A SLOW TRIANGLE MAP WITH A SEGMENT OF INDIFFERENT FIXED POINTS AND A COMPLETE TREE OF RATIONAL PAIRS

CLAUDIO BONANNO, ALESSIO DEL VIGNA, AND SARA MUNDAY

Abstract. We study the two-dimensional continued fraction algorithm introduced in [6] and the associated *triangle map* T, defined on a triangle $\triangle \subseteq \mathbb{R}^2$. We introduce a slow version of the triangle map, the map S, which is ergodic with respect to the Lebesgue measure and preserves an infinite Lebesgue-absolutely continuous invariant measure. We discuss the properties that the two maps T and S share with the classical Gauss and Farey maps on the interval, including an analogue of the weak law of large numbers and of Khinchin's weak law for the digits of the triangle sequence, the expansion associated to T. Finally, we confirm the role of the map S as a two-dimensional version of the Farey map by introducing a complete tree of rational pairs, constructed using the inverse branches of S, in the same way as the Farey tree is generated by the Farey map, and then, equivalently, generated by a generalised mediant operation.

1. Introduction

The theory of (regular) continued fractions has received much attention from researchers in ergodic theory in the last decades, most recently thanks to the development of infinite ergodic theory ([1, 9, 10, 11]). For instance, the general results of ergodic theory have been applied to the Gauss and the Farey maps to obtain new proofs of the Gauss-Kuzmin Theorem, Khinchin's weak law and other metric results first obtained by Khinchin and Lévy.

One of the most notable results in the theory of continued fractions is Lagrange's Theorem, which states that a real number has an eventually periodic continued fraction expansion if and only if it is a quadratic irrational. In a letter to Jacobi, Hermite asked whether it was possible to obtain a similar classification for the algebraic irrationals of higher degree. It was for this reason that Jacobi developed what is now called the Jacobi-Perron algorithm, and the theory of multidimensional continued fractions began. Unfortunately, despite numerous attempts and the introduction of many different algorithms, Hermite's question remains unanswered. We refer the reader to [4] for a geometric description of the theory of multidimensional continued fractions and to [23] for some applications of ergodic theory in this area.

In this paper we consider the two-dimensional version of the continued fraction algorithm introduced in [6]. The algorithm, which we describe in Section 2.3, is based on the iteration of a map T defined on a triangle $\triangle \subseteq \mathbb{R}^2$, and for this reason, T is referred to as the *triangle map* and the expansions obtained through this method are called *triangle sequences*. The ergodic properties of T are studied in [16, 7]; in particular, it is shown that the map T is ergodic with respect to the Lebesgue measure on $\triangle$ and preserves a Lebesgue-absolutely continuous probability measure. The triangle map behaves similarly to the Gauss map in many ways, for instance, the triangle map acts on triangle sequences by left-shifting the digits, exactly as the Gauss map does for the regular continued fraction expansions.

The similarity between the two maps is strengthened by the results of this paper. We introduce a map S on the triangle $\triangle$, which plays for T the same role that the Farey map plays for the Gauss map. For this reason we call S a *slow triangle map*. From the point of view of ergodic theory, it is interesting to notice that the map S is a piecewise linear fractional map on a finite partition with a segment of indifferent fixed

We thank the referee for useful comments and for pointing out a mistake in the previous version. This research is part of the authors' activity within the DinAmicI community, see www.dinamici.org. The authors are partially supported by the research project PRA_2017_22 "Dynamical systems in analysis, geometry, mathematical logic and celestial mechanics" of the University of Pisa. C. Bonanno and S. Munday are partially supported by the Istituto Nazionale di Alta Matematica and its division Gruppo Nazionale di Fisica Matematica. C. Bonanno is partially supported also by the research project PRIN 2017S35EHN_004 "Regular and stochastic behaviour in dynamical systems" of the Italian Ministry of Education and Research.

points, that is, points for which the determinant of the Jacobian is 1, and that it is non-uniformly expanding elsewhere. We show that, similarly to the Farey map, S preserves an infinite Lebesgue-absolutely continuous measure and it is ergodic with respect to the Lebesgue measure on $\triangle$. It follows that the statistical behaviour of summable observables along orbits of S is non-standard. This phenomenon, for the Farey map, makes it impossible to improve Khinchin's weak law for the coefficients of the regular continued fraction expansion to a strong law. However, we are able to exploit certain results from infinite ergodic theory to show that the system generated by S is pointwise dual ergodic and, under a further assumption, prove a weak law of large numbers for S, from which we obtain an analogue of Khinchin's weak law for the digits of the triangle sequences.

The connection between the Gauss and the Farey maps and the regular continued fractions can be studied also through the Farey tree, a binary tree which contains all the rational numbers in $(0, 1)$ (see *e.g.* [5]). The Farey tree is strongly related to the Farey map, but it can also be defined through the mediant operation on fractions. We recall the definition of the Farey tree and its basic properties in Section 5. Analogously, in this paper we define a tree of rational pairs, first by using a suitable modification of the map S limited to the set of indifferent fixed points, and then by using a generalised mediant operation defined on pairs of rational numbers. We prove that the two trees are in fact identical level by level, and that the tree is complete, that is, it contains every pair of rational numbers in $\bar{\triangle}$. This last result improves on the results of [2], where the authors study different trees generated by the triangle map and its generalisations, but show that none of them are complete.

The paper is organised as follows. In Section 2 we recall the definition of the triangle map T and the associated two-dimensional continued fraction algorithm. We also introduce the map S and study its basic ergodic properties. Lastly, we define a dynamical system on an infinite strip, which is isomorphic to the action of S on $\bar{\triangle}$. This isomorphism gives a useful intuitive representation of the action of S and simplifies some computations. Section 3 contains the main results on the ergodic properties of S. We prove that the map S is pointwise dual ergodic with respect to a sequence $a_n(S) \asymp \frac{n}{\log^2 n}$, and use various results from Infinite Ergodic Theory (see [1, 11]) to show that if the sequence $a_n(S)$ is regularly varying (see (3.1)) then we have the weak law of large numbers for summable observables (Theorem 3.2), and a Khinchin-type weak law for the triangle sequences (Corollary 3.3). The technical results are proved in Appendix A and B. In Section 4 we apply a result from [15] to our map S. Recalling that the behaviour of Birkhoff sums of summable observables drastically changes in infinite ergodic theory, following [13], in [15] the authors give a Birkhoff Ergodic Theorem for non-summable observables for infinite-measure-preserving dynamical systems. We use the version of S defined on the strip and prove a pointwise convergence theorem for non-summable observables. Finally, in Section 5 we introduce the tree of rational pairs produced by the counterimages of $\tilde{S}$, a slightly modified version of S. In Theorem 5.4 we prove that the tree is complete and that each pair of rationals appears exactly once. Then we introduce an algorithm on the triangle $\triangle$, based on the notion of mediant of two fractions, and show in Theorem 5.10 that the tree can be generated also by this algorithm. This concludes the similarity between the slow triangle map S and the Farey map. For these reasons S may be considered a *two-dimensional Farey map*. Many interesting questions remain open about the tree and its connections with the map S and with the approximation of irrational pairs by rational pairs. These problems will be subject of future research.

2. The setting

As anticipated in the introduction, the main goal of this paper is to investigate a two-dimensional map related to the triangle map T, as introduced in [6]. Let us first recall the definition of the map $T : \triangle \to \bar{\triangle}$, where $\triangle$ is the triangle

$$\triangle := \left\{ (x, y) \in \mathbb{R}^2 \, : \, 1 \geq x \geq y > 0 \right\}.$$

Consider the countable partition $\{\triangle_k\}_{k\geq 0}$ of $\triangle$ into disjoint triangles

$$\triangle_k := \left\{ (x, y) \in \triangle \, : \, 1 - x - ky \geq 0 > 1 - x - (k + 1)y \right\},$$

shown in Figure 1, and the segment $\Lambda \coloneqq \{0 \le x \le 1,\, y = 0\}$. Note that $\bar{\triangle} = \bigcup_{k\ge 0} \triangle_k \cup \Lambda$. The triangle map $T : \triangle \to \bar{\triangle}$ is then defined to be

$$T(x,y) \coloneqq \left(\frac{y}{x}, \frac{1-x-ky}{x}\right) \quad \text{for } (x,y) \in \triangle_k.$$

We now define a map $S : \bar{\triangle} \to \bar{\triangle}$ that can be thought of as a "slow version" of the map T. Let us start with the partition $\{\Gamma_0, \Gamma_1\}$ of $\bar{\triangle}$ (see Figure 1), where

$$\Gamma_0 \coloneqq \triangle_0 = \left\{(x,y) \in \mathbb{R}^2 \,:\, 1 \ge x \ge y > 1-x\right\},$$

and

$$\Gamma_1 \coloneqq \bar{\triangle} \setminus \Gamma_0 = \bigcup_{k\ge 1} \triangle_k \cup \Lambda = \left\{(x,y) \in \mathbb{R}^2 \,:\, 1-y \ge x \ge y \ge 0\right\}.$$

We define $S : \bar{\triangle} \to \bar{\triangle}$ by setting

$$S(x,y) \coloneqq \begin{cases} \left(\frac{y}{x}, \frac{1-x}{x}\right) & \text{if } (x,y) \in \Gamma_0 \\ \left(\frac{x}{1-y}, \frac{y}{1-y}\right) & \text{if } (x,y) \in \Gamma_1 \end{cases}. \tag{2.1}$$

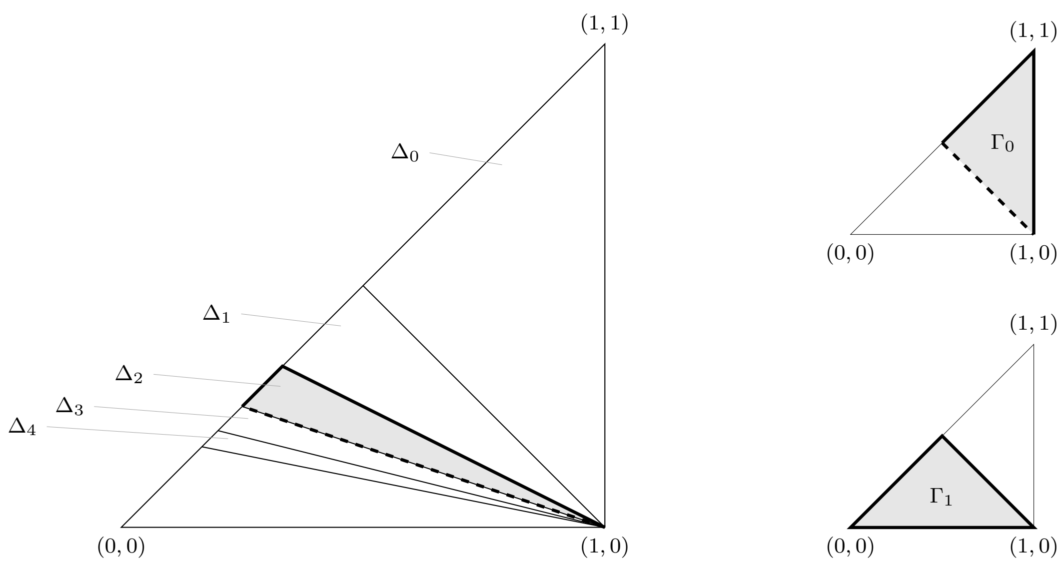

FIGURE 1. *Left.* Partition of $\triangle$ into $\{\triangle_k\}_{k\ge 0}$. *Right.* Partition of $\bar{\triangle}$ into Γ_0 and Γ_1.

The relation between these two maps is that the triangle map T is the jump transformation of S on the set Γ_0. In other words, if we introduce the first passage time function

$$\tau(x,y) \coloneqq 1 + \min\left\{k \ge 0 \,:\, S^k(x,y) \in \Gamma_0\right\},$$

then it can be readily calculated that $T(x,y) = S^{\tau(x,y)}(x,y)$ for each $(x,y) \in \triangle$. Notice that $S(\triangle_k) = \triangle_{k-1}$ for $k \ge 1$, and that $S(\Gamma_0) \cup \{x = y,\, 0 \le x \le 1\} = S(\Gamma_1) = \bar{\triangle}$. Moreover the segment Λ consists of fixed points, that is $S(x,0) = (x,0)$ for $0 \le x \le 1$. The determinant of the Jacobian of S turns out to be

$$JS(x,y) = \begin{cases} \frac{1}{x^3} & \text{if } (x,y) \in \Gamma_0 \\ \frac{1}{(1-y)^3} & \text{if } (x,y) \in \Gamma_1 \end{cases},$$

and it follows that $JS(x,0) = 1$ for $0 \le x \le 1$. Thus the segment Λ consists of indifferent fixed points.

2.1. **Invariant measure and the transfer operator.** In [16] it is shown that the map T is ergodic, and from [7] we know that the unique ergodic, Lebesgue-absolutely continuous T-invariant probability measure on $\triangle$ is given by the density

$$k(x,y) = \frac{12}{\pi^2 x(1+y)}.$$

Applying classical results from ergodic theory ([1, 11]), the existence of an ergodic, Lebesgue-absolutely continuous S-invariant measure immediately follows. One way to find the density $h(x,y)$ of this measure is to look for a fixed point of the transfer operator $\mathcal{P}$ associated to S. Let

$$\phi_0 \coloneqq (S|_{\Gamma_0})^{-1} : \bar{\triangle} \setminus \{x=y,\, 0 \le x \le 1\} \to \Gamma_0, \quad \phi_0(x,y) = \left(\frac{1}{1+y},\, \frac{x}{1+y} \right)$$

and

$$\phi_1 \coloneqq (S|_{\Gamma_1})^{-1} : \bar{\triangle} \to \Gamma_1, \quad \phi_1(x,y) = \left(\frac{x}{1+y},\, \frac{y}{1+y} \right)$$

be the local invers maps of S. The transfer operator $\mathcal{P}$ is then defined for each measurable function f on $\bar{\triangle}$ by setting

$$\begin{aligned}(\mathcal{P}f)(x,y) &= |J\phi_0(x,y)| f(\phi_0(x,y)) + |J\phi_1(x,y)| f(\phi_1(x,y)) = \\ &= \frac{1}{(1+y)^3} f\left(\frac{1}{1+y},\, \frac{x}{1+y} \right) + \frac{1}{(1+y)^3} f\left(\frac{x}{1+y},\, \frac{y}{1+y} \right).\end{aligned}$$

A straightforward computation shows that $\mathcal{P}h = h$ for $h(x,y) = \frac{1}{xy}$.

Proposition 2.1. *The system $(\bar{\triangle}, \mu, S)$ is conservative, and the map S admits a unique, up to multiplicative constants, ergodic invariant measure μ, absolutely continuous with respect to the Lebesgue measure m, given by the density $h(x,y) = \frac{1}{xy}$. The measure μ is σ-finite and $\mu(\bar{\triangle}) = +\infty$.*

Proof. For conservativity, in light of Maharam's Recurrence Theorem [11, Theorem 2.2.14], it is enough to observe that

$$\triangle = \bigcup_{n=0}^{\infty} S^{-n}(\Gamma_0) \pmod{\mu},$$

which is a consequence of the fact that $\triangle_k \subseteq S^{-k}(\Gamma_0)$, for each $k \ge 0$. We have already discussed the existence of the measure μ above. That μ is unique follows, for example, from [11, Theorem 2.4.35], on noting that S is conservative, ergodic, and is certainly non-singular with respect to m. Finally, that μ is σ-finite follows from computing the measure of the triangles $\triangle_k$ as in [7]. □

2.2. **An equivalent system on a strip.** For later use, we now introduce another system, isomorphic to the map S. Using the change of coordinates defined on $\triangle$ by

$$(x,y) \mapsto (u,v) \in \Sigma \coloneqq (0,1] \times [0,+\infty), \qquad u(x,y) = \frac{y}{x}, \quad v(x,y) = \frac{1-x}{y},$$

one can show that the system $(\triangle, \mu, S)$ is isomorphic mod μ to the system (Σ, ρ, F) given by

$$F(u,v) = \begin{cases} \left(v, \frac{1}{v}(\frac{1}{u} - 1)\right) & \text{if } (u,v) \in \Pi_0 \coloneqq \{(u,v) \in \Sigma \,:\, v < 1\} \\ (u, v-1) & \text{if } (u,v) \in \Pi_1 \coloneqq \{(u,v) \in \Sigma \,:\, v \ge 1\} \end{cases},$$

with $d\rho(u,v) = \frac{1}{1+uv} du\, dv$. The sets Π_0 and Π_1 partition the strip Σ and correspond mod μ to Γ_0 and Γ_1, respectively. We also introduce the countable partition $\{\Sigma_k\}_{k \ge 0}$, where

$$\Sigma_k \coloneqq \{(u,v) \in \Sigma \,:\, k \le v < k+1\}$$

is a unit squares, as shown in Figure 2. Note that $\Sigma = \bigcup_{k \ge 0} \Sigma_k$. This is the analogue of the partition $\{\triangle_k\}_{k \ge 0}$ of the triangle $\triangle$. The local inverses of the map F are given by

$$(F|_{\Pi_0})^{-1}(u,v) = \left(\frac{1}{uv+1}, u \right) \quad \text{and} \quad (F|_{\Pi_1})^{-1}(u,v) = (u, v+1),$$

so that the transfer operator $\mathcal{P}_F$ associated to F turns out to be

$$(\mathcal{P}_F g)(u,v) = \frac{u}{(uv+1)^2} g\left(\frac{1}{uv+1}, u\right) + g\left(u, v+1\right).$$

It can be immediately verified that the density of the measure ρ is a fixed point of $\mathcal{P}_F$.

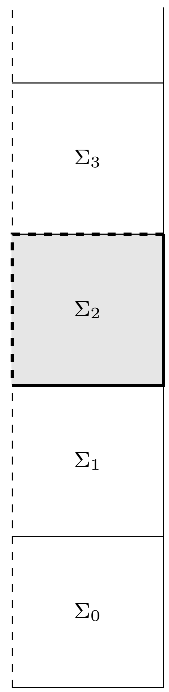

FIGURE 2. Partition of the strip Σ into $\{\Sigma_k\}_{k\geq 0}$.

2.3. **Triangle sequences.** Let us now recall the definition of the *triangle sequence* associated to a point (x,y) in $\triangle$ and certain results concerning their digits from [6]. We start by setting $d_{-2} := 1$, $d_{-1} := x$ and $d_0 := y$ and, supposing that $d_{k-3} > d_{k-2} > d_{k-1} > 0$, we recursively define $\alpha_k := \alpha_k(x,y)$ to be the non-negative integer such that

$$d_{k-3} - d_{k-2} - \alpha_k d_{k-1} \geq 0$$

and

$$d_{k-3} - d_{k-2} - (\alpha_k + 1) d_{k-1} < 0.$$

Then set $d_k := d_{k-3} - d_{k-2} - \alpha_k d_{k-1} \in \mathbb{R}^+$. If at any stage we find that $d_k = 0$, the process stops. We shall write $(x,y) = (\alpha_1, \alpha_2, \ldots)$ to denote the triangle sequence of (x,y). Another way of defining the triangle sequence is to note that $\alpha_k(x,y) = m$ if and only if $T^{k-1}(x,y) \in \triangle_m$, and the process stops if $T^n(x,y) \in \Lambda$ for some $n \geq 1$. From this way of looking at the triangle sequence digits, it immediately follows that if $(x,y) = (\alpha_1, \alpha_2, \ldots)$, then $T(x,y) = (\alpha_2, \alpha_3, \ldots)$. In other words, the triangle map acts on triangle sequences as the shift map, exactly as the Gauss map does for the continued fraction expansions. We also have the following relation between the digits α_k and the first passage time:

$$\tau(T^{k-1}(x,y)) = 1 + \alpha_k(x,y). \tag{2.2}$$

In [6], the following results for the triangle sequence are given.

- If (x,y) is a pair of rational numbers in $\mathbb{Q}^2 \cap \bar{\triangle}$, then the triangle sequence associated to (x,y) is finite. However, the converse is not true: non-rational points can also have finite triangle sequences.

- Every infinite sequence of non-negative integers $(\alpha_1, \alpha_2, \ldots)$ has a pair $(x, y) \in \triangle$ that has this sequence as its triangle sequence.
- If an integer k appears infinitely often in a given sequence of integers, there is a unique pair $(x, y) \in \triangle$ that has this sequence as its triangle sequence.

Note that there are entire line segments with every point having identical infinite triangle sequences. This is essentially due to the fact that the refinements of the partition $\{\triangle_k\}_{k\geq 0}$ with respect to the map T do not have diameters shrinking to 0. Thus, whilst the triangle sequence can usefully be thought of as a two-dimensional generalisation of the continued fraction expansion, in certain respects it behaves rather differently. However, this behaviour is not in contrast with the ergodicity of the map, since as shown in [16] for Lebesgue almost every point the refinements of the partition $\{\triangle_k\}_{k\geq 0}$, along the triangle sequence of the point, shrink to the point. In the language of multidimensional continued fraction expansions (see [4]), this means that the triangle sequence is *weakly convergent* at Lebesgue almost every point.

3. A weak law of large numbers

For dynamical systems with an infinite invariant measure, in general it is only possible to establish weaker statistical properties than those for systems with an invariant probability measure. For example, if (X, μ, R) is a conservative and ergodic measure-preserving system such that $\mu(X) = \infty$, then Birkhoff's Ergodic Theorem becomes the weak statement that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (f \circ R^k)(x) = 0,$$

for μ-almost every $x \in X$ and for all $f \in L^1(X, \mu)$. Moreover, the exact asymptotic pointwise behaviour cannot be recovered for all $f \in L^1(X, \mu)$ by changing the normalising sequence, due to Aaronson's Ergodic Theorem [1, Theorem 2.4.2], which basically states that for any sequence of positive real numbers, the growth rate of the Birkhoff sums will be either over- or under-estimated infinitely often. Nevertheless, it is possible to obtain distributional limit laws for the ergodic sums of some classes of dynamical systems with an infinite invariant measure (see [1, Chapter 3]).

A first step is to show that that the system $(\bar{\triangle}, \mu, S)$ is pointwise dual ergodic, which means that there exists a sequence $(a_n(S))_{n\geq 0}$ such that

$$\lim_{n\to\infty} \frac{1}{a_n(S)} \sum_{k=0}^{n-1} (\mathcal{P}^k f)(x, y) = \int_{\bar{\triangle}} f d\mu$$

for μ-almost every $(x, y) \in \bar{\triangle}$ and for all $f \in L^1(\bar{\triangle}, \mu)$, where $\mathcal{P}$ is the transfer operator of the system. We prove it for $(\bar{\triangle}, \mu, S)$.

Theorem 3.1. *The system $(\bar{\triangle}, \mu, S)$ is pointwise dual ergodic, and the sequence $(a_n(S))_{n\geq 0}$ satisfies*[1] $a_n(S) \asymp \frac{n}{\log^2 n}$.

Distributional limit laws follow from pointwise dual ergodicity under the assumption that the sequence $(a_n(S))_{n\geq 0}$ is regularly varying. We recall that a sequence $(a_n)_{n\geq 0}$ is said to be *regularly varying of index* $\alpha \in \mathbb{R}$ if for all $c > 0$ we have that

$$\lim_{n\to\infty} \frac{a_{\lfloor cn \rfloor}}{a_n} = c^\alpha. \tag{3.1}$$

If $\alpha = 0$ the sequence is called *slowly varying.*

Theorem 3.2 (Weak law of large numbers)**.** *Let* Prob *be a probability measure on $\bar{\triangle}$, absolutely continuous with respect to the Lebesgue measure. If the sequence $(a_n(S))_{n\geq 0}$ in Theorem 3.1 is regularly varying of index*

[1]We say that $a_n \asymp b_n$ if and only if $a_n = O(b_n)$ and $b_n = O(a_n)$.

$\alpha = 1$, then for all $f \in L^1(\bar{\triangle}, \mu)$ and for all $\varepsilon > 0$

$$\lim_{n\to\infty} \text{Prob}\left(\left| \frac{1}{a_n(S)} \sum_{k=0}^{n-1} (f \circ S^k)(x,y) - \int_{\bar{\triangle}} f d\mu \right| > \varepsilon \right) = 0.$$

Corollary 3.3 (Khinchin weak law)**.** *Let* Prob *be a probability measure on $\bar{\triangle}$, absolutely continuous with respect to the Lebesgue measure. If the sequence $(a_n(S))_{n\geq 0}$ in Theorem 3.1 is regularly varying of index $\alpha = 1$, then there exists a sequence $(b_n)_{n\geq 0}$ such that for all $\varepsilon > 0$*

$$\lim_{n\to\infty} \text{Prob}\left(\left| \frac{1}{b_n} \sum_{k=0}^{n-1} \alpha_k(x,y) - 1 \right| > \varepsilon \right) = 0,$$

and $b_n \asymp n \log^2 n$. In particular, for m-almost every $(x,y) \in \bar{\triangle}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \alpha_k(x,y) = +\infty.$$

Proof. Applying Theorem 3.2 to the function $f = \mathbf{1}_{\triangle_0}$, the proof follows from a standard duality argument between Birkhoff sums and the return time function (see for example [26, pag. 22]), which is related to the triangle sequence by (2.2). □

Given Theorem 3.1, the proof of Theorem 3.2 is then completed by appealing to the Darling-Kac theorem, which implies that the distributional limit of the Birkhoff sums $\frac{1}{a_n(S)} \sum_{k=0}^{n-1} f \circ S^k$ is $\left(\int_{\bar{\triangle}} f d\mu\right) \mathcal{M}_1$ for all $f \in L^1(\bar{\triangle}, \mu)$, where $\mathcal{M}_1$ is the random variable with normalised Mittag-Leffler distribution of order $\alpha = 1$, (see [1, Corollary 3.7.3]). In particular, since $\mathcal{M}_1$ is constant, the Birkhoff sums converge in probability and Theorem 3.2 is proved.

3.1. **Proof of Theorem 3.1.** We first recall the results we use to prove that the system $(\bar{\triangle}, \mu, S)$ is pointwise dual ergodic.

Definition 3.4. Let V be a measure-preserving transformation of the probability space $(\Omega, \mathcal{A}, \nu)$ and let $\mathcal{C} \subseteq \mathcal{B}$ be a countable measurable partition which is generating for V. Let us denote by $\mathcal{C}^k$, $k \geq 1$, the iterated partitions, that is $\mathcal{C}^k := \bigvee_{j=0}^{k-1} V^{-j}\mathcal{C}$. The system $(\Omega, \mathcal{A}, \nu, V, \mathcal{C})$ is said to be *ψ-mixing* if the sequence

$$\psi_n := \sup_{\substack{C \in \mathcal{C}^k \\ B \in \mathcal{A},\, \nu(B)>0}} \frac{\left|\nu\left(C \cap V^{-(k+n)}B\right) - \nu(C)\nu(B)\right|}{\nu(C)\nu(B)}$$

satisfies $\psi_n \to 0$ as $n \to \infty$.

Remark 3.5. The property defined above as ψ-mixing is often referred to as *continued fraction mixing* since in particular it is satisfied by the Gauss map, see [11, Theorem 5.2.7].

Proposition 3.6 ([1], Lemma 3.7.4 and Proposition 3.7.5)**.** *Let R be a conservative, ergodic measure-preserving transformation of the space $(X, \mathcal{B}, \mu)$, and let $A \in \mathcal{B}$ with $0 < \mu(A) < +\infty$. Define*

(i) *$\varphi_A(x) := \inf\{j \geq 1 \,:\, R^j(x) \in A\}$, the first return time function to A, which is finite for μ-almost every $x \in A$;*

(ii) *the induced map[2] $R_A : A \to A$ as $R_A(x) := R^{\varphi_A(x)}(x)$ for μ-almost every $x \in A$.*

Let $\mathcal{C} \subseteq \mathcal{B} \cap A$ be a countable measurable partition which generates $\mathcal{B}$ under R_A, such that φ_A is $\mathcal{C}$-measurable. If the induced system $(A, \mathcal{B} \cap A, R_A, \mu|_A, \mathcal{C})$ is ψ-mixing, then the original system (X, μ, R) is pointwise dual ergodic.

[2]The induced map R_A is an ergodic measure-preserving transformation of the probability space $(A, \mathcal{B} \cap A, \mu|_A)$. See, for instance, [1, Proposition 1.5.2 and 1.5.3].

To prove the pointwise dual ergodicity of $(\bar{\triangle}, \mu, S)$ it is thus enough to find a set $A \subseteq \bar{\triangle}$ of finite positive measure that satisfies the assumptions of Proposition 3.6. The key point of the previous result is to prove that the induced system is ψ-mixing. To this end, we exploit the properties of the fibred systems introduced by Schweiger in [22] and proved by Nakada to be ψ-mixing under some additional conditions [20, Theorem 2].

Definition 3.7. Let A be a compact and connected subset of $\mathbb{R}^d$, with the Borel σ-algebra $\mathcal{B}$ and let m denote the d-dimensional normalised Lebesgue measure on A. Let V be a measurable map of A onto itself. The pair (A, V) is called a *fibred system* if it satisfies the following properties.

(h1) There exists a finite or countable measurable partition $\mathcal{C} = \{C_i\}_{i\in\mathfrak{J}}$ of A such that the restriction of V to C_i is injective for all $i \in \mathfrak{J}$.

(h2) The map V is differentiable[3] and non-singular.

For $i \in \mathfrak{J}$, we denote by ψ_i the inverse of the restriction $V|_{C_i}$. The cylinder sets of the iterated partition $\mathcal{C}^n = \bigvee_{j=0}^{n-1} V^{-j}\mathcal{C}$ are defined inductively to be

$$C_{i_1,\dots,i_n} = C_{i_1} \cap V^{-1} C_{i_2,\dots,i_n},$$

and we denote by $\psi_{i_1,\dots,i_n}$ the local inverse of V^n restricted to $C_{i_1,\dots,i_n}$. Note that $\psi_{i_1,\dots,i_n} = \psi_{i_1} \circ \cdots \circ \psi_{i_n}$. In order to state the result about the ψ-mixing property of fibred systems, we introduce the following further conditions.

(h3) There exists a sequence $(\sigma(n))_{n\geq 0}$ with $\sigma(n) \to 0$ as $n \to \infty$ and such that

$$\sup_{(i_1,\dots,i_n)\in\mathfrak{J}^n} \operatorname{diam} C_{i_1,\dots,i_n} \leq \sigma(n).$$

(h4) There exist a finite number of measurable subsets $U_1, \dots, U_N$ of A such that for any cylinder $C_{i_1,\dots,i_n}$ of positive measure, there exists U_j with $1 \leq j \leq N$ such that $V^n(C_{i_1,\dots,i_n}) = U_j$ up to measure-zero sets.

(h5) There exists a constant $\lambda \geq 1$ such that

$$\operatorname*{ess\,sup}_{V^n(C_{i_1,\dots,i_n})} |J\psi_{i_1,\dots,i_n}| \leq \lambda \operatorname*{ess\,inf}_{V^n(C_{i_1,\dots,i_n})} |J\psi_{i_1,\dots,i_n}|$$

where $J\psi_{i_1,\dots,i_n}$ denotes the Jacobian determinant of $\psi_{i_1,\dots,i_n}$.

(h6) For any $1 \leq j \leq N$, U_j contains a proper cylinder.

(h7) There is a constant $r_1 > 0$ such that

$$|J\psi_{i_1,\dots,i_n}(p_1) - J\psi_{i_1,\dots,i_n}(p_2)| \leq r_1 m(C_{i_1,\dots,i_n}) \|p_1 - p_2\|$$

for any $p_1, p_2 \in U_j$ and all j.

(h8) There is a constant $r_2 > 0$ such that

$$\|\psi_{i_1,\dots,i_n}(p_1) - \psi_{i_1,\dots,i_n}(p_2)\| \leq r_2 \sigma(n) \|p_1 - p_2\|$$

for any $p_1, p_2 \in U_j$ and all j.

(h9) Let $\mathcal{F}$ be a finite partition generated by $U_1, \dots, U_N$ and denote by $\mathcal{F}_m^c$ the cylinders in $\mathcal{C}^m$ that are not contained in any element of $\mathcal{F}$. Then, as $m \to \infty$

$$\gamma(m) := \sum_{C(i_1,\dots,i_m)\in\mathcal{F}_m^c} m(C(i_1, \dots, i_m)) \to 0.$$

Proposition 3.8 ([20], Theorem 2). *A system $(A, \mathcal{B}, V, \mathcal{C})$ satisfying* (h1)-(h9) *admits an invariant probability measure ν and is ψ-mixing.*

[3] In [22] it is only assumed that V is measurable. We assume differentiability to simplify the approach to the system $(\bar{\triangle}, \mu, S)$.

The strategy to prove the pointwise dual ergodicity of our system $(\bar{\triangle}, \mu, S)$ is therefore to find a set $A \subseteq \bar{\triangle}$ of finite positive measure in such a way that the induced system satisfies (h1)-(h9). We set

$$A \coloneqq \{(x,y) \in \Gamma_0 \,:\, S(x,y) \in \Gamma_0\}\,. \tag{3.2}$$

By definition of S, the set A is the triangle with vertices $Q_1 = \left(\frac{1}{2}, \frac{1}{2}\right)$, $Q_2 = \left(\frac{2}{3}, \frac{1}{3}\right)$ and $Q_3 = (1,1)$, with the sides Q_1Q_2 and Q_2Q_3 not included. Furthermore, notice that every point in the interior of A has triangle sequence of the form $(0,\,0,\,\alpha_3,\,\ldots)$. Let V be the induced map of S on A, that is

$$V(x,y) \coloneqq S^{\varphi_A(x,y)}(x,y),$$

defined for m-almost $(x,y) \in A$, and where $\varphi_A(x,y) = \min\{j \geq 1 \,:\, S^j(x,y) \in A\}$ is finite for m-almost $(x,y) \in A$. Let us first introduce the partition of A given by the level sets of the function φ_A, that is $\tilde{\mathcal{C}} = \{\tilde{C}_k\}_{k\in\mathbb{N}}$ with

$$\tilde{C}_k \coloneqq \{(x,y) \in A \,:\, \varphi_A(x,y) = k\}.$$

Note that $\tilde{C}_1$ is the open triangle with vertices Q_1, Q_2 and $\left(\frac{3}{4}, \frac{1}{2}\right)$, whereas $\tilde{C}_2 = \emptyset$. For each set $\tilde{C}_k$ we introduce the sub-partition $\{C_{k,\sigma} \,:\, \sigma \in \{0,1\}^k\}$, where σ is the symbolic representation of the orbit $\{(x,y),\, S(x,y),\, \ldots,\, S^{k-1}(x,y)\}$ of a point $(x,y) \in \tilde{C}_k$, with respect to the partition $\{\Gamma_0, \Gamma_1\}$. Thus we consider the countable partition

$$\mathcal{C} \coloneqq \{C_{k,\sigma} \,:\, k \geq 1,\ \sigma \in \{0,1\}^k\}.$$

The partition $\mathcal{C}$ is measurable and V is clearly injective on each cylinder, because $S|_{\Gamma_0}$ and $S|_{\Gamma_1}$ are injective and points in the same cylinder have the same symbolic orbit in $\bar{\triangle}$ up to their first return to A. Thus, assumptions (h1) and (h2) are satisfied by the system $(A, \mathcal{B}, V, \mathcal{C})$. Moreover, by the standard results for induced maps recalled above, the transformation V preserves the measure $\mu|_A$, which can be normalised to be a probability measure ν. Some of the remaining assumptions (h3)-(h9) are trivially verified. By definition, V^n maps each cylinder $C_{i_1,\,\ldots,\,i_n}$ from the iterated partition $\mathcal{C}^n$ onto A, thus we can choose $N = 1$ and $U_1 = A$ in order to satisfy assumptions (h4), (h6) and (h9).

It remains to show that the conditions (h3), (h5), (h7) and (h8) also hold for our choice of A, V and $\mathcal{C}$. To this end, we prove some properties of the local inverses of V^n. Let $i = (k,\sigma)$, with $k \geq 1$ and $\sigma = (\sigma_1,\,\ldots,\,\sigma_k) \in \{0,1\}^k$, and let C_i be a cylinder of our partition. A local inverse $\psi_i : A \to C_i$ is given by

$$\psi_i = \phi_{\sigma_1} \circ \phi_{\sigma_2} \circ \cdots \circ \phi_{\sigma_k}.$$

Note that, for $k = 1$, the only possible index is given by $\sigma = (0)$, and the corresponding local inverse is simply ϕ_0. Moreover, by the definition of A, the indices $i = (k,\sigma)$ with $k \geq 3$ satisfy $\sigma_1 = \sigma_2 = 0$, so that $\psi_i = \phi_0 \circ \phi_0 \circ \phi_{\sigma_3} \circ \cdots \circ \phi_{\sigma_k}$. In this way all local inverses ψ_i are of the form

$$\psi_i(x,y) = \left(\frac{r_1 + s_1 x + t_1 y}{r + sx + ty},\ \frac{r_2 + s_2 x + t_2 y}{r + sx + ty}\right), \tag{3.3}$$

with non-negative integer coefficients r_1, r_2, r, s_1, s_2, s, t_1, t_2, t (where the dependence on i has been dropped to simplify the notation). A straightforward computation shows that

$$D\psi_i(x,y) = \begin{pmatrix} \dfrac{(rs_1 - r_1 s) + (s_1 t - st_1)y}{(r+sx+ty)^2} & \dfrac{(rt_1 - r_1 t) - (s_1 t - st_1)x}{(r+sx+ty)^2} \\[2ex] \dfrac{(rs_2 - r_2 s) + (s_2 t - st_2)y}{(r+sx+ty)^2} & \dfrac{(rt_2 - r_2 t) - (s_2 t - st_2)x}{(r+sx+ty)^2} \end{pmatrix}. \tag{3.4}$$

Proposition 3.9. *Let $\psi_{i_1,\,\ldots,\,i_n} : A \to C_{i_1,\,\ldots,\,i_n}$ be a local inverse of V^n and let $D\psi_{i_1,\,\ldots,\,i_n}$ be its Jacobian matrix. Then there exists a sequence $(d(n))_{n\geq 0}$ such that $\lim_{n\to\infty} d(n) = 0$ and*

$$\max\left\{\sup_A \left(|(D\psi_{i_1,\,\ldots,\,i_n})_{11}| + |(D\psi_{i_1,\,\ldots,\,i_n})_{21}|\right),\ \sup_A \left(|(D\psi_{i_1,\,\ldots,\,i_n})_{12}| + |(D\psi_{i_1,\,\ldots,\,i_n})_{22}|\right)\right\} \leq d(n).$$

Proposition 3.10. *Let $\psi_i : A \to \mathbb{R}^2$ be a local inverse of V given by $\psi = \phi_{\sigma_1} \circ \phi_{\sigma_2} \circ \cdots \circ \phi_{\sigma_k}$. Then $r + s + t > 0$ and*

$$J\psi_i(x,y) = \frac{1}{(r+sx+ty)^3}.$$

Proposition 3.9 follows from results used to prove the main result in [16]. We give a proof of the proposition in Appendix A for completeness and also because we obtain an explicit estimate for the sequence $d(n)$. Concerning Proposition 3.10, it is immediate from the construction of the local inverses of V that $r+s+t>0$. The formula for the Jacobian determinant is a particular case of a result in [24] (see also Proposition 2 in [23]). We are now in a position to prove that conditions (h3), (h5), (h7) and (h8) hold for our system.

Proof of (h3). Using Proposition 3.9 we have

$$\begin{aligned}\|\psi_{i_1,\dots,i_n}(x_1,y_1)-\psi_{i_1,\dots,i_n}(x_2,y_2)\| &\le \\ &\le \sup_A \left(|(D\psi_{i_1,\dots,i_n})_{11}|+|(D\psi_{i_1,\dots,i_n})_{21}|\right)|x_1-x_2|+ \\ &\qquad + \sup_A \left(|(D\psi_{i_1,\dots,i_n})_{12}|+|(D\psi_{i_1,\dots,i_n})_{22}|\right)|y_1-y_2| \le \\ &\le d(n)\left(|x_1-x_2|+|y_1-y_2|\right) \le \sqrt{2}\, d(n)\|(x_1,y_1)-(x_2,y_2)\|.\end{aligned}$$

Since $C_{i_1,\dots,i_n}=\psi_{i_1,\dots,i_n}(A)$, we have

$$\operatorname{diam} C_{i_1,\dots,i_n} \le \sqrt{2}d(n)\cdot \operatorname{diam} A = \frac{\sqrt{10}}{3}d(n),$$

so that (h3) is satisfied with $\sigma(n)=\frac{\sqrt{10}}{3}d(n)$. □

Proof of (h8). The above proof of (h3) also shows that (h8) is satisfied with $r_2=\frac{3}{\sqrt{5}}$. □

Proof of (h5). We have $V^n(C_{i_1,\dots,i_n})=A$ for all n and all cylinders $C_{i_1,\dots,i_n}$. For all $(x,y)\in A$ holds

$$\frac{1}{27}(r+s+t)^3 \le (r+sx+ty)^3 \le (r+s+t)^3$$

if the coefficients r, s, and t are non-negative. Then from Proposition 3.10 it follows that condition (h5) holds with $\lambda=27$. □

Proof of (h7). Using Proposition 3.10 we have

$$|J\psi_{i_1,\dots,i_n}(x_1,y_1)-J\psi_{i_1,\dots,i_n}(x_2,y_2)| \le 3\sqrt{2}\left(\max_{(x,y)\in A}\frac{s+t}{(r+sx+ty)^4}\right)\|(x_1,y_1)-(x_2,y_2)\|.$$

Arguing as above

$$\max_{(x,y)\in A}\frac{s+t}{(r+sx+ty)^4} \le 3\,\frac{s+t}{r+s+t}\max_{(x,y)\in A} J\psi_{i_1,\dots,i_n}(x,y) \le 81 \inf_{(x,y)\in A} J\psi_{i_1,\dots,i_n}(x,y) \le 81 m(C_{i_1,\dots,i_n})$$

where m denotes the normalised Lebesgue measure on A. It follows that (h7) holds with $r_1=243\sqrt{2}$. □

We have thus proved that the induced map V of S on the triangle A satisfies the assumptions of Proposition 3.8, hence the induced map V is ψ-mixing. As a consequence, our system $(\bar{\triangle},\mu,S)$ satisfies the assumptions of Proposition 3.6, hence it is pointwise dual ergodic.

The second part of Theorem 3.1 concerns the return sequence $a_n(S)$. To achieve the conclusion, we use [1, Lemma 3.7.4] and [26, Proposition 7]: these results imply that, given the *wandering rate* $w_n(A)$ of the set A, it holds

$$a_n(S) \asymp \frac{n}{w_n(A)}$$

In Appendix B we recall the definition of the wandering rate $w_n(A)$ of the set A and show that $(w_n(A))_{n\ge1}$ satisfies $w_n(A)\asymp \log^2 n$ (see Propositions B.1 and B.4). This completes the proof of Theorem 3.1.

Remark 3.11. It is known (see e.g. [1, Lemma 3.7.4] and [26, Proposition 7]) that if $w_n(A)$ is regularly varying of index $1-\alpha$, then

$$a_n(S) \sim \frac{1}{\Gamma(2-\alpha)\Gamma(1+\alpha)}\frac{n}{w_n(A)},$$

and $a_n(S)$ is a regularly varying sequence of index α. Hence to obtain that $a_n(S)$ is regularly varying of index $\alpha=1$ it is enough to show that $w_n(A)$ is slowly varying. The last is the additional assumption we need in Theorem 3.2 and Corollary 3.3.Unfortunately we don't have a proof that $w_n(A)$ is slowly varying. At the end of Appendix B we discuss this property.

4. Pointwise convergence of Birkhoff averages for a class of non-summable observables

As already mentioned at the beginning of Section 3, for infinite-measure-preserving systems (like our map S, or equivalently the map F on the strip as described in Section 2.2), the strict analogue of Birkhoff's Ergodic Theorem is trivial, in the sense that it tells us only that for every observable $f \in L^1(\mu)$, the Birkhoff averages of f for a system (X, μ, R)

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ R^k(x)$$

converge μ-almost everywhere to zero. In a recent paper [15], the question of convergence of Birkhoff sums for "global observables", which were first introduced by Lenci [13, 14] in the context of infinite mixing, is considered. In [15], a global observable is rather vaguely defined to be any L^∞ function for which a Birkhoff-like theorem could in principle be shown to hold. We would like to apply one of the results of this paper to give certain examples of L^∞ observables for our map F for which the Birkhoff average can shown to be almost everywhere constant. In order to state this result, first we need to recall a certain dynamically-defined partition.

Assume that $(X, \mathcal{B}, \mu, R)$ is a conservative and ergodic system. Given a set L_0 with $0 < \mu(L_0) < +\infty$, we have that

$$\bigcup_{k \geq 0} R^{-k} L_0 = X \pmod{\mu},$$

that is, L_0 is a sweep-out set. Now recursively define, for each $k \geq 1$,

$$L_k \coloneqq \left(R^{-1} L_{k-1} \right) \setminus L_0.$$

Then the collection $\{L_k\}_{k \geq 0}$ forms a partition of X.

Theorem 4.1 ([15])**.** *Let $(X, \mathcal{B}, \mu, R)$ be an infinite-measure-preserving, conservative, ergodic dynamical system, endowed with the partition $\{L_k\}_{k \geq 0}$, as described above. Let $f \in L^\infty(X, \mu)$ admit $f^* \in \mathbb{C}$ with the following property: $\forall \varepsilon > 0$, $\exists N, K \in \mathbb{N}$ such that $\forall x \in \bigcup_{k \geq K} L_k$,*

$$\left| \frac{1}{N} \sum_{k=0}^{N-1} f \circ R^k(x) - f^* \right| \leq \varepsilon.$$

Then for μ-almost every $x \in X$,

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ R^k(x) = f^*.$$

Let us recall the system (Σ, ρ, F) defined in Section 2.2, which is isomorphic to $(\bar{\triangle}, \mu, S)$, and let us describe a suitable partition of the strip Σ for the application of Theorem 4.1. We let $L_0 \coloneqq \Pi_0$, and then, as above, define recursively the sets

$$L_k \coloneqq \left(F^{-k} L_{k-1} \right) \setminus L_0 = \{ (u, v) \in \Sigma \, : \, k \leq v < k+1 \} = \Sigma_k.$$

As a class of observables we consider the set $\mathcal{G}$ of functions $f(u, v) \coloneqq g(u) \cdot h(v)$, where $g : (0, 1) \to \mathbb{R}$ is a bounded function and $h : \mathbb{R} \to \mathbb{R}$ is a continuous α-periodic function, with $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

Theorem 4.2. *Let (Σ, ρ, F) be the system defined in Section 2.2, and $f : \Sigma \to \mathbb{C}$ a function $f(u, v) = g(u) h(v)$ in the space $\mathcal{G}$ defined above, with g constant if $\int_0^\alpha h dv \neq 0$. Then there exists a constant $f^* \in \mathbb{C}$ such that for ρ-almost every $(u, v) \in \Sigma$*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ F^k(u, v) = f^*.$$

Proof. For all $N \in \mathbb{N}$ and for all $(u, v) \in \Sigma$ with $v > N$ we have

$$\frac{1}{N} \sum_{k=0}^{N-1} f \circ F^k(u, v) = \frac{1}{N} \sum_{k=0}^{N-1} f(u, v-k) = \\ = \frac{1}{N} g(u) \sum_{k=0}^{N-1} h(v-k) = g(u) \cdot \frac{1}{N} \sum_{k=0}^{N-1} h \circ \tau^k(v),$$

where $\tau : \mathbb{R}/\alpha\mathbb{Z} \to \mathbb{R}/\alpha\mathbb{Z}$ is defined by $\tau(x) \coloneqq x - 1 \pmod{\alpha}$. Note that τ preserves the Lebesgue measure and is topologically conjugate to the rotation $R_{\frac{1}{\alpha}} : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$, $R_{\frac{1}{\alpha}}(x) \coloneqq x - \frac{1}{\alpha} \pmod 1$. The map $R_{\frac{1}{\alpha}}$ is uniquely ergodic with respect to the Lebesgue measure since $\alpha \in \mathbb{R} \setminus \mathbb{Z}$, and thus so is τ. It then follows, since h is continuous and $\mathbb{R}/\alpha\mathbb{Z}$ is compact, that

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} h \circ \tau^k(v) = \frac{1}{\alpha} \int_0^\alpha h(v)\, dv =: h^*$$

uniformly on $\mathbb{R}$. If $h^* = 0$, there exists $N_\star \in \mathbb{N}$ such that for all $n \geq N_\star$

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} h \circ \tau^k(v) \right| < \frac{\varepsilon}{\|g\|_\infty}.$$

If we choose $N = N_\star$, $K > N$, and $f^* = 0$, then for all $(u, v) \in \Sigma$ with $v \geq K$

$$\left| \frac{1}{N} \sum_{k=0}^{N-1} f \circ F^k(u, v) - f^* \right| = \left| \frac{1}{N} \sum_{k=0}^{N-1} g(u) h(v-k) \right| \leq \|g\|_\infty \left| \frac{1}{N} \sum_{k=0}^{N-1} h \circ \tau^k(v) \right| < \varepsilon,$$

and we can apply Theorem 4.1 with $f^* = 0$. In the case $h^* \neq 0$, we can repeat the argument when $g(u) \equiv \bar{g}$ is a constant function. In that case there exists $N_\star \in \mathbb{N}$ such that for all $n \geq N_\star$

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} h \circ \tau^k(v) - h^* \right| < \frac{\varepsilon}{\bar{g}}\,,$$

and we can apply Theorem 4.1 as above with $f^* = \bar{g}\, h^*$. ☐

5. A complete triangular tree of rational pairs

Our aim in this section is to construct a tree that contains every pair of rational numbers in $\mathbb{Q}^2 \cap \bar{\triangle}$, first by using a modified version of the map S and then, equivalently, by giving a geometric contruction by way of a mediant operation defined on pairs of rational numbers. The construction mimics that of the Farey tree, generated by the Farey map, which we now recall.

Firstly, the *Farey map* is the map $F : [0, 1] \to [0, 1]$ defined by setting

$$F(x) \coloneqq \begin{cases} \frac{x}{1-x} & \text{if } 0 \leq x \leq \frac{1}{2} \\ \frac{1-x}{x} & \text{if } \frac{1}{2} \leq x \leq 1 \end{cases}$$

We generate a binary tree using the map F by defining the levels $\mathcal{L}_n \coloneqq F^{-n}\left(\frac{1}{2}\right)$, with the vertices connected as shown in Figure 3. Note that the two "children" of each vertex are not simply the inverse images of that vertex. If we label a step down to the left with "0" and a step down to the right with "1", the position of a given rational number $\frac{p}{q} \in \mathcal{L}_n$ is described by a path $\omega_1 \cdots \omega_n$, with each $\omega_i \in \{0, 1\}$, and such that if F_0 and F_1 denote the inverse branches of F, then $\frac{p}{q} = F_{\omega_1} \circ \cdots \circ F_{\omega_n}\left(\frac{1}{2}\right)$. One of the most important properties of the Farey tree is that it contains all the rational numbers in the interval $(0, 1)$, and each rational number appears in the tree exactly once[4]. In other words, $\bigcup_{k=0}^{+\infty} F^{-k}\left(\frac{1}{2}\right) = \mathbb{Q} \cap (0, 1)$.

[4]In particular, a rational number $\frac{p}{q} \in \mathcal{L}_n$ if and only if its continued fraction expansion $\frac{p}{q} = [a_1, \ldots, a_r]$ with $a_r > 1$ is such that $\sum_{i=1}^r a_i = n + 2$.

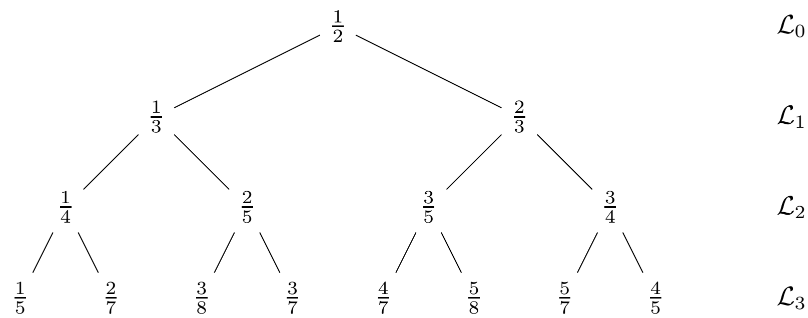

FIGURE 3. The first four levels of the Farey tree.

Another way to define the levels of the Farey tree is by considering the Stern-Brocot sets $(\mathcal{F}_n)_{n\geq -1}$, where we define $\mathcal{F}_{-1} \coloneqq \left\{\frac{0}{1}, \frac{1}{1}\right\}$, and for all $n \geq 0$, $\mathcal{F}_n$ is defined recursively from $\mathcal{F}_{n-1}$ by inserting the mediant of each pair of neighbouring fractions. Recall that the mediant of two fractions $\frac{p}{q}$ and $\frac{r}{s}$ is defined to be

$$\frac{p}{q} \oplus \frac{r}{s} \coloneqq \frac{p+r}{q+s}.$$

It is easy to verify that the mediant falls between the two rational numbers it is computed from, that is if $\frac{p}{q} < \frac{r}{s}$ then $\frac{p}{q} < \frac{p}{q} \oplus \frac{r}{s} < \frac{r}{s}$. The first few of the Stern-Brocot sets are as follows:

$$\mathcal{F}_0 = \left\{\frac{0}{1}, \frac{1}{2}, \frac{1}{1}\right\}, \quad \mathcal{F}_1 = \left\{\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}\right\}, \quad \mathcal{F}_2 = \left\{\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}\right\}.$$

It is also straightforward to prove that $\#\mathcal{F}_n = 2^{n+1} + 1$, and that $\mathcal{L}_n = \mathcal{F}_n \setminus \mathcal{F}_{n-1}$ for all $n \geq 0$. For more details on the Farey tree, we refer to [5].

We now describe the construction of our two-dimensional Farey-like tree. We use the local inverse $\phi_0 : \bar{\triangle} \setminus \{x = y\} \to \Gamma_0$, the restricted local inverse $\phi_1 : \bar{\triangle} \setminus \Lambda \to \Gamma_1 \setminus \Lambda$ (which we will continue to call ϕ_1), and a new map $\phi_2 : \{x = y \,:\, 0 \leq x \leq 1\} \to \Lambda$ defined to be $\phi_2(x, x) \coloneqq (x, 0)$. The geometric action of ϕ_0 and ϕ_1 is shown in Figure 4. These three maps are the local inverses of the map

$$\tilde{S} : \bar{\triangle} \to \bar{\triangle}, \quad \tilde{S}(x, y) \coloneqq \begin{cases} S(x, y) & \text{if } (x, y) \in \bar{\triangle} \setminus \Lambda \\ (x, x) & \text{if } (x, y) \in \Lambda \end{cases}.$$

The map $\tilde{S}$ is a modified version of the map S defined in (2.1).

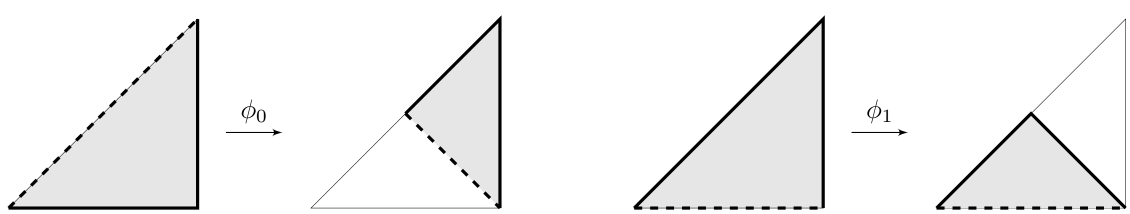

FIGURE 4. Geometric action of the maps ϕ_0 and ϕ_1.

We now define the sequence $(\mathcal{T}_n)_{n\geq -1}$ of the levels of the tree associated to $\tilde{S}$. First set

$$\mathcal{T}_{-1} \coloneqq \{(0, 0), (1, 0), (1, 1)\} \quad \text{and} \quad \mathcal{T}_0 \coloneqq \left\{ \left(\frac{1}{2}, 0\right), \left(1, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}\right) \right\},$$

which include the vertices of the triangle $\triangle$ and the middle points of the sides, respectively.

Definition 5.1. For each $n \geq -1$, define $\mathcal{B}_n := \mathcal{T}_n \cap \partial\triangle$ and $\mathcal{I}_n := \mathcal{T}_n \cap \mathring{\triangle}$ to be, respectively, the *boundary points* and the *interior points* of the n-th level of the tree. Moreover, we respectively denote with $\mathcal{B}_{\leq n} := \bigcup_{k=-1}^{n} \mathcal{B}_k$ and $\mathcal{I}_{\leq n} := \bigcup_{k=-1}^{n} \mathcal{I}_k$ the boundary and interior points of the tree up to level n.

Clearly $\mathcal{B}_{-1} = \mathcal{T}_{-1}$ and $\mathcal{B}_0 = \mathcal{T}_0$. We now define precisely how the levels of the tree are constructed, by showing all the possibilities for taking counterimages depending on the location of the point in $\bar{\triangle}$. Let $n \geq 0$.

(R1) An interior point $\left(\frac{p}{q}, \frac{r}{q}\right) \in \mathcal{I}_n$ generates the two interior points $\left(\frac{q}{r+q}, \frac{p}{r+q}\right)$ and $\left(\frac{p}{r+q}, \frac{r}{r+q}\right)$ in $\mathcal{I}_{n+1}$, through the application of ϕ_0 and ϕ_1, respectively.

(R2) A boundary point $\left(\frac{p}{q}, \frac{p}{q}\right) \in \mathcal{B}_n$ generates the point $\left(\frac{p}{q}, 0\right) \in \mathcal{B}_n$ through the application of ϕ_2 and the boundary point $\left(\frac{p}{p+q}, \frac{p}{p+q}\right) \in \mathcal{B}_{n+1}$ through the application of ϕ_1.

(R3) A boundary point $\left(\frac{p}{q}, 0\right) \in \mathcal{B}_n$ generates the point $\left(1, \frac{p}{q}\right) \in \mathcal{B}_n$ through the application of ϕ_0.

(R4) A boundary point $\left(1, \frac{p}{q}\right) \in \mathcal{B}_n$ generates the boundary point $\left(\frac{q}{p+q}, \frac{q}{p+q}\right) \in \mathcal{B}_{n+1}$ and the interior point $\left(\frac{q}{p+q}, \frac{p}{p+q}\right) \in \mathcal{I}_{n+1}$, through the application of ϕ_0 and ϕ_1, respectively.

The basic portions of the counterimages tree generated from a boundary point and from an interior point are shown in Figure 5. Note that the points of the tree always have rational coordinates, since we start from points with rational coordinates and ϕ_0, ϕ_1 and ϕ_2 are linear fractional maps. Furthermore, taking a counterimage does not necessarily implies that the level in the tree changes. Indeed, applying rules (R1), (R4), and rule (R2) with ϕ_1 makes the level to increase, whereas applying the other rules does not change the level. The levels $\mathcal{T}_0$, $\mathcal{T}_1$ and $\mathcal{T}_2$ of the tree are shown in Figure 10 at the end of the paper. Note that we always write the two fractions of each pair reduced to their least common denominator, apart from the pairs containing 0 and/or 1. This choice also has a geometric motivation, as we shall remark after Definition 5.5.

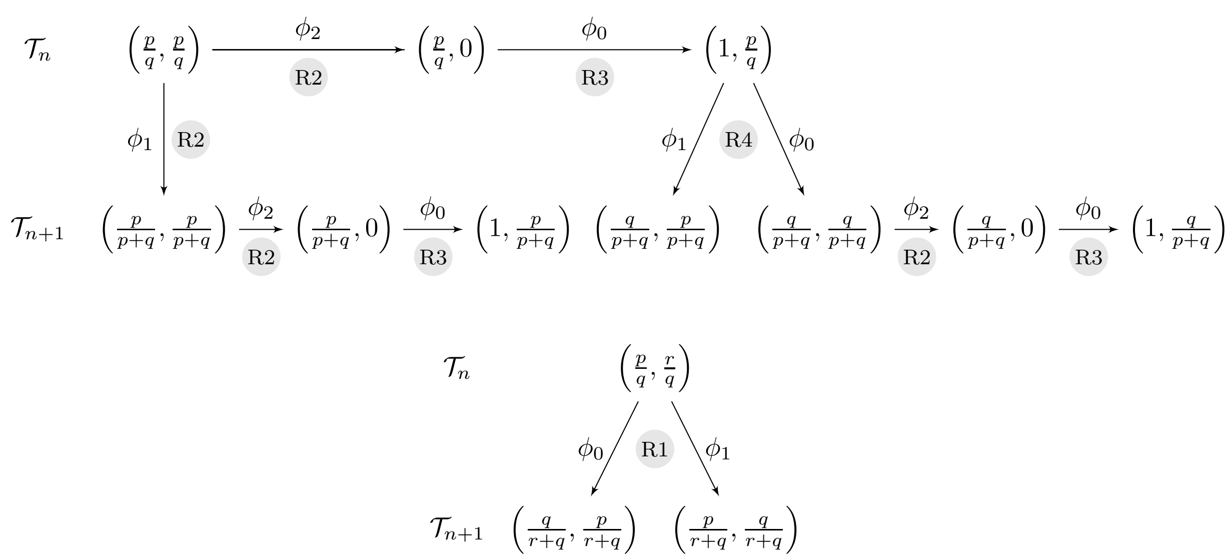

FIGURE 5. *Above.* Basic portion of the tree generated from a boundary point $\left(\frac{p}{q}, \frac{p}{q}\right) \in \mathcal{B}_n$. *Below.* Basic portion of the tree generated from an interior point $\left(\frac{p}{q}, \frac{r}{q}\right) \in \mathcal{I}_n$.

Lemma 5.2. *For all $n \geq 0$ we have*

$$\#\mathcal{B}_n = 3 \cdot 2^n \quad \text{and} \quad \#\mathcal{I}_n = n2^{n-1},$$

with the points of $\mathcal{B}_n$ equally distributed on the three sides of $\triangle$. As a consequence, the number of points of each level of the tree is given by

$$\#\mathcal{T}_n = 3 \cdot 2^n + n2^{n-1}.$$

Proof. We argue by induction on $n \geq 0$. If $n = 0$ we have $\mathcal{B}_0 = \mathcal{T}_0$ and $\mathcal{I}_0 = \emptyset$. Thus $\#\mathcal{B}_0 = 3$, with one point on each side of $\triangle$, and $\#\mathcal{I}_0 = 0$: the base case is proved. Suppose $\#\mathcal{B}_n = 3 \cdot 2^n$, with 2^n points on each side of $\triangle$, $\#\mathcal{I}_n = n2^{n-1}$ for some $n \geq 0$ and consider the subsequent level of the tree. The points of $\mathcal{B}_{n+1}$ can be obtained from those of $\mathcal{B}_n$ as follows (refer to Figure 5):

- $\mathcal{B}_n$ contains 2^n points on $\bar{\triangle} \cap \{x = y\}$, each of which gives 3 points in $\mathcal{B}_{n+1}$, one per side, applying (R2) with ϕ_1, followed by (R2) with ϕ_2, and (R3);
- $\mathcal{B}_n$ contains 2^n points on $\bar{\triangle} \cap \{x = 1\}$, each of which gives 3 points in $\mathcal{B}_{n+1}$, one per side, applying (R4), followed by (R2) with ϕ_2, and (R3).

Note that the other 2^n points contained in $\mathcal{B}_n$ lie on the line $\{y = 0\}$, and these points are all mapped back inside $\mathcal{B}_n$, in light of rule (R3). Therefore, $\#\mathcal{B}_{n+1} = 2^n \cdot 3 + 2^n \cdot 3 = 3 \cdot 2^{n+1}$. Furthermore, by construction, we have 2^{n+1} points of $\mathcal{B}_{n+1}$ on each side of the triangle, so that the points of $\mathcal{B}_{n+1}$ are equally distributed on the three sides of $\triangle$. The points of $\mathcal{I}_{n+1}$ are obtained from those of $\mathcal{T}_n$ in this way:

- each point in $\mathcal{I}_n$ generates two points in $\mathcal{I}_{n+1}$, according to rule (R1);
- each point in $\mathcal{B}_n \cap \{x = 1\}$ gives one point in $\mathcal{I}_{n+1}$, using rule (R4).

As a consequence $\#\mathcal{I}_{n+1} = n2^{n-1} \cdot 2 + 2^n = (n+1)2^n$. The inductive step is proved and the proof is complete. □

Lemma 5.3. *Let $F : [0,1] \to [0,1]$ be the Farey map. For all $n \geq 0$ we have*

$$\mathcal{B}_n = \left\{ \left(\frac{p}{q}, \frac{p}{q}\right), \left(\frac{p}{q}, 0\right), \left(1, \frac{p}{q}\right) \; : \; \frac{p}{q} \in F^{-n}\left(\frac{1}{2}\right) \right\}.$$

Proof. We claim that it is suffices to prove that for all $n \geq 0$

$$\left\{ \left(\frac{p}{q}, \frac{p}{q}\right) \; : \; \frac{p}{q} \in F^{-n}\left(\frac{1}{2}\right) \right\} \subseteq \mathcal{B}_n. \tag{5.1}$$

Indeed by (R2) and (R3), it easily follows that $\left\{ \left(\frac{p}{q}, \frac{p}{q}\right), \left(\frac{p}{q}, 0\right), \left(1, \frac{p}{q}\right) \; : \; \frac{p}{q} \in F^{-n}\left(\frac{1}{2}\right) \right\} \subseteq \mathcal{B}_n$ for all $n \geq 0$, and since from Lemma 5.2 we have $\#\mathcal{B}_n = 3 \cdot 2^n$ for all $n \geq 0$, the claim is proved. We now argue by induction to prove that (5.1) holds for all $n \geq 0$. If $n = 0$ we have $\left(\frac{1}{2}, \frac{1}{2}\right) \in \mathcal{B}_0$, thus the base case is proved. Now suppose that (5.1) holds for some $n \geq 0$ and let $\frac{r}{s} \in F^{-(n+1)}\left(\frac{1}{2}\right)$, so that $F\left(\frac{r}{s}\right) \in F^{-n}\left(\frac{1}{2}\right)$. In order to prove that $\left(\frac{r}{s}, \frac{r}{s}\right) \in \mathcal{B}_{n+1}$ we distinguish between two cases.

- If $0 \leq \frac{r}{s} \leq \frac{1}{2}$ then $F\left(\frac{r}{s}\right) = \frac{r}{s-r}$, and by the induction hypothesis $\left(\frac{r}{s-r}, \frac{r}{s-r}\right) \in \mathcal{B}_n$. By rule (R2) we then have that $\phi_1\left(\frac{r}{s-r}, \frac{r}{s-r}\right) = \left(\frac{r}{s}, \frac{r}{s}\right) \in \mathcal{B}_{n+1}$.
- If $\frac{1}{2} < \frac{r}{s} \leq 1$ then $F\left(\frac{r}{s}\right) = \frac{s-r}{r}$. By the induction hypothesis we have that $\left(\frac{s-r}{r}, \frac{s-r}{r}\right) \in \mathcal{B}_n$ and, by construction, we also have that $\left(1, \frac{s-r}{r}\right) \in \mathcal{B}_n$. Hence rule (R4) yields that $\phi_0\left(1, \frac{s-r}{r}\right) = \left(\frac{r}{s}, \frac{r}{s}\right) \in \mathcal{B}_{n+1}$.

□

We are now in a position to prove the first main result of this section.

Theorem 5.4. *The tree defined by the level sets $\mathcal{T}_n$ is complete, that is,*

$$\bigcup_{n\geq -1} \mathcal{T}_n = \mathbb{Q}^2 \cap \bar{\triangle},$$

and every pair of rational numbers appears in the tree exactly once.

Proof. We have shown in Lemma 5.3 that every pair of rational numbers of the form $\left(1, \frac{a}{b}\right)$, with $0 \leq \frac{a}{b} \leq 1$, appears in some set $\mathcal{B}_n$. Let $\left(\frac{p}{q}, \frac{r}{q}\right)$ be an arbitrary pair of rational numbers in the interior of the triangle $\triangle$. Then, as noted in Section 2.3, this point has a terminating triangle sequence, say $(\alpha_1, \ldots, \alpha_m)$. Then it follows that $S^{\sum_{i=1}^m \alpha_i + m - 1}\left(\frac{p}{q}, \frac{r}{q}\right) = \left(1, \frac{a}{b}\right)$, for some rational number $\frac{a}{b}$. Thus, as a backward image of the point $\left(1, \frac{a}{b}\right)$ under S, our original, arbitrary, point must lie in the tree. Moreover, that each point appears exactly once is clear, considering the geometric action of the maps ϕ_i. □

At this point, we will begin the description of a tree of points from the triangle using a mediant operation on pairs of rational numbers. This will then be shown to be equivalent to the description of our two-dimensional Farey tree given above in terms of counterimages.

Definition 5.5. Let $\left(\frac{p}{q}, \frac{r}{q}\right)$ and $\left(\frac{p'}{q'}, \frac{r'}{q'}\right)$ be two couples of fractions. We define their *mediant* to be

$$\left(\frac{p}{q}, \frac{r}{q}\right) \oplus \left(\frac{p'}{q'}, \frac{r'}{q'}\right) := \left(\frac{p}{q} \oplus \frac{p'}{q'}, \frac{r}{q} \oplus \frac{r'}{q'}\right) = \left(\frac{p+p'}{q+q'}, \frac{r+r'}{q+q'}\right).$$

Note that we require that the two fractions of each couple have the same denominator, in order that the mediant of two points lies on the open segment joining the two points. In what follows we always assume that the two fractions of each couple are reduced to their least common denominator.

Definition 5.6. Let $\mathfrak{S} \subseteq \mathbb{R}^2$ be a line segment, and let $\mathfrak{R} \subseteq \mathfrak{S} \cap \mathbb{Q}^2$ be a finite subset of rational points on $\mathfrak{S}$. Let $r := \#\mathfrak{R}$, with $2 \leq r < +\infty$, and write $\mathfrak{R} = \{\mathfrak{r}_i \,:\, i = 1, \ldots, r\}$ with $\mathfrak{r}_i \leq_{lex} \mathfrak{r}_{i+1}$ for all $i = 1, \ldots, r-1$, where $\leq_{lex}$ is the lexicographic order on $\mathbb{R}^2$. We define the *Farey sum of* $\mathfrak{R}$ to be the set $\mathfrak{R}^{\oplus}$, where

$$\mathfrak{R}^{\oplus} := \{\mathfrak{r}_i \oplus \mathfrak{r}_{i+1} \,:\, i = 1, \ldots, r-1\} \cup \mathfrak{R}.$$

To simplify the presentation, in what follows the maps ϕ_0 and ϕ_1 are extended to $\bar{\triangle}$.

Lemma 5.7. *The maps ϕ_0 and ϕ_1 preserve the mediant of any two rational pairs in their respective domains.*

Proof. Let $\left(\frac{p}{q}, \frac{r}{q}\right)$ and $\left(\frac{p'}{q'}, \frac{r'}{q'}\right)$ be two rational points in the domain of ϕ_0. Then

$$\phi_0\left(\left(\frac{p}{q}, \frac{r}{q}\right) \oplus \left(\frac{p'}{q'}, \frac{r'}{q'}\right)\right) = \phi_0\left(\frac{p+p'}{q+q'}, \frac{r+r'}{q+q'}\right) = \left(\frac{q+q'}{r+q+r'+q'}, \frac{p+p'}{r+q+r'+q'}\right)$$

and

$$\phi_0\left(\frac{p}{q}, \frac{r}{q}\right) \oplus \phi_0\left(\frac{p'}{q'}, \frac{r'}{q'}\right) = \left(\frac{q}{r+q}, \frac{p}{r+q}\right) \oplus \left(\frac{q'}{r'+q'}, \frac{p'}{r'+q'}\right) = \left(\frac{q+q'}{r+q+r'+q'}, \frac{p+p'}{r+q+r'+q'}\right).$$

An analogous computation can be done for the map ϕ_1: we leave the details to the reader. □

Let us observe here that, since ϕ_0 and ϕ_1 preserve the mediant operation and are also monotonic along line segments (with respect to the lexicographic order), we have that for $i = 0, 1$,

$$\begin{aligned} \phi_i(\mathfrak{R})^{\oplus} &= \{\phi_i(\mathfrak{r}_i) \oplus \phi_i(\mathfrak{r}_{i+1}) \,:\, i = 1, \ldots, r-1\} \cup \phi_i(\mathfrak{R}) = \\ &= \{\phi_i(\mathfrak{r}_i \oplus \mathfrak{r}_{i+1}) \,:\, i = 1, \ldots, r-1\} \cup \phi_i(\mathfrak{R}) = \phi_i\left(\mathfrak{R}^{\oplus}\right). \end{aligned} \tag{5.2}$$

In order to give the definition of a tree of mediants, we first define a sequence of measurable partitions $(\mathscr{P}_n)_{n\geq 0}$ of $\bar{\triangle}$, such that $\mathscr{P}_n$ consists of 2^n subtriangles of $\bar{\triangle}$ and each $\mathscr{P}_n$ is a refinement of the previous $\mathscr{P}_{n-1}$. Let $\mathscr{P}_0$ be the whole triangle $\bar{\triangle}$. The three vertices of $\bar{\triangle}$ are labelled with "0", "1" and "2" as follows:

$$v_0 = (0,0) = \left(\frac{0}{1}, \frac{0}{1}\right), \quad v_1 = (1,0) = \left(\frac{1}{1}, \frac{0}{1}\right), \quad v_2 = (1,1) = \left(\frac{1}{1}, \frac{1}{1}\right).$$

Taking the Farey sum between v_0 and v_2 one obtains

$$v_0 \oplus v_2 = \left(\frac{1}{2}, \frac{1}{2}\right).$$

We partition the triangle $\bar{\triangle}$ into two subtriangles by the line segment joining v_1 and $v_0 \oplus v_2$. This determines the partition $\mathscr{P}_1$. Moreover, we label the vertices of the two subtriangles according to the geometric rule shown in Figure 6: that is, the new vertex is labelled "2" in both the subtriangles, the other vertices of the subtriangle containing the old vertex "0" remain as they were, whereas in the subtriangle containing the old vertex "2", this "2" becomes a "1" and the remaining vertex is labelled "0" (note that, in this second subtriangle, this can be seen as a rotation of the old labels). We now proceed inductively. Suppose we have the partition $\mathscr{P}_n$, consisting of 2^n triangles. Each triangle of $\mathscr{P}_n$ is partitioned into two subtriangles by the line segment joining the vertex labelled "1" with the mediant of the vertex "0" and the vertex "2". This gives us the next partition $\mathscr{P}_{n+1}$. Figure 7 shows the partitions $\mathscr{P}_0$, $\mathscr{P}_1$, and $\mathscr{P}_2$.

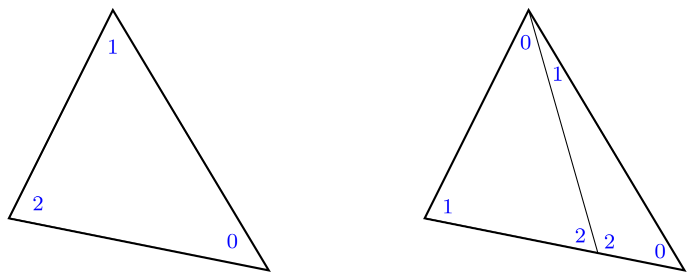

FIGURE 6. Partition of a triangle of $\mathscr{P}_n$ into two subtriangles and relabelling of the vertices.

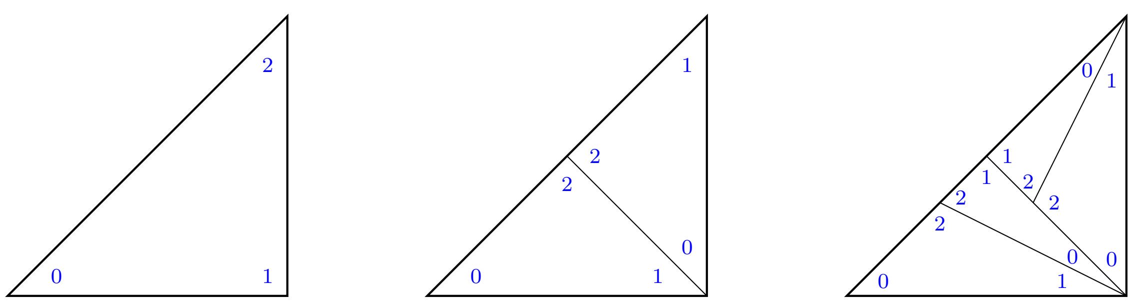

FIGURE 7. From left to right: partitions $\mathscr{P}_0$, $\mathscr{P}_1$, and $\mathscr{P}_2$, along with the labelling of the vertices.

For the definition of the tree, in a Farey-like way, we recursively define a sequence $(\mathcal{S}_n)_{n\geq -1}$ of nested sets of pairs of rationals. First set $\mathcal{S}_{-1} \coloneqq \{(0,0), (1,0), (1,1)\}$ and consider the partition $\mathscr{P}_0$. The basic idea is to insert the mediant of each pair of neighbouring points along each side of the partition $\mathscr{P}_0$. For instance, at the first step we have three sides (the sides of $\bar{\triangle}$), and we add to $\mathcal{S}_{-1}$ one point along each side, providing

$$\mathcal{S}_0 = \left\{(0,0), \left(\frac{1}{2}, 0\right), (1,0), \left(1, \frac{1}{2}\right), (1,1), \left(\frac{1}{2}, \frac{1}{2}\right)\right\}.$$

Again we proceed inductively. Suppose we have the set $\mathcal{S}_n$ and consider the partition $\mathscr{P}_{n+1}$. The set $\mathcal{S}_{n+1}$ is obtained from $\mathcal{S}_n$ inserting the mediant of each pair of neighbouring points along each side of the partition $\mathscr{P}_{n+1}$. In other words, for $n \geq -1$,

$$\mathcal{S}_{n+1} \coloneqq \bigcup_{\mathfrak{S} \in \mathscr{S}_{n+1}} (\mathfrak{S} \cap \mathcal{S}_n)^{\oplus},$$

where $\mathscr{S}_0$ is the set of the three sides of $\bar{\triangle}$, and $\mathscr{S}_n$ is obtained from $\mathscr{S}_{n-1}$ by adding the line segments used to partition the triangles of $\mathscr{P}_{n-1}$ to obtain $\mathscr{P}_n$.

Now we will start to work towards showing that our two trees are in fact identical. Figure 8 at the end of the section may ease the understanding of the argument. Let $n \geq 0$ and let $\omega \in \{0,1\}^n$ be a word of length $|\omega| = n$ over the two symbols "0" and "1". We define

$$\phi_\omega \coloneqq \phi_{\omega_1} \circ \phi_{\omega_2} \circ \cdots \circ \phi_{\omega_n} \quad \text{and} \quad \triangle_\omega \coloneqq \phi_\omega(\bar{\triangle}).$$

Lemma 5.8. *Let ℓ be the open line segment joining $(1,0)$ and $\left(\frac{1}{2},\frac{1}{2}\right)$, that is $\ell \coloneqq \left\{(x, 1-x) \,:\, \frac{1}{2} < x < 1\right\}$. For $n \geq 0$ the following holds.*

(i) *$\mathscr{P}_n = \{\triangle_\omega \,:\, |\omega| = n\}$, and the labelling of the vertices of each $\triangle_\omega$ is such that the vertex "k" of $\triangle_\omega = \phi_\omega(\bar{\triangle})$ is $\phi_\omega(v_k)$, for $k = 0,1,2$.*

(ii) *$\mathscr{S}_n = \mathscr{S}_0 \cup \left\{\overline{\phi_\omega(\ell)} \,:\, |\omega| \leq n-1\right\}$.*

Proof. (i) The statement is trivially true when $n = 0$. We argue by induction on $n \geq 1$. From the definition of the maps ϕ_0 and ϕ_1 it is straightforward to see that

$$\mathscr{P}_1 = \left\{\phi_0(\bar{\triangle}), \phi_1(\bar{\triangle})\right\}$$

and that the relabelling of the vertices agrees with the geometric action of the two maps. This proves the case $n = 1$. We also observe that the partition $\mathscr{P}_1$ is obtained through the line segment $\bar{\ell}$. For the inductive step, suppose that, for a certain $n \geq 1$, $\mathscr{P}_n = \{\triangle_\omega \,:\, |\omega| = n\}$ and that the labelling of the vertices of each $\triangle_\omega$ is induced by ϕ_ω as in (i). Consider a triangle $\triangle_\omega = \phi_\omega(\bar{\triangle}) \in \mathscr{P}_n$, and note that

$$\triangle_\omega = \phi_\omega(\bar{\triangle}) = \phi_\omega(\phi_0(\bar{\triangle}) \cup \phi_1(\bar{\triangle})) = \phi_{\omega 0}(\bar{\triangle}) \cup \phi_{\omega 1}(\bar{\triangle}).$$

This partition is obtained through $\overline{\phi_\omega(\ell)}$ and we now prove that it agrees with the definition of $\mathscr{P}_{n+1}$. Indeed, from Lemma 5.7, ϕ_ω preserves the mediant, so that $\overline{\phi_\omega(\ell)}$ joins $\phi_\omega(v_1)$ with $\phi_\omega(v_0 \oplus v_2) = \phi_\omega(v_0) \oplus \phi_\omega(v_2)$. Thus $\phi_{\omega 0}(\bar{\triangle})$, $\phi_{\omega 1}(\bar{\triangle}) \in \mathscr{P}_{n+1}$. This proves $\{\triangle_\omega \,:\, |\omega| = n+1\} \subseteq \mathscr{P}_{n+1}$, and the two sets are in fact the same since they have the same cardinality. It remains to show that the labelling of the vertices of $\phi_{\omega 0}(\bar{\triangle})$ and $\phi_{\omega 1}(\bar{\triangle})$ according to the definition of $\mathscr{P}_{n+1}$ is induced by $\phi_{\omega 0}$ and $\phi_{\omega 1}$. This immediately follows by computing the images of the vertices of $\bar{\triangle}$ under $\phi_{\omega 0}$ and $\phi_{\omega 1}$.

(ii) From (i) we have that $\left\{\overline{\phi_\omega(\ell)} \,:\, |\omega| = n\right\}$ contains the line segments needed to pass from $\mathscr{P}_n$ to $\mathscr{P}_{n+1}$. $\square$

In light of (ii) of the previous Lemma we have, for $n \geq -1$,

$$\mathcal{S}_{n+1} = \bigcup_{\mathfrak{S} \in \mathscr{S}_{n+1}} (\mathfrak{S} \cap \mathcal{S}_n)^\oplus = \bigcup_{\mathfrak{S} \in \mathscr{S}_0} (\mathfrak{S} \cap \mathcal{S}_n)^\oplus \cup \bigcup_{|\omega| \leq n} (\overline{\phi_\omega(\ell)} \cap \mathcal{S}_n)^\oplus.$$

By Lemma 5.3 and by the characterisation of the levels of the Farey tree in terms of Stern-Brocot sets, it is easy to verify that

$$\bigcup_{\mathfrak{S} \in \mathscr{S}_0} (\mathfrak{S} \cap \mathcal{S}_n)^\oplus = \mathcal{B}_{\leq n+1}.$$

Hence

$$\mathcal{S}_{n+1} = \mathcal{B}_{\leq n+1} \cup \bigcup_{|\omega| \leq n} (\overline{\phi_\omega(\ell)} \cap \mathcal{S}_n)^\oplus, \tag{5.3}$$

and this leads us towards studying the interior points of the counterimages tree in order to prove that the two trees coincide level by level.

Proposition 5.9. *For $n \geq 1$, the following properties hold.*

(i) *$\ell \cap \mathcal{I}_n = \phi_1(\{x = 1\} \cap \mathcal{B}_{n-1})$ and $\#(\ell \cap \mathcal{I}_n) = 2^{n-1}$.*

(ii) *Let $\mathfrak{p}$ and $\mathfrak{q}$ be the two endpoints of ℓ, then*

$$\{\mathfrak{p}, \mathfrak{q}\} \cup (\ell \cap \mathcal{I}_{\leq n}) = \left(\{\mathfrak{p}, \mathfrak{q}\} \cup (\ell \cap \mathcal{I}_{\leq n-1})\right)^\oplus,$$

that is, the interior points up to level n on ℓ are obtained from those up to level $n-1$ by inserting mediants of neighbouring points, where here the endpoints of ℓ are also included.

Let $\omega \in \{0,1\}^*$ be a binary word of finite length $|\omega| \le n-1$. Then the following properties hold.

(iii) $\phi_\omega(\ell) \cap \mathcal{I}_n = \phi_\omega\left(\ell \cap \mathcal{I}_{n-|\omega|}\right)$ *and* $\#\left(\phi_\omega(\ell) \cap \mathcal{I}_n\right) = 2^{n-|\omega|-1}$.

(iv) *Let* $\mathfrak{p}_\omega := \phi_\omega(\mathfrak{p})$ *and* $\mathfrak{q}_\omega := \phi_\omega(\mathfrak{q})$ *be the two endpoints of* $\phi_\omega(\ell)$. *Then*

$$\{\mathfrak{p}_\omega, \mathfrak{q}_\omega\} \cup \left(\phi_\omega(\ell) \cap \bigcup_{k=|\omega|+1}^{n} \mathcal{I}_k\right) = \left(\{\mathfrak{p}_\omega, \mathfrak{q}_\omega\} \cup \left(\phi_\omega(\ell) \cap \bigcup_{k=|\omega|+1}^{n-1} \mathcal{I}_k\right)\right)^{\oplus}.$$

Proof. (i) The function ϕ_1 bijectively maps the open vertical side $\{(1,y) \,:\, 0<y<1\}$ of $\bar{\triangle}$ onto ℓ. Moreover, since $n \ge 1$, by (R4) ϕ_1 sends points of $\mathcal{B}_{n-1}$ to points in $\mathcal{I}_n$. The cardinality computation immediately follows from the first part and Lemma 5.3.

(ii) From Lemma 5.3 we know that, for $n \ge 1$, the set $\{x=1\} \cap \mathcal{B}_{\le n-1}$ corresponds to the Stern-Brocot set of level $n-1$. More precisely, $\{x=1\} \cap \mathcal{B}_{\le n-1} = \{(1,y) \,:\, y \in \mathcal{F}_{n-1}\}$. Thus on the vertical side $\{x=1\}$ of $\bar{\triangle}$, the points up to level $n-1$ are obtained from those up to level $n-2$ by taking mediants between neighbouring points. In other words

$$\{x=1\} \cap \mathcal{B}_{\le n-1} = \left(\{x=1\} \cap \mathcal{B}_{\le n-2}\right)^{\oplus}.$$

We now apply ϕ_1 to both sides of the previous equality, getting

$$\phi_1\left(\{x=1\} \cap \mathcal{B}_{\le n-1}\right) = \bigcup_{k=-1}^{n-1} \phi_1\left(\{x=1\} \cap \mathcal{B}_k\right) \overset{\text{(i)}}{=} \{\mathfrak{p}, \mathfrak{q}\} \cup \left(\ell \cap \mathcal{I}_{\le n}\right)$$

and, using (5.2),

$$\phi_1\left(\left(\{x=1\} \cap \mathcal{B}_{\le n-2}\right)^{\oplus}\right) = \left(\phi_1\left(\{x=1\} \cap \mathcal{B}_{\le n-2}\right)\right)^{\oplus} = \left(\{\mathfrak{p}, \mathfrak{q}\} \cup \left(\ell \cap \mathcal{I}_{\le n-1}\right)\right)^{\oplus}.$$

(iii) In case $|\omega| = 0$ the first part is trivial and the second one has been proved in (i). Thus we can consider $1 \le |\omega| \le n-1$. The function ϕ_ω bijectively maps the open segment ℓ onto $\phi_\omega(\ell)$. Moreover, by applying (R4) $|\omega|$ times, ϕ_ω maps points of $\mathcal{I}_{n-|\omega|}$ to points in $\mathcal{I}_n$. For the second part, (i) implies that $\#\left(\phi_\omega(\ell) \cap \mathcal{I}_n\right) = \#\left(\ell \cap \mathcal{I}_{n-|\omega|}\right) = 2^{n-|\omega|-1}$.

(iv) From (ii) we know that for all $n \ge 1$ the interior points up to level n on ℓ are obtained from those up to level $n-1$ by inserting mediants of neighbouring points, also considering the endpoints of ℓ. Since ϕ_ω preserves mediants, we can conclude applying ϕ_ω to both sides of the equality in (ii). □

The above proposition characterises the location in $\bar{\triangle}$ of the interior points of our tree. In particular, it holds that

$$\mathcal{I}_n = \bigcup_{|\omega| \le n-1} \left(\phi_\omega(\ell) \cap \mathcal{I}_n\right),$$

that is, the interior points of level n are located along the backward images of ℓ under compositions of ϕ_0 and ϕ_1 of length $\le n-1$. To prove this, note that the inclusion "$\supseteq$" is trivial and that the two sets have the same cardinality. Indeed, using Proposition 5.9-(iii),

$$\#\left(\bigcup_{|\omega| \le n-1} \left(\phi_\omega(\ell) \cap \mathcal{I}_n\right)\right) = \sum_{s=0}^{n-1} \sum_{|\omega|=s} \#\left(\phi_\omega(\ell) \cap \mathcal{I}_n\right) = \sum_{s=0}^{n-1} \sum_{|\omega|=s} 2^{n-s-1} = n 2^{n-1}.$$

As last step, we now write the set of the interior points up to level $n+1$ in a convenient way. For $n \ge 1$,

$$\begin{aligned} \mathcal{I}_{\le n+1} &= \bigcup_{|\omega| \le n} \left(\phi_\omega(\ell) \cap \mathcal{I}_{\le n+1}\right) = \bigcup_{|\omega| \le n} \left(\{\mathfrak{p}_\omega, \mathfrak{q}_\omega\} \cup \left(\phi_\omega(\ell) \cap \mathcal{I}_{\le n}\right)\right)^{\oplus} = \\ &= \bigcup_{|\omega| \le n} \left(\overline{\phi_\omega(\ell)} \cap \left(\{\mathfrak{p}_\omega, \mathfrak{q}_\omega\} \cup \mathcal{I}_{\le n}\right)\right)^{\oplus}. \end{aligned} \tag{5.4}$$

Theorem 5.10. *For all $n \geq 0$ we have $\mathcal{T}_n = \mathcal{S}_n \setminus \mathcal{S}_{n-1}$, that is the tree defined by counterimages and that defined by Farey sums coincide level by level.*

Proof. It suffices to show that, for all $n \geq 0$, $\mathcal{S}_n = \mathcal{T}_{\leq n} = \mathcal{B}_{\leq n} \cup \mathcal{I}_{\leq n}$. We argue by induction on $n \geq 0$. Since $\mathcal{S}_0 = \mathcal{T}_{-1} \cup \mathcal{T}_0$, the base case is proved. Now suppose that, for a certain $n \geq 0$, $\mathcal{S}_n = \mathcal{B}_{\leq n} \cup \mathcal{I}_{\leq n}$. By applying the inductive hypothesis we have

$$\mathcal{S}_{n+1} \overset{(5.3)}{=} \mathcal{B}_{\leq n+1} \cup \bigcup_{|\omega| \leq n} (\overline{\phi_\omega(\ell)} \cap \mathcal{S}_n)^{\oplus} = \mathcal{B}_{\leq n+1} \cup \bigcup_{|\omega| \leq n} \left(\overline{\phi_\omega(\ell)} \cap (\mathcal{B}_{\leq n} \cup \mathcal{I}_{\leq n})\right)^{\oplus} =$$

$$= \mathcal{B}_{\leq n+1} \cup \bigcup_{|\omega| \leq n} \left(\overline{\phi_\omega(\ell)} \cap (\{\mathfrak{p}_\omega, \mathfrak{q}_\omega\} \cup \mathcal{I}_{\leq n})\right)^{\oplus},$$

where the last equality holds since, along $\overline{\phi_\omega(\ell)}$, the points are all in $\mathcal{I}_{\leq n}$, with the possible exception of the endpoints of $\phi_\omega(\ell)$, namely $\mathfrak{p}_\omega$ and $\mathfrak{q}_\omega$. Equation (5.4) allows us to conclude that $\mathcal{S}_{n+1} = \mathcal{B}_{\leq n+1} \cup \mathcal{I}_{\leq n+1}$, and the inductive step is proved. □

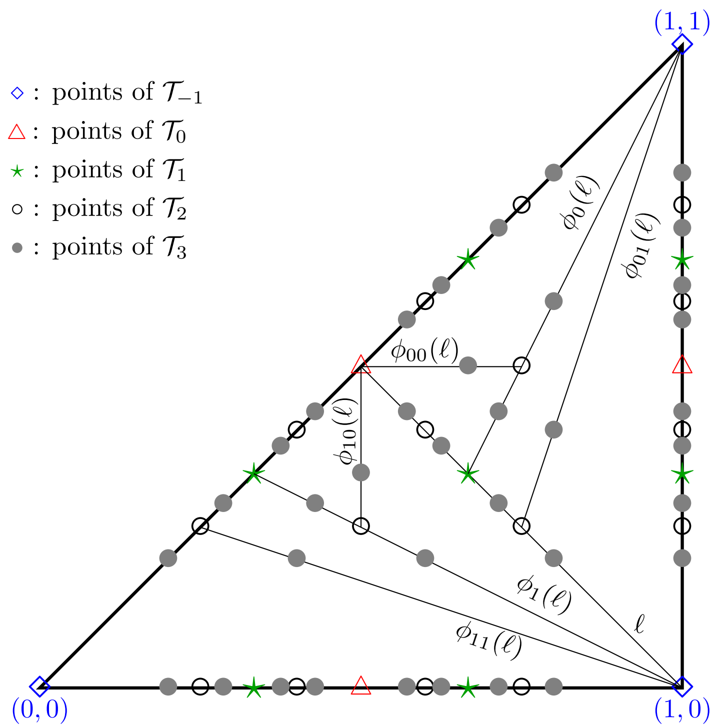

FIGURE 8. The first four levels of the tree generated through the local inverses of the map $\tilde{S}$ represented as set of points of $\overline{\triangle}$.

Remark 5.11. At this point it is natural to ask, given the map S and the tree defined here, whether or not a version of the Minkowski question mark function could be defined in this setting. We recall, briefly, that the original Minkowski question mark function was introduced as another way of demonstrating the Lagrange property of continued fractions, in that it maps every rational number to the subset of dyadic rationals (that is, those having denominators containing only powers of 2) and every quadratic irrational to the remaining rational numbers (see [17, 12], and for other 1-dimensional analogues, [18, 19]). These functions are now known as *slippery Devil's staircases* for the fact that they are strictly increasing but nevertheless singular with respect to the Lebesgue measure.

This natural question has been studied in [8] for the map S and many possible generalisations. A higher dimensional version of the Minkowski function for a different map has been introduced in [21].

Appendix A. Some results on the local inverses of V

In this appendix we prove some properties of the local inverses of the map V, needed for the argument of Section 3.1. We recall that V is the induced map of S on the set $A = \{(x, y) \in \Gamma_0 \,:\, S(x, y) \in \Gamma_0\}$, and that each local inverse of V is a linear fractional map, as in (3.3). In general, a linear fractional map ψ of the form

$$\psi(x, y) = \left(\frac{r_1 + s_1 x + t_1 y}{r + sx + ty},\ \frac{r_2 + s_2 x + t_2 y}{r + sx + ty} \right),$$

where the cofficients are non-negative integers, can be expressed in projective coordinates by the 3×3 matrix

$$M_\psi := \begin{pmatrix} r & s & t \\ r_1 & s_1 & t_1 \\ r_2 & s_2 & t_2 \end{pmatrix}$$

by associating a point $(\frac{x}{z}, \frac{y}{z}) \in \mathbb{R}^2$ to a vector $v = (z,\, x,\, y)^t$, so that $\psi\left(\frac{x}{z}, \frac{y}{z}\right)$ is associated to the vector $M_\psi v$. For instance, the two inverse maps ϕ_0 and ϕ_1 have matrices

$$M_0 := M_{\phi_0} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad M_1 := M_{\phi_1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that the composition of linear fractional maps translates into the left multiplication of their matrices. As a consequence, since both M_0 and M_1 have unit determinant, every product involving these two matrices also has unit determinant. To every linear fractional map as above, we associate the vectors $v_1(\psi)$, $v_2(\psi)$, $v(\psi) \in \mathbb{R}^3$ corresponding to the rows of the associated matrix M_ψ. In other words,

$$v_1(\psi) = \begin{pmatrix} r_1 \\ s_1 \\ t_1 \end{pmatrix}, \qquad v_2(\psi) = \begin{pmatrix} r_2 \\ s_2 \\ t_2 \end{pmatrix}, \qquad v(\psi) = \begin{pmatrix} r \\ s \\ t \end{pmatrix}.$$

In what follows, we use the notation $\|\cdot\|$ for the Euclidean norm and $\|\cdot\|_1$ for the 1-norm on $\mathbb{R}^3$. The two norms are equivalent and, in particular, for all $v \in \mathbb{R}^3$ holds

$$\|v\| \le \|v\|_1 \le \sqrt{3}\|v\|. \tag{A.1}$$

Moreover, to each $v \in \mathbb{R}^3 \setminus \{0\}$ with non-negative components we associate the normalised vector P_v given by $P_v := \frac{v}{\|v\|_1}$.

Lemma A.1. *For any $v, w \in \mathbb{R}^3 \setminus \{0\}$ with non-negative components, $\|v \times w\| \le \sqrt{3}\|v\|\|w\|\|P_v - P_w\|$.*

Proof. Let $\theta_{v,w} \in \left[0, \frac{\pi}{2}\right]$ be the angle between the two vectors v and w, so that $\|v \times w\| = \|v\|\|w\| \sin\theta_{v,w}$, and between the two vectors P_v and P_w. Let $\lambda_{w,u} := \|P_w - (P_w \cdot P_v)P_v\|$, the modulus of the component of P_w orthogonal to P_v (see Figure 9). Then by simple geometric considerations we have

$$\sin\theta_{v,w} = \frac{\lambda_{w,v}}{\|P_w\|} = \frac{\lambda_{w,v}}{\|w\|}\|w\|_1 \overset{(A.1)}{\le} \sqrt{3}\lambda_{w,v} \le \sqrt{3}\|P_v - P_w\|,$$

□

Lemma A.2. *Let Φ be an arbitrary composition of the maps ϕ_0 and ϕ_1. Then the matrix M_Φ satisfies $\|v(\Phi)\|_1 \ge \|v_1(\Phi)\|_1$ and $\|v(\Phi)\|_1 \ge \|v_2(\Phi)\|_1$.*

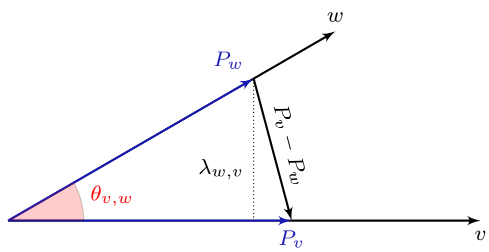

FIGURE 9. Graphical representation of vectors P_v and P_w, along with the quantities involved in the proof of Lemma A.1.

Proof. We argue by induction on the length $l \geq 1$ of Φ as a composition of maps. If $l = 1$, then Φ is either ϕ_0 or ϕ_1, and in both cases the thesis is true. For the inductive step, let $l \geq 1$ and suppose that the thesis is true for a certain Φ of length l. Let

$$M_\Phi = \begin{pmatrix} r & s & t \\ r_1 & s_1 & t_1 \\ r_2 & s_2 & t_2 \end{pmatrix}$$

be the matrix of Φ. We have

$$M_{\phi_0 \circ \Phi} = M_0 M_\Phi = \begin{pmatrix} r + r_2 & s + s_2 & t + t_2 \\ r & s & t \\ r_1 & s_1 & t_1 \end{pmatrix} \quad \text{and} \quad M_{\phi_1 \circ \Phi} = M_1 M_\Phi = \begin{pmatrix} r + r_2 & s + s_2 & t + t_2 \\ r_1 & s_1 & t_1 \\ r_2 & s_2 & t_2 \end{pmatrix}.$$

For the first matrix it holds that

$$\|v(\phi_0 \circ \Phi)\|_1 = r + s + t + r_2 + s_2 + t_2 \geq r + s + t = \|v_1(\phi_0 \circ \Phi)\|_1$$

since r_2, s_2, $t_2 \geq 0$, and that

$$\|v(\phi_0 \circ \Phi)\|_1 \geq r + s + t \geq r_1 + s_1 + t_1 = \|v_2(\phi_0 \circ \Phi)\|_1$$

by the inductive assumption. Analogous estimates hold for $M_{\phi_1 \circ \Phi}$. □

Lemma A.3. *Let Φ be an arbitrary composition of the maps ϕ_0 and ϕ_1.*

(i) *If $D\Phi$ denotes the Jacobian matrix of Φ, then*

$$\max \left\{ \sup_A \left(|(D\Phi)_{11}| + |(D\Phi)_{21}| \right), \sup_A \left(|(D\Phi)_{12}| + |(D\Phi)_{22}| \right) \right\} \leq$$
$$\leq 27\sqrt{3} \left(\left\| P_{v(\Phi)} - P_{v_1(\Phi)} \right\| + \left\| P_{v(\Phi)} - P_{v_2(\Phi)} \right\| \right).$$

(ii) *For $k = 0, 1$, let D_k be the Jacobian matrix of $\phi_k \circ \Phi$, then*

$$\max \left\{ \sup_A \left(|(D_k)_{11}| + |(D_k)_{21}| \right), \sup_A \left(|(D_k)_{12}| + |(D_k)_{22}| \right) \right\} \leq$$
$$\leq 27\sqrt{3} \left(\left\| P_{v(\Phi)+v_2(\Phi)} - P_{v_1(\Phi)} \right\| + \left\| P_{v(\Phi)+v_2(\Phi)} - P_{v_2(\Phi)} \right\| \right).$$

In particular the worst case is realised for $k = 0$.

(iii) *We have that*

$$\left\| P_{v(\Phi)+v_2(\Phi)} - P_{v_1(\Phi)} \right\| + \left\| P_{v(\Phi)+v_2(\Phi)} - P_{v_2(\Phi)} \right\| \leq \left\| P_{v(\Phi)} - P_{v_1(\Phi)} \right\| + \left\| P_{v(\Phi)} - P_{v_2(\Phi)} \right\|.$$

Proof. (i) The map Φ is a composition of ϕ_0 and ϕ_1, thus of the form (3.4). Since we are interested in the local inverses of V, we look at the supremum of the Jacobian matrix of Φ on A. Hence we have

$$\begin{aligned}\sup_A \left(|(D\Phi)_{11}| + |(D\Phi)_{21}|\right) &\le 9 \cdot \frac{|r_1 s - r s_1| + |s_1 t - s t_1| + |r_2 s - r s_2| + |s_2 t - s t_2|}{(r+s+t)^2} \le \\ &\le 9 \cdot \frac{\|v(\Phi)\times v_1(\Phi)\|_1 + \|v(\Phi)\times v_2(\Phi)\|_1}{\|v(\Phi)\|^2} \le \\ &\overset{\text{(A.1)}}{\le} 9\sqrt{3}\cdot \frac{\|v(\Phi)\times v_1(\Phi)\| + \|v(\Phi)\times v_2(\Phi)\|}{\|v(\Phi)\|^2} \le \\ &\overset{\text{Lem. A.1}}{\le} 27 \cdot \frac{\|v_1(\Phi)\| \left\|P_{v(\Phi)} - P_{v_1(\Phi)}\right\| + \|v_2(\Phi)\| \left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\|}{\|v(\Phi)\|} \le \\ &\overset{\text{(A.1)}}{\le} 27\sqrt{3}\cdot \frac{\|v_1(\Phi)\|_1 \left\|P_{v(\Phi)} - P_{v_1(\Phi)}\right\| + \|v_2(\Phi)\|_1 \left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\|}{\|v(\Phi)\|_1}.\end{aligned}$$

From Lemma A.2 we have $\|v_1(\Phi)\|_1 \le \|v(\Phi)\|_1$ and $\|v_2(\Phi)\|_1 \le \|v(\Phi)\|_1$, so that

$$\sup_A \left(|(D\Phi)_{11}| + |(D\Phi)_{21}|\right) \le 27\sqrt{3}\cdot \left(\left\|P_{v(\Phi)} - P_{v_1(\Phi)}\right\| + \left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\|\right)$$

The same estimate holds for $\sup_A \left(|(D\Phi)_{12}| + |(D\Phi)_{22}|\right)$ and thus (i) is proved.

(ii) The matrices associated to the maps $\phi_k \circ \Phi$ for $k = 0, 1$ are

$$M_0 M_\Phi = \begin{pmatrix} v(\Phi) + v_2(\Phi) \\ v(\Phi) \\ v_1(\Phi) \end{pmatrix} \quad \text{and} \quad M_1 M_\psi = \begin{pmatrix} v(\Phi) + v_2(\Phi) \\ v_1(\Phi) \\ v_2(\Phi) \end{pmatrix}.$$

Applying (i) to the map $\phi_k \circ \Phi$ we have

$$\begin{aligned}\max\left\{\sup_A \left(|(D_0)_{11}| + |(D_0)_{21}|\right), \sup_A \left(|(D_0)_{12}| + |(D_0)_{22}|\right)\right\} &\le \\ \le 27\sqrt{3}\left(\left\|P_{v(\Phi)+v_2(\Phi)} - P_{v(\Phi)}\right\| + \left\|P_{v(\Phi)+v_2(\Phi)} - P_{v_1(\Phi)}\right\|\right)\end{aligned}$$

and

$$\begin{aligned}\max\left\{\sup_A \left(|(D_1)_{11}| + |(D_1)_{21}|\right), \sup_A \left(|(D_1)_{12}| + |(D_1)_{22}|\right)\right\} &\le \\ \le 27\sqrt{3}\left(\left\|P_{v(\Phi)+v_2(\Phi)} - P_{v_1(\Phi)}\right\| + \left\|P_{v(\Phi)+v_2(\Phi)} - P_{v_2(\Phi)}\right\|\right).\end{aligned}$$

To finish the proof it suffices to show that

$$\left\|P_{v(\Phi)+v_2(\Phi)} - P_{v(\Phi)}\right\| \le \left\|P_{v(\Phi)+v_2(\Phi)} - P_{v_2(\Phi)}\right\|. \tag{A.2}$$

To this end, note that $P_{v(\Phi)+v_2(\Phi)}$ is a convex combination of $P_{v(\Phi)}$ and $P_{v_2(\Phi)}$, in particular

$$P_{v(\Phi)+v_2(\Phi)} = \frac{\|v(\Phi)\|_1}{\|v(\Phi)\|_1 + \|v_2(\Phi)\|_1} P_{v(\Phi)} + \frac{\|v_2(\Phi)\|_1}{\|v(\Phi)\|_1 + \|v_2(\Phi)\|_1} P_{v_2(\Phi)},$$

and since from Lemma A.2 we have $\|v_2(\Phi)\|_1 \le \|v(\Phi)\|_1$, (A.2) easily follows.

(iii) The three points $P_{v(\Phi)}$, $P_{v_1(\Phi)}$ and $P_{v_2(\Phi)}$ belong to the standard 2-symplex in $\mathbb{R}^3$ and define a triangle $\triangle_\Phi$ since they are linearly independent. Furthermore, (ii) implies that $P_{v(\Phi)+v_2(\Phi)} = \lambda P_{v(\Phi)} + (1-\lambda) P_{v_2(\Phi)}$ for some $\lambda \in \left(\frac{1}{2}, 1\right)$. Also $P_{v(\Phi)+v_2(\Phi)}$, $P_{v_1(\Phi)}$ and $P_{v_2(\Phi)}$ define a triangle $\triangle'_\Phi$, which is a subtriangle of $\triangle_\Phi$. In particular, $\triangle_\Phi$ and $\triangle'_\Phi$ have a common side and the non-common vertex $P_{v(\Phi)+v_2(\Phi)}$ belongs to the side of $\triangle_\Phi$ with vertices $P_{v(\Phi)}$ and $P_{v_2(\Phi)}$. The inequality to prove easily follows from this geometric interpretation, since perimeter of the subtriangle $\triangle'_\Phi$ is less than or equal to the perimeter of $\triangle_\Phi$. Besides this geometrical approach, an analytic estimate easily follows from the triangle inequality:

$$\begin{aligned}\left\|P_{v(\Phi)+v_2(\Phi)} - P_{v_1(\Phi)}\right\| + \left\|P_{v(\Phi)+v_2(\Phi)} - P_{v_2(\Phi)}\right\| =&\\ &= \left\|\lambda P_{v(\Phi)} + (1-\lambda)P_{v_2(\Phi)} - P_{v_1(\Phi)}\right\| + \lambda\left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\| =\\ &= \left\|-(1-\lambda)(P_{v(\Phi)} - P_{v_2(\Phi)}) + P_{v(\Phi)} - P_{v_1(\Phi)}\right\| + \lambda\left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\| \le\\ &\le (1-\lambda)\left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\| + \left\|P_{v(\Phi)} - P_{v_1(\Phi)}\right\| + \lambda\left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\| =\\ &= \left\|P_{v(\Phi)} - P_{v_1(\Phi)}\right\| + \left\|P_{v(\Phi)} - P_{v_2(\Phi)}\right\|.\end{aligned}$$

□

Lemma A.4. *Let $\psi_{i_1,\dots,i_n} : A \to C_{i_1,\dots,i_n}$ be a local inverse of V^n. Then*

$$\left\|P_{v(\psi_{i_1,\dots,i_n})} - P_{v_1(\psi_{i_1,\dots,i_n})}\right\| + \left\|P_{v(\psi_{i_1,\dots,i_n})} - P_{v_2(\psi_{i_1,\dots,i_n})}\right\| \le \tilde{d}(n)$$

where

$$\tilde{d}(n) \coloneqq \left\|P_{v(\phi_0^n)} - P_{v_1(\phi_0^n)}\right\| + \left\|P_{v(\phi_0^n)} - P_{v_2(\phi_0^n)}\right\|.$$

Proof. As outlined in Section 3.1, $\psi_{i_1,\dots,i_n} = \psi_{i_1} \circ \cdots \circ \psi_{i_n}$ and, for $h = 1, \dots, n$, $\psi_{i_h} = \phi_0 \circ \Phi_{i_h}$, where Φ_{i_h} is empty or a composition of the maps ϕ_0 and ϕ_1 beginning with ϕ_0. In Proposition A.3 (iii) we proved that the estimation function introduced in (i) is decreasing with respect to the number of compositions of the maps ϕ_0 or ϕ_1. The inequality of this lemma follows, since ϕ_0^n contains the least possible number of compositions of the maps ϕ_0 and ϕ_1 compatible with the definition of the local inverses, and by Proposition A.3-(ii) realises the worst case. □

Lemma A.5. *Let $(\tilde{d}(n))_{n\ge 0}$ be the real sequence introduced in Lemma A.4. Then* $\lim_{n\to+\infty} \tilde{d}(n) = 0$.

Proof. Arguing by induction on $n \ge 0$, it is easy to prove that

$$M_{\phi_0^n} = M_0^n = \begin{pmatrix} f_{n+4} & f_{n+2} & f_{n+3} \\ f_{n+3} & f_{n+1} & f_{n+2} \\ f_{n+2} & f_n & f_{n+1} \end{pmatrix},$$

where $(f_n)_{n\ge 0}$ is recursively defined to be

$$\begin{cases} f_0 = 0 \\ f_1 = 1 \\ f_2 = 0 \\ f_{n+3} = f_{n+2} + f_n & \text{for } n \ge 1 \end{cases}.$$

The sequence $(\nu_n)_{n\ge 0}$, $\nu_n \coloneqq f_{n+4}$, is also referred to as the Narayana's cows sequence. It is known that this sequence has a ratio limit, *i.e.* there exists $\lim_{n\to\infty} \frac{\nu_{n+1}}{\nu_n} =: \gamma < +\infty$ [25][5]. Note that for $n \ge 4$ and for $r \ge 1$ we have $\frac{f_{n+r}}{f_n} = \prod_{j=0}^{r-1} \frac{f_{n+1+j}}{f_{n+j}}$, so that

$$\lim_{n\to+\infty} \frac{f_{n+r}}{f_n} = \gamma^r. \tag{A.3}$$

[5]More precisely, γ is the only real root of the characteristic equation $x^3 - x^2 - 1 = 0$.

For $n \geq 0$ we have

$$P_{v_1(\phi_0^n)} = \frac{1}{f_{n+1}+f_{n+2}+f_{n+3}} \begin{pmatrix} f_{n+3} \\ f_{n+1} \\ f_{n+2} \end{pmatrix} = \frac{1}{f_{n+5}} \begin{pmatrix} f_{n+3} \\ f_{n+1} \\ f_{n+2} \end{pmatrix}$$

$$P_{v_2(\phi_0^n)} = \frac{1}{f_n+f_{n+1}+f_{n+2}} \begin{pmatrix} f_{n+2} \\ f_n \\ f_{n+1} \end{pmatrix} = \frac{1}{f_{n+4}} \begin{pmatrix} f_{n+2} \\ f_n \\ f_{n+1} \end{pmatrix},$$

$$P_{v(\phi_0^n)} = \frac{1}{f_{n+2}+f_{n+3}+f_{n+4}} \begin{pmatrix} f_{n+4} \\ f_{n+2} \\ f_{n+3} \end{pmatrix} = \frac{1}{f_{n+6}} \begin{pmatrix} f_{n+4} \\ f_{n+2} \\ f_{n+3} \end{pmatrix}$$

so that

$$\tilde{d}(n) \leq \left\| P_{v(\phi_0^n)} - P_{v_1(\phi_0^n)} \right\|_1 + \left\| P_{v(\phi_0^n)} - P_{v_2(\phi_0^n)} \right\|_1 = \\ = \sum_{k=0}^{2} \left(\left| \frac{f_{n+k+2}}{f_{n+6}} - \frac{f_{n+k+1}}{f_{n+5}} \right| + \left| \frac{f_{n+k+2}}{f_{n+6}} - \frac{f_{n+k}}{f_{n+4}} \right| \right).$$

Using (A.3), for each $k = 0,\, 1,\, 2$ we have $\frac{f_{n+k+2}}{f_{n+6}} - \frac{f_{n+k+1}}{f_{n+5}} \to \gamma^{4-k} - \gamma^{4-k} = 0$ and analogously $\frac{f_{n+k+2}}{f_{n+6}} - \frac{f_{n+k}}{f_{n+4}} \to 0$ as $n \to +\infty$. This proves that $\lim_{n \to +\infty} \tilde{d}(n) = 0$. ☐

Proof of Proposition 3.9. It follows directly from Lemma A.3-(i), A.4 and A.5 with $d(n) = 27\sqrt{3}\,\tilde{d}(n)$. ☐

Appendix B. The wandering rate of the set A

The set A is defined in (3.2) and it is the triangle with vertices $Q_1 = (\frac{1}{2}, \frac{1}{2})$, $Q_2 = (\frac{2}{3}, \frac{1}{3})$ and $Q_3 = (1, 1)$, with the sides Q_1Q_2 and Q_2Q_3 not included. We consider the wandering rate $w_n(A)$ for $n \geq 1$, which is defined to be

$$w_n(A) \coloneqq \sum_{k=0}^{n-1} \mu(A \cap \{\varphi > k\}),$$

where φ is the first-return time function in A. Extending the function φ to all $\bar{\triangle}$ by

$$\varphi(x, y) \coloneqq \min \{n \geq 1 \,:\, S^n(x, y) \in A\}$$

we obtain the *hitting time* function of A, which is well-defined and finite μ-almost everywhere since the system $(\bar{\triangle}, \mu, S)$ is conservative and ergodic. We now recall that, for $k \geq 1$,

$$\mu(A \cap \{\varphi > k\}) = \mu(A^{\complement} \cap \{\varphi = k\}),$$

where $A^{\complement} \coloneqq \bar{\triangle} \setminus A$ [26, Lemma 1]. We thus study the diverging sequence $\sum_{k=1}^{n} \mu(A^{\complement} \cap \{\varphi = k\})$. The first step is to study the structure of $A^{\complement} \cap \{\varphi = k\}$ for $k \geq 1$, the set of points in $A^{\complement}$ which hit A for the first time after exactly k iterations of the map S. This set can be expressed in terms of the local inverse of S as follows. Let

$$\Omega_k \coloneqq \left\{ \omega \in \{0,1\}^k \,:\, \pi_{\omega_i \omega_{i+1}} = 1 \ \forall i = 0,\, \ldots,\, k-2,\ \omega_{k-1} = 1 \right\}, \quad \text{where } \Pi = (\pi_{ij})_{i,j=0,1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

In this way, Ω_k is the set of binary words of length k, which all end with a "1", and in which the string "00" never appears. Then

$$A^{\complement} \cap \{\varphi = k\} = \bigcup_{\omega \in \Omega_k} \phi_\omega(A) = \bigcup_{\omega \in \Omega_k} \phi_{\omega_0} \circ \phi_{\omega_1} \circ \cdots \circ \phi_{\omega_{k-2}} \circ \phi_1(A).$$

Indeed, a point in $\phi_{\omega_0} \circ \phi_{\omega_1} \circ \cdots \circ \phi_{\omega_{k-2}} \circ \phi_1(A)$ has symbolic code given by $\omega_0\omega_1 \ldots \omega_{k-2}100$ with the word "00" not appearing in the first k symbols. This is equivalent to saying that such a point does not visit A in the first $k-1$ iterations, hence the point is in $A^{\complement} \cap \{\varphi = k\}$. The converse also obviously holds. Note that, in case $k = 1$, we have $A^{\complement} \cap \{\varphi = 1\} = \phi_1(A)$.

We first obtain an estimate from above for the wandering rate. In what follows, we write $a_n \lesssim b_n$ if and only if $a_n = O(b_n)$.

Proposition B.1. *The wandering rate $w_n(A)$ satisfies $w_n(A) \lesssim \log^2 n$.*

Proof. Using the properties of the map S and its local inverses, one immediately verifies that

$$\bigcup_{k=1}^{n} \left(A^{\complement} \cap \{\varphi = k\} \right) \subseteq \bigcup_{k=0}^{n-1} \triangle_k,$$

where $\{\triangle_k\}_{k\geq 0}$ is the partition represented in Figure 1. Hence

$$w_n(A) \leq \sum_{k=0}^{n-1} \mu(\triangle_k).$$

Using now the dynamical system defined in Section 2.2 on the strip Σ, we have $\mu(\triangle_k) = \rho(\Sigma_k)$ for all $k \geq 0$, so that

$$w_n(A) \leq \sum_{k=0}^{n-1} \mu(\triangle_k) = \sum_{k=0}^{n-1} \rho(\Sigma_k) = \sum_{k=0}^{n-1} \int_k^{k+1} \left(\int_0^1 \frac{1}{1+uv} du \right) dv = \int_0^n \frac{\log(1+v)}{v} dv \lesssim \log^2 n.$$

□

To obtain an estimate from below, we use the matrix representation of the local inverses defined in Appendix A.

Lemma B.2. *For a map $\psi = \phi_{\omega_0} \circ \phi_{\omega_1} \circ \cdots \circ \phi_{\omega_{k-2}} \circ \phi_1$ with matrix representation*

$$M_\psi = \begin{pmatrix} r & s & t \\ r_1 & s_1 & t_1 \\ r_2 & s_2 & t_2 \end{pmatrix}$$

it holds that

$$\frac{m(A)}{(r_1+s_1+t_1)(r_2+s_2+t_2)(r+s+t)} \leq \mu(\psi(A)) \leq \frac{27 m(A)}{(r_1+s_1+t_1)(r_2+s_2+t_2)(r+s+t)}$$

Proof. By definition of μ, denoting $\psi(x,y) = (\psi_1(x,y), \psi_2(x,y))$,

$$\mu(\psi(A)) = \iint_{\psi(A)} \frac{1}{xy} dxdy = \iint_A \frac{1}{\psi_1(x,y)\psi_2(x,y)} \left| J\psi(x,y) \right| dxdy.$$

Moreover by Proposition 3.10, we have

$$\psi_1(x,y) = \frac{r_1 + s_1 x + t_1 y}{r + sx + ty}, \quad \psi_2(x,y) = \frac{r_2 + s_2 x + t_2 y}{r + sx + ty}, \quad J\psi(x,y) = \frac{1}{(r+sx+ty)^3}$$

hence

$$\mu(\psi(A)) = \iint_A \frac{1}{(r_1+s_1x+t_1y)(r_2+s_2x+t_2y)(r+sx+ty)} dxdy$$

Since for $(x,y) \in A$ we can use $\frac{1}{2} \leq x \leq 1$ and $\frac{1}{3} \leq y \leq 1$, the proof is complete. □

We are then led to study the terms

$$t_{\omega_0\omega_1\ldots\omega_{k-2}1} := \frac{1}{(r_1+s_1+t_1)(r_2+s_2+t_2)(r+s+t)} \tag{B.1}$$

for the maps $\psi = \phi_{\omega_0} \circ \phi_{\omega_1} \circ \cdots \circ \phi_{\omega_{k-2}} \circ \phi_1$ with $\omega \in \Omega_k$. We shall also write simply t_ψ to shorten the notation. Thus we consider the sequence

$$\tau_n := \sum_{k=1}^{n} \sum_{\omega \in \Omega_k} t_{\omega_0\omega_1\ldots\omega_{k-2}1},$$

which by Lemma B.2 satisfies

$$m(A) \sum_{\omega\in\Omega_k} t_{\omega_0\omega_1\dots\omega_{k-2}1} \le \mu(A^{\complement} \cap \{\varphi = k\}) \le 27 m(A) \sum_{\omega\in\Omega_k} t_{\omega_0\omega_1\dots\omega_{k-2}1}. \tag{B.2}$$

and then

$$m(A)\,\tau_n \le w_n(A) \le 27\, m(A)\,\tau_n\,. \tag{B.3}$$

Moreover, given a linear fractional map ψ with matrix representation

$$M_\psi = \begin{pmatrix} r & s & t \\ r_1 & s_1 & t_1 \\ r_2 & s_2 & t_2 \end{pmatrix}$$

if we introduce the vector

$$V_\psi := \begin{pmatrix} r+s+t \\ r_1+s_1+t_1 \\ r_2+s_2+t_2 \end{pmatrix},$$

the term t_ψ in (B.1) is the inverse of the product of the components of V_ψ. We also use the notation t_{V_ψ} for t_ψ.

We now define a tree $\mathcal{V}$ of vectors, in such a way that the k-th level of $\mathcal{V}$ is associated to the set $A^{\complement} \cap \{\varphi = k\}$. We first make a small modification in order to simplify the argument. For each $k \ge 1$, we consider the subsets

$$\Phi_k := A^{\complement} \cap \{\varphi = k\} \cap \Gamma_1,$$

so that

$$\Phi_k = \bigcup_{\omega\in\Omega_k} \phi_1 \circ \phi_{\omega_1} \circ \cdots \circ \phi_{\omega_{k-2}} \circ \phi_1(A)$$

Obviously $\Phi_1 = A^{\complement} \cap \{\varphi = 1\} = \{\phi_1(A)\}$, whereas for example

$$\Phi_2 = \{\phi_1 \circ \phi_1(A)\} \subsetneq A^{\complement} \cap \{\varphi = 2\} = \{\phi_0 \circ \phi_1(A),\, \phi_1 \circ \phi_1(A)\}\,.$$

We are now ready to introduce the levels of our tree $\mathcal{V}$. For each $k \ge 1$ we define

$$L_k := \left\{ V_\psi \,:\, \psi = \phi_1 \circ \phi_{\omega_1} \circ \cdots \circ \phi_{\omega_{k-2}} \circ \phi_1 \right\} \quad \text{and} \quad \lambda_k := \sum_{V\in L_k} t_V,$$

where t_V is the inverse of the product of the components of the vector V. The k-th row of $\mathcal{V}$ is the set L_k. We have then associated two objects to each set $A^{\complement} \cap \{\varphi = 1\}$: the list of vectors L_k and the quantity λ_k. For instance, corresponding to Φ_1 we obtain

$$V_1 := V_{\phi_1} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

and $\lambda_1 = t_1 = \frac{1}{2}$. The vector V_1 is the root of our tree $\mathcal{V}$. Then

$$L_2 = \left\{ V_{\phi_1\circ\phi_1} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right\} \quad \text{and} \quad L_3 = \left\{ V_{\phi_1\circ\phi_1\circ\phi_1} = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}, V_{\phi_1\circ\phi_0\circ\phi_1} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} \right\},$$

as follows by writing the matrix representation of the involved maps. Furthermore, for the first rows, one easily finds $\lambda_2 = t_{V_{\phi_1\circ\phi_1}} = \frac{1}{3}$, $\lambda_3 = t_{V_{\phi_1\circ\phi_1\circ\phi_1}} + t_{V_{\phi_1\circ\phi_0\circ\phi_1}} = \frac{1}{4} + \frac{1}{8}$, and so on.

Moreover the tree $\mathcal{V}$ can be generated from the root vector V_1 by the following algorithm, without using the maps ψ. Let us consider the matrices

$$M_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad M_{10} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

that are the matrix representations of the maps ϕ_1 and $\phi_1 \circ \phi_0$ respectively. Let them act on the vectors of the tree to generate new vectors. When we apply M_1 to a vector $V \in L_k$, we obtain a vector in L_{k+1}, and when we apply M_{10} we obtain a vector in L_{k+2}. Hence, vectors in the k-th row of $\mathcal{V}$ are generated by applying M_1 to all vectors in the $(k-1)$-th row and M_{10} to all vectors in the $(k-2)$-th row. Applying this algorithm starting from $L_1 = \{V_1\}$, we immediately obtain for the first rows

$$L_2 = \left\{ M_1 V_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right\}$$

$$L_3 = \left\{ M_1(M_1 V_1) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix},\ M_{10} V_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} \right\},$$

as above.

Lemma B.3. *For $n \geq 1$ define*

$$\tilde{\tau}_n \coloneqq \sum_{k=1}^{n} \lambda_k = \sum_{k=1}^{n} \sum_{V \in L_k} t_V .$$

Then $\tilde{\tau}_n < \tau_n < \tilde{\tau}_n + \frac{\mu(\Gamma_0)}{m(A)}$ and $\tilde{\tau}_n \gtrsim \log^2 n$.

Proof. The difference between $\tilde{\tau}_n$ and τ_n is that for each $k = 1, \dots, n$, in $\tilde{\tau}_n$ we are not considering the terms t_{V_ψ} for the maps $\psi = \phi_{\omega_0} \circ \phi_{\omega_1} \circ \cdots \circ \phi_{\omega_{k-2}} \circ \phi_1$ with $\omega_0 = 0$. Recalling that for such maps $t_{V_\psi} \leq \frac{\mu(\psi(A))}{m(A)}$ by Lemma B.2, that $\psi(A) \subseteq \Gamma_0$ if $\omega_0 = 0$, and that the sets $\psi(A)$ are disjoint for different maps ψ by definition, for all $n \geq 1$ we have that

$$\tilde{\tau}_n < \tau_n < \tilde{\tau}_n + \frac{\mu(\Gamma_0)}{m(A)}.$$

We prove by induction that each row L_k with $k \geq 2$ contains the vectors

$$\begin{pmatrix} k+1 \\ j \\ 1 \end{pmatrix} \qquad j = 1, \dots, k-1.$$

By the definition of λ_k, this implies that $\tilde{\tau}_n \coloneqq \sum_{k=1}^n \lambda_k \geq \sum_{k=1}^n \frac{1}{k+1} \sum_{j=1}^{k-1} \frac{1}{j} \gtrsim \log^2 n$. For $k = 2$, the row L_2 contains only the vector $M_1 V_1$, and the base case is proved. Let us assume that the statement is true for $r = 2, \dots, k$, then using the algorithm to construct $\mathcal{V}$, we have that L_{k+1} contains the vectors

$$M_1 \begin{pmatrix} k+1 \\ j \\ 1 \end{pmatrix} = \begin{pmatrix} k+2 \\ j \\ 1 \end{pmatrix} \qquad j = 1, \dots, k-1 \quad \text{and} \quad M_{10} \begin{pmatrix} k \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} k+2 \\ k \\ 1 \end{pmatrix}.$$

Hence the statement is true for L_{k+1}. □

Proposition B.4. *The wandering rate $w_n(A)$ satisfies $w_n(A) \gtrsim \log^2 n$.*

Proof. It follows from (B.3) and Lemma B.3. □

Finally we discuss the property of regular variation for $w_n(A)$. The first remark is that if $w_n(A)$ is regularly varying then it is slowly varying. By (3.1), if $w_n(A)$ is regularly varying then there exists $\alpha \in \mathbb{R}$ such that

$$\lim_{n \to \infty} \frac{w_{cn}}{w_n} = c^\alpha$$

for all $c \in \mathbb{N}$. However by Propositions B.1 and B.4, there exist two constants k_1, k_2 with $0 < k_1 < 1 < k_2$ such that

$$k_1 \frac{\log^2(cn)}{\log^2(n)} \leq \frac{w_{cn}}{w_n} \leq k_2 \frac{\log^2(cn)}{\log^2(n)}$$

and passing to the limit we obtain

$$k_1 \le c^\alpha \le k_2$$

for all $c \in \mathbb{N}$. Hence $\alpha = 0$.

A second remark is that we have a sufficient condition on the sequence λ_k from Lemma B.3 for $w_n(A)$ being slowly varying. Since $\tilde{\tau}_n(A) \gtrsim \log^2 n$, it is immediate that

$$\liminf_{k\to\infty} \frac{k\,\lambda_k}{\log^2 k} = 0$$

To have that $w_n(A)$ is slowly varying it is enough that also the limsup vanishes.

Lemma B.5. *If $\lambda_k = o(\frac{\log^2 k}{k})$ then $w_n(A)$ is slowly varying.*

Proof. From (B.2) we obtain that if τ_n is slowly varying the same holds for $w_n(A)$. Indeed

$$w_{2n}(A) - w_n(A) = \sum_{k=n+1}^{2n} \mu(A^{\complement} \cap \{\varphi = k\}) \le 27m(A) \sum_{k=n+1}^{2n} \sum_{\omega\in\Omega_k} t_{\omega_0\omega_1\dots\omega_{k-2}1} = 27m(A)(\tau_{2n} - \tau_n)$$

and

$$w_n(A) = \sum_{k=1}^{n} \mu(A^{\complement} \cap \{\varphi = k\}) \ge m(A) \sum_{k=1}^{n} \sum_{\omega\in\Omega_k} t_{\omega_0\omega_1\dots\omega_{k-2}1} = m(A)\tau_n.$$

In conclusion

$$1 \le \frac{w_{2n}(A)}{w_n(A)} = 1 + \frac{w_{2n}(A) - w_n(A)}{w_n(A)} \le 1 + 27\frac{\tau_{2n} - \tau_n}{\tau_n}.$$

If $(\tau_n)_{n\ge1}$ is slowly varying the term $\frac{\tau_{2n}-\tau_n}{\tau_n}$ is vanishing, and the result follows.

Moreover from Lemma B.3 it is immediate that if $\tilde{\tau}_n$ is slowly varying then the same is true for τ_n. We are thus reduced to study $\tilde{\tau}_n$. We first claim that it is enough to show that (3.1) holds with $\alpha = 0$ only for $c = 2$ (see for example [3, Proposition 1.10.1]). Indeed, for $1 < c < 2$ we write

$$1 \le \frac{\tilde{\tau}_{\lfloor cn \rfloor}}{\tilde{\tau}_n} \le \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n}.$$

For $c > 2$, let $k \ge 1$ such that $c \le 2^k$, then we write

$$1 \le \frac{\tilde{\tau}_{\lfloor cn \rfloor}}{\tilde{\tau}_n} \le \frac{\tilde{\tau}_{2^k n}}{\tilde{\tau}_{2^{k-1} n}} \frac{\tilde{\tau}_{2^{k-1} n}}{\tilde{\tau}_{2^{k-2} n}} \cdots \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n}$$

and for all $j = 1, \dots, k$ we have $\frac{\tilde{\tau}_{2^j n}}{\tilde{\tau}_{2^{j-1} n}} \to 1$, because it is a subsequence of $\frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n}$. Hence (3.1) follows again with $\alpha = 0$ for $c > 2$. We can proceed analogously for the case $0 < c < 1$, which completes the proof of the claim.

Moreover we follow the proof of [3, Theorem 1.5.4] to show that (3.1) holds with $\alpha = 0$ for $c = 2$. Let $\alpha > 0$, then by definition the sequence $\phi(n) \coloneqq n^\alpha\, \tilde{\tau}_n$ is non-decreasing. We also show that the sequence $\psi(n) \coloneqq n^{-\alpha}\, \tilde{\tau}_n$ is eventually non-increasing. Indeed

$$\psi(n) - \psi(n+1) = \psi(n) \left(1 - \frac{\tilde{\tau}_{n+1}}{\tilde{\tau}_n} \frac{n^\alpha}{(n+1)^\alpha}\right) = \psi(n) \left(1 - \frac{1 + \frac{\lambda_{n+1}}{\tilde{\tau}_n}}{(1+\frac{1}{n})^\alpha}\right)$$

and $\lambda_{n+1} = o(\frac{\log^2(n+1)}{n+1})$ together with $\tilde{\tau}_n \gtrsim \log^2 n$ implies

$$1 + \frac{\lambda_{n+1}}{\tilde{\tau}_n} = o\left(\frac{1}{n}\right).$$

Hence we have that

$$1 + \frac{\lambda_{n+1}}{\tilde{\tau}_n} < \left(1 + \frac{1}{n}\right)^\alpha = 1 + \alpha\frac{1}{n} + o\left(\frac{1}{n}\right)$$

for n big enough. It follows that $\psi(n) - \psi(n+1) \geq 0$ eventually. For n big enough we can then write

$$2^{-\alpha} = \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n} \frac{\phi(n)}{\phi(2n)} \leq \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n} \leq \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n} \frac{\psi(n)}{\psi(2n)} = 2^{\alpha}$$

hence

$$2^{-\alpha} \leq \liminf_{n\to\infty} \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n} \leq \limsup_{n\to\infty} \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n} \leq 2^{\alpha}.$$

Since the previous argument can be repeated for all $\alpha > 0$ it follows that

$$\lim_{n\to\infty} \frac{\tilde{\tau}_{2n}}{\tilde{\tau}_n} = 1\,.$$

□

Finally, we recall from [1] and [26] that for the pointwise dual ergodic system $(\bar{\triangle}, \mu, S)$ the renormalising sequence $a_n(S)$ is defined in terms of the wandering rate of a subset on which the induced map is ψ-mixing (see Proposition 3.6), and $a_n(S)$ is asymptotically independent on the chosen subset with this property. In particular this implies that if Γ_0 satisfies Proposition 3.6 then we have

$$a_n(S) \asymp \frac{n}{w_n(\Gamma_0)}$$

where

$$w_n(\Gamma_0) = \sum_{k=0}^{n-1} \mu(\triangle_k) = \int_0^n \frac{\log(1+v)}{v}\, dv$$

as shown in Proposition B.1. Since it is not difficult to show that $w_n(\Gamma_0)$ is slowly varying, then we would have

$$a_n(S) \sim \frac{n}{\int_0^n \frac{\log(1+v)}{v}\, dv}$$

as discussed in Remark 3.11. Unfortunately it is not known whether Γ_0 is a *good set* to which apply Proposition 3.6.

References

[1] J. Aaronson, "An introduction to infinite ergodic theory". Mathematical Surveys and Monographs, 50. American Mathematical Society, Providence, RI, 1997.

[2] I. Amburg *et al.*, *Stern sequences for a family of multidimensional continued fractions: TRIP-Stern sequences*, J. Integer Seq. **20** (2017), no. 1, Article 17.1.7

[3] N.H. Bingham, C.M. Goldie, J.L. Teugels, "Regular variation", Encyclopedia of Mathematics and its Applications, 27. Cambridge University Press, Cambridge, 1989.

[4] A.J. Brentjes, "Multidimensional continued fraction algorithms". Mathematical Centre Tracts, 145. Mathematisch Centrum, Amsterdam, 1981.

[5] C. Bonanno, S. Isola, *Orderings of the rationals and dynamical systems*, Colloq. Math. **116** (2009), no. 2, 165–189.

[6] T. Garrity, *On periodic sequences for algebraic numbers*, J. Number Theory **88** (2001), no. 1, 86–103.

[7] T. Garrity, *On Gauss-Kuzmin statistics and the transfer operator for a multidimensional continued fraction algorithms: the Triangle map*, arXiv: 1509.01840v1 [math.NT]

[8] T. Garrity, P. Mcdonald, *Generalizing the Minkowski question mark function to a family of multidimensional continued fractions*, Int. J. Number Theory **14** (2018), no. 9, 2473–2516.

[9] M. Iosifescu, C. Kraaikamp, "Metrical theory of continued fractions". Mathematics and its Applications, 547. Kluwer Academic Publishers, Dordrecht, 2002.

[10] S. Isola, *From infinite ergodic theory to number theory (and possibly back)*, Chaos Solitons Fractals **44** (2011), no. 7, 467–479.

[11] M. Kesseböhmer, S. Munday, B.O. Stratmann, "Infinite ergodic theory of numbers". De Gruyter Graduate. De Gruyter, Berlin, 2016.

[12] M. Kesseböhmer, B. O. Stratmann, *Fractal analysis for sets of non-differentiability of Minkowski's question mark function,* J. Number Theory, **128** (2008), no. 9, 2663–2686.

[13] M. Lenci, *On infinite-volume mixing*, Comm. Math. Phys. **298** (2010), no. 2, 485–514.

[14] M. Lenci *Exactness, K-property and infinite mixing*, Publ. Mat. Urug. **14** (2013), 159–170.

[15] M. Lenci, S. Munday, *Pointwise convergence of Birkhoff averages for global observables*, Chaos **28** (2018), no. 8, 083111.

[16] A. Messaoudi, A. Nogueira, F. Schweiger, *Ergodic properties of triangle partitions*, Monatsh. Math. **157** (2009), no. 3, 283–299.
[17] H. Minkowski, *Geometrie der Zahlen*, Gesammelte Abhandlungen, Vol. 2, 1911; reprinted by Chelsea, New York, (1967), 43–52.
[18] J. J. Miao, S. Munday, *Derivatives of slippery Devil's staircases*, Discrete Contin. Dyn. Syst. Ser. S **10** (2017), no. 2, 353–365.
[19] S. Munday, *On the derivative of the α-Farey-Minkowski function*, Discrete Contin. Dyn. Syst. **34** (2014), no. 2, 709–732.
[20] H. Nakada, R. Natsui, *On the metrical theory of continued fraction mixing fibred systems and its application to Jacobi-Perron algorithm*, Monatsh. Math. **138** (2003), no. 4, 267–288.
[21] G. Panti, *Multidimensional continued fractions and a Minkowski function*, Monatsh Math **154** (2008), no. 3, 247–264.
[22] F. Schweiger, *Kuzmin's theory revisited*, Ergodic Theory Dynam. Systems **20** (2000), no. 2, 557–565.
[23] F. Schweiger, "Multidimensional continued fractions". Oxford Science Publications. Oxford University Press, Oxford, 2000.
[24] W.A. Veech, *Interval exchange transformations*, J. Analyse Math. **33** (1978), 222–272
[25] R. Patrick Vernon, *Relationships between Fibonacci-type sequences and Golden-type ratios*, Notes on Number Theory and Discrete Mathematics **24** (2018), no. 2, 85–89.
[26] R. Zweimuller, "Surrey notes on infinite ergodic theory", http://mat.univie.ac.at/%7Ezweimueller/MyPub/SurreyNotes.pdf

Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, Italy
E-mail address: `claudio.bonanno@unipi.it`

Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, Italy
E-mail address: `delvigna@mail.dm.unipi.it`

Dipartimento di Architettura, Università degli studi Roma Tre, via Madonna dei Monti 40, 00184, Roma
E-mail address: `saraann.munday@uniroma3.it`

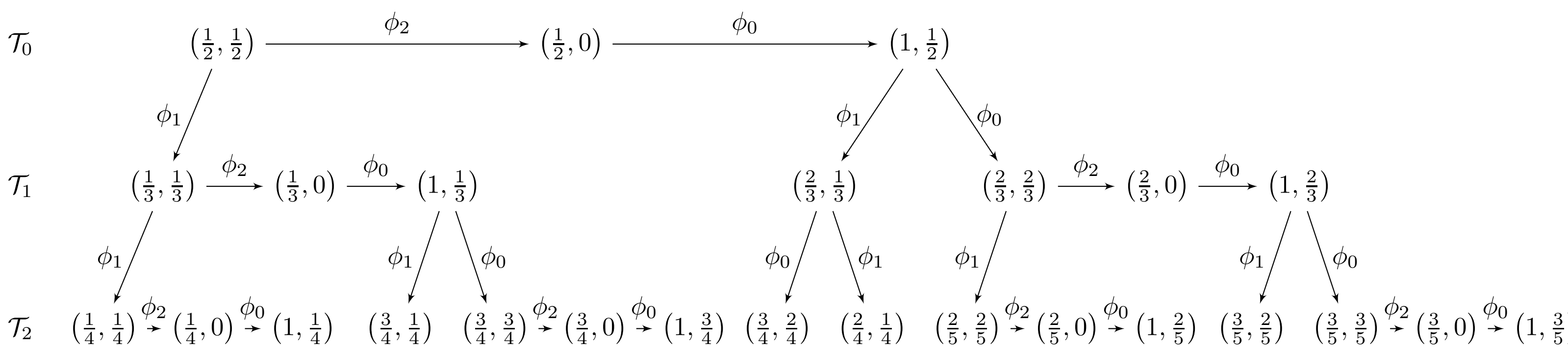

FIGURE 10. The first three levels of the tree generated through the local inverses of the map $\tilde{S}$.