x & 7-2x & 5+2x \\
1+x & 2-x & 1-x & 2-x & 2-x & 1
\end{pmatrix}
$$

thus we obtain that $R(x) - 1 = P_1(x)$ on $[2/7, 1/2]$, with

$$
\begin{aligned}
P_1(x) \quad := \quad & 2^{-7}3^{-19}(2+5x)(4-5x)(5-7x)(5-x)(5-4x)(2-x)^8(4+x)^3(1+x)^6 \\
& (7-5x)^3(7-2x)^3(5+2x)^3 x(1-x)^2 - 1
\end{aligned}
$$

Using the command `roots` of Scilab 4.1.2 which computes numerically the roots of a given polynomial function, we obtained the roots of $P_1(x)$ depicted as crosses in Figure 2.

Figure 2: Roots of $P_1(x)$

We see that none of these roots is in the segment $[2/7, 1/2]$ (also drawn as the thick horizontal segment in the middle of the picture). Indeed looking more closely at the roots found by the computer, the one on the left (respectively the right) of the segment is 0.0006183 (resp. 0.6297289). The fact that $R > 1$ on $[2/7, 1/2]$ comes from the computation of $R(1/2) \approx 12.989284$. We also estimate that $R(2/7) \approx 99.311045$.

- On the interval $(1/5, 2/7)$.

The first above code has only to be slightly modified: the definition of `z2` (line 4) is now

```
z2=(3*x)/(2-x);
```

It appears that the corresponding matrix A is also polynomial and the computer gives us

$$A = \begin{pmatrix} 4+10x & 4+4x & 10-14x & 10-2x & 10-8x & 8-4x \\ 4+x & 4-5x & 4-2x & 7-5x & 7-2x & 5+2x \\ 6-3x & 6-12x & 3+3x & 9-9x & 9 & 6+6x \\ 8+2x & 8-10x & 8-4x & 14-10x & 14-4x & 10+4x \\ 8+2x & 8-10x & 8-4x & 14-10x & 14-4x & 10+4x \\ 6+6x & 6-3x & 9-9x & 12-6x & 12-6x & 9 \end{pmatrix}$$

Then we compute the roots of the polynomial function $P_2(x)$ representing $R(x) - 1$ on $[1/5, 2/7]$ and as it can be seen in Figure 3, none of them belongs to this segment (drawn as thick line, the closest root is 0.4569663).

Figure 3: Roots of $P_2(x)$

Thus R is strictly positive on $[1/5, 2/7]$, since we already knew that $R(2/7) > 0$. We estimate that $R(1/5) \approx 418.66239$.

- On the interval $(0, 1/5)$.

The additional modification is the definition of `z3` (line 5),

```
z3=(1+x)/(3-3*x);
```

Again we obtain a polynomial matrix A,

$$A = \begin{pmatrix} 4+10x & 4+4x & 8-4x & 10-2x & 10-8x & 8-4x \\ 4+x & 4-5x & 5-7x & 7-5x & 7-2x & 5+2x \\ 6-3x & 6-12x & 6-12x & 9-9x & 9 & 6+6x \\ 8+2x & 8-10x & 10-14x & 14-10x & 14-4x & 10+4x \\ 8+2x & 8-10x & 10-14x & 14-10x & 14-4x & 10+4x \\ 6+6x & 6-3x & 9-9x & 12-6x & 12-6x & 9 \end{pmatrix}$$

and Figure 4 gives a picture of the roots of $P_3(x)$ representing $R(x)-1$ on $[0,1/5]$. None of them belongs to this segment (the closest one is 0.3995404).

Figure 4: Roots of $P_3(x)$

Thus R is also strictly positive on $[0,1/5]$, since we have already seen that $R(1/5)>0$. We estimate that $R(0)\approx 13496.561$.

In conclusion, the function F defined in (11) is positive on $[0,1/2]$, as it is illustrated in Figure 5.

There is a more analytical approach to justify the positivity of F. Indeed, the above polynomial expressions for the three matrices A (one on each of the segments $[0,1/5]$, $[1/5,2/7]$ and $[2/7,1/2]$) enable to get a simple uniform bound C on F' on $\in (0,1/2)\setminus\{1/5,2/7\}$ via the formula

$$\forall\, x\in(0,1/2)\setminus\{1/5,2/7\},\qquad \left|F'(x)\right| = \frac{1}{36}\left|\sum_{i,j\in[\![1,6]\!]}\frac{A'_{i,j}(x)}{A_{i,j}(x)}\right| \leq \max_{i,j\in[\![1,6]\!]}\left|\frac{A'_{i,j}(x)}{A_{i,j}(x)}\right|$$

Working out the details, we get that we can take $C=14/3$ (obtained for $i=3$, $j=2$ and $x=2/7$ for the matrix A valid on the interval $[1/5,2/7]$). Thus if there exists $N\in\mathbb{N}^*$ such that

$$\min_{i\in[\![0,N]\!]} F(i/N) - 14/(6N) > 0$$

we can conclude to the positivity of F on $[0,1/2]$. We tried with $N=\lceil 14/(3F(1/2))\rceil = 66$ and it works (except that we have not taken into account the preciseness of the computer).

Figure 5: Graph of function F

The fact that F and R are decreasing on [0,1/2] can also be checked numerically. Indeed, using the command `derivat` of Scilab 4.1.2, we can compute first the derivatives of the polynomial functions P_1, P_2 and P_3 and next their roots. It appears that they do not belong respectively to the segments $[2/7, 1/2]$, $[1/5, 2/7]$ and $[0, 1/5]$. Since we have already seen that $F(0) > F(1/5) > F(2/7) > F(1/2)$, it follows that F is decreasing on $[0, 1/2]$ and thus bounded below by $F(1/2)$.

As it was mentioned to us by André Galligo, there is also a totally algebraic approach to this problem via the use of Sturm's sequences. Since we are dealing with polynomial functions with rational coefficients, this method can be made rigorous, if we are allowed to work on the computer with numbers with a large number of digits (here of order $42 = \lceil \log_{10}(36!) \rceil$, since we have to consider the values of all the derivatives of the polynomial function of interest at the boundary of the interval where we are looking for roots), in order to deal only with true "integers". So this approach would require to work with another mathematical software, such as Maple.

Finally let us also draw the graph of the function

$$\widetilde{F} \, : \, [0, 1/2] \ni x \quad \mapsto \quad \frac{1}{6} \sum_{i \in [\![1,6]\!]} \ln(G(i, x))$$

to show that it was necessary to iterate the kernel M to get a lower bound as the one in Proposition 11 (or to be able to apply the result of Barnsley and Elton [1]).

Indeed, $\widetilde{F}$ is negative in a neighborhood of 1/2, which means that the isoperimetric value of the triangles near this neighborhood have the tendency to decrease. But fortunately, the abscisse of the characteristic point of new triangle obtained by the barycentric subdivision has a chance of 4 over 6 (see Figure 1) to be close to zero, zone where the isoperimetric functional has a strong tendency to increase.

7.2 On the Lipschitz constants of Remark 18

We implement on the computer the verification of the criterion (24) from Diaconis and Freedman [4] in our particular situation. We adpot an analytical approach, despite the fact we are essentially

Figure 6: Graph of function $\widetilde{F}$

dealing with polynomial functions. Let us denote (recall the notation of (9))

$$\forall\, i \in [\![1,6]\!],\, \forall\, x \in [0,1/2], \qquad D(i,x) \;\coloneqq\; \left|z_i'\right|(x) \;=\; \frac{1}{G^2(i,x)}$$

and for any $n \in \mathbb{N}^*$ and $(i,j) \in [\![1,6]\!]^2$,

$$K_{i,j}(n) \;\coloneqq\; \sup_{x \in S_n} D(i, z_j(x)) D(j,x)$$

where $S_n \coloneqq [\![0,n]\!]/(2n)$ is a mesh of $[0,1/2]$. Then for any $(i,j) \in [\![1,6]\!]^2$, the Lipschitz constant of $z_i \circ z_j$ is given by

$$\begin{aligned} K_{i,j} \;&\coloneqq\; \sup_{x \in [0,1/2]} D(i, z_j(x)) D(j,x) \\ &=\; \sup_{n \in \mathbb{N}^*} K_{i,j}(n) \end{aligned}$$

Thus to show that criterion (24) is not fulfilled by M^2, it is sufficient to check that for some $n \in \mathbb{N}^*$, $\mathcal{K}_2(n) \coloneqq \prod_{(i,j)\in[\![1,6]\!]^2} K_{i,j}(n) > 1$.
Similarly for M^3, we define for any $n \in \mathbb{N}^*$ and any $(i,j,k) \in [\![1,6]\!]^3$,

$$K_{i,j,k}(n) \;\coloneqq\; \sup_{x \in S_n} D(i, z_j \circ z_k(x)) D(j, z_k(x)) D(k,x)$$

so that for any $(i,j,k) \in [\![1,6]\!]^3$, the Lipschitz constant of $z_i \circ z_j \circ z_k$ is given by $K_{i,j,k} = \sup_{n\in\mathbb{N}} K_{i,j,k}(n)$. We compute that

$$\begin{aligned} \max_{i\in[\![1,6]\!], x\in[0,1/2]} D(i,x) \;&=\; \frac{8}{3} \\ \max_{i\in[\![1,6]\!], x\in[0,1/2]} \left|\frac{dD(i,x)}{dx}\right| \;&=\; \frac{32}{3} \end{aligned}$$

so we easily deduce that

$$\max_{(i,j,k)\in[\![1,6]\!]^3, x\in[0,1/2]} \left|\frac{d}{dx} D(i, z_j \circ z_k(x)) D(j, z_k(x)) D(k,x)\right| \;\leq\; \frac{195552}{243} \;\leq\; 805$$

It follows that

$$\forall\, n \in \mathbb{N}^*,\, \forall\, (i,j,k) \in [\![1,6]\!]^3, \qquad K_{i,j,k} \;\; \leq \;\; K_{i,j,k}(n) + \frac{202}{n}$$

Thus to see that criterion (24) is satisfied by M^3, it is sufficient to check that for some $n \in \mathbb{N}^*$,

$$\mathcal{K}_3(n) \;\; \coloneqq \prod_{(i,j,k)\in[\![1,6]\!]^3} \left(K_{i,j,k}(n) + \frac{202}{n}\right) < 1$$

The following program (where the index set $[\![1,6]\!]$ has been replaced by the more convenient set $[\![0,5]\!]$) gives that $\mathcal{K}_2(1) =$ `k2` ≈ 92.067225 and $\mathcal{K}_3(10000) =$ `k3` ≈ 0.0116774, values which justify the assertions made in Remark 18.

```
function z=Z(x,i)
select i
case 0 then z=3/2*x/(1+x);
case 1 then z=3*bool2s((x<2/7))*x/(2-x)+bool2s((x>=2/7))*(2-4*x)/(2-x);
case 2 then z=bool2s((x<1/5)).*(1+x)./(3-3*x)+bool2s((x>=1/5)).*(2-4*x)./(3-3*x);
case 3 then z=(1+x)/(4-2*x);
case 4 then z=(1-2*x)/(4-2*x);
case 5 then z=(1-2*x)/3;
end; endfunction

function d=D(x,i)
select i
case 0 then d=3/2 ./(1+x).^2;
case 1 then d=6 ./(2-x).^2;
case 2 then d=2/3 ./(1-x).^2;
case 3 then d=3/2 ./(2-x).^2;
case 4 then d=3/2 ./(2-x).^2;
case 5 then d=2/3
end; endfunction

function d=DD(x,i)
j=floor(i/6); k=i-6*j;
d=D(Z(x,k),j)*D(x,j);
endfunction

function d=DDD(x,i)
j=floor(i/36); k=floor((i-36*j)/6); l=i-36*j-6*k ;
d=D(Z(Z(x,l),k),j)*D(Z(x,l),k)*D(x,l)+202/10000;
endfunction

r2=[0 1/2]; I2=0:35; K2=feval(r2,I2,DDD);
k2=prod(max(K2,'r'))

r3=0:1/(2*10000):.5; I3=0:215; K3=feval(r3,I3,DDD);
k3=prod(max(K3,'r'))
```

7.3 Approximation of the invariant measure μ

Proposition 17 leads to an approximation of the invariant measure μ through the strong law of large numbers. In practice, we run the following code, which simulates the chain $(Z_n)_{n\in[\![1,N]\!]}$ with

$N = 100000$ and starting from an initial variable Z_1 chosen uniformly over $[0, 1/2]$.

```
n=100000;
z(1)=.5*rand(); s(1)=0;
for i=2:n
j=floor(6*rand());
z(i)=Z(z(i-1),j);
s(i)=s(i-1)-log(D(z(i-1),j))/2;
end
L=s(n)/n
```

The program takes into account the functions `Z` and `D` defined in the code of the previous subsection. Figure 7 shows the approximation we got of μ through a histogram using 100 bars.

Figure 7: An approximation of μ

So we have the impression that μ is absolutely continuous, despite the fact that it is difficult to judge of that property from such histograms.

Finally, the code provides an approximation of the rate L defined in (17). Running several times the program, we always got values between 0.07 and 0.08, suggesting that Theorem 1 should be satisfied with $\chi = 0.07$.

Aknowledgments:

We are indepted to David Blackwell and Curtis T. McMullen for kindly providing us their manuscripts [2, 5] in preparation on the subject. We are grateful to the ANR's Chaire d'Excellence program for providing financial support and to the Laboratoire J. A. Dieudonné (UMR 6621) in Nice and to the Institut de Mathématiques (UMR 5219) in Toulouse, where parts of this work were done, for their hospitality.

References

[1] Michael F. Barnsley and John H. Elton. A new class of Markov processes for image encoding. *Adv. in Appl. Probab.*, 20(1):14–32, 1988.

[2] David Blackwell. Barycentric subdivision. Article in preparation, private communication, 2008.

[3] Claude Dellacherie and Paul-André Meyer. *Probabilités et potentiel. Chapitres V à VIII*, volume 1385 of *Actualités Scientifiques et Industrielles [Current Scientific and Industrial Topics]*. Hermann, Paris, revised edition, 1980. Théorie des martingales. [Martingale theory].

[4] Persi Diaconis and David Freedman. Iterated random functions. *SIAM Rev.*, 41(1):45–76 (electronic), 1999.

[5] Persi Diaconis and Curtis T. McMullen. Barycentric subdivision. Article in preparation, 2008.

[6] Lester E. Dubins and David A. Freedman. Invariant probabilities for certain Markov processes. *Ann. Math. Statist.*, 37:837–848, 1966.

[7] Robert Osserman. The isoperimetric inequality. *Bull. Amer. Math. Soc.*, 84(6):1182–1238, 1978.

[8] William F. Stout. *Almost sure convergence.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Probability and Mathematical Statistics, Vol. 24.

†diaconis@math.stanford.edu

Department of Statistics
Sequoia Hall
390 Serra Mall
Stanford University
Stanford, CA 94305-4065

‡miclo@latp.univ-mrs.fr

Laboratoire d'Analyse, Topologie, Probabilités
Centre de Mathématiques et Informatique
Université de Provence
39, rue Frédéric Joliot-Curie
13453 Marseille Cedex 13, France