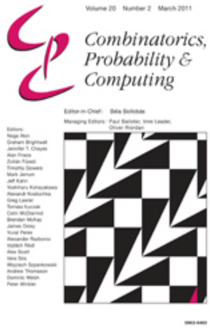

On Barycentric Subdivision

PERSI DIACONIS and LAURENT MICLO

Combinatorics, Probability and Computing / Volume 20 / Issue 02 / March 2011, pp 213 - 237
DOI: 10.1017/S0963548310000441, Published online: 24 November 2010

Combinatorics, Probability and Computing (2011) 20, 213–237. © Cambridge University Press 2010
doi:10.1017/S0963548310000441

On Barycentric Subdivision

PERSI DIACONIS[1] and LAURENT MICLO[2]

[1]Departments of Statistics and Mathematics, Stanford University, USA
(e-mail: diaconis@math.stanford.edu)

[2]Institut de Mathématiques de Toulouse, UMR 5219, Université de Toulouse and CNRS, France
(e-mail: miclo@math.univ-toulouse.fr)

Received 2 March 2009; revised 20 July 2010; first published online 24 November 2010

We dedicate this paper to the memory of David Blackwell

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 go to infinity with time. Nevertheless, if the triangles are renormalized through a similitude to have their longest edge equal 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]. The stationary distribution of this limit chain is particularly important in our study.

1. Introduction

Let $\triangle$ be a given triangle in the plane (to avoid triviality the vertices will always be assumed not to be all the same). The three medians of $\triangle$ intersect at the barycentre: this cuts it into six small triangles, say $\triangle_1$, $\triangle_2$, $\triangle_3$, $\triangle_4$, $\triangle_5$, $\triangle_6$. Next, each $\triangle_i$, for $i \in [\![1, 6]\!]$ (which denotes the set $\{1, 2, \ldots, 6\}$), can itself be subdivided 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 well known numerically (we learned it from Blackwell [3]; see also the survey by Butler and Graham [4]) and it has been proved (see Bárány, Beardon and Carne [1], Diaconis and McMullen [7] and Hough [10]) that as the barycentric subdivision goes on, most of the triangles become flat. The original motivation for this kind of result was to show that the barycentric subdivision is not a good procedure for constructing nice triangularizations of surfaces. For more information on other kinds of triangle subdivisions, we refer to a recent manuscript of Butler and Graham [4]. The goal of this paper is to propose a probabilistic approach to this phenomenon.

First, we adopt a Markovian point of view. Let $\triangle(0) := \triangle$, and throw a fair die to choose $\triangle(1)$ among the six triangles $\triangle_i$, $i \in [\![1,6]\!]$. Continuing in the same way, we get a Markov chain $(\triangle(n))_{n\in\mathbb{N}}$: if the first n triangles have been constructed, the next one is obtained by choosing uniformly (and independently from what was done before) one of the six triangles of the barycentric subdivision of the last-obtained triangle. Of course, at any time $n \in \mathbb{N}^*$ ($\mathbb{N}^*$ stands for $\mathbb{N}\setminus\{0\}$), 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 our results more analytically, 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 longest (respectively the shortest) 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 considered, otherwise we would have to consider positive similitudes and x would have to belong to $[0,1]$). Any time we are 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 implicitly be enforced throughout this 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 $\triangle(n)$. The first result justifies the assertion that as the barycentric subdivision goes on, the triangles become flat.

Theorem 1.1. *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.*

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

It can be shown that this is true with $\chi = 0.035$ (but this is not the best constant: numerical experiments from [8] suggest that the above bound should hold with $\chi \approx 0.07$); nevertheless the previous result remains asymptotical. Contrary to Blackwell [3] (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 previously proved by Bárány, Beardon and Carne [1], Diaconis and McMullen [7] (using Furstenberg's theorem on products of random matrices in $SL_2(\mathbb{R})$) and Hough [10], who used dynamical systems arguments (via an identification with a random walk on $SL_2(\mathbb{R})$).

There is a stronger notion of convergence to flatness that asks for the triangles to have an angle which is almost equal to π. With the preceding notation, 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 1.2. *The sequence* $(\triangle(n))_{n\in\mathbb{N}}$ *becomes strongly flat in probability:*

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

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: there are isosceles triangles that become flatter and flatter, but their maximum angle converges to $\pi/2$. Indeed, Theorem 1.2 is more difficult to obtain than Theorem 1.1 because $(X_n)_{n\in\mathbb{N}}$ does not converge, as the following result shows. Define the limit set of this sequence as the the intersection over $p \in \mathbb{N}$ of the closures of the sets $\{X_n \,:\, n \geqslant p\}$.

Theorem 1.3. *Almost surely, the limit set of* $(X_n)_{n\in\mathbb{N}}$ *is* $[0, 1/2]$.

It follows from Theorem 1.1 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 limiting 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. 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, $Z := (Z_n)_{n\in\mathbb{N}}$ 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 asymptotic behaviour. 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 do not know if μ is absolutely continuous).

The plan of the paper is as follows. Section 2 contains some preliminaries; in particular we will show, by studying the evolution of a convenient variant of the isoperimetric value, that the triangle Markov chain returns as close as we want to the set of flat triangles infinitely often. This is a first step in the direction of Theorem 1.1. In Section 3, we begin our investigation of the limiting Markov chain Z, to obtain some information valid in a neighbourhood of the set of flat triangles. Then in Section 4 we put together the previous global and local results to prove Theorem 1.1. Ergodicity and the attracting invariant measure μ of the Markov chain Z are studied in Section 5, using results of Dubins and Freedman [9], Barnsley and Elton [2] and Diaconis and Freedman [6] on iterated random functions. This will lead to the proofs of Theorem 1.2 and Theorem 1.3 in Section 6. Corresponding numerical experiments can be found in the appendix of [8], which is an extended version of this paper.

2. A weak result on attraction to flatness

The purpose of this section is to give some preliminary information and bounds on the triangle Markov chain obtained by barycentric subdivisions. By themselves, these results are not sufficient to conclude the a.s. convergence toward the set of flat triangles, but at least they give a heuristic hint for this behaviour: a quantity comparable to the isoperimetry value of the triangle has a tendency to increase after the barycentric subdivision and so to diverge to infinity with time, in the mean.

To measure the separation between a given triangle $\triangle$ and the set of flat triangles, we use the quantity $J(\triangle)$, which is the sum of the squares of the lengths of the edges divided by the area (this is well defined in $(0,+\infty]$, since the vertices are assumed not to be all the same). We have $J(\triangle)=+\infty$ if and only if $\triangle$ is flat. Furthermore, the functional J is invariant under similitude, so it depends only on the characteristic point (x,y) of $\triangle$, and we have

$$J(\triangle)=2\frac{1+x^2+y^2+(1-x)^2+y^2}{y}=4\frac{x^2+y^2-x+1}{y},$$

and under the restriction $x^2+y^2\leqslant(1-x)^2+y^2\leqslant 1$, we get $3\leqslant J(\triangle)y\leqslant 6$, namely

$$\frac{y}{6}\leqslant(J(\triangle))^{-1}\leqslant\frac{y}{3}, \tag{2.1}$$

so that the convergence of y to zero is equivalent to the divergence of $J(\triangle)$ to $+\infty$. Note that $J(\triangle)$ is comparable with the isoperimetric value $I(\triangle)$ of $\triangle$, defined as the square of the perimeter of $\triangle$ divided by its area of $\triangle$:

$$\frac{1}{3}I(\triangle)\leqslant J(\triangle)\leqslant I(\triangle). \tag{2.2}$$

With the notation of the Introduction, for $n\in\mathbb{N}$, write $J_n:=J(\triangle(n))$. Our first goal is to show the following result.

Proposition 2.1. *Almost surely, we have* $\limsup_{n\to\infty}J_n=+\infty$.

The proof will be based on elementary considerations of one step of the barycentric subdivision. Consider $\triangle$, a triangle in the normalized form given in the Introduction. For simplicity, we let A, B and C denote 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 barycentre of $\triangle$. We index the small triangles obtained by the barycentric subdivision as

$$\begin{aligned}\triangle_1&:=\{A,D,G\}, &\triangle_2&:=\{D,B,G\}, &\triangle_3&:=\{B,E,G\},\\ \triangle_4&:=\{E,C,G\}, &\triangle_5&:=\{C,F,G\}, &\triangle_6&:=\{F,A,G\}.\end{aligned} \tag{2.3}$$

It is well known that all the triangles $\triangle_i$, for $i\in[\![1,6]\!]$, have the same area (one straightforward way to see it is to note that this property is invariant by affine transformations and to consider the equilateral case).

Next define, with $|\cdot|$ denoting the length,

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

An immediate computation gives that

$$l_1^2=\frac{x^2}{4}+\frac{y^2}{4}-x+1,\quad l_2^2=\frac{x^2}{4}+\frac{y^2}{4}+\frac{x}{2}+\frac{1}{4},\quad l_3^2=x^2+y^2-x+\frac{1}{4}, \tag{2.4}$$

and since we also have

$$L_1^2 + L_2^2 + L_3^2 = 2(x^2 + y^2 - x + 1), \tag{2.5}$$

we get that

$$\frac{l_1^2 + l_2^2 + l_3^2}{L_1^2 + L_2^2 + L_3^2} = \frac{3}{4}.$$

These ingredients imply the following probabilistic statement.

Lemma 2.2. *For any $n \in \mathbb{N}$, we have*

$$\mathbb{E}[J_{n+1}|\mathcal{T}_n] = \frac{4}{3}J_n,$$

where the left-hand side is a conditional expectation with respect to $\mathcal{T}_n$, the σ-algebra generated by $\triangle(n), \triangle(n-1), \ldots, \triangle(0)$.

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

$$\mathbb{E}[J_{n+1}|\triangle(n)] = \frac{4}{3}J_n.$$

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 notation introduced above. Because the small triangles have the same area and the barycentre cuts the median segments into a ratio (1/3,2/3), we get that

$$\begin{aligned}\mathbb{E}[J(\triangle(1))|\triangle(0) = \triangle] &= \frac{1}{6}\sum_{i\in[\![1,6]\!]} J(\triangle_i)\\ &= \frac{1}{6}\left(\frac{L_1^2}{2} + \frac{L_2^2}{2} + \frac{L_3^2}{2} + \frac{10}{9}(l_1^2 + l_2^2 + l_3^2)\right)\frac{6}{\mathcal{A}(\triangle)}\\ &= \frac{4}{3}J(\triangle),\end{aligned}$$

where $\mathcal{A}(\triangle)$ is the area of $\triangle$. □

In general the previous submartingale information is not enough to deduce a.s. convergence. Taking expectations, for any $n \in \mathbb{N}$, $\mathbb{E}[J_{n+1}] \geqslant (4/3)\mathbb{E}[J_n]$, thus $\mathbb{E}[J_n] \geqslant (4/3)^n J(\triangle)$, so we can just deduce $\mathbb{L}^1$-divergence of J_n for large $n \in \mathbb{N}$, but this is not a very useful result.

Proof of Proposition 2.1. Note that the numbers J_n, $n \in \mathbb{N}$, are uniformly bounded below by a positive constant. This is a consequence of the usual isoperimetric inequality (see, for instance, Osserman [11]), or more simply, we can get directly from (2.1) that

$$\forall\ n \in \mathbb{N}, \quad J_n \geqslant 2\sqrt{3}.$$

But from Lemma 2.2 we see that

$$\forall\ n \in \mathbb{N}, \quad \mathbb{P}[J_{n+1} \geqslant (4/3)J_n|\mathcal{T}_n] \geqslant \frac{1}{6},$$

and consequently

$$\forall\ n, m \in \mathbb{N}, \quad \mathbb{P}[J_{n+m} \geqslant (4/3)^m 2\sqrt{3}|\mathcal{T}_n] \geqslant \frac{1}{6^m}. \tag{2.6}$$

Let $R > 1$ be an arbitrary large number and consider $m \in \mathbb{N}^*$ such that $(4/3)^m 2\sqrt{3} \geqslant R$. The $\{0,1\}$-valued sequence $(\mathbb{1}_{J_{m(n+1)} \geqslant R})_{n \in \mathbb{N}}$ stochastically dominates a sequence of independent Bernoulli variables of parameter $1/6^m$. It follows that a.s. we have

$$\limsup_{n \to \infty} J_n \geqslant R,$$

and since R can be chosen arbitrarily large, Proposition 2.1 is proved. □

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

Lemma 2.3. *There exist two constants* $0 < a < b < +\infty$ *such that*

$$\forall\ n \in \mathbb{N}, \quad aJ_n \leqslant J_{n+1} \leqslant bJ_n.$$

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, with the above notation,

$$\forall\ i \in [\![1, 6]\!], \quad aJ(\triangle) \leqslant J(\triangle_i) \leqslant bJ(\triangle).$$

Such inequalities are obvious for flat triangles, so assume that $\triangle$ is not flat. Since the areas are easy to compare, we just need to consider the diameters (whose squares are comparable, within the range [1,3], with the sums of the squares of the lengths of the edges), denoted by d. We have clearly $d(\triangle) = 1$ and $d(\triangle_i) \leqslant 1$ for $i \in [\![1, 6]\!]$. The reverse bound $d(\triangle_i) \geqslant 1/4$, for $i \in [\![1, 6]\!]$, is a consequence of the equalities $|[A, F]| = |[F, C]| = 1/2$, $|[B, E]| = |[E, C]| \geqslant 1/2$ and $|[D, G]| = |[G, C]|/2 \geqslant 1/4$. □

3. Near the limit flat Markov chain

Our goal here is twofold. First we show that the kernel of the triangle Markov chain converges nicely to the kernel of the flat triangle Markov chain as the triangle becomes 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 neighbourhood 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 (2.3). 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

$$\forall\ x \in [0, 1/2], \quad z_i(x) := x_i(x, 0). \tag{3.1}$$

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

$$\forall\ x \in [0, 1/2], \quad M(x, \cdot\,) = \frac{1}{6} \sum_{i \in [\![1,6]\!]} \delta_{z_i(x)}. \tag{3.2}$$

The next result gives bounds on the discrepancy between Q and M as the triangles become flat.

Lemma 3.1. *There exists a constant* $K > 0$ *such that*

$$\forall\ i \in [\![1,6]\!], \forall\ (x, y) \in \mathcal{D}, \quad \max(|x_i(x, y) - z_i(x)|, |y_i(x, y)|) \leqslant Ky.$$

Proof. We first check that for any fixed $i \in [\![1,6]\!]$, the map

$$(x, y^2) \mapsto (x_i(x, y), y_i^2(x, y)) \tag{3.3}$$

is (uniformly) Lipschitz on $\mathcal{D}_2$, where $\mathcal{D}_2$ is the image of $\mathcal{D}$ under $(x, y) \mapsto (x, y^2)$.

Indeed, denote by $0 \leqslant L_{i,1} \leqslant L_{i,2} \leqslant L_{i,3}$ the ordered lengths of the triangle $\triangle_i$. Applying (2.5) to this triangle, it appears that

$$y_i^2 = \frac{L_{i,1}^2 + L_{i,2}^2 + L_{i,3}^2}{2L_{i,3}^2} - x_i^2 + x_i - 1. \tag{3.4}$$

Let h_i be the height of $\triangle_i$ orthogonal to the edge of length $L_{i,3}$; we have $L_{i,1}^2 = h_i^2 + (x_i L_{i,3})^2$ and $L_{i,2}^2 = h_i^2 + ((1 - x_i)L_{i,3})^2$. It follows that

$$x_i = \frac{L_{i,3}^2 - L_{i,2}^2 + L_{i,1}^2}{2L_{i,3}^2}. \tag{3.5}$$

Finally, notice that (2.4) implies that the mappings $(x, y^2) \mapsto L_{i,j}^2$, for $j \in [\![1,3]\!]$, are uniformly Lipschitz on $\mathcal{D}_2$. Furthermore, as seen in the proof of Lemma 2.3, on $\mathcal{D}_2$, the mapping $(x, y^2) \mapsto L_{i,3}^2$ is bounded below by $1/16$, so (3.5) and (3.4) imply that the mapping described in (3.3) is uniformly Lipschitz.

The bounds given in Lemma 3.1 are an easy consequence of this Lipschitz property and of the boundedness of $\mathcal{D}$. □

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. Here we define I_n as the isoperimetric value of $\triangle(n)$. This amounts to evaluating the sign of $1/6 \sum_{i \in [\![1,6]\!]} \ln(I(\triangle_i)/I(\triangle))$, by the Markov property. Of course, the previous ratios are not rigorously defined if the triangle $\triangle$ is 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 same ratio 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),

$$
\begin{aligned}
G(1, x) &= \sqrt{\frac{2}{3}}(1+x), \quad & G(2, x) &= \sqrt{\frac{1}{6}}(2-x), \\
G(3, x) &= \sqrt{\frac{3}{2}}(1-x), \quad & G(4, x) &= \sqrt{\frac{2}{3}}(2-x), \\
G(5, x) &= \sqrt{\frac{2}{3}}(2-x), \quad & G(6, x) &= \sqrt{\frac{3}{2}}.
\end{aligned}
\tag{3.6}
$$

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

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

So, to prove that $\mathbb{E}[\ln(I_{n+1}/I_n) | \triangle(n) = \triangle] > 0$ for nearly flat triangles $\triangle$, it would suffice to show that the mapping $x \mapsto \sum_{i \in [\![1,6]\!]} \ln(G(i, x))$ only takes positive values on $[0, 1/2]$. Unfortunately, this is not true, since it takes negative values in a neighbourhood of $1/2$ (see the appendix of [8]). To get around this problem, we iterate the barycentric subdivision one more step.

Proposition 3.2. *There exist a constant $\gamma > 0$ and a neighbourhood $\mathcal{N}$ of the set of the flat triangles, such that*

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

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

Proof. Coming back to the notation at the beginning of the Introduction, we want to find $\mathcal{N}$ and γ as above and satisfying

$$\forall\, \triangle \in \mathcal{N}, \quad \frac{1}{36} \sum_{i,j \in [\![1,6]\!]} \ln\left(\frac{6\mathcal{P}(\triangle_{i,j})}{\mathcal{P}(\triangle)} \right) \geqslant \gamma. \tag{3.7}$$

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

$$F(x) := \frac{1}{36} \sum_{i,j \in [\![1,6]\!]} \ln(G(j, z_i(x)) G(i, x)), \tag{3.8}$$

where the $z_i(x)$, for $i \in [\![1, 6]\!]$, were defined in (3.1). More explicitly, we will compute in Section 5 (see Lemma 5.1) 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 piecewise rational fraction

$$
\begin{aligned}
R(x) &:= \exp(36F(x)) \\
&= \prod_{i,j \in [\![1,6]\!]} G(i, z_j(x)) G(j, x).
\end{aligned}
\tag{3.9}
$$

After computations (see the appendix of [8]), it appears that this is indeed a piecewise polynomial function. By numerically studying the zeros of $R-1$ of the three underlying polynomial functions, we show that F does not vanish on $[0,1/2]$. So, by continuity, we get that $\gamma := \min_{[0,1/2]} F/2 > 0$. Then, using the above uniform convergence, we can find a neighbourhood $\mathcal{N}$ of the set of flat triangles so that (3.7) is fulfilled. □

In the appendix of [8], it is checked 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 a position to prove Theorem 1.1. The principle behind the proof is that there is a neighbourhood $\mathcal{N}'$ 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 neighbourhood and then to converge exponentially fast to the set of flat triangles. This event will eventually occur, since triangle Markov chains always return to $\mathcal{N}'$.

In order to see that the triangle Markov chain has a positive chance of remaining trapped in a neighbourhood of the set of flat triangles, we will use a general martingale argument. To do so, we introduce some notation. 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 $R \geqslant R_0 > 0$, 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 $n \in \mathbb{N}$,

$$|N_{n+1}^{(R)} - N_n^{(R)}| \leqslant A, \tag{4.1}$$

$$V_{n+1}^{(R)} - V_n^{(R)} \geqslant \gamma + N_{n+1}^{(R)} - N_n^{(R)}. \tag{4.2}$$

The next result shows that if R is large enough, with high probability $V^{(R)}$ will never go below $R/2$. This is classical, but without a precise reference at hand, we recall the underlying arguments.

Lemma 4.1. *We have*

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

and furthermore, a.s.,

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

Proof. The first estimate is an immediate consequence of the Hoeffding–Azuma inequality, which, applied to the bounded difference martingale $(-N_n^{(R)})_{n\in\mathbb{N}}$ starting from 0, asserts that, for any $t \in \mathbb{R}_+$,

$$\forall\ n \in \mathbb{N}^*, \quad \mathbb{P}[-N_n^{(R)} > t] \leqslant \exp(-t^2/(2nA^2)).$$

In particular, since for any $n \in \mathbb{N}$ we have

$$V_n^{(R)} \geqslant R + n\gamma + N_n^{(R)}, \tag{4.3}$$

we get

$$\begin{aligned}\mathbb{P}[V_n^{(R)} < R/2] &\leqslant \mathbb{P}[-N_n^{(R)} > R/2 + n\gamma] \\ &\leqslant \exp\left(-\frac{R}{4nA^2} - \frac{R\gamma}{2A^2} - \frac{n\gamma^2}{2A^2}\right) \\ &\leqslant \exp\left(-\frac{R\gamma}{2A^2}\right)\exp\left(-\frac{n\gamma^2}{2A^2}\right),\end{aligned}$$

and the first stated bound follows by summation over $n \in \mathbb{N}^*$.

The second bound is also due to the fact that the increments of the martingale $N^{(R)}$ are bounded, which implies the validity of the iterated logarithm law (see, for instance, Stout [12]): almost surely,

$$\limsup_{n\to\infty} \frac{|N_n^{(R)}|}{\sqrt{n\ln(\ln(n))}} \leqslant A.$$

Thus (4.3) enables us to conclude. □

Lemma 4.1 will be applied with $V^{(R)}$, the logarithm of isoperimetric values, or rather with a sequence of the kind $(\ln(I_{2n}))_{n\in\mathbb{N}}$.

More precisely, consider the neighbourhood $\mathcal{N}$ obtained in Proposition 3.2. There exists a small constant $\epsilon > 0$ such that $\mathcal{N}$ contains $\{(x, y) \in \mathcal{D} : 0 \leqslant y < \epsilon\}$ and so taking into account (2.1), there exists $R_1 > 1$ such that $\{\triangle : \ln(I(\triangle)) > R_1\} \subset \mathcal{N}$ (again we are slightly abusing notation 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 $R := \ln(I(\triangle(T)))$ satisfies $R \geqslant 2R_1$. Define a stopping time τ for the shifted chain $(\triangle(T + 2n))_{n\in\mathbb{N}}$ by

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

which is infinite if the set on the right-hand side is empty. Let $\gamma > 0$ be the constant appearing in Proposition 3.2. We construct a stochastic chain $V^{(R)}$ in the following way:

$$\forall\, n \in \mathbb{N}, \quad V_n^{(R)} := \begin{cases} \ln(I(\triangle(T + 2n))) & \text{if } n \leqslant \tau, \\ \ln(I(\triangle(T + 2\tau))) + \gamma(n - \tau) & \text{otherwise.} \end{cases}$$

Let us check that the assumptions for Lemma 4.1 are satisfied. Following the traditional Doob–Meyer semi-martingale decomposition (see, for instance, Dellacherie and Meyer [5]), we define

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

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 notation, this is the σ-algebra $\mathcal{T}_{T+n}$, where the filtration $(\mathcal{T}_m)_{m\in\mathbb{N}}$ was introduced in Lemma 2.2. After conditioning on $\mathcal{F}_0$ and taking advantage of the strong Markov property, we can apply Lemma 2.3 to see that (4.1) is satisfied with $A = (b/a)^2$ (we even have $N_{n+1}^{(R)} - N_n^{(R)} = 0$ for $n \geqslant \tau$). Furthermore, we have for any $n \in \mathbb{N}$

$$\begin{aligned}
V_{n+1}^{(R)} - V_n^{(R)} &= \mathbb{E}[V_{n+1}^{(R)}|\mathcal{F}_n] - V_n^{(R)} + V_{n+1}^{(R)} - \mathbb{E}[V_{n+1}^{(R)}|\mathcal{F}_n] \\
&= \mathbb{E}[V_{n+1}^{(R)} - V_n^{(R)}|\mathcal{F}_n] + N_{n+1}^{(R)} - N_n^{(R)} \\
&= \mathbb{E}[\ln(I_{T+2(n+1)}/I_{T+2n})|\triangle(T+2n)]\mathbb{1}_{n\leqslant\tau} + \gamma\mathbb{1}_{n>\tau} + N_{n+1}^{(R)} - N_n^{(R)} \\
&\geqslant \gamma + N_{n+1}^{(R)} - N_n^{(R)},
\end{aligned}$$

where the last inequality comes from Proposition 3.2. Then Lemma 4.1 implies the following result.

Proposition 4.2. *Let $\mathcal{N}' := \{\triangle \,:\, \ln(I(\triangle)) > R_1\}$. There exists a large enough constant $R_2 \geqslant 2R_1$ such that, for any finite stopping time T for the triangle Markov chain $(\triangle(n))_{n\in\mathbb{N}}$ satisfying $\ln(I(\triangle(T))) \geqslant R_2$, 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.

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

Indeed, Lemma 4.1 shows that we can find $R_2 \geqslant 2R_1$ such that

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

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

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

Lemma 2.3 permits extending these results to the statement of Proposition 4.2 (up to replacement of R_2 by bR_2/a).

Now the proof of Theorem 1.1 is clear. By iteration, introduce two sequences $(S_n)_{n\in\mathbb{N}}$ and $(T_n)_{n\in\mathbb{N}}$ of stopping times for the triangle Markov chain: start with $S_0 = 0$ and for any $n \in \mathbb{N}$, if S_n has been defined, take

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

Of course, if for some $n \in \mathbb{N}$, $S_n = \infty$, then for any $m \geqslant n$, $S_m = T_m = \infty$. Conversely, via Proposition 2.1, 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

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

Up to conditioning on $\{S_n < \infty\}$, Lemma 4.2 shows that

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

thus it follows easily that $\mathbb{P}[\cup_{n\in\mathbb{N}} E_n] = 1$. Lemma 4.2 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. Now the bound (2.1) implies the validity of Theorem 1.1 with $\chi = \gamma/2$.

Remark 4.3. Let $\gamma_2 := F(1/2) = \min_{x\in[0,1/2]} F(x)$. A closer look at the proof of Proposition 3.2 shows that, for any $\gamma < \gamma_2$, we can find a neighbourhood $\mathcal{N}$ of the set of flat triangles such that the lower bound of Proposition 3.2 is satisfied. By the above arguments, it follows that Theorem 1.1 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

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

where $(\iota_n)_{n\in\mathbb{N}^*}$ 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 ($Z_0 := x$) constructed from $(\iota_n)_{n\in\mathbb{N}^*}$ through the relations

$$\forall\ n \in \mathbb{N}, \quad Z_{n+1} := z_{\iota_{n+1}}(Z_n). \tag{4.4}$$

Then define

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

An easy extension of the previous proof shows that Theorem 1.1 holds with $\chi = \gamma_N/N$ and consequently with $\chi = \lim_{N\to\infty} \gamma_N/N$. The quantity γ_N/N converges due to the weak convergence of the Markov chain $(Z_n)_{n\in\mathbb{N}}$, uniformly in its initial distribution, as we will show in the next section. Indeed, if μ is the attracting invariant probability associated to $(Z_n)_{n\in\mathbb{N}}$, we will see that for any $x \in [0, 1/2]$,

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

with

$$L := \frac{1}{6} \sum_{i\in[\![1,6]\!]} \int \ln(G(i, x))\, \mu(dx). \tag{4.5}$$

It follows from Cesaro's lemma that

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

Since this convergence holds uniformly in $x \in [0, 1/2]$, we get that Theorem 1.1 is satisfied with $\chi = L$. In the appendix of [8], we get a numerical evaluation of $L \approx 0.07$.

5. Ergodicity of the limit flat Markov chain

This section studies the limit flat Markov chain $Z := (Z_n)_{n\in\mathbb{N}}$. First we will see that it admits a unique invariant probability μ and that it converges exponentially fast to μ 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 the kernel of Z given in (3.2) in the language of iterated random functions.

Lemma 5.1. *With the notation of the previous sections, we have, for all $x \in [0, 1/2]$,*

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

Proof. These are immediate computations, based on the fact that for any flat triangle, the abscissa of the characteristic point is the ratio of the shortest edge by the longest edge. For instance, the lengths of the edges of the triangle $\triangle_2$ are $L_1/2$, $l_1/3$ and $2l_3/3$ with $L_1 = x$, $l_1 = 1 - \frac{x}{2}$ and $l_3 = \frac{1}{2} - x$, which leads to the above expression for $z_2(x)$. □

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 [2]. 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

$$\forall\, x \in S, \quad N(x, \cdot) := \sum_{i\in[\![1,n]\!]} p_i \delta_{f_i(x)}. \tag{5.1}$$

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

$$\forall\, x \neq y \in S, \quad \sum_{i\in[\![1,n]\!]} p_i \ln\left(\frac{|f_i(y) - f_i(x)|}{|y - x|}\right) \leqslant r, \tag{5.2}$$

the kernel N is ergodic: it admits a unique invariant and attracting probability μ, 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 [2] show that there exists $q \in (0, 1]$ and $\rho \in (0, 1)$ such that

$$\forall\, x, y \in S, \quad \sum_{i\in[\![1,n]\!]} p_i |f_i(y) - f_i(x)|^q \leqslant \rho |y - x|^q. \tag{5.3}$$

Let us rewrite this bound in a more probabilistic way. Let $(\iota_n)_{n\in\mathbb{N}}$ 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 follows: $U_0^x = x$

and for any $n \in \mathbb{N}$, $U_{n+1}^x = f_{\iota_{n+1}}(U_n^x)$. This 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 as

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

and admits an immediate extension:

$$\forall\ n \in \mathbb{N}, \forall\ x, y \in S, \quad \mathbb{E}[|U_n^y - U_n^x|^q] \leqslant \rho^n |y - x|^q.$$

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 than (or equal to) 1. We can now show the following result.

Lemma 5.2. *Under the above assumption* (5.2)*, we have for any* $n \in \mathbb{N}$ *and any* $x \in S$,

$$D(N^n(x, \cdot), \mu) \leqslant \operatorname{diam}(S) \rho^n,$$

where ρ *and* q *are as in* (5.3) *and* $\operatorname{diam}(S)$ *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.*

$$\lim_{N \to \infty} \frac{1}{N+1} \sum_{n \in [\![0, N]\!]} f(U_n^x) = \mu[f]. \tag{5.4}$$

Proof. Let $f \in \mathcal{L}(1)$. We compute that

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

The stated 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

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

since the right-hand side 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 a solution to the Poisson equation

$$\begin{cases} \forall\ x \in S, \quad \varphi(x) - N(x, \varphi) = f(x) - \mu[f], \\ \qquad\qquad\qquad\qquad \mu[\varphi] = 0. \end{cases}$$

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

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

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. It is also true for all Lipschitz functions f. Next, given a continuous function f on S and $m \in \mathbb{N}^*$, by usual approximations, it is possible to find a Lipschitz function $\widetilde{f}_m$ on S such that $\|f - \widetilde{f}_m\|_{S,\infty} \leqslant 1/m$, where $\|\cdot\|_{S,\infty}$ is the uniform norm on S. It follows that on a measurable set Ω_m of probability 1,

$$\limsup_{N\to\infty} \frac{1}{N+1} \sum_{n\in[\![0,N]\!]} f(U_n^x) \leqslant \mu[f] + 2/m,$$

$$\liminf_{N\to\infty} \frac{1}{N+1} \sum_{n\in[\![0,N]\!]} f(U_n^x) \geqslant \mu[f] - 2/m.$$

Thus on the set $\cap_{m\in\mathbb{N}^*}\Omega_m$ of full probability, (5.4) is true. □

Let us discuss condition (5.2). Note that since the functions f_i, for $i \in [\![1, n]\!]$, are Lipschitz, they are absolutely continuous. Let us write f_i' for their respective weak derivatives. By letting y and x become close in criterion (5.2), we get that almost everywhere in $x \in S$,

$$\sum_{i\in[\![1,n]\!]} p_i \ln(|f_i'(x)|) \leqslant r. \tag{5.5}$$

Condition (5.5) is not sufficient to ensure 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

$$\forall\ x \in [0, 1], \quad f_i(x) := \begin{cases} \min(2x, 1) & \text{if } i = 1, \\ \max(0, -1 + 2x) & \text{if } i = 2. \end{cases}$$

In this case (5.5) 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 (5.5) is also sufficient. This is 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 $x \in S$,

$$|f_i'(x)| = (a_i x + b_i)^{-2} \tag{5.6}$$

(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 $x < y \in S$,

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

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

$$\ln\left(\frac{|f_i(y) - f_i(x)|}{|y-x|}\right) \leqslant \frac{\ln(|f_i'(y)|) + \ln(|f_i'(x)|)}{2}, \tag{5.7}$$

and consequently

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

a formula which enables passing from (5.5) to (5.2). It only has to be checked for a.e. $x, y \in S$.

It is now time to come back to the flat triangle Markov chain. 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 $[\![1,6]\!]$. Now condition (5.6) is satisfied. Since $z_2'(2/7-) = -z_2'(2/7+)$ and $z_3'(1/5-) = -z_3'(1/5+)$, we see that $|z_i'|(x)$ can be defined everywhere (a convention that we will adopt from now on). Indeed, we compute that, for any $x \in [0, 1/2]$,

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

Unfortunately (5.5) is not true and surprisingly it is a computation we have already encountered: comparing with (3.6), we see that

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

and thus, by the observation before Proposition 3.2, 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 the kernel M one more time (this trick was also used by Barnsley and Elton in one example of their paper [2]). 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 $[\![1,6]\!]^2$. The advantage is that we have for any $i, j \in [\![1,6]\!]$ and any $x \in [0,1/2]$,

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

Thus

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

and in particular the left-hand side is negative due to Proposition 3.2 (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 (5.6) is no longer satisfied by the functions $f_{i,j}$. To avoid this problem, we come back directly to the bound (5.7): for $i, j \in [\![1,6]\!]$ and $y > x \in [0,1/2]$, we write

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

In this situation we can also come back from (5.6) to (5.5) and the results of Barnsley and Elton [2] ensure that the iterated Markov kernel M^2 is ergodic. To come back from M^2 to M is not difficult, as is shown in the following result.

Proposition 5.3. *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 in $\{\mu, \mu M\} = \{\mu\}$, 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 $(Z_{1+2n})_{n\in\mathbb{N}}$. □

Let us mention that there exists a cruder way to deduce the ergodicity of M.

Remark 5.4. Diaconis and Freedman [6] consider a simpler criterion for ergodicity of a random function Markov kernel (5.1): for $i \in [\![1,n]\!]$, let $K_i := \sup_{x\neq y} |f_i(y) - f_i(x)|/|y-x|$ be the Lipschitz constant of f_i and assume that there exists a constant $r < 0$ such that

$$\forall\ x,y \in S, \qquad \sum_{i\in[\![1,n]\!]} p_i \ln(K_i) \leqslant r. \tag{5.8}$$

Then the kernel N is ergodic, and Diaconis and Freedman [6] show 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 (5.8) implies (5.2). Since, for $i \in [\![1,n]\!]$, K_i is the essential supremum of $|f_i'(x)|$, (5.8) corresponds to the exchange of essential supremum and sum in (5.2). Let us now come back to our flat triangle Markov chain. From the previous considerations, (5.8) cannot be satisfied with $N = M$. It does not work with $N = M^2$ either, so this is an example where the criterion (5.2) is fulfilled while (5.8) is not. But condition (5.8) 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 computations, we refer to the appendix of [8].

To finish we prove two properties of μ which will be needed in the following section.

Lemma 5.5. *The probability μ contains no atom, in particular $\mu(\{0\}) = 0$.*

Proof. The proof needs a few steps and notation. Let us define

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

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

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

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

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 consider the point $1/2$. We get

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

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

Step 2. 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 the appendix of [8]), we get that

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

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

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

it appears that equality is possible only if $\mu(\{y\}) = \mu^*$ for all $y \in S(x)$, namely $S(x) \subset S^*$.

We now study 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

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

  and since we know that $\mu(\{0\}) \leqslant \mu^*/2$, 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 $\tilde{z}_2$ 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 its inverse by $\tilde{z}_2^{-1}$.

Step 3. For $x \in (0, 1/2]$, the set $\{\tilde{z}_2^{-n}(x)\ :\ n \in \mathbb{N}\}$ 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 in 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 of 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 [9]. The second piece of information we need about μ is a direct consequence of a result of the latter paper.

Lemma 5.6. *The support of μ is the whole segment* $[0, 1/2]$.

Proof. By Theorem 4.9 of Dubins and Freedman [9], the support of μ is the whole segment $[0, 1/2]$ if we can cover it with 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 behaviour

Our main goal here is to prove Theorems 1.2 and 1.3. 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 information we have on the chain $(Z_n)_{n\in\mathbb{N}}$.

A natural coupling between the above chains is based on the construction alluded to in Remark 4.3. Assume that (X_0, Y_0) and Z_0 are given and let $(\imath_n)_{n\in\mathbb{N}^*}$ 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 (4.4), and similarly we iteratively define $(X_n, Y_n)_{n\in\mathbb{N}}$ via

$$\forall\ n \in \mathbb{N}, \quad (X_{n+1}, Y_{n+1}) := (x_{\imath_{n+1}}(X_n), y_{\imath_{n+1}}(Y_n)).$$

In these relations, the indices refer to the conventions made in (2.3) and (3.1). A first simple property of this coupling is given in the following result.

Lemma 6.1. *The random variables $|X_n - Z_n|$ converge in probability to zero as n goes to infinity:*

$$\forall\ \epsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}[|X_n - Z_n| > \epsilon] = 0.$$

Proof. First we iterate Lemma 3.1 to show that with $K' := K^2 + 8K/3 > 0$ we have

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

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

$$\begin{aligned}
&|x_i(x_j(x,y), y_j(x,y)) - z_i(z_j(x))| \\
&\quad \leqslant |x_i(x_j(x,y), y_j(x,y)) - z_i(x_j(x,y))| + |z_i(x_j(x,y)) - z_i(z_j(x))| \\
&\quad \leqslant K|y_j(x,y)| + \frac{8}{3}|x_j(x,y) - z_j(x)| \\
&\quad \leqslant K^2 y + \frac{8K}{3} y.
\end{aligned}$$

Let $q \in (0,1]$ and $\rho \in (0,1)$ as in (5.3) but relative to the kernel $N = M^2$. Then, for any $n \in \mathbb{N}$ we can write

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

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

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

After integration, the above bound leads to

$$\forall\ n \in \mathbb{N}, \quad a_{n+2} \leqslant \rho a_n + b_n.$$

We deduce that

$$\forall\ n \in \mathbb{N}, \quad a_{2n} \leqslant a_0 \rho^n + \sum_{m \in [\![0, n-1]\!]} b_{2m} \rho^{n-1-m}, \tag{6.1}$$

where $\lim_{n\to\infty} a_{2n} = 0$ 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.*, in the end we will be assured of

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

and thus of the claimed 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 1.2 follows quite easily.

Proof of Theorem 1.2. For $n \in \mathbb{N}$, let A'_n (resp. A''_n) be the angle between $[(0,0),(X_n,Y_n)]$ and $[(X_n,Y_n),(X_n,0)]$ (resp. $[(X_n,0),(X_n,Y_n)]$ and $[(X_n,Y_n),(1,0)]$), so that $A_n = A'_n + A''_n$. Since the length of $[(X_n,0),(1,0)]$ 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 $n \in \mathbb{N}$

$$\begin{aligned}\mathbb{P}[Y_n/X_n \geqslant \epsilon] &= \mathbb{P}[Y_n/X_n \geqslant \epsilon, X_n > 2\eta] + \mathbb{P}[Y_n/X_n \geqslant \epsilon, X_n \leqslant 2\eta] \\ &\leqslant \mathbb{P}[Y_n \geqslant 2\epsilon\eta] + \mathbb{P}[X_n \leqslant 2\eta] \\ &\leqslant \mathbb{P}[Y_n \geqslant 2\epsilon\eta] + \mathbb{P}[|X_n - Z_n| \geqslant \eta] + \mathbb{P}[Z_n \leqslant \eta]. \end{aligned} \tag{6.2}$$

By letting n go to infinity, taking into account that the stationary distribution μ of Z is continuous, we get

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{P}[Y_n/X_n \geqslant \epsilon] &\leqslant \lim_{n\to\infty} \mathbb{P}[Z_n \leqslant \eta] \\ &= \mu([0,\eta]),\end{aligned}$$

because as n goes to infinity, the law of Z_n converges weakly to μ and this probability gives weight 0 to the boundary $\{\eta\}$ of $(-\infty, \eta]$. Using again Lemma 5.5 and letting η go to zero, we obtain that $\lim_{n\to\infty} \mathbb{P}[Y_n/X_n \geqslant \epsilon] = 0$, and consequently the stated convergence in probability. □

Remark 6.2. We do not know if $(A_n)_{n\in\mathbb{N}}$ converges to π a.s. One way to deduce this result, via the Borel–Cantelli lemma, would be to show that, for any given $\epsilon > 0$,

$$\sum_{n\in\mathbb{N}} \mathbb{P}[Y_n/X_n \geqslant \epsilon] < +\infty. \tag{6.3}$$

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 want 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 the appendix of [8] would suggest that this property holds with $\alpha = 1$ (and $\lim_{\eta\to 0_+} \mu([0,\eta])/\eta \leqslant 1$).

In order to prove Theorem 1.3, we need two technical results. In all that follows, let us fix some $a \in [0, 1/2]$ and $\epsilon > 0$ and define $\mathcal{O} := [a-\epsilon, a+\epsilon] \cap [0, 1/2]$.

Lemma 6.3. *There exist* $\eta > 0$ *and* $N \in \mathbb{N}^*$ *such that*

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

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

Proof. This is a consequence of Lemma 5.2 applied to M^2: there exists $\rho \in (0,1)$ such that, for any $z \in [0,1/2]$ and $n \in \mathbb{N}$, $D(M^n(z,\cdot),\mu) \leqslant \rho^{\lfloor n/2\rfloor}/2$. Let φ be the function vanishing outside $(a-\epsilon, a+\epsilon)$, affine on $[a-\epsilon, a]$ and $[a, a+\epsilon]$ such that $\varphi(a) = \epsilon$. By definition of D, we have

$$\forall\, z \in [0,1/2],\ \forall\, n \in \mathbb{N}, \quad |M^n(z,\varphi) - \mu[\varphi]| \leqslant \frac{\rho^{\lfloor n/2\rfloor}}{2}.$$

Since the support of μ is [0,1/2], we have $\eta := \mu[\varphi] > 0$. So if we choose $N \in \mathbb{N}$ such that $\rho^{\lfloor N/2 \rfloor} < \eta$, we get that for any $z \in [0, 1/2]$, $\mathbb{P}_z[X_N \in \mathcal{O}] \geqslant M^N(z, \varphi) > \eta/2$. □

For the second technical result, we need further notation: for $(x, y) \in \mathcal{D}$ and $(i_1, i_2, \ldots, i_N) \in [\![1, 6]\!]^N$, we let $x_{i_1,i_2,\ldots,i_N}(x, y)$, $y_{i_1,i_2,\ldots,i_N}(x, y)$ and $z_{i_1,i_2,\ldots,i_N}(x)$ denote the values of X_n, Y_n and Z_n when $(X_0, Y_0, Z_0) = (x, y, z)$ and $(\imath_1, \imath_2, \ldots, \imath_N) = (i_1, i_2, \ldots, i_N)$.

Lemma 6.4. *There exists a constant K_N such that*

$$\forall\ (i_1, i_2, \ldots, i_N) \in [\![1, 6]\!]^N,\ \forall\ (x, y) \in \mathcal{D}, \quad \begin{cases} |x_{i_1,i_2,\ldots,i_N}(x, y) - z_{i_1,i_2,\ldots,i_N}(x)| \leqslant K_N y, \\ |y_{i_1,i_2,\ldots,i_N}(x, y)| \leqslant K_N y. \end{cases}$$

Proof. For $N = 1$ this is just Lemma 3.1. The general case is proved by an easy iteration, similar to the one already used in the proof of Lemma 6.1, starting with

$$\begin{aligned} &|x_{i_1,i_2,\ldots,i_N}(x, y) - z_{i_1,i_2,\ldots,i_N}(x)| \\ &\quad \leqslant |x_{i_N}(x_{i_1,i_2,\ldots,i_{N-1}}(x, y), y_{i_1,i_2,\ldots,i_{N-1}}(x, y) - z_{i_N}(x_{i_1,i_2,\ldots,i_{N-1}}(x, y))| \\ &\qquad + |z_{i_N}(x_{i_1,i_2,\ldots,i_{N-1}}(x, y)) - z_{i_N}(z_{i_1,i_2,\ldots,i_{N-1}}(x))|. \end{aligned}$$ □

We can now come to our last task.

Proof of Theorem 1.3. Let $\mathcal{O}' := [a - 2\epsilon, a + 2\epsilon] \cap [0, 1/2]$. We want to show an analogous result to Lemma 6.3 but for the chain $(X_n, Y_n)_{n\in\mathbb{N}}$, namely to find $\eta' > 0$ and $N' \in \mathbb{N}^*$ such that

$$\inf_{(x,y)\in\mathcal{D}} \mathbb{P}_{(x,y)}[X_{N'} \in \mathcal{O}'] \geqslant \eta' \tag{6.4}$$

(let us recall that under $\mathbb{P}_{(x,y)}$, $(X_0, Y_0) = (x, y)$). To do so, we first consider η and N as in Lemma 6.3 and consider $\delta > 0$ sufficiently small such that $K_N \delta^{1/2^{N-1}} < \epsilon$. Then, according to Lemmas 6.3 and 6.4, we have

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

To extend this estimate to the whole domain $\mathcal{D}$, we come back to (2.1) and (2.6), which enables us to find $N'' \in \mathbb{N}$ such that

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

Now the Markov property implies (6.4) with $\eta = \eta'\eta''$ and $N' = N + N''$.

Since this bound is uniform over $(x, y) \in \mathcal{D}$, the sequence $(\mathbb{1}_{\mathcal{O}'}(X_{nN'}))_{n\in\mathbb{N}}$ is stochastically bounded below by an independent family of Bernoulli variables of parameter η', and we deduce that a.s.

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

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

The details of the above proof are necessary because in general one cannot deduce from the convergence in probability of $|X_n - Z_n|$ to zero as n goes to infinity the a.s. equality of the limit sets of $(X_n)_{n\in\mathbb{N}}$ and of $(Z_n)_{n\in\mathbb{N}}$. This property instead requires the a.s. convergence of $|X_n - Z_n|$ to zero, and this leads to the following observations.

Remark 6.5. Coming back to Remark 6.2, to prove (6.3) via (6.2), we are also missing an estimate of the kind

$$\exists\, K, p, \chi > 0 \;:\; \forall\, n \in \mathbb{N}, \quad \mathbb{E}[Y_n^p] \leqslant K \exp(-\chi n). \tag{6.5}$$

Blackwell [3] succeeded in obtaining such a bound (with $p = 1/2$) by exhibiting an appropriate supermartingale with the help of the computer; see also the survey by Butler and Graham [4]. His result can be seen to imply Theorem 1.1, with $\chi = 0.04$.

Furthermore, it allows for a more direct proof of Theorem 1.3. Indeed, if (6.5) is satisfied for some $p > 0$, then for any $q > 0$,

$$\sum_{n\in\mathbb{N}} \mathbb{E}[Y_n^q] < +\infty$$

(this is immediate for $q = p$, and use the Hölder inequality for $0 < q < p$ and the elementary bound $y^q \leqslant (\sqrt{3}/2)^{q-p} y^p$ for $y \in [0, \sqrt{3}/2]$ and $q > p$). The arguments of the proof of Lemma 6.1 (especially (6.1) and a similar relation for odd integers) then show that

$$\sum_{n\in\mathbb{N}} \mathbb{E}[|X_n - Z_n|^q] < +\infty,$$

and consequently that $|X_n - Z_n|$ converges a.s. to zero. This is sufficient to deduce that almost surely the limit set of $(X_n)_{n\in\mathbb{N}}$ coincides with that of $(Z_n)_{n\in\mathbb{N}}$, thus the law of large numbers for Z and Lemma 5.6 implies Theorem 1.3.

In the same spirit, one can go further toward justifying the assertion made in the Introduction that asymptotically $(X_n)_{n\in\mathbb{N}}$ is almost Markovian. Let us introduce the supremum distance S on $[0, 1/2]^{\mathbb{N}}$, seen as the set of trajectories from $\mathbb{N}$ to $[0, 1/2]$:

$$\forall\, x := (x_n)_{n\in\mathbb{N}},\ z := (z_n)_{n\in\mathbb{N}} \in [0, 1/2]^{\mathbb{N}}, \quad S(x, z) := \sup_{n\in\mathbb{N}} |x_n - z_n|.$$

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

$$s_m := \inf \mathbb{E}[S(X_{[\![m,\infty[\![}, Z)],$$

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 $\lim_{m\to\infty} s_m = 0$. 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 (6.5) and (6.1) 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.

Acknowledgements

We are indebted to David Blackwell, Curtis T. McMullen and Bob Hough for kindly providing us their manuscripts [3, 7, 10] in preparation on the subject. We acknowledge the generous support of the ANR's Chaire d'Excellence program of the CNRS. We are very grateful to an editor and to the referees, who helped us to improve the paper by simplifying its arguments. In particular, the suggestion to use the simpler functional J instead of I in Section 2 is due to one of them.

References

[1] Bárány, I., Beardon, A. F. and Carne, T. K. (1996) Barycentric subdivision of triangles and semigroups of Möbius maps. *Mathematika* **43** 165–171.
[2] Barnsley, M. F. and Elton, J. H. (1988) A new class of Markov processes for image encoding. *Adv. Appl. Probab.* **20** 14–32.
[3] Blackwell, D. (2008) Barycentric subdivision. Private communication.
[4] Butler, S. and Graham, R. (2010) Iterated triangle partitions. Preprint available at: `http://www.math.ucla.edu/~butler/PDF/iterated_triangles.pdf`.
[5] Dellacherie, C. and Meyer, P.-A. (1980) *Probabilités et Potentiel*, Chap. V–VIII, Vol. 1385 of *Actualités Scientifiques et Industrielles* (*Current Scientific and Industrial Topics*), revised edition, Hermann.
[6] Diaconis, P. and Freedman, D. (1999) Iterated random functions. *SIAM Rev.* **41** 45–76.
[7] Diaconis, P. and McMullen, C. T. (2010) Barycentric subdivision. Article in preparation.
[8] Diaconis, P. and Miclo, L. (2010) On barycentric partitions, with simulations. Preprint available at: `http://hal.archives-ouvertes.fr/` and `http://arxiv.org/`.
[9] Dubins, L. E. and Freedman, D. A. (1966) Invariant probabilities for certain Markov processes. *Ann. Math. Statist.* **37** 837–848.
[10] Hough, B. (2009) Tessellation of a triangle by repeated barycentric subdivision. *Electron. Commun. Probab.* **14** 270–277.
[11] Osserman, R. (1978) The isoperimetric inequality. *Bull. Amer. Math. Soc.* **84** 1182–1238.
[12] Stout, W. F. (1974) *Almost Sure Convergence*, Vol. 24 of *Probability and Mathematical Statistics*, Academic Press.