Exponentially fast convergence to flat triangles in the iterated barycentric subdivision

by

MATTHIAS KLÖCKNER

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

Exponentially fast convergence to flat triangles in the iterated barycentric subdivision

submitted by **Matthias Klöckner** in partial fulfillment of the requirements for the degree of **Master of Science** in **Mathematics**.

Examining Committee:

Gordon Slade, Professor of Mathematics

Supervisor

William Casselman, Professor of Mathematics

Supervisory Committee

Abstract

The barycentric subdivision dissects a triangle along its three medians into six *children* triangles. The children of a flat triangle (i.e. the vertices are collinear) are flat. For any triangle Δ let the *shape* $S(\Delta)$ be the unique complex point z in the first quadrant such that Δ is similar to the triangle $-1, 1, z$ in which the edge between -1 and 1 has maximal length. Only flat triangles' shapes lie on the real line. If $\Delta^{(n)}$ is a Markov chain of triangles with $\Delta^{(n)}$ chosen uniformly amongst the children of $\Delta^{(n-1)}$, then we call the Markov chain $S(\Delta^{(n)})$ a *shape chain* and we call it *(non-)flat*, if $\Delta^{(0)}$ and therefore each $\Delta^{(n)}$ is (non-)flat. Let Z_n be a non-flat and X_n be a flat shape chain. We say that a sequence of random variables W_n taking values in $\mathbb{C} \setminus \{0\}$ *decays exactly* or *at least with rate* χ, if $\chi > 0$ and almost surely $\lim_n \frac{1}{n} \ln |W_n| = -\chi$ or $\limsup_n \frac{1}{n} \ln |W_n| \leq -\chi$, resp. In a paper from 2011, P. Diaconis and L. Miclo show that $\Im Z_n$ decays at least with rate χ' for some universal constant χ' and that X_n has an invariant measure μ. We prove that $\Im Z_n$ decays exactly with rate χ for a universal constant χ which we express as an integral w.r.t. μ. The above paper also shows the convergence of $Z_n - X_n$ to 0 in probability for a specific coupling (X_n, Z_n). For this coupling we prove that $Z_n - X_n$ decays exactly with rate χ.

Lay Summary

In any triangle Δ the three lines between each vertex and the middle point of the opposite edge intersect in one point, the *barycentre* of Δ, and cut Δ into six smaller triangles which we call the *children* of Δ. We randomly choose one of the six children of $\Delta^{(1)} = \Delta$ by throwing a die and call the new triangle $\Delta^{(2)}$. Then we choose $\Delta^{(3)}$ amongst the children of $\Delta^{(2)}$ by throwing another die, and so on. To any triangle Δ we assign a certain number $\mathrm{flat}(\Delta)$ which is 0, if Δ is flat (i.e. the three vertices lie on the same line), and positive otherwise. The smaller $\mathrm{flat}(\Delta)$ is, the "flatter" Δ looks. It has been known that the sequence $\mathrm{flat}(\Delta^{(1)})$, $\mathrm{flat}(\Delta^{(2)})$, $\ldots$ converges to 0 exponentially fast. We prove the existence of and a formula for the exact exponential rate.

Preface

The dissertation is original, unpublished, independent work by the author, M. Klöckner.

Table of Contents

List of Figures

List of symbols and abbreviations

The abbreviation "a.s." stands for "almost surely". We use the following notations:

- $\Re z$ and $\Im z$ denote the real and imaginary part of a complex number z.
- $\mathbb{H} = \{z \in \mathbb{C} : \Im z > 0\}$ and $\mathbb{H}_0 = \mathbb{H} \cup \mathbb{R}$.
- $\overline{\mathbb{H}}_0 = \mathbb{H}_0 \cup \{\infty\}$ and $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ are subsets of the extended complex plane (Riemann sphere).
- $\mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$ denotes the extended real number line, where we distinguish between $-\infty$ and ∞. We define intervals in a usual way, e.g. $(0, \infty] := (0, \infty) \cup \{\infty\}$.
- Sequences are indexed by $n \in \mathbb{N}_0$ or $n \in \mathbb{N}$ and we simply write s_n for a sequence $(s_n)_n$. Moreover, any limit (inferior, superior) is implicitly understood w.r.t. $n \to \infty$.
- We omit the $\circ$ symbol for compositions of functions. For a function $f : C \to C$ we write f^n for the n-fold composition $f \dots f : C \to C$.
- δ_x denotes the Dirac measure at x.

Acknowledgements

I owe thanks to my supervisor Prof. Slade who has been an excellent and very supportive mentor during my Master's program. His supervision fostered my passion for research by providing the perfect balance between freedom and guidance. My research and presentation skills benefited from the very helpful feedback during our meetings, particularly those in company of Prof. Casselman who kindly joined the supervisory committee. I am deeply grateful that the UBC was in particular a place for me where two great mathematicians with decades of experience were repeatedly meeting with me and giving me invaluable advice. I also want to thank my mentor Prof. Ritter from my Bachelor's program, my friends in Kaiserslautern and Vancouver, as well as my family who has always been a great support for me.

Dedication

Für Oma Aloisia.

1 Introduction

1.1 Model

Consider a triangle Δ in the complex plane. If Δ is degenerate (i.e. its three vertices are collinear), then following the terminology of [1] and [6] we call Δ *flat*. We allow the special case of two vertices being identical, but we exclude the case of all three vertices being identical. The three medians of any triangle Δ (which connect each vertex to the midpoint of the opposite edge) intersect in the *barycentre* and dissect Δ into six smaller triangles which we call the *children* of Δ. This construction is called the *barycentric subdivision*. We can even make sense of it for flat triangles whose children are also flat. Now we repeat this construction by applying the barycentric subdivision to each child of Δ creating a total of 6^2 *grand-children* of Δ. Repeating this procedure iteratively results in 6^n *descendants* of Δ in the n-th generation for $n \in \mathbb{N}_0$.

We introduce a Markov chain $\Delta^{(n)}$. Starting with $\Delta^{(0)}$ we randomly choose one of its six children by throwing a die and we call the new triangle $\Delta^{(1)}$. Now we randomly choose $\Delta^{(2)}$ amongst the children of $\Delta^{(1)}$ by throwing another die, and so on. In each step the die is thrown independently of all the previous dice. In case of a flat starting triangle $\Delta^{(0)}$ we obtain a Markov chain of flat triangles $\Delta^{(n)}$.

For a flat triangle Δ we set $\mathrm{flat}(\Delta) := 0$. For any other triangle Δ we define

$$\mathrm{flat}(\Delta) := \frac{(\text{shortest height in } \Delta)^2}{\text{area of } \Delta} = 2 \cdot \frac{\text{shortest height in } \Delta}{\text{longest edge in } \Delta}. \tag{1.1}$$

According to the formula for the area of a triangle, a shortest height is orthogonal to a longest edge. Here we use the indefinite article "a" because z might have several shortest heights or several longest edges. If Δ and Δ' are similar, then $\mathrm{flat}(\Delta) = \mathrm{flat}(\Delta')$. The smaller $\mathrm{flat}(\Delta)$ is, the "flatter" Δ looks. The value $\mathrm{flat}(\Delta)$ is maximal, namely $\sqrt{3}$, if and only if Δ is equilateral.

For two similar triangles Δ, Δ' there is a (not necessarily unique) one-to-one correspondence between the children of Δ and the children of Δ' such that corresponding children are

similar. Hence, if we can characterize the "shape" of Δ by some complex number $S(\Delta)$ — i.e. two triangles Δ and Δ' are similar if and only if their S-values $S(\Delta)$, $S(\Delta')$ are equal — then $S(\Delta) = S(\Delta')$ implies that the multiset of S-values of the children of Δ equals the multiset of S-values of the children of Δ'. Hence $S(\Delta)$ uniquely determines the multiset of S-values of the children of Δ. We obtain the following Markov chain which we call a *shape chain*:

$$Z_n := S(\Delta^{(n)}), \quad n \in \mathbb{N}_0. \tag{1.2}$$

Unless otherwise stated, we always suppose that z and z' lie in the set $\mathbb{H}_0$ of complex numbers with non-negative imaginary part, so $\mathbb{H}_0$ is the union of the upper half plane $\mathbb{H}$ with the real line. We denote by Δ_z the triangle with vertices -1, 1 and z and we say that Δ_z is the *associated triangle of* z. We call the edge between -1 and 1 the 0-edge, the edge between -1 and z the 1-edge and the edge between 1 and z the 2-edge of Δ_z, see Figure 1.

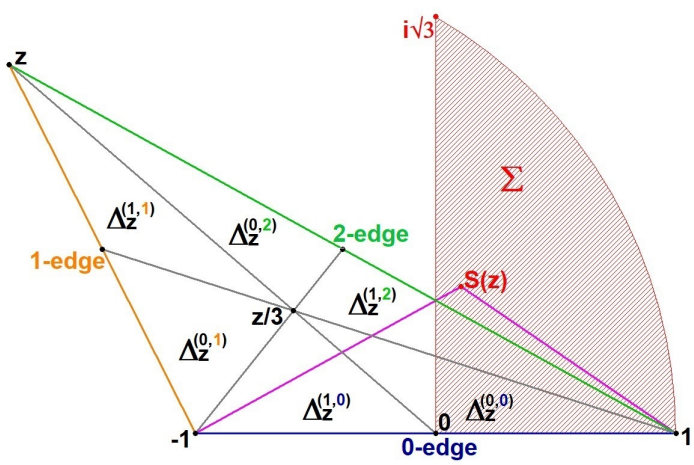

Figure 1: A triangle Δ_z, its shape $S(z) \in \Sigma$ and its children $\Delta_z^{(l,k)}$. Their second index k is depicted in the same colour as the corresponding k-edge of z.

From now on we usually think of a point z in terms of its associated triangle, so we speak for example of the 0-edge of z and we call z, z' similar, if Δ_z and $\Delta_{z'}$ are similar. Moreover, we set $\mathrm{flat}(z) := \mathrm{flat}(\Delta_z)$.

Let Σ be the intersection of the first quadrant and the closed disk around -1 of radius 2, see Figure 1. Equivalently, Σ is the set of all points z such that the 0-edge of z is a longest edge and the 2-edge of z is a shortest edge. For $z \in \Sigma$ the length of a longest edge of z is 2 and the length of a shortest height is $\Im z$, hence

$$\mathrm{flat}(z) = \Im z. \tag{1.3}$$

For any triangle Δ there is a unique point $z' \in \Sigma$ such that Δ and $\Delta_{z'}$ are similar. We call $\Delta_{z'}$ the *standardized version* of Δ and z' the *shape* of Δ. Both Δ and $\Delta_{z'}$ have the flat-value $\Im z'$. Let us denote z' by $S(\Delta)$. So $\mathrm{flat}(\Delta) = \Im S(\Delta)$. We see that $S(z) := S(\Delta_z)$ defines a projection from $\mathbb{H}_0$ into Σ, where $S(z)$ is the unique point in Σ that is similar to z, see Figure 1. We call this projection the *shape function* S. The restriction of S to $\mathbb{R}$ is a projection from $\mathbb{R}$ to $\Sigma \cap \mathbb{R} = [0, 1]$. Using (1.2) as a definition for Z_n, we see that

$$\Im Z_n = \mathrm{flat}(\Delta^{(n)}). \tag{1.4}$$

In Section 1.2 we formulate our result that a.s. $\Im Z_n = \mathrm{flat}(\Delta^{(n)})$ converges to 0 "exponentially fast" and we give a formula for this exponential rate.

The barycentre of z is the arithmetic mean $\frac{1}{3}z$ of its vertices -1, 1, z. We denote the six children of z by $\Delta_z^{(0,0)}$, $\Delta_z^{(1,0)}$, $\Delta_z^{(0,1)}$, $\Delta_z^{(1,1)}$, $\Delta_z^{(0,2)}$, $\Delta_z^{(1,2)}$ in clockwise order, where $\Delta_z^{(0,0)}$ is the triangle with vertices $0, 1, z/3$, see Figure 1. We index the six children Δ_z^j by

$$j = (l, k) \in \mathcal{J} := \{0, 1\} \times \{0, 1, 2\}. \tag{1.5}$$

Unless otherwise stated, we always view the indices l, l' as elements of the cyclic group $\mathbb{Z}_2$ and the indices k, k' as elements of the cyclic group $\mathbb{Z}_3$. We always assume that $j, j' \in \mathcal{J}$.

For $j = (l, k)$ and the child $\Delta := \Delta_z^j$ we define the *outer vertex* of Δ as the vertex of Δ that is also a vertex of Δ_z, and we define the *outer edge* of Δ as the edge that is one half of an edge (namely the k-edge) of Δ_z.

Let us define the six *child functions* $a_j : \mathbb{H}_0 \to \Sigma$ that describe the shapes of the children of Δ_z:

$$a_j(z) := S\left(\Delta_z^j\right). \tag{1.6}$$

The triangle $\Delta^{(n)}$ is chosen uniformly amongst the children of $\Delta^{(n-1)}$ and independently of $\Delta^{(0)}, \dots, \Delta^{(n-2)}$. Since $\Delta^{(n-1)}$ is similar to its standardized version $\Delta_{Z_{n-1}}$, the multiset of the S-values of the children of $\Delta^{(n-1)}$ equals the multiset of the S-values of $\Delta^j_{Z_{n-1}}$ with $j \in \mathcal{J}$. Thus $Z_n = S(\Delta^{(n)})$ is chosen uniformly amongst $S(\Delta^j_{Z_{n-1}})$ with $j \in \mathcal{J}$ and independently of $Z_0, \dots, Z_{n-2}$. We find a sequence $j_1, j_2, \dots$ of independent random indices such that each j_n is uniformly distributed on $\mathcal{J}$ and $Z_n = S(\Delta^{j_n}_{Z_{n-1}})$. Using the child functions this can be written as

$$Z_n = a_{j_n}(Z_{n-1}). \tag{1.7}$$

In (1.2) we defined the Markov chain Z_n as the sequence of S-values of the Markov chain $\Delta^{(n)}$. From now on we assume $Z_0 \in \mathbb{H}$, i.e. $\Delta^{(0)}$ is not flat. Hence $Z_n \in \mathbb{H}$ for any n. Analogously to the non-flat setting we define X_n as the sequence of S-values of a corresponding Markov chain of flat triangles: We fix $X_0 \in \Sigma \cap \mathbb{R} = [0,1]$ and set

$$X_n := a_{j_n}(X_{n-1}). \tag{1.8}$$

Notice that we use the same random indices $j_1, \dots, j_n$ for X_n and Z_n. In Section 1.2 we present our main result which describes the a.s. behaviour of this specific coupling (X_n, Z_n). We see that X_n is a Markov chain with state space $[0,1]$ and that Z_n is a Markov chain with state space $\Sigma \setminus \mathbb{R}$. The Markov kernel M_{flat} of X_n is given by

$$M_{\text{flat}}(x, \cdot) = \tfrac{1}{6} \sum_{j \in \mathcal{J}} \delta_{a_j(x)}, \quad x \in [0,1], \tag{1.9}$$

and the kernel of Z_n has the same formula with x replaced by $z \in \Sigma \setminus \mathbb{R}$.

1.2 Main result

We deal with the following two notions of asymptotic exponential decay of a sequence of complex random variables:

Definition 1.1 Let $\chi > 0$. We formally write $\ln(0) = -\infty$ and say that a sequence W_n of complex random variables *decays exactly with rate* χ or *decays at least with rate* χ if a.s.

$$\lim_n \frac{\ln |W_n|}{n} = -\chi \quad \text{or} \quad \limsup_n \frac{\ln |W_n|}{n} \le -\chi, \quad \text{resp.}$$

An equivalent formulation for this definition is the following: A.s. for any $\epsilon > 0$ there is some $n_0 \in \mathbb{N}$ such that any $n \geq n_0$ satisfies

$$e^{(-\chi-\epsilon)n} \leq |W_n| \leq e^{(-\chi+\epsilon)n} \quad \text{or} \quad |W_n| \leq e^{(-\chi+\epsilon)n}, \quad \text{resp.} \tag{1.10}$$

In both cases we note that in particular, W_n converges to 0 a.s.

Definition 1.2 An *upcircle* is a circle in $\mathbb{H}_0$ with tangent line $\mathbb{R}$. For any $x \in \mathbb{R}$ and $z \in \mathbb{H}$ we define the *radius* $R(x, z) = |z - x|^2/(2\Im z)$.

We show in Section 2.5 that $R(x, z)$ is the radius of the unique upcircle containing x and z.

Definition 1.3 Let $b : D_b \to \mathbb{R}$ be a function with $D_b \subseteq \mathbb{R}$. We say that b is *up-to-sign differentiable* if there is a continuous function $\beta : D_b \to [0, \infty)$ such that for all but finitely many $x \in D_b$ the function b is differentiable in x with $|b'(x)| = \beta(x)$. Here, β is unique and we call β the *up-to-sign derivative* of b and we formally write $|b'| := \beta$.

Before we state our main results, let us summarize Section 1.1: We always assume that the index j lies in $\mathcal{J} = \{0, 1\} \times \{0, 1, 2\} = \mathbb{Z}_2 \times \mathbb{Z}_3$. The points $a_j(z) \in \Sigma$ are the shapes of the six children of Δ_z. We fix arbitrary points $Z_0 \in \Sigma \setminus \mathbb{R}$ and $X_0 \in [0, 1]$. The Markov chains Z_n and X_n with state spaces $\Sigma \setminus \mathbb{R}$ and $[0, 1]$ have the recursive formulas $Z_n = a_{j_n}(Z_{n-1})$ and $X_n = a_{j_n}(X_{n-1})$, where $j_1, j_2, \ldots$ are independent indices which are uniformly distributed on $\mathcal{J}$.

Theorem 1.4 The kernel M_{flat} of X_n has an invariant measure μ, i.e. $\mu M_{\text{flat}} = \mu$, and μ is unique. Each child function a_j viewed as a function $[0, 1] \to [0, 1]$ is Lipschitz continuous and has an up-to-sign derivative $|a_j'| : [0, 1] \to (0, \infty)$. The constant

$$\chi := -\tfrac{1}{6} \sum_j \int \ln |a_j'| \, d\mu$$

is bounded from below by $\chi_{\min} := \frac{1}{3} \ln \frac{3}{2} \approx 0.1352$ and bounded from above by $\chi_{\max} := \frac{1}{3} \ln \frac{91}{2} - \ln 3 \approx 0.1740$.

Proof. Postponed to Theorem 4.18 (on page 34). ∎

Theorem 1.5 Let R_n be the radius $R(X_n, Z_n)$. The following three sequences decay exactly with rate χ:

$$R_n, \quad Z_n - X_n, \quad \Im Z_n = \mathrm{flat}(\Delta^{(n)}).$$

The sequence $\Re Z_n - X_n$ decays at least with rate χ.

Proof. Postponed to Theorem 5.17 (on page 46). ∎

1.3 Relation to earlier work

Let us briefly discuss how our work relates to previous papers that use dynamical systems arguments:

- The first proof of the a.s. convergence to flat triangles in the iterated barycentric subdivision was given by Bárány, Beardon and Carne [1] using Furstenberg's theorem on random products of matrices in $SL_2(\mathbb{R})$, see [8].

- Wilkinson [10] explains in Section 1.7 the measurably hyperbolic cocycle (over a measure-preserving dynamical system) which "lurks behind random barycentric subdivision". With the same approach as in [1], Furstenberg's theorem leads to the observation that the aspect ratio defined by

$$\alpha_n := \frac{\text{area of } \Delta^{(n)}}{(\text{longest edge in } \Delta^{(n)})^2} = \tfrac{1}{2} \cdot \frac{\text{shortest height of } \Delta^{(n)}}{\text{longest edge in } \Delta^{(n)}} \tag{1.11}$$

  decays exactly with some rate 2γ, where γ is the (upper) Lyapunov constant. A comparison to our definition (1.1) shows that $4\alpha_n = \mathrm{flat}(\Delta^{(n)})$. Thus Theorem 1.5 implies $\chi = 2\gamma$.

- Hough [9] presents an application of Azuma's inequality resulting in lower and upper bounds $\delta_- \approx 0.05857$ and $\delta_+ \approx 0.09461$ for the (upper) Lyapunov constant γ. These bounds show the inaccuracy of the numerical estimate $\gamma \approx 0.0446945$ in [10, Theorem 0.1]. The "bridge" $\chi = 2\gamma$ to the dynamical systems setting gives us the bounds

$0.1171 \approx 2\delta_- \leq \chi \leq 2\delta_+ \approx 0.1892$. Notice that our bounds $0.1352 \approx \chi_{\min} \leq \chi \leq \chi_{\max} \approx 0.1740$ are sharper.

Our main reference is Diaconis and Miclo [6] who we refer to as "they". We have

$$X_n = -2Z_n + 1, \quad Z_n = 2(-X_n + iY_n) + 1, \tag{1.12}$$

where the left-hand sides are our objects and the right-hand sides are theirs. Notice that $\Im Z_n$ in our notation and Y_n in their notation only differ by a constant factor and therefore have the same behaviour in terms of asymptotic exponential decay. We translate some of their results into our notation and compare them to our work:

- Their Theorem 1.1 states that $\Im Z_n$ decays at least with rate χ' for some universal constant χ'. Our Theorem 1.5 states that $\Im Z_n$ decays exactly with rate χ for some universal constant χ.

- Their Lemma 6.1 states that $\Re Z_n - X_n$ converges to 0 in probability. Together with the a.s. convergence of $\Im Z_n$ to 0 it follows that $Z_n - X_n$ converges to 0 in probability. Our Theorem 1.5 shows that $Z_n - X_n$ converges a.s. to 0.

- Their Theorem 1.3 states that a.s. the limit set of the sequence $\Re Z_n$ is $[0, 1]$. In Remark 6.5, they show under the assumption of an unproven estimate (6.5) that $\Re Z_n - X_n$ converges a.s. to 0 and they point out how this result (which follows from our Theorem 1.5) allows for a much easier proof of their Theorem 1.3.

- Their Proposition 5.3 states that the kernel $M := M_{\text{flat}}$ of X_n is ergodic and that the Markov chain X_n satisfies the law of large numbers: There is a unique invariant probability measure μ on $[0, 1]$, i.e. $\mu M = \mu$, and μ is attracting, i.e. for any probability measure ν on $[0, 1]$ the sequence νM^n converges weakly to μ. For any continuous function $\phi : [0, 1] \to \mathbb{R}$ we have the a.s. convergence
$$\lim_n \frac{1}{n} \sum_{t=1}^{n} \phi(X_t) = \int \phi \, d\mu. \tag{1.13}$$
Little is known about the invariant measure μ. They show that μ is continuous and has support $[0, 1]$, but neither they nor we have been able to prove that μ is absolutely continuous.

- Their proof of Proposition 5.3 relies on a result of Barnsley and Elton [2] which requires the technical assumption that the restricted child functions $a_j : [0,1] \to (0,\infty)$ are "contractions on average". We formulate and verify this assumption in Sections 4 and 11 with a new method that makes the calculations much easier and also yields the bounds $\chi_{\min} \leq \chi \leq \chi_{\max}$ as a byproduct.

1.4 New approach

Remember that we call points z, z' similar, if their associated triangles Δ_z, $\Delta_{z'}$ are similar.

Definition 1.6 We call a stochastic process of points $W_n \in \mathbb{H}_0$ a *barycentric process*, if W_n is similar to $a_{i_n}(W_{n-1})$ for some sequence $i_1, i_2, \ldots$ of independent random indices, each of which is uniformly distributed on $\mathcal{J}$. We call a barycentric process W_n *flat* (or *non-flat*), if W_0 and therefore each W_n is real (not real).

Shape chains are special barycentric processes for which "similar" in the above definition can be replaced by "equal". In particular, shape chains have state space Σ. We easily verify that W_n is a barycentric process if and only if $S(W_n)$ is a shape chain. All we need is that there is a one-to-one correspondence between the children of W_{n-1} and the children of $S(W_{n-1})$ such that corresponding children are similar.

Diaconis and Miclo [6] only consider the coupling

$$(X_n, Z_n) = (a_{j_n}(X_{n-1}), a_{j_n}(Z_{n-1})) \tag{1.14}$$

of a flat and a non-flat shape chain, whereas our approach consists of the analysis and the comparison of different couplings of flat and non-flat barycentric processes. The rest of this subsection gives a glimpse into the main ideas behind this new approach, but may be skipped at first reading as the proofs in the following sections do not formally require this discussion.

In Definition 2.10 we introduce the set $\mathcal{F}$ containing Möbius transformations f with real coefficients and the set $\mathcal{G} := \mathcal{F} \cup \{mf : f \in \mathcal{F}\}$, where $m(z) := -\overline{z}$ is the reflection across the imaginary axis. We show in Remark 3.12 for each j that there is a set $\mathcal{G}_j$ of six functions

in $\mathcal{G}$ such that for any z the six (not necessarily pairwise different) evaluations $g(z)$ with $g \in \mathcal{G}_j$ are exactly those points that are similar to $a_j(z)$. Hence for fixed z there is a (not necessarily unique) function $a_j^z \in \mathcal{G}_j$ such that

$$a_j(z) = a_j^z(z). \tag{1.15}$$

Notice that a_j^z maps z into Σ. For any other point z' the image $a_j^z(z')$ is similar to $a_j(z')$, but not necessarily equal. In particular, $a_j^z(z')$ does not necessarily lie in Σ.

For $g \in \mathcal{G}$ let us write component-wise $g(x,z) := (g(x), g(z))$ for $(x,z) \in \mathbb{R} \times \mathbb{H}$. In Sections 3.2 and 3.3 we quote parts of Definitions 8.1 and 8.5, where we introduce a set $D \subset \mathbb{R} \times \mathbb{H}$ and *generalized child functions* $a_{1,j}, a_{2,j} : D \to D$ in such a way that "most" pairs $(x,z) \in D$ satisfy

$$a_{1,j}(x,z) = a_j^x(x,z), \quad a_{2,j}(x,z) = a_j^z(x,z). \tag{1.16}$$

This implies that the first component of $a_{1,j}(x,z) = (a_j^x(x), a_j^x(z))$ is equal to $a_j(x) \in \Sigma$, whereas the second component $a_j^x(z)$ is just similar to $a_j(z)$. Analogously, (1.16) implies that the second component of $a_{2,j}(x,z)$ is equal to $a_j(z) \in \Sigma$, whereas the first component is just similar to $a_j(x)$.

For each component index $c = 1, 2$ we define a Markov chain $V_{c,n} = (X_{c,n}, Z_{c,n})$ by $V_{c,0} := (X_0, Z_0)$ and

$$V_{c,n} := a_{c,j_n}(V_{c,n-1}). \tag{1.17}$$

We see that $X_{1,n}$ is exactly our flat shape chain X_n, whereas $Z_{1,n}$ is just similar to $a_{j_n}(Z_{1,n-1})$ which means that $Z_{1,n}$ is a non-flat barycentric process. Similarly, $X_{2,n}$ is a flat barycentric process and $Z_{2,n}$ is our non-flat shape chain Z_n. So each Markov chain $V_{c,n}$ is a coupling of a flat and a non-flat barycentric process.

The definition of $V_{c,n}$ may first seem a bit obscure, so let us reformulate it in the hope of making it more clear. For any $n \in \mathbb{N}$ we define the random functions

$$g_{1,n} := a_{j_n}^{X_n}, \quad g_{2,n} := a_{j_n}^{Z_n}. \tag{1.18}$$

By design, the functions $g_{1,n}$ and $g_{2,n}$ only "make sure" that X_n and Z_n are mapped into Σ. A comparison to the definitions above shows that (1.17) is equivalent to $V_{c,n} = g_{c,n}(V_{c,n-1})$. By

the component-wise definition $g(x,z) := (g(x), g(z))$ this means that in any step we obtain $(X_{c,n}, Z_{c,n})$ by applying the **same** function $g_{c,n} \in \mathcal{G}$ to both component of $(X_{c,n-1}, Z_{c,n-1})$. This is the main difference to the coupling (X_n, Z_n) where in any step we obtain $X_n = g_{1,n}(X_{n-1})$ and $Z_n = g_{2,n}(Z_{n-1})$ by applying potentially **different** functions $g_{1,n}, g_{2,n} \in \mathcal{G}$ to the components of (X_{n-1}, Z_{n-1}).

Applying in every step the same function $g \in \mathcal{G}$ to both components of $V_{1,n-1}$ makes it easier to describe the upcircle containing $X_{c,n}$, $Z_{c,n}$ than to describe the upcircle containing X_n, Z_n. The essential reason, as discussed in Section 2.5, is that any function $g \in \mathcal{G}$ preserves upcircles, and in a rather nice way: Given the radius of a pair (x,z), see Definition 1.2, we can compute the radius of the pair $g(x,z)$ with the following easy formula, where g' denotes the derivative of g restricted to the real line:

$$Rg(x,z) = |g'(x)| \cdot R(x,z). \tag{1.19}$$

For each component index $c = 1, 2$ we define the radius

$$R_{c,n} := R(V_{c,n}). \tag{1.20}$$

The first milestone in the proof of Theorem 5.17 (which restates Theorem 1.5) is that $R_{1,n}$ decays exactly with the rate χ from Theorem 1.4. Showing this is most of the work in this thesis. Afterwards, we conclude that $R_{2,n}$ decays exactly with rate χ and a comparison between $V_{1,n}$ and $V_{2,n}$ finally leads to the heart of Theorem 1.5, namely that the radius $R_n = R(X_n, Z_n)$ decays exactly with rate χ.

How do we prove that $R_{1,n}$ decays exactly with rate χ? As a consequence of formula (1.19) we show that $Ra_{1,j}(x,z) = |a_j'(x)| \cdot R(x,z)$ with $|a_j'|$ denoting the up-to-sign derivative of a_j. Hence

$$R_{1,n} = |a_{j_n}'(X_{n-1})| \cdot R_{1,n-1}. \tag{1.21}$$

Our main technical tool is the law of large numbers (1.13) which we modify in Section 4.5 for a two-dimensional Markov chain resulting in the a.s. convergence

$$\lim_n \frac{1}{n} \sum_{t=1}^{n} \ln |a_{j_n}'(X_{n-1})| = \chi. \tag{1.22}$$

With (1.21) it follows that $R_{1,n}$ decays exactly with rate χ.

Our method to prove the existence of and a formula for χ is new, even though the law of large numbers (1.13) has been shown in [6]. However, they only use it for the proofs of their Theorems 1.2 and 1.3 and not for their Theorem 1.1 which addresses asymptotic exponential decay.

2 Geometric preliminaries

This section presents auxiliary geometric results required for the proofs of our main results in Sections 4 and 5. Almost all proofs are postponed to Sections 6 and 9. The objects in this section are non-random and "static" (we do not consider any sequences).

In Sections 2.1 and 2.2 we describe the barycentric subdivision in a new geometric setting with explicit formulas that involve Möbius transformations. In Section 2.5 we first introduce a new coordinate system that uniquely describes pairs $(x, z) \in \mathbb{R} \times \mathbb{H}$ by their so-called upcircle coordinates (x, r, d). In Section 2.6 we describe Möbius transformation in upcircle coordinates.

2.1 New geometric setting

Unless stated otherwise, we assume throughout this subsection that z is a point in $\mathbb{H}_0$. So far, given a triangle Δ_z and one of its six children $\Delta := \Delta_z^j$, we consider the standardized version $\Delta_{z'}$ of Δ, i.e. z' is the shape $S(\Delta) = a_j(z)$, see definition (1.6) of a_j. In order to construct $a_j(z)$, we first rotate, rescale and translate Δ such that a longest edge of Δ is mapped to the line segment between -1 and 1. This (orientation preserving) similarity transformation maps z to some new point w and therefore Δ to Δ_w. Hence the 0-edge of Δ_w is a longest edge of Δ_w. If w lies in the first quadrant, then $z' = w$. Otherwise we obtain $\Delta_{z'}$ by reflecting Δ_w across the imaginary axis. In both cases the point z' lies in Σ and $\Delta_{z'}$ is similar to Δ. Thus z' is the shape $a_j(z)$ of Δ.

Let us introduce a new setting: Instead of a similarity transformation that maps the child Δ to some new triangle $\Delta_{z'}$ and a **longest** edge of Δ to the 0-edge of $\Delta_{z'}$, we now consider a similarity transformation that maps Δ to some new triangle $\Delta_{z'}$ and the **outer** edge of Δ to the 0-edge of $\Delta_{z'}$.

Definition 2.1 Let $\Delta := \Delta_z^j$ be some child of Δ_z. If the outer edge of Δ has length 0, then we set $A_j(z) := \infty$. Otherwise, there is a unique similarity transformation that maps the outer edge of Δ to the line segment between -1 and 1, the outer vertex of Δ to 1 and

the inner vertex of Δ, namely the barycentre $\frac{1}{3}z$ of Δ_z, to some new point $z' \in \mathbb{H}_0$. We set $A_j(z) := z'$. This defines a function $A_j : \mathbb{H}_0 \to \overline{\mathbb{H}}_0$.

Figure 2 compares the above definition to the child functions.

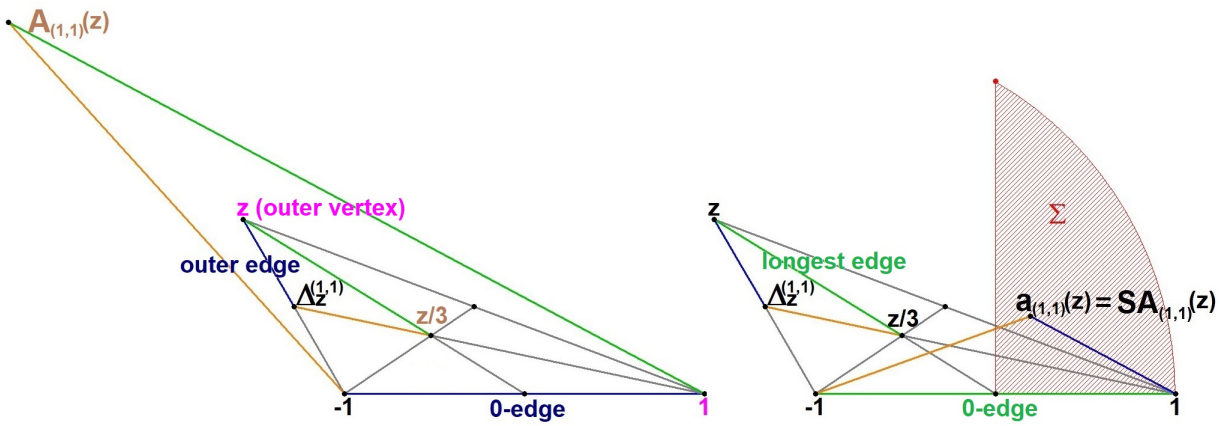

Figure 2: Comparison of $A_{(1,1,)}(z)$ and $a_{(1,1,)}(z)$. Corresponding edges in similar triangles have the same colour.

The following definition extends the shape function $S : \mathbb{H}_0 \to \Sigma$ to a function defined on all of $\overline{\mathbb{H}}_0$. We use the same notation for this extended function.

Definition 2.2 We write $S(\infty) := 1 \in \Sigma$.

Lemma 2.3 Any $z \in \mathbb{H}_0$ satisfies

$$a_j(z) = SA_j(z). \tag{2.1}$$

Proof. We set $z' := a_j(z)$ and $z'' := A_j(z)$. If $z'' = \infty$ then two vertices in the child Δ_z^j are identical, hence $z' = S(\Delta_z^j) = 1 = S(z'')$. Now suppose $z'' \neq \infty$. Then the triangles $\Delta_{z'}$ and Δ_z^j are similar. By Definition 2.1 the triangles $\Delta_{z''}$ and Δ_z^j are similar, too. Hence z' and z'' are similar, i.e. $z' = S(z') = S(z'')$. ∎

Remark 2.4 Given $A_j(z)$ we can compute $a_j(z) \in \Sigma$, but not the other way around, i.e. $A_j(z)$ contains more information. The reason is that A_j is bijective, whereas a_j is not. Furthermore, the explicit formula for A_j is quite easy (as it turns out in Section 2.3), whereas a_j involves comparisons of lengths and therefore inconvenient square roots and squares.

2.2 Möbius transformations

Definition 2.5 Let $\mathcal{F}$ be the set of all *Möbius transformations* $f : \overline{\mathbb{H}}_0 \to \overline{\mathbb{H}}_0$ with coefficients $c_1, c_2, c_3, c_4 \in \mathbb{R}$, i.e.

$$f(z) = \frac{c_1 z + c_2}{c_3 z + c_4}, \quad c_1 c_4 - c_2 c_3 > 0. \tag{2.2}$$

For any such f and $x, y \in \mathbb{R}$ we set

$$f'(x, y) := \frac{c}{(c_3 x + c_4)(c_3 y + c_4)} \in [-\infty, \infty], \quad c := c_1 c_4 - c_2 c_3 > 0, \tag{2.3}$$

and we formally write $f'(x) := f'(x, x) \in (0, \infty]$.

Remark 2.6 Let us present some (well-known or easily verifiable) facts about the above definition: $(\mathcal{F}, \circ)$ is a group and any $f \in \mathcal{F}$ has the following properties:

(i) f is well-defined. With coefficients as in (2.2) we have $f(\infty) = c_1/c_3 \in \overline{\mathbb{R}}$ and $f^{-1}(\infty) = -c_4/c_3 \in \overline{\mathbb{R}}$.

(ii) f is bijective and maps $\mathbb{R}$ to $\mathbb{R}$ and $\mathbb{H}$ to $\mathbb{H}$.

(iii) Let $x \in \mathbb{R}$. If $f(x) = \infty$ then $f'(x) = \infty$. If $f(x) \neq \infty$ then f is differentiable in x and the derivative is indeed $f'(x) = f'(x, x) \in (0, \infty)$.

(iv) Let $x, y \in \mathbb{R}$ with $x \neq y$. Then

$$f'(x, y) = \frac{f(x) - f(y)}{x - y}.$$

Definition 2.7 We define:

(i) the *mirror function* $m : \overline{\mathbb{H}}_0 \to \overline{\mathbb{H}}_0$ by $m(z) := -\Re z + i\Im z = -\overline{z}$ for $z \in \mathbb{H}_0$ and $m(\infty) := \infty$,

(ii) $\mathcal{G} := \{m^l f : l \in \{0,1\}, f \in F\} = \mathcal{F} \cup \{mf : f \in \mathcal{F}\}$,

(iii) $(mf)'(x) := -f'(x) \in (0,\infty]$ for $f \in \mathcal{F}$ and any $x \in \mathbb{R}$,

(iv) $g(v) := (g(v_1), g(v_2))$ component-wise for $g \in \mathcal{G}$ and any pair $v = (v_1, v_2) \in \overline{\mathbb{H}}_0 \times \overline{\mathbb{H}}_0$.

Remark 2.8 Let us comment on parts of the above definition:

(i) The mirror function reflects any point z, as well as its associated triangle Δ_z across the imaginary axis.

(iii) Definitions 2.5 and 2.7 define for any $g \in \mathcal{G}$ the (formal) derivative $g' : \mathbb{R} \to [-\infty, \infty] \setminus \{0\}$.

(iv) It is always clear from the context whether g denotes a function from $\overline{\mathbb{H}}_0$ to itself or a function from $\overline{\mathbb{H}}_0 \times \overline{\mathbb{H}}_0$ to itself.

Lemma 2.9 $(\mathcal{G}, \circ)$ is a group with subgroup $(\mathcal{F}, \circ)$. Any $g \in \mathcal{G}$ has the following properties:

(i) g is bijective and maps $\mathbb{R}$ to $\mathbb{R}$ and $\mathbb{H}$ to $\mathbb{H}$.

(ii) For $x, y \in \mathbb{R}$ with $x \neq y$ the square of the difference quotient can be expressed in terms of two derivatives via

$$\frac{\Big(g(x) - g(y)\Big)^2}{(x-y)^2} = |g'(x)| \cdot |g'(y)|.$$

Proof. In order to see that $\mathcal{G}$ is a group we first notice that $\widetilde{f} := mfm \in \mathcal{F}$ because $\widetilde{f}(z) = (c_1 z - c_2)/(-c_3 z + c_4)$ and $c_1 c_4 - (-c_2)(-c_3) > 0$. Since $m^2 = \mathrm{id}$ it follows that $\widetilde{f} m = mf$ and $m\widetilde{f} = \widetilde{f} m$. Now the verification of the group properties of $\mathcal{G}$ is straightforward.

(i) follows from Remark 2.6.(ii).

(ii): Let $g = m^l f$ with $f \in \mathcal{F}$. Then the left-hand side of the formula in (ii) is $f'(x,y)^2 = f'(x)^2 f'(y)^2$, see Definition 2.5. This equals the right-hand side because $|g'(x)| = |f'(x)|$ and $|g'(y)| = |f'(y)|$. ∎

2.3 Explicit formula for the barycentric subdivision

Definition 2.10 We define:

(i) the *rotation function* $h \in \mathcal{F}$ by

$$h(z) := \frac{z-3}{z+1} = 1 - \frac{4}{z+1},$$

(ii) the six functions $S_j \in \mathcal{G}$ by

$$S_{(l,k)} := m^l h^k, \quad (l,k) \in \mathcal{J},$$

(iii) the *rescaling function* $\alpha \in \mathcal{F}$ by

$$\alpha(z) := \tfrac{2}{3}z - 1.$$

Remark 2.11 Let us explain the reason for the name "rotation function": Lemma 6.7.(i) shows that the (associated triangles of) z and $h(z)$ are similar with the same orientation and that the k-edge of z corresponds with the $(k-1)$-edge of $h(z)$.

Proposition 2.12 For any $z \in \mathbb{H}_0$ the explicit formula for the barycentric subdivision is

$$A_j(z) = \alpha S_j(z).$$

Proof. Postponed to Proposition 6.9 (on page 52). ∎

Let us use the formulas in the above proposition and in Lemma 2.3 for the following extensions of the functions $A_j : \mathbb{H}_0 \to \overline{\mathbb{H}}_0$ and $a_j : \mathbb{H}_0 \to \Sigma$ to functions defined on all of $\overline{\mathbb{H}}_0$. We use the same notation A_j and a_j for these extended functions.

Definition 2.13 We define:

(i) $A_j := \alpha S_j \in \mathcal{G}$,

(ii) the *child function* $a_j := SA_j : \overline{\mathbb{H}}_0 \to \Sigma$.

2.4 Shape sets

Definition 2.14 We define the *shape sets*

$$\Sigma_{0,k} := \{z : \text{ the } k\text{-edge is a longest edge and the } (k-1)\text{-edge is a shortest edge of } z\},$$

$$\Sigma_{1,k} := \{z : \text{ the } k\text{-edge is a longest edge and the } (k+1)\text{-edge is a shortest edge of } z\}.$$

We include ∞ in $\Sigma_{(0,1)}$ and $\Sigma_{(1,2)}$, but not in the other shape sets.

Remark 2.15 Let us comment on the above definition:

(i) By definition, the set $\Sigma_{(0,0)}$ is just Σ.

(ii) Figure 3 shows how the imaginary axis and the two circles of radius 2 around -1 and 1 cut $\overline{\mathbb{H}}_0$ into the six (closed) shape sets. Notice that the only unbounded shape sets $\Sigma_{(0,1)}$ and $\Sigma_{(1,2)}$ are also the only shape sets that include the element ∞.

(iii) We show in Lemma 6.11.(i) that the six (not necessarily pairwise different) evaluations $S_j(z)$ with $j \in \mathcal{J}$ are exactly those points that are similar to z. In part (iii) of the same lemma we prove the characterization

$$\Sigma_j = \{z \in \overline{\mathbb{H}}_0 : S(z) = S_j(z)\}. \tag{2.4}$$

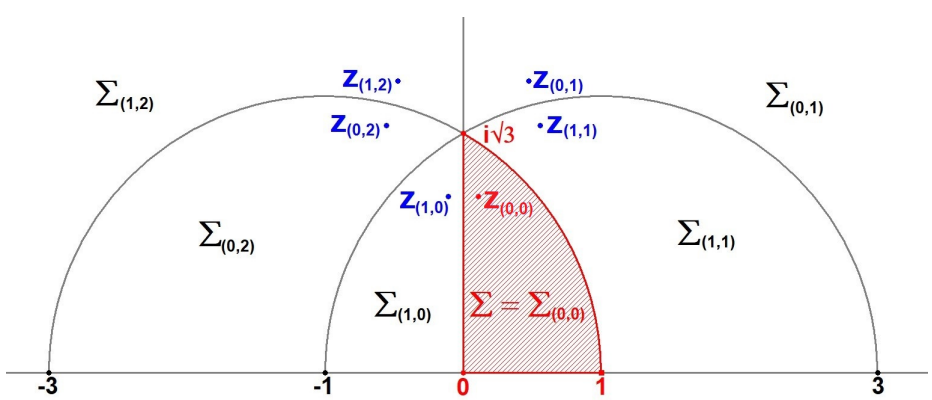

Figure 3: The shape sets Σ_j and elements $z_j \in \Sigma_j$ with $z_{(0,0)} = S(z_j) = S_j(z_j)$.

2.5 Upcircle coordinates

Throughout this subsection we assume $(x, r, d), (x', r', d') \in \mathbb{R} \times (0, \infty) \times \mathbb{R}$ and $z \in \mathbb{H}$. In Definition 1.2 we defined an *upcircle* as a circle in $\mathbb{H}_0$ with tangent line $\mathbb{R}$, see Figure 4, and we defined the *radius* of (x, z). Now we extend this definition.

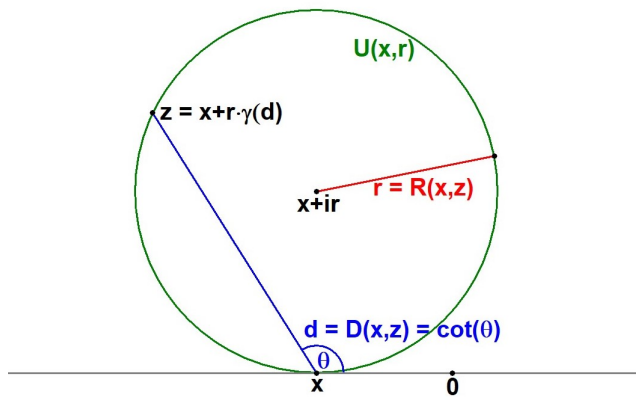

Figure 4: Upcircle $U(x, r)$.

Definition 2.16 We define the

(i) *upcircle* $U(x, r) := \{z : |z - (x + ir)| = r\} \subseteq \mathbb{H}_0$,

(ii) *radius function* $R(x, z) := |z - x|^2/(2\Im z) \in (0, \infty)$,

(iii) *angle function* $D(x, z) := (\Re z - x)/\Im z \in \mathbb{R}$,

(iv) *unit upcircle function* $\gamma(d) := 2/(d - i) \in \mathbb{H}$,

(v) *upcircle coordinates function* $u(x, z) := (x, R(x, z), D(x, z))$.

Remark 2.17 Let us comment on two parts of the above definition:

(i) The upcircle $U(x, r)$ is the circle of radius r that intersects the real line exactly in x. We call $U(0, 1)$ the *unit upcircle* and notice that $U(x, r) = x + r \cdot U(0, 1)$.

(iii) The *angle* $D(x,z)$ is the cotangent of the actual angle $\theta = \angle(x+1,x,z)$, see Figure 4.

Proposition 2.18 The objects from Definition 2.16 have the following properties:

(i) $\gamma : \mathbb{R} \to U(0,1) \setminus \{0\}$ is bijective.

(ii) u is bijective with inverse $u^{-1}(x,r,d) = (x, x + r \cdot \gamma(d)) \in \{x\} \times U(x,r)$.

Proof. Postponed to Proposition 9.3 (on page 70). ∎

Remark 2.19 Let us comment on the above proposition:

(i) Since $u^{-1}u(x,z) = (x,z)$ we see that $z = x + R(x,z) \cdot \gamma(D(x,z)) \in U(x, R(x,z))$. So $R(x,z)$ is the (unique) radius r such that z lies on the upcircle $U(x,r)$ and the position of z on that upcircle is uniquely determined by $D(x,z)$.

(ii) Since $uu^{-1}(x,r,d) = (x,r,d)$ we see that $x + r \cdot \gamma(d)$ is the (unique) point $z \in \mathbb{H}$ such that $R(x,z) = r$ and $D(x,z) = d$. Our proof in Section 9.1 uses the following lemma in the special case $x' = x$, whereas the general case with arbitrary $x' \in \mathbb{R}$ is needed for the proof of Lemma 5.9 in Section 5.3.

Lemma 2.20 Let $z := x + r \cdot \gamma(d)$ and $x' \in \mathbb{R}$. For $c := d^2 + 1$ and $y := (x - x')/(2r)$ we have the formulas

$$\frac{R(x',z)}{r} = cy^2 + 2dy + 1, \quad D(x',z) - d = cy.$$

2.6 Möbius transformations in upcircle coordinates

Definition 2.21 For $g \in \mathcal{G}$ we define $g^\sharp : \mathbb{R} \to [-\infty, \infty] \setminus \{0\}$ by

$$g^\sharp(x) := \frac{2}{x - g^{-1}(\infty)}.$$

Remark 2.22 Let us show the connection between $g^\sharp$ and the (formal) derivative $g' : \mathbb{R} \to [-\infty, \infty] \setminus \{0\}$, see Remark 2.8.(iii) (even though we do not need this in the following): Suppose that $f \in \mathcal{F}$ is not affine linear. With the notation from Definition 2.5 we have

$$f'(x) = \frac{c}{c_3^2} \cdot \frac{1}{(x - f^{-1}(\infty))^2}. \tag{2.5}$$

Hence f' and $(f^\sharp)^2$ are identical up to a constant (positive) factor. Due to Definition 2.7.(iii) of the formal derivative $(mf)'$ it follows for any $g \in \mathcal{G}$ that g' and $(g^\sharp)^2$ are identical up to a constant (real) factor.

Proposition 2.23 Let $g \in \mathcal{G}$ and $(x, z) \in \mathbb{R} \times \mathbb{H}$ such that $g(x) \neq \infty$ (which is equivalent to $|g'(x)| \neq \infty$). Suppose we know the upcircle coordinates $(x, r, d) := u(x, z)$.

(i) The upcircle coordinates $(x', r', d') := ug(x, z)$ have the formulas

$$x' = g(x), \quad r' = |g'(x)| \cdot r, \quad |d'| = \left| d + g^\sharp(x) \cdot r \right|.$$

(ii) g preserves upcircles in the following way:

$$g(U(x, r)) = U(x', r').$$

Proof. Postponed to Proposition 9.6 (on page 71). The essential reason for (ii) is that in (i) the formulas for x' and r' do not depend on d. ∎

Remark 2.24 The formula in the above proposition expresses the absolute value of the new angle d' in terms of g and the previous upcircle coordinates (x, r, d). Let us explain the geometric meaning of the absolute value of an angle d (even though we do not need this in the following): We easily verify that $\gamma(|d|) = \mathcal{Q}\gamma(d)$ for the projection $\mathcal{Q}$ from $\mathbb{H}_0$ into the first quadrant, i.e. $\mathcal{Q}(z) = |\Re z| + i\Im z$. So if a pair (x, z) has upcircle coordinates (x, r, d), or equivalently $z - x$ has upcircle coordinates $(0, r, d)$, then $\mathcal{Q}(z - x)$ has upcircle coordinates $(0, r, |d|)$.

3 Other preliminaries

This section presents auxiliary results about the (generalized) child functions required for the proofs of our main results in Sections 4 and 5. All proofs are postponed to Sections 7, 8 and 9.

3.1 Up-to-sign derivatives of the child functions

Let us introduce the following convenient notation for iterated compositions.

Definition 3.1 Let C be a set and $(f_i)_{i\in\mathcal{I}}$ be a family of functions $f_i : C \to C$. For $n \in \mathbb{N}_0$ and a multiindex $I = (i_1, \dots, i_n) \in \mathcal{I}^n$ we set $f_I := \mathrm{id}$, if $n = 0$, and otherwise

$$f_I := f_{i_n} \dots f_{i_1}.$$

For the rest of this subsection we assume that x, y are real variables with $x \neq y$. Let us restate Definition 1.3 (on page 5) in part (i) of the following definition.

Definition 3.2 Let $b : D_b \to \mathbb{R}$ be a function with $D_b \subseteq \mathbb{R}$.

(i) We say that b is *up-to-sign differentiable* if there is a continuous function $\beta : D_b \to [0, \infty)$ such that for all but finitely many $x \in D_b$ the function b is differentiable in x with $|b'(x)| = \beta(x)$. In that case, β is unique. We call β the *up-to-sign derivative* of b and we formally write $|b'| := \beta$.

(ii) We formally write $|b'(x, y)| := |b(x) - b(y)|/|x - y|$ for $x, y \in D_b$.

Proposition 3.3 We restrict the child functions to functions $a_j : [0, 1] \to [0, 1]$. Let $I \in \mathcal{J}^n$ be a multiindex with $n \in \mathbb{N}_0$. Then the composition a_I is Lipschitz-continuous and has an up-to-sign derivative $|a_I'| : [0, 1] \to (0, \infty)$. Moreover, this up-to-sign derivative satisfies:

(i) the chain rule $|a_I'(x)| = |a_i'(a_J(x))| \cdot |a_J'(x)|$, if we write $I = (J, i)$ with $J \in \mathcal{I}^{n-1}$,

(ii) the "Cauchy-Schwarz inequality" $|a'_I(x,y)| \leq |a'_I(x)|^{1/2}|a'_I(y)|^{1/2}$.

In the "$n = 1$" case $I = j \in \mathcal{J}$ we have:

(iii) $|a'_j(x)| = |S_{j'}A_j(x)|$, if $a_j(x) = S_{j'}A_j(x)$,

(iv) $|a'_j(x)| = \min_{j'} |(S_{j'}A_j)'(x)|$.

Proof. Postponed to Proposition 7.10 (on page 61). ∎

Remark 3.4 If we interpret $|a'_I(x)|$ as $|a'_I(x,x)|$, then inequality (ii) is reminiscent of the actual Cauchy-Schwarz inequality. Inequality (ii) is crucial for Section 4, where we prove the ergodicity of the flat shape chain.

3.2 Generalized shape functions

Remark 3.5 According to Remark 2.15.(iv), on each shape set Σ_j the shape function S equals the upcircle preserving function $S_j \in \mathcal{G}$, but unfortunately, S itself does not preserve upcircles. Why? Imagine we fix $x \in \mathbb{R} \cap \Sigma_j$ and let r grow. Then $C := U(x,r)$ grows and will eventually contain an element $z \in \Sigma_{j'} \setminus \Sigma_j$. But then applying S pointwise to C means that we apply the different upcircle preserving functions S_j and $S_{j'}$ to different parts of C, thus $S(C)$ is not necessarily an upcircle anymore ($S(C)$ is definitely not an upcircle, if $x \notin \Sigma_{j'}$).

From now on we write $v = (v_1, v_2)$ for (x, z) and consider a fixed component index $c \in \{1, 2\}$.

Remark 3.6 If we defined $S(v)$ component-wise, then we would have $S(v) = (S_j(v_1), S_{j'}(v_2))$ for potentially **different** upcircle preserving functions S_j and $S_{j'}$. Instead, we construct in the following a *generalized shape function* $S^c : D \to D$ for some set $D \subseteq \mathbb{R} \times \mathbb{H}$ which for any input v applies the **same** upcircle preserving function S_j to both components of v, where j is chosen such that $v_c \in \Sigma_j$. This choice ensures that the cth component

of $S^c(v) = (S_j(v_1), S_j(v_2))$, namely $S_j(v_c)$, equals $S(v_c)$ and therefore lies in Σ (unfortunately, the other component of $S^c(v)$ does not necessarily lie in Σ). With the projection $\pi_c(v_1, v_2) := v_c$ onto the cth component we can express this "design feature" of S^c as

$$\pi_c S^c = S\pi_c : D \to \Sigma. \tag{3.1}$$

In Definition 8.1 we define the *exceptional set* $E \subseteq \mathbb{R}$ and the above mentioned set D as

$$D := (\mathbb{R} \setminus E) \times \mathbb{H}. \tag{3.2}$$

The details do not matter here, all we need to know for now is that the following lemma and proposition hold.

Lemma 3.7 The sets E and D have the following properties:

(i) E is countable.

(ii) If $(x, z) \in D$, then $A_j(x, z) \in D$.

Proof. Postponed to Lemma 8.2 (on page 62). ∎

Proposition 3.8 There is a function $S^c : D \to D$ such that any $v \in D$ satisfies the following two conditions:

(i) $S^c(v) = S_j(v)$ for some j such that $v_c \in \Sigma_j$,

(ii) $S^c S_j(v) = S^c(v)$ for any j.

Proof. Postponed to Proposition 8.3 (on page 63). ∎

Corollary 3.9 Any $v \in D$ satisfies the following two conditions:

(i) The multiset of the six evaluations $S_j S^c(v)$ equals the multiset of the six evaluations $S_j(v)$.

(ii) $S^{c'} S^c(v) = S^{c'}(v)$ for any $c, c' \in \{1, 2\}$.

Proof. Postponed to Corollary 8.4 (on page 63). ∎

The functions S^1, S^2 satisfying the conditions in the above proposition are not unique. In the following, we fix a specific choice of such functions S^1, S^2 and call them the *generalized shape functions*.

3.3 Definition of the generalized child functions

Definition 3.10 For $c \in \{1, 2\}$ we define the *generalized child function*

$$a_{c,j} := S^c A_j : D \to D.$$

Indeed, $a_{c,j}$ maps D to D because (the two-dimensional version of) A_j maps D to D, see Lemma 3.7.(ii). Only for the following two remarks we set

$$g_{j',j} := S_{j'} A_j \in \mathcal{G}, \quad j', j \in \mathcal{J}, \tag{3.3}$$

and notice for any j and z that there exists an index j' such that

$$a_j(z) = g_{j',j}(z). \tag{3.4}$$

The following remark is an almost literal copy of Remark 3.6 with the (generalized) shape function replaced by a (generalized) child function.

Remark 3.11 If we defined $a_j(v)$ component-wise, then we would have $a_j(v) = (g_{j',j}(v_1), g_{j'',j}(v_2))$ for potentially **different** upcircle preserving functions $g_{j',j}$ and $g_{j'',j}$. Instead, each generalized child function $a_{c,j}$ applies the **same** upcircle preserving function $g_{j',j}$ to both components of v, where j' is chosen such that $A_j(v_c) \in \Sigma_{j'}$. This choice ensures that the cth component of $a_{c,j}(v)$, namely $g_{j',j}(v_c) = S_{j'} A_j(v_c) = S A_j(v_c) = a_j(v_c)$, lies in Σ (unfortunately, the other component of $a_{c,j}(v)$ does not necessarily lie in Σ). With the projection $\pi_c(v_1, v_2) := v_c$ onto the cth component we can express this "design feature" of $a_{c,j}$ as

$$\pi_c a_{c,j} = a_j \pi_c : D \to \Sigma. \tag{3.5}$$

We also could have shown the last line with the "design feature" (3.1) of S^c, that is $\pi_c S^c = S\pi_c : D \to \Sigma$. We compose from the right side with the function $A_j : D \to D$. Together with $\pi_c A_j = A_j \pi_c$, which is a consequence of the component-wise definition of $A_j(v)$, it follows (3.5).

Remark 3.12 Let us show consistency with Section 1.4, where we alluded to the generalized child functions. This remark may be skipped at first reading as the following proofs do not formally require this discussion. In (1.16) we presented the "characteristic feature" of $a_{c,j}$: For "most" $v = (x, z) \in D$ we obtain $a_{c,j}(v)$ by applying the same function $a_j^{v_c} \in \mathcal{G}_j \subseteq \mathcal{G}$ to both components of v, that is

$$a_{c,j}(v) = a_j^{v_c}(v). \tag{3.6}$$

Let us now clarify the meaning of "most" and reveal the definitions of those objects that were just informally introduced in Section 1.4. Consider the subset $\mathcal{G}_j := \{g_{j',j} : j' \in \mathcal{J}\}$ of $\mathcal{G}$. The six (not necessarily pairwise different) evaluations $g(z)$ with $g \in \mathcal{G}_j$ are exactly those points that are similar to $A_j(z)$, or equivalently, that are similar to $a_j(z)$. For any fixed z there is a (not necessarily unique) index j' such that $a_j(z) = g_{j',j}(z)$ and for one such index j' we set

$$a_j^z := g_{j',j} \in \mathcal{G}_j. \tag{3.7}$$

Hence (1.15) is satisfied, that is $a_j(z) = a_j^z(z)$. We claimed "most" elements (x, z) of D satisfy (1.16), that is

$$a_{1,j}(x, z) = a_j^x(x, z), \quad a_{2,j}(x, z) = a_j^z(x, z). \tag{3.8}$$

Let us discuss what "most" means. Fix j and a component index $c \in \{1, 2\}$. Let us define D_j^c as the set of all pairs $v \in D$ such that the pair $v' := A_j(v)$ has the following property: Its cth component $v'_c = A_j(v_c)$ lies in exactly one shape set, which means that there is just one possible choice for the index j' in the above definition of $a_j^{v_c}$. But by Proposition 3.8 this also means that there is just one possible choice j' such that $S^c(v') = S_{j'}(v')$. Thus any $v \in D_j^c$ satisfies

$$a_{c,j}(v) = S^c A_j(v) = S_{j'} A_j(v) = g_{j',j}(v) = a_j^{v_c}(v). \tag{3.9}$$

This is the cth equation in (3.8). In Definition 8.1 we formally introduce D and it is then clear that D_j^1 is actually the same as D and that $D \setminus D_j^2$ has Lebesgue measure 0. So it is fair to say that "most" elements of D lie in D_j^c.

3.4 Generalized child functions in upcircle coordinates

In this subsection we assume that x, y are real variables with $x \neq y$. We consider the generalized child functions $a_{c,j} : D \to D$ in the case $c = 1$. In Definition 9.7 we define a function $a_j^\sharp : \mathbb{R} \setminus E \to \mathbb{R}$ such that the following lemma holds which is a consequence of Proposition 2.23.(i).

Lemma 3.13 The function $a_j^\sharp$ has the following properties:

(i) $a_j^\sharp$ is well-defined and bounded.

(ii) Given the upcircle coordinates $(x, r, d) := u(x, z)$, we have the following formula for the upcircle coordinates $(x', r', d') := ua_{1,j}(x, z)$:

$$x' = a_j(x), \quad r' = |a_j'(x)| \cdot r, \quad |d'| = |d + a_j^\sharp(x) \cdot r|.$$

Proof. Postponed to Lemma 9.9 (on page 73). ∎

Let us introduce the following functions that are central for Sections 4 and 5.

Definition 3.14 Let $n \in \mathbb{N}_0$ and $J_n := (j_1, \ldots, j_n)$ be a random multiindex. We define (random) functions x_n, r_n, d_n and a (non-random) function Φ_n:

(i) $x_n : [0, 1] \to [0, 1] : \ x \mapsto a_{J_n}(x)$,

(ii) $r_n : [0, 1] \to (0, \infty) : \ x \mapsto |x_n'(x)|$,

(iii) $\Phi_n : [0, 1] \to \mathbb{R} : \ x \mapsto \mathbb{E} \ln r_n(x)$,

(iv) $d_n : D \to \mathbb{R} : \ v \mapsto D(a_{1,J_n}(v))$.

Note that $x_0(x) = x$ and $r_0(x) = 1$, as well as $\Phi_0(x) = 0$ and $d_0(v) = D(v)$. We use the same notation r_n as for the up-to-sign derivative and define for $x, y \in [0,1]$ the difference quotient

$$r_n(x,y) := \frac{|x_n(x) - x_n(y)|}{|x-y|}. \tag{3.10}$$

Remark 3.15 Let us translate parts (i) and (ii) of the (non-random) Proposition 3.3 into this new (random) framework: The function x_n is Lipschitz-continuous and has the up-to-sign derivative $r_n : [0,1] \to (0,\infty)$. Moreover, r_n satisfies

(i) the chain rule $r_n(x) = |a'_{j_n} x_{n-1}(x)| \cdot r_{n-1}(x)$,

(ii) the "Cauchy-Schwarz inequality" $r_n(x,y) \leq r_n(x)^{1/2} r_n(y)^{1/2}$.

Lemma 3.16 Let $v \in D$ and $(x, r, d) := u(v)$.

(i) For $n \in \mathbb{N}_0$ the pair $a_{1,J_n}(v)$ has the upcircle coordinates

$$ua_{1,J_n}(v) = \Big(x_n(x), r_n(x) \cdot r, d_n(v)\Big).$$

(ii) For $n \in \mathbb{N}$ we have the recursive formulas

$$r_n(x) = |a'_{j_n} x_{n-1}(x)| \cdot r_{n-1}(x), \tag{3.11}$$

$$|d_n(v)| = |d_{n-1}(v) + a^\sharp_{j_n}(x) \cdot r_{n-1}(x) \cdot r|. \tag{3.12}$$

Proof. (i): A simple induction using Remark 3.15.(i) and Lemma 3.13.(ii) shows that the first two components of the upcircle coordinates ua_{1,J_n} are indeed $x_n(x)$ and $r_n(x) \cdot r$.

(ii) folllows from (i) by using Lemma 3.13.(ii) a second time. ∎

4 Ergodicity of the flat shape chain

Theorem 1 in [2] shows the ergodicity of the flat shape chain X_n under the assumption of the "average contractivity between points" condition. This technical assumption requires the geometric mean of the difference quotients $|a_j'(x,y)|$ to be uniformly bounded in x and y by a number strictly less than 1. In Section 4.1 we prove the equivalence of average contractivity to the boundedness of a one-dimensional function which we later verify in Section 4.3. In Section 4.2 we discuss the ergodicity and the law of large numbers for X_n. The existence of an invariant measure μ leads to the proof of the first theorem (Theorem 1.4) in Section 4.4. The law of large numbers for X_n is modified in Section 4.5 for a two-dimensional Markov chain whose first component is X_n. This modified law of large numbers provides the main technical tool for the proof of the second theorem (Theorem 1.5) in Section 5.

The ergodic results of Sections 4.2 and 4.5 are not specific to the iterated barycentric subdivision. We take account of this by formulating them in a more general setting in order to emphasize the relevant underlying structure. We "weave in" explanations (called *applications*) how the general framework applies to our specific setting of the iterated barycentric subdivision.

4.1 Contractivity on average

We remind ourselves of the notation from Definition 3.1 for iterated compositions $f_I = f_{i_n} \dots f_{i_1}$ with $I = (i_1, \dots, i_n)$.

Setting 4.1 We impose the following general setting:

(i) Let $d \in \mathbb{N}$ and C be a compact subset of $(\mathbb{R}^d, |\cdot|)$, where $|\cdot|$ is some norm. We fix a finite index set $\mathcal{I}$ and always assume $i \in \mathcal{I}$. Let $(f_i)_i$ be a family of Lipschitz-continuous functions $f_i : C \to C$.

(ii) Consider the Markov kernel M from C to C given by

$$M(x,\cdot) := \frac{1}{|\mathcal{I}|}\sum_i \delta_{f_i(x)}, \quad x \in C. \tag{4.1}$$

(iii) Let $i_1, i_2, \ldots$ be a sequence of independent random indices such that each i_n is uniformly distributed on $\mathcal{I}$. For $n \in \mathbb{N}_0$ we define the random multiindex $I_n := (i_1, \ldots, i_n)$ which is uniformly distributed on $\mathcal{I}^n$.

Remark 4.2 If each function f_i is not only Lipschitz-continuous, but even a contraction, then $(f_i)_i$ contracts on average. However, there are families $(f_i)_i$ which contract on average without f_i being (global) contractions.

Application 4.3 The setting of the flat barycentric subdivision is a special case of Setting 4.1 with $d = 1$, $C = [0,1]$, $\mathcal{I} = \mathcal{J}$ and $f_i : C \to C$ as the restricted child function a_i. Hence $M = M_{\text{flat}}$. We also identify $i_n = j_n$ and $I_n = J_n$.

For the rest of this subsection we always assume that $x \neq y$. We remind ourselves of the notation $|b'(x,y)| = |b(x) - b(y)|/|x - y|$, see Definition 3.2.

Definition 4.4 Let us define the following notions of contractivity on average:

(i) $(f_i)_i$ *contracts on average in step* $N \in \mathbb{N}$, if

$$\sup_{x,y\in C} \mathbb{E} \ln |f'_{I_N}(x,y)| < 0,$$

(ii) $(f_i)_i$ *strongly contracts on average in step* $N \in \mathbb{N}$, if there exists $\epsilon > 0$ such that

$$\sup_{x,y\in C} \mathbb{E} \ln \Big(\max\{|f'_{I_N}(x,y)|, \epsilon\} \Big) < 0.$$

(iii) $(f_i)_i$ (*strongly*) *contracts on average*, if it (strongly) contracts on average in step 1.

Application 4.5 The family $(a_j)_j$ contracts on average in step N, if and only if

$$\sup_{x,y\in[0,1]} \mathbb{E} \ln r_N(x,y) < 0. \tag{4.2}$$

Remark 4.6 The very technical assumption (ii) in the above definition is tailor-made for the modified law of large numbers stated in Proposition 4.25. It also obviously implies assumption (i) on which we comment now:

(i) The family $(f_i)_{i\in\mathcal{I}}$ contracts on average in step N if and only if the family $(f_I)_{I\in\mathcal{I}^N}$ contracts on average. The reason is that I_N is uniformly distributed on $\mathcal{I}^N$.

(ii) If $(f_i)_i$ contracts on average, then it also contracts on average in any step $N \in \mathbb{N}$. Even though we do not need this observation in the following, let us briefly justify it: We immediately verify that

$$|f'_{I_N}(x,y)| = |f'_{i_n}(f_{I_{n-1}}(x), f_{I_{N-1}}(y))| \cdot |f'_{I_{n-1}}(x,y)|. \tag{4.3}$$

Applying $\sup_{x,y} \mathbb{E}\ln(\cdot)$ to both sides and using the independence of i_n and I_{n-1} shows that $(f_i)_i$ contracts on average in step N, if $(f_i)_i$ contracts on average in step 1 and $N-1$. Now we use induction over $N \in \mathbb{N}$.

(iii) The family $(f_i)_i$ contracts on average in step N if and only if there exists $\kappa < 0$ such that any $x, y \in C$ satisfy

$$\frac{1}{|\mathcal{I}|}\sum_i \ln|f'_i(x,y)| \le \kappa.$$

Application 4.7 Let $N \in \mathbb{N}$. The following conditions are equivalent:

(i) $(a_j)_j$ strongly contracts on average in step N.

(ii) $(a_j)_j$ contracts on average in step N.

(iii) Φ_N is bounded from above by a negative number.

Proof. Remember that $\Phi_N = \mathbb{E}\ln r_N : [0,1] \to \mathbb{R}$, where r_N denotes the up-to-sign derivative of $x_N = a_{J_N}$.

"(i)⇒(ii)" is trivial.

"(ii)⇒(iii)": Since $r_N(x) = \lim_{y\to x} r_N(x,y)$ we have $\Phi(x) = \lim_{y\to x}\mathbb{E}\ln r_N(x,y)$. Thus

$$\sup_{x\in[0,1]}\Phi(x) \le \sup_{x,y\in[0,1]}\mathbb{E}\ln r_N(x,y). \tag{4.4}$$

"(iii)⇒(i)": Any $x, y \in [0,1]$ satisfy the "Cauchy-Schwarz inequality" in Remark 3.15, namely

$$r_N(x,y) \leq r_N(x)^{1/2} r_N(y)^{1/2}. \tag{4.5}$$

The image set C' of r_N is a subset of $(0,\infty)$ and compact in $\mathbb{R}$. Let $\epsilon := \min C' > 0$. Thus the right-hand side of inequality (4.5) is at least ϵ. Applying the non-decreasing function $\ln(\max\{\cdot,\epsilon\}) : \mathbb{R} \to \mathbb{R}$ to the inequality shows that

$$\ln\left(\max\left\{r_N(x,y),\epsilon\right\}\right) \leq \tfrac{1}{2}\ln r_N(x) + \tfrac{1}{2}\ln r_N(y). \tag{4.6}$$

Taking first the expectation and then the supremum shows that (iii) implies (i). ∎

Remark 4.8 Let us comment on the equivalence in Application 4.7:

(i) Condition (iii) is the easiest to verify because Φ_N is one-dimensional, whereas each difference quotient $r_N(\cdot,\cdot)$ is a two-dimensional function.

(ii) Using this equivalence, we show in Remark 4.17 that $(a_j)_j$ (strongly) contracts on average in some step $N \in \mathbb{N}$. For our purpose the existence of such N is all we need.

(iii) Without using this equivalence, [6] show via numerical computations[1] that $(a_j)_j$ contracts on average in step 2, but not in step 1.

4.2 Law of large numbers

Let us impose Setting 4.1 in this subsection.

Definition 4.9 We define:

(i) M is *ergodic*, if it has a unique invariant probability measure μ which is also attracting, i.e. $\mu M = \mu$ and νM^n converges weakly to μ for any probability measure ν on C,

[1] included in the appendix of [7], but not included in the published version [6].

(ii) For any $x \in C$ we define the Markov chain

$$U_n^x := f_{I_n}(x), \quad n \in \mathbb{N}_0,$$

which starts in $U_0^x = x$ and has kernel M.

(iii) M satisfies the *law of large numbers*, if M is ergodic with invariant measure μ and for any continuous function $\phi : C \to \mathbb{R}$ and $x \in C$ we have a.s.

$$\lim_n \frac{1}{n+1} \sum_{t=0}^{n} \phi(U_t^x) = \int \phi \, d\mu.$$

Application 4.10 As a consequence of $f_i = a_i$ and $I_n = J_n$ we have $U_n^x = a_{J_n}(x) = x_n(x)$, see Definition 3.14.(i).

Proposition 4.11 If $(f_i)_i$ contracts on average, then M satisfies the law of large numbers.

Proof. We apply Theorem 1 from [2] in the following special case: The complete metric space (X, d) is $(C, |\cdot|)$, the Lipschitz-continuous function w_i is f_i for an index i and the probability p_i to choose w_i in any given step is (uniformly) $1/|\mathcal{I}|$. The assumption of this theorem is exactly the equivalent reformulation of $(f_i)_i$ contracting on average in Remark 4.6.(iii). The first part of the theorem states that M is ergodic.

Using the ergodicity of M, Lemma 5.2 from [6] shows the law of large numbers for M. Again, we work with the uniform distribution $p_i = 1/|\mathcal{I}|$. Their lemma is presented in the special case $d = 1$, but they emphasize that the compact set C (or S in their notation) can be any complete, separable metric space. ∎

Lemma 4.12 Let $N \in \mathbb{N}$. The kernel $M^N = M \dots M$ has the formula

$$M^N(x, \cdot) = \frac{1}{|\mathcal{I}^N|} \sum_{I \in \mathcal{I}^N} \delta_{f_I(x)}, \quad x \in C, \tag{4.7}$$

and satisfies the following properties:

(i) If M^N is ergodic, then M is ergodic.

(ii) If M^N satisfies the law of large numbers, then M satisfies the law of large numbers.

Proof. Postponed to Lemma 10.4 (on page 79) where we present the arguments in the proof of Proposition 5.3 from [6] in greater detail. ∎

Proposition 4.13 If $(f_i)_i$ contracts on average in some step N, then M satisfies the law of large numbers

Proof. As mentioned in Remark 4.6.(i) the family $(f_I)_{I\in\mathcal{I}^N}$ contracts on average, so by Proposition 4.11 the kernel M^N satisfies the law of large numbers. Applying Lemma 4.12.(ii) finishes the proof. ∎

The following lemma presents other formulas for the right-hand side in Definition 4.9.(iii).

Lemma 4.14 Suppose M has an invariant measure μ. Let $\phi : C \to \mathbb{R}$ be continuous and $N \in \mathbb{N}_0$. Then

$$\int \phi \, d\mu = \mathbb{E} \int \phi f_{I_N} \, d\mu.$$

Proof. Since $\mu = \mu M^N$ we know

$$\int \phi \, d\mu = \int \int \phi \, dM^N(x, \cdot) \, \mu(dx). \tag{4.8}$$

Due to formula (4.7) for M^N, the inner integral in the above double integral equals

$$\frac{1}{|\mathcal{I}^N|} \sum_{I\in\mathcal{I}^N} \sum_I \phi f_I(x) = \mathbb{E}\phi(U_N^x). \tag{4.9}$$

It only remains to interchange the outer integral with the expectation $\mathbb{E}$ which can be written as a finite sum. ∎

4.3 Shape integrand

Where do the bounds $\chi_{\min} = \frac{1}{3}\ln\frac{3}{2}$ and $\chi_{\max} = \frac{1}{3}\ln\frac{91}{2} - \ln 3$ in Theorem 1.4 come from? In Definition 11.1 we introduce (non-random) continuous functions $\Psi, \Psi_n : \overline{\mathbb{R}} \to \mathbb{R}$ with

$$\Psi_n(x) := \mathbb{E}\Psi x_n(x), \quad n \in \mathbb{N}_0. \tag{4.10}$$

We call Ψ the *shape integrand* and Ψ_n the *shape integrand of order* n. The details of Ψ are not relevant for this section; all we need to know for now is that the following two propositions hold.

Proposition 4.15 For any $n \in \mathbb{N}_0$ and $x \in [0,1]$ we have the following approximation of $\Phi_n(x)$ by the sum of the shape integrands of orders up to $n-1$:

$$\left| \Phi_n(x) - \sum_{t=0}^{n-1} \Psi_t(x) \right| \leq \ln \tfrac{4}{3}.$$

Proof. Postponed to Proposition 11.3 (on page 86). ∎

Proposition 4.16 The shape integrand has the following extrema:

$$\max_{x\in[0,1]} \Psi(x) = \Psi(0) = -\chi_{\min}, \qquad \min_{x\in[0,1]} \Psi(x) = \Psi(1) = -\chi_{\max}.$$

Proof. Postponed to Proposition 11.5 (on page 87). ∎

Remark 4.17 By Propositions 4.15 and 4.16 we have that

$$\sup_{x\in[0,1]} \Phi_n(x) \leq \sum_{t=0}^{n-1} \sup_{x\in[0,1]} \Psi_t(x) + \ln \tfrac{4}{3} \leq -\chi_{\min} \cdot n + \ln \tfrac{4}{3}, \tag{4.11}$$

$$\inf_{x\in[0,1]} \Phi_n(x) \geq \sum_{t=0}^{n-1} \inf_{x\in[0,1]} \Psi_t(x) - \ln \tfrac{4}{3} \geq -\chi_{\max} \cdot n - \ln \tfrac{4}{3}. \tag{4.12}$$

4.4 Proof of the first theorem

Let us restate and prove Theorem 1.4 (on page 5).

Theorem 4.18 The kernel M_{flat} of X_n has an invariant measure μ, i.e. $\mu M_{\text{flat}} = \mu$, and μ is unique. Each child function a_j viewed as a function $[0,1] \to [0,1]$ is Lipschitz-continuous and has an up-to-sign derivative $|a_j'| : [0,1] \to (0,\infty)$. The constant

$$\chi := -\tfrac{1}{6} \sum_j \int \ln |a_j'| \; d\mu$$

is bounded from below by $\chi_{\min} := \frac{1}{3} \ln \frac{3}{2} \approx 0.1352$ and bounded from above by $\chi_{\max} := \frac{1}{3} \ln \frac{91}{2} - \ln 3 \approx 0.1740$.

Proof. Since $\chi_{\min} > 0$ the upper bound (4.11) implies for some $N \in \mathbb{N}$ that the function Φ_N is bounded from above by a negative number. According to Application (4.7) this is equivalent to $(a_j)_j$ contracting on average in step N. It follows with Proposition 4.13 that

M_{flat} satisfies the law of large numbers. In particular, M_{flat} has an invariant measure μ. Choosing $\phi = \Psi : [0,1] \to \mathbb{R}$ shows that a.s.

$$\lim_n \frac{1}{n} \sum_{t=0}^{n-1} \Psi x_t(x) = \int \Psi \, d\mu =: \kappa. \tag{4.13}$$

By Proposition 4.16 we see that κ and each term $\frac{1}{n}\sum_{t=0}^{n-1} \Psi x_t(x)$ with $n \in \mathbb{N}$ are bounded from below by $-\chi_{\max}$ and from above by $-\chi_{\min}$. Taking the expectation and using the definition $\Psi_t(x) = \mathbb{E}\Psi x_t(x)$, as well as the dominated convergence theorem leads to the a.s. convergence

$$\lim_n \frac{1}{n} \sum_{t=0}^{n-1} \Psi_t(x) = \kappa. \tag{4.14}$$

It follows with Proposition 4.15 that a.s. $\lim_n \frac{1}{n}\Phi_n(x) = \kappa$. According to Remark 4.17 the random variables $\frac{1}{n}\Phi_n(x)$ are uniformly bounded in $n \in \mathbb{N}$. Taking the integral w.r.t. the probability measure μ and applying again the dominated convergence theorem leads to the a.s. convergence

$$\lim_n \frac{1}{n} \int \Phi_n \, d\mu = \kappa. \tag{4.15}$$

It remains to show that $\int \Phi_n \, d\mu = -\chi \cdot n$ for any $n \in \mathbb{N}_0$ because this implies that $\chi = -\kappa \in [\chi_{\min}, \chi_{\max}]$. Lemma 4.14 in the case $\phi = \ln|a'_j|$ yields $\int \ln|a'_j| \, d\mu = \mathbb{E}\int \ln|a'_j x_{n-1}(x)| \, \mu(dx)$. Together with the definition of χ and the independence of j_n, $x_{n-1}(x)$ it follows that

$$-\chi = \tfrac{1}{6}\sum_j \int \ln|a'_j| \, d\mu = \mathbb{E}\int \ln|a'_{j_n}(x)| \, \mu(dx) = \mathbb{E}\int \ln|a'_{j_n} x_{n-1}(x)| \, \mu(dx). \tag{4.16}$$

In the last term we can interchange the expectation $\mathbb{E}$ and the integral w.r.t. μ. The implication $\Psi_n(x) = \mathbb{E}\ln|a'_{j_n} x_{n-1}(x)| + \Psi_{n-1}(x)$ of the chain rule $r_n(x) = |a'_{j_n} x_{n-1}(x)| \cdot r_{n-1}(x)$, see Remark 3.15.(i), results in

$$-\chi = \int \Phi_n(x) \, d\mu - \int \Phi_{n-1}(x) \, d\mu. \tag{4.17}$$

Together with $\Phi_0(x) = 0$ it follows that indeed $\int \Phi_n \, d\mu = -\chi \cdot n$. ∎

Remark 4.19 The above proof shows that $-\chi = \int \Psi \, d\mu$. It follows with Lemma 4.14 that

$$-\chi = \int \Psi_n \, d\mu, \quad n \in \mathbb{N}_0. \tag{4.18}$$

This explains the origin of the name "shape integrand of order n". Numerical simulations suggest that the (non-random) functions Ψ_n converge uniformly to a constant which must be $-\chi$. In principal, this provides an opportunity to approximate χ arbitrarily close.

Definition 4.20 Let $C'_{\text{flat}} := \bigcup_j \{|a'_j(x)| : x \in [0,1]\}$.

Remark 4.21 The set C'_{flat} is a subset of $(0,\infty)$ and compact in $\mathbb{R}$. Why? In the above Theorem 4.18 we state that $|a'_j|$ only takes values in $(0,\infty)$. By definition, each up-to-sign derivative $|a'_j|$ is continuous and defined on a compact set. Hence C'_{flat} is compact.

4.5 Modified law of large numbers

Setting 4.22 We extend Setting 4.1 in the following way:

(i) Let $(f'_i)_i$ be a family of Lipschitz-continuous functions $f'_i : C \to C'$ where $C' \subseteq \mathbb{R}$ is compact. Thus $\widetilde{C} := C \times C' \subseteq \mathbb{R}^{d+1}$ is compact, too.

(ii) For any $n \in \mathbb{N}$ and multiindices $I = (i, J) \in \mathcal{I}^n$ with $J \in \mathcal{I}^{n-1}$ let us define $F_I : C \to \widetilde{C}$ by

$$F_I := (f_I, f'_i f_J). \tag{4.19}$$

In particular, $F_i = (f_i, f'_i)$.

(iii) In Definition 4.9.(ii) we defined for any $x \in C$ the Markov chain $U^x_n = f_{I_n}(x)$ with state space C. Now we define for any $x \in C$ a stochastic process $\widetilde{U}^x_n$ with state space $\widetilde{C}$ by

$$\widetilde{U}^x_n := F_{I_n}(x), \quad n \in \mathbb{N}_0. \tag{4.20}$$

Remark 4.23 A priori, there is no connection between f'_i and f_i in the above general setting. The prime symbol is merely a notation and does not necessarily stand for any kind of derivative. However, in the following application we choose f'_i to be the up-to-sign derivative of f_i.

Application 4.24 We remind ourselves of Application 4.3 where we chose f_i to be the restricted child function $a_i : [0,1] \to [0,1]$. In particular, $f_{I_n} = a_{J_n} = x_n$. Now we choose f'_i to be the up-to-sign derivative $|a'_i| : [0,1] \to C'$ with the compact set $C' := C'_{\text{flat}}$, see Definition 4.20. In the above setting we have $\widetilde{C} = [0,1] \times C'$. By definition and by the chain rule (3.11) any $n \in \mathbb{N}$ fulfills

$$F_{I_n} : [0,1] \to \widetilde{C} : x \mapsto \Big(x_n(x), |a'_{j_n} x_{n-1}(x)|\Big) = \left(x_n(x), \frac{r_n(x)}{r_{n-1}(x)}\right). \tag{4.21}$$

Let us present the modified law of large numbers.

Proposition 4.25 Suppose that $(f_i)_i$ strongly contracts on average in some step N. Let μ denote the invariant measure of M, see . For any continuous function $\phi : \widetilde{C} \to \mathbb{R}$ and $x \in C$ we have a.s.

$$\lim_n \frac{1}{n} \sum_{t=1}^{n} \phi(\widetilde{U}_t^x) = \frac{1}{|\mathcal{I}|} \sum_i \int \phi F_i \, d\mu. \tag{4.22}$$

Proof. Postponed to Proposition 10.8 (on page 81). ∎

Remark 4.26 In the situation of the above proposition, the existence of the invariant measure μ of M is ensured by Proposition 4.13 and the fact that in particular, $(f_i)_i$ contracts in step N.

We remind ourselves of Definition 1.1 which defines for any sequence of complex random variables W_n the meaning of W_n decaying exactly (or at least) with some (positive) rate.

Application 4.27 For any $x \in [0, 1]$ the sequence $r_n(x)$ decays exactly with rate χ.

Proof. We remind ourselves of the previous Application 4.24. Let μ denote the invariant measure of M_{flat}, see Theorem 4.18. We apply Proposition 4.25 for the continuous function $\phi : \widetilde{C} \to \mathbb{R} : (x, x') \mapsto \ln(x')$. Then the right-hand side of (4.22) is just $-\chi$ by definition. According to (4.21) the left-hand side of (4.22) is

$$\lim_n \frac{1}{n} \sum_{t=1}^{n} \ln \left(\frac{r_t(x)}{r_{t-1}(x)} \right) = \lim_n \frac{1}{n} \ln r_n(x). \tag{4.23}$$

This means that $r_n(x)$ decays exactly with rate χ. ∎

5 Asymptotic exponential decay

The auxiliary results in Sections 5.1 - 5.5 pave the way to the proof of the second theorem in Section 5.6. Some of these results are presented in a slightly more general framework. We continue to "weave in" explanations (called *applications*) how the general framework applies to our specific setting of the iterated barycentric subdivision.

5.1 Asymptotic exponential rates

Remark 5.1 In the following definition we distinguish between the subset $\overline{\mathbb{R}} = \{z \in \mathbb{C} : \Im(z) = 0\} \cup \{\infty\}$ of $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ and the extended real number line $[-\infty, \infty] = \mathbb{R} \cup \{-\infty, \infty\}$. We only view $[-\infty, \infty]$ as a partially (and totally) ordered set. So in particular, we distinguish between the elements $-\infty$ and ∞ of $[-\infty, \infty]$, whereas we regard the elements $-\infty$ and ∞ of $\overline{\mathbb{R}}$ as identical. We always interpret absolute values as elements of $[0, \infty] \subseteq [-\infty, \infty]$.

Definition 5.2 Let $\kappa, \kappa' \in \mathbb{R}$ and w_n be a sequence in $\overline{\mathbb{C}}$ and v_n be a sequence in $\mathbb{R} \times \mathbb{H}$. We formally write $\ln(0) = -\infty$ and say that a sequence

(i) w_n *has at most rate* κ if $\frac{1}{n} \limsup_n \ln |w_n| \leq \kappa$,

(ii) w_n *has at least rate* κ if $\liminf_n \frac{1}{n} \ln |w_n| \geq \kappa$,

(iii) w_n *has exactly rate* κ if $\lim_n \frac{1}{n} \ln |w_n| = \kappa$,

(iv) v_n *has upcircle rates* κ, κ' if $R(v_n)$ has exactly rate κ and $|D(v_n)| - d_\infty$ has at most rate κ' for some number $d_\infty \in [0, \infty)$.

Lemma A.1 in the appendix collects several elementary facts about asymptotic exponential rates which we use in the following without explicitly referring to them.

Remark 5.3 Let us reformulate Definition 1.1 with the new terminology from the above definition: A sequence of complex random variables W_n decays exactly (at least) with rate κ, if and only if $\kappa > 0$ and W_n has a.s. exactly (at most) rate $-\kappa$.

Application 5.4 Application 4.27 states for any $x \in [0,1]$ that $r_n(x)$ has exactly rate $-\chi < 0$.

5.2 Series of complex numbers

Proposition 5.5 Let w_n, w'_n be sequences in $\mathbb{C}$. Suppose that w_n has at most some rate $\kappa \in \mathbb{R}$.

(i) If $\kappa \geq 0$ and $w'_n = \sum_{t=1}^{n} w_t$ then w'_n has at most rate κ.

(ii) If $\kappa < 0$ and w'_n satisfies the recursion $|w'_n| = |w'_{n-1} + w_n|$ then $|w'_n| - w_\infty$ has at most rate κ for some number $w_\infty \in [0, \infty)$.

Proof. For both parts it suffices to show that w'_n resp. $|w'_n| - w_\infty$ has at most rate κ' for a fixed but arbitrary $\kappa' > \kappa$. There is some $n_0 \in \mathbb{N}$ such that any $n \geq n_0$ satisfies $|w_n| \leq e^{\kappa' n}$.

(i): Since $\kappa' > 0$ we have

$$\sum_{t=n_0}^{n} |w_t| \leq \sum_{t=0}^{n} e^{\kappa' t} = \frac{e^{\kappa'(n+1)} - 1}{e^{\kappa'} - 1} \leq \frac{e^{\kappa'}}{e^{\kappa'} - 1} e^{\kappa' n}. \tag{5.1}$$

Together with $|w'_n| \leq \sum_{t=1}^{n_0-1} |w_t| + \sum_{t=n_0}^{n} |w_t|$ it follows that w'_n has at most rate κ.

(ii): We may assume that $\kappa' < 0$. We can write the recursion as $s_n w'_n = w'_{n-1} + w_n$ for certain numbers $s_n \in \mathbb{C}$ such that $|s_n| = 1$. Let $\sigma_n := \prod_{t=1}^{n} s_t$ for $n \in \mathbb{N}_0$ (in particular $\sigma_0 = 1$) and notice that $|\sigma_n| = 1$. We have $\sigma_n w'_n = \sigma_{n-1} w'_{n-1} + \sigma_{n-1} w_n$. Hence

$$\sigma_n w'_n = w'_0 + \sum_{t=1}^{n} \sigma_{t-1} w_t. \tag{5.2}$$

The limit $w := w'_0 + \sum_{t=1}^{\infty} \sigma_{t-1} w_t \in \mathbb{C}$ exists because $\sum_{t=n_0}^{\infty} |\sigma_{t-1} w_t| \leq \sum_{t=n_0}^{\infty} e^{\kappa' t} < \infty$. Let $w_\infty := |w|$. We conclude for $n \geq n_0$ that

$$\Big| |w'_n| - w_\infty \Big| = \Big| |\sigma_n w'_n| - |w| \Big| \leq |\sigma_n w'_n - w| \leq \left| \sum_{t=n+1}^{\infty} \sigma_{t-1} w_t \right| \tag{5.3}$$

$$\leq \sum_{t=n+1}^{\infty} |w_t| \leq e^{\kappa'(n+1)} \sum_{t=0}^{\infty} e^{\kappa' t} = \frac{e^{\kappa'}}{1 - e^{\kappa'}} \cdot e^{\kappa' n}. \tag{5.4}$$

Thus $|w'_n| - w_\infty$ has at most rate κ'. ∎

Application 5.6 Let $v \in D$ and $x, y \in [0,1]$.

(i) $a_{1,J_n}(v)$ has a.s. upcircle rates $-\chi$, $-\chi$.

(ii) $x_n(x) - x_n(y)$ has a.s. at most rate $-\chi$.

Proof. (i): Let $(x, r, d) := u(v)$. Lemma 3.16.(i) states that

$$ua_{1,J_n}(v) = (x_n(x), r_n(x) \cdot r, d_n(v)). \tag{5.5}$$

Application 4.27 states that $r_n(x)$ and therefore also $r_n(x) \cdot r$ have a.s. exactly rate $-\chi$. Since the function $a_j^\sharp$ is bounded, see Lemma 3.13.(i), it follows that $w_n := a_{j_n}^\sharp(x) \cdot r_{n-1}(x) \cdot r$ has a.s. at most rate $-\chi$. Equation (3.12) states that $|d_n(v)| = |d_{n-1}(v) + w_n|$. Proposition 5.5.(ii) ensures that $|d_n(v)| - d_\infty(v)$ has a.s. at most rate $-\chi$.

(ii): The case $x = y$ is trivial. Let $x \neq y$. The "Cauchy-Schwarz inequality" in Remark 3.15.(ii) states that

$$\frac{|x_n(x) - x_n(y)|}{|x - y|} = r_n(x, y) \leq r_n(x)^{1/2} r_n(x)^{1/2}. \tag{5.6}$$

Hence $x_n(x) - x_n(y)$ has at most rate $\frac{1}{2}(-\chi) + \frac{1}{2}(-\chi) = -\chi$. ∎

5.3 Sequences of upcircle coordinates

Only in this subsection, we ignore Definition 3.14 and instead denote (non-random) upcircle coordinates by (x_n, r_n, d_n).

Remark 5.7 If a sequence v_n in $\mathbb{R} \times \mathbb{H}$ has upcircle rates κ, 0, then in particular, $D(v_n)$ has at most rate 0.

Lemma 5.8 Let $\kappa \in \mathbb{R}$ and $v_n = (x_n, z_n)$ be a sequence in $\mathbb{R} \times \mathbb{H}$ with upcircle rates κ, 0.

(i) The sequences $z_n - x_n$ and $\Im z_n$ have exactly rate κ.

(ii) The sequence $\Re z_n - x_n$ has at most rate κ.

Proof. We know by Proposition 2.18 that $z_n = x_n + r_n \cdot \gamma(d_n)$ and that the unit upcircle function γ is bounded. Thus the sequences $\gamma(d_n)$, $\Im\gamma(d_n)$, $\Re\gamma(d_n)$ are bounded and therefore have at most rate 0. Hence the following sequences have at most rate κ:

$$z_n - x_n = r_n \cdot \gamma(d_n), \quad \Im z_n = r_n \cdot \Im\gamma(d_n), \quad \Re z_n - x_n = r_n \cdot \Re\gamma(d_n) \tag{5.7}$$

It remains to show that $\gamma(d_n)$ and $\Im\gamma(d_n)$ have at least rate 0 because this implies that $z_n - x_n$ and $\Im z_n$ have at least rate κ. Since d_n has at most rate 0 it follows that $d_n - i = 2/\gamma(d_n)$ and $d_n^2 + 1 = 2/\Im\gamma(d_n)$ have at most rate 0, too. Hence $\gamma(d_n)$ and $\Im\gamma(d_n)$ have at least rate 0. ∎

Lemma 5.9 Let $v_n = (x_n, z_n)$ and $v'_n = (x'_n, z_n)$ be sequences in $\mathbb{R} \times \mathbb{H}$ with the same second components. Suppose that $x_n - x'_n$ has at most rate κ and that v_n has upcircle rates κ, 0. Then v'_n has upcircle rates κ, 0, too.

Proof. Let $c_n := d_n^2 + 1$ and $y_n := (x_n - x'_n)/(2r_n)$. Lemma 2.20 states that

$$q_n := \frac{r'_n}{r_n} = c_n \cdot y_n^2 + 2d_n \cdot y_n + 1, \quad d'_n - d_n = c_n \cdot y_n. \tag{5.8}$$

Since d_n has at most rate 0 and $x_n - x'_n$ has at most rate κ and r_n has exactly rate κ, it follows that c_n and y_n have at most rate 0. This results in the following two observations:

- $d'_n - d_n$ has at most rate 0. Since d_n has at most rate 0, it follows that d'_n has at most rate 0.
- q_n has at most rate 0 because all three summands have at most rate 0.

We also have

$$q_n = c_n \cdot \left(y_n + \frac{d_n}{c_n}\right)^2 + \frac{1}{c_n} \geq \frac{1}{c_n}. \tag{5.9}$$

Since c_n has at most rate 0, its reciprocal and therefore also q_n have at least rate 0. So q_n has exactly rate 0. Thus $r'_n = r_n \cdot q_n$ has exactly rate κ. ∎

5.4 Comparison of both generalized shape functions

We remind ourselves of Corollary 3.9.(ii) which states for any component indices $c, c' \in \{1, 2\}$ that

$$S^{c'} S^{c} = S^{c'}. \tag{5.10}$$

The case $c = 2$ of the following proposition is illustrated in Figure 5. We obtain an illustration of the case $c = 1$ by interchanging the labels x, r, z with x', r', z'.

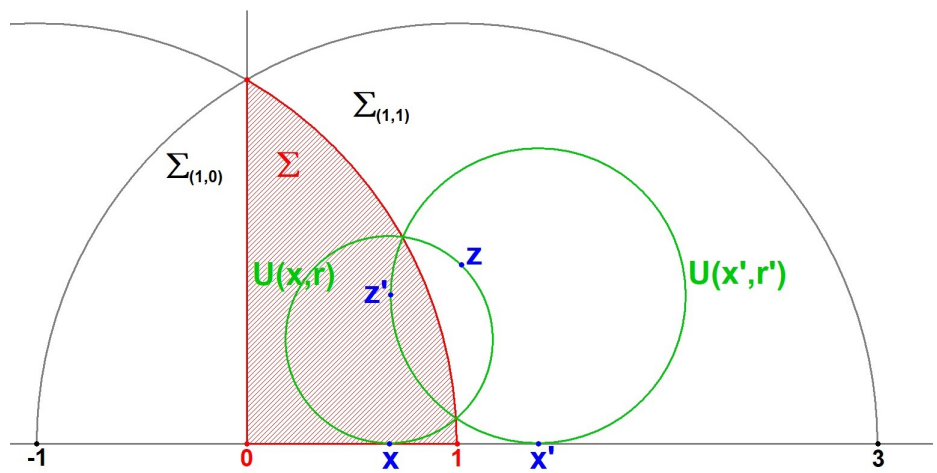

Figure 5: Upcircles $U(x, r)$ and $U(x', r') = S_{(1,1)}(U(x, r))$ with elements z and $z' = S_{(1,1)}(z)$.

Proposition 5.10 Let $v = (x, z) \in D$ and $v' := S^c(v)$. If $c = 2$ we assume $x \in \Sigma$, and if $c = 1$ we assume $z \in \Sigma$. We define the upcircle coordinates $(x, r, d) := u(v)$ and $(x', r', d') := u(v')$. Suppose r is small enough such that $U(x, r)$ intersects at most two shape sets. Then the following bounds hold:

$$|x - x'| \leq 4r, \quad \tfrac{1}{4} r \leq r' \leq 4r, \quad \Big| |d| - |d'| \Big| \leq 2r. \tag{5.11}$$

Proof. Postponed to Proposition 9.10 (on page 74). ∎

Corollary 5.11 Let $\widetilde{v} \in D$ and $\{c, c'\} \in \{1, 2\}$. We set $v := S^c(\widetilde{v})$ and $v' := S^{c'}(\widetilde{v})$, as well as $(x, r, d) := u(v)$ and $(x', r', d') := u(v')$. Suppose that $U(x, r)$ intersects at most two shape sets. Then the bounds in (5.11) hold.

Proof. By (5.10) we have $v' = S^{c'}(v)$. The "design feature" (3.1) of S^c states that the cth component of v lies in Σ. If $c' = 2$ then $c = 1$ and $x \in \Sigma$. If $c' = 1$ then $c = 2$ and $z \in \Sigma$. So we can apply the above proposition with c replaced by c'. ∎

Remark 5.12 The above corollary means that the (potentially different) upcircle coordinates of $S^1(\widetilde{v})$ and $S^2(\widetilde{v})$ become more and more similar, if r goes to 0. This is the key for Proposition 5.13. Its two parts (i) and (ii) seem rather unrelated at first glance, but we present them together as one proposition in order to avoid redundancy in the proofs. Application 5.14 below is crucial for the proof of the second theorem in Section 5.6.

Proposition 5.13 For both component indices $c = 1, 2$ let $\widetilde{v}_{c,n}$ be a sequence in D and $(x_{c,n}, r_{c,n}, d_{c,n})$ be the upcircle coordinates of $v_{c,n} := S^c(\widetilde{v}_{c,n})$. Suppose $r_{1,n}$ has exactly some rate $\kappa < 0$.

(i) Suppose $\widetilde{v}_{1,n} = \widetilde{v}_{2,n}$. Then $r_{2,n}$ has exactly rate κ. Moreover, if $|d_{1,n}| - d_\infty$ has at most rate κ for some number $d_\infty \in [0, \infty)$, then $|d_{2,n}| - d_\infty$ has at most rate κ, too.

(ii) Suppose there are indices $i_1, i_2, \ldots$ in $\mathcal{J}$ such that $\widetilde{v}_{c,n} = A_{i_n}(v_{c,n-1})$ for both $c = 1, 2$. If $r_{2,n}$ has at most rate κ, then $x_{1,n} - x_{2,n}$ has at most rate κ.

Application 5.14 We fix $\kappa < 0$ and $v \in D$. Let us frame parts (i) and (ii) from the above (non-random) proposition in the following (random) context:

(i) If $S^1 A_{J_n}(v)$ has a.s. upcircle rates κ, κ, then $S^2 A_{J_n}(v)$ has a.s. upcircle rates κ, κ, too.

(ii) Let $V_{c,n} := a_{1,J_n}(v)$ for $c = 1, 2$. If both radii $R(V_{1,n})$ and $R(V_{2,n})$ have a.s. exactly rate κ, then the difference of the first components of $V_{1,n}$ and $V_{2,n}$ has a.s. at most rate κ.

Proof (of the proposition). We first present parts of the proofs of (i) and (ii) simultaneously. For that purpose we set $c = 1$, $c' = 2$ for part (i) and $c = 2$, $c' = 1$ for part (ii). The sequence $r_{c,n}$ converges to 0 because it has at most rate $\kappa < 0$. So there is some $n_0 \in \mathbb{N}$ such that for any $n \geq n_0$ the upcircle $U(x_{c,n}, r_{c,n})$ intersects at most two shape sets, see the illustration of

the shape sets in Figure 3. We may assume w.l.o.g. that $n_0 = 1$ (otherwise we always add n_0 to the sequence index n). Let

$$v_n' := S^{c'}(v_{c,n}) = S^{c'}(\widetilde{v}_{c,n}), \tag{5.12}$$

where we used the definition $v_{c,n} = S^c(\widetilde{v}_{c,n})$ and $S^{c'}S^c = S^{c'}$, see (5.10). Corollary 5.11 yields

$$|x_n' - x_{c,n}| \le 4r_{c,n}, \quad \tfrac{1}{4}r_{c,n} \le r_n' \le 4r_{c,n}, \quad \Big||d_n'| - |d_{c,n}|\Big| \le 2r_{c,n}. \tag{5.13}$$

(i): According to (5.12) we have $v_n' = S^2(\widetilde{v}_{1,n}) = v_{2,n}$. So (5.13) shows that $r_{2,n}$ has exactly rate κ and $|d_{2,n}| - |d_{1,n}|$ has at most rate κ. If $|d_{1,n}| - d_\infty$ has at most rate κ, then $|d_{2,n}| - d_\infty$ has at most rate κ, too.

(ii): Let

$$R_n := (r_{1,n} \cdot r_{2,n})^{1/2}, \quad y_n := \frac{|x_{1,n} - x_{2,n}|}{R_n}. \tag{5.14}$$

The sequence R_n has at most rate $\frac{1}{2}(\kappa + \kappa) = \kappa$, so we only need to show that y_n has at most rate 0. We have

$$v_{1,n} = a_{1,i_n}(v_{1,n-1}), \quad v_n' = a_{1,i_n}(v_{2,n-1}), \quad v_{2,n} = S_j A_{i_n}(v_{2,n-1}) \tag{5.15}$$

for some $j \in \mathcal{J}$ (that depends on n). Lemma 3.13.(ii) describes a_{1,i_n} in upcircle coordinates and Proposition 2.23.(i) describes $S_j A_{i_n}$ in upcircle coordinates. In particular, we know the formulas

$$(x_{1,n}, r_{1,n}) = \Big(a_{i_n}(x_{1,n-1}),\ |(a_{i_n})'(x_{1,n-1})| \cdot r_{1,n-1}\Big), \tag{5.16}$$

$$(x_n', r_n') = \Big(a_{i_n}(x_{2,n-1}),\ |(a_{i_n})'(x_{2,n-1})| \cdot r_{2,n-1}\Big), \tag{5.17}$$

$$r_{2,n} = |(S_j A_{i_n})'(x_{2,n-1})| \cdot r_{2,n-1}. \tag{5.18}$$

The "Cauchy-Schwarz inequality" in Proposition 3.3.(ii) yields

$$|x_{1,n} - x_n'|^2 = |a_{i_n}(x_{1,n-1}) - a_{i_n}(x_{2,n-1})|^2 \le \frac{r_{1,n}}{r_{1,n-1}} \cdot \frac{r_n'}{r_{2,n-1}} \cdot |x_{1,n-1} - x_{2,n-1}|^2. \tag{5.19}$$

Part (iv) from the same Proposition 3.3 states that $|(a_{i_n})'| \le |(S_j A_{i_n})'|$, thus $r_n' \le r_{2,n}$. Together with inequality (5.19) it follows that

$$\frac{|x_{1,n} - x_n'|}{R_n} \le \frac{|x_{1,n-1} - x_{2,n-1}|}{R_{n-1}} = y_{n-1}. \tag{5.20}$$

By (5.13) we have $|x'_n - x_{2,n}| \leq 3r_{2,n}$ and therefore

$$\frac{|x'_n - x_{2,n}|}{R_n} \leq 3\left(\frac{r_{2,n}}{r_{1,n}}\right)^{1/2} =: q_n. \tag{5.21}$$

Combining both inequalities with the triangle inequality yields $y_n \leq y_{n-1} + q_n$ and therefore

$$y_n \leq y_0 + \sum_{t=1}^{n} q_t. \tag{5.22}$$

Since $r_{2,n}$ has at most rate κ and $r_{1,n}^{-1}$ has at most rate $-\kappa$ we see that q_n has at most rate 0. By Proposition 5.5.(i) it follows that $\sum_{t=1}^{n} q_t$ has at most rate 0. Therefore y_n has at most rate 0. ∎

5.5 Equality of distributions

The following proposition shows for fixed $v \in D$ that the Markov chain $a_{c,J_n}(v) = S^c A_{j_n} \dots S^c A_{j_1}(v)$, which applies S^c in every step, is essentially the same as the Markov chain $S^c A_{J_n}(v) = S^c A_{j_n} \dots A_{j_1}(v)$, which applies S^c only at the very end.

Proposition 5.15 Let $v \in D$. The two stochastic processes $a_{c,J_n}(v)$ and $S^c A_{J_n}(v)$ have the same distribution in the Borel space of sequences in D.

Proof. Postponed to Proposition 8.7 (on page 67). Let us only give a heuristic argument here: The key is Corollary 8.4.(i) which states that the multiset of the six evaluations $S_j(v)$ remains the same if we replace v by $w := S^c(v)$. With the formula $A_j = \alpha S_j$ we conclude that

$$\{A_j(v) : j \in \mathcal{J}\} = \{A_j(w) : j \in \mathcal{J}\}. \tag{5.23}$$

So we can replace v by $w = S^c(v)$ without changing the next "generation" $\{A_j(v) : j \in \mathcal{J}\}$. This makes it plausible that "smuggling in" S^c between any two functions in the random composition $A_{J_n}(v) = A_{j_n} \dots A_{j_1}(v)$ results in a random composition $A_{j_n} a_{c,J_{n-1}}(v) = A_{j_n} S^c A_{j_{n-1}} \dots A_{j_2} S^c A_{j_1}(v)$ with the same distribution as $A_{J_n}(v)$. Thus $S^c A_{J_n}(v)$ has the same distribution as $a_{c,J_n}(v)$. ∎

Corollary 5.16 Let $v \in D$ and $\kappa, \kappa' \in \mathbb{R}$. Then $a_{c,J_n}(v)$ has a.s. upcircle rates κ, κ', if and only if $S^c A_{J_n}(v)$ has a.s. upcircle rates κ, κ'.

5.6 Proof of the second theorem

Let us restate and prove Theorem 1.5 (on page 6).

Theorem 5.17 Let R_n be the radius $R(X_n, Z_n)$. The following three sequences decay exactly with rate χ:

$$R_n, \quad Z_n - X_n, \quad \Im Z_n = \mathrm{flat}(\Delta^{(n)}).$$

The sequence $\Re Z_n - X_n$ decays at least with rate χ.

Proof. Let $\kappa := -\chi < 0$. We prove an equivalent, but more convenient reformulation of this theorem, where the formulations "decay exactly with rate χ" and "decay at least with rate χ" are replaced by "have exactly rate κ" and "decay at most with rate κ".

In Section 1 we fixed points $X_0 \in [0,1]$, $Z_0 \in \Sigma \setminus [0,1]$ and introduced the Markov chains $X_n = a_{j_n}(X_{n-1}) \in [0,1]$ and $Z_n = a_{j_n}(Z_{n-1}) \in \Sigma \setminus [0,1]$. Hence $X_n = a_{J_n}(X_0)$ and $Z_n = a_{J_n}(Z_0)$. Since E is countable, see Lemma 3.7.(i), there exists $X_0' \in [0,1] \setminus E$. Let $v := (X_0', Z_0) \in (\mathbb{R} \setminus E) \times \mathbb{H} = D$.

We study the following five two-dimensional stochastic processes, where the arrows indicate the chronological order of our investigation:

$$V_{1,n} := a_{1,J_n}(v) \;\rightarrow\; V_{-1,n} := S^1 A_{J_n}(v) \tag{5.24}$$

$$\downarrow$$

$$V_{0,n} := (X_n, Z_n) \;\leftarrow\; V_{2,n} := a_{2,J_n}(v) \;\leftarrow\; V_{-2,n} := S^2 A_{J_n}(v) \tag{5.25}$$

We first show that $V_{t,n}$ has a.s. upcirle rates κ, κ for $t = 1, -1, -2, 2$ and afterwards we show that $V_{0,n}$ has a.s. upcircle rates κ, 0. Then Lemma 5.8 finishes the proof. The numbers in the brackets of the following list indicate the current value of t.

(1) Application 5.6.(i) states that $V_{1,n}$ has a.s. upcircle rates κ, κ.

(-1) Corollary 5.16 states that $V_{-1,n}$ has a.s. upcircle rates κ, κ.

(-2) Application 5.14.(i) states that $V_{-2,n}$ has a.s. upcircle rates κ, κ.

(2) Corollary 5.16 states that $V_{2,n}$ has a.s. upcircle rates κ, κ.

(0) Application 5.14.(ii) states that $X_{2,n} - X_{1,n}$ has at most rate κ. According to the "design feature" (3.5) of a_{1,J_n} the first component $X_{1,n}$ of $V_{1,n}$ equals $a_{J_n}(X_0') = x_n(X_0')$. By Application 5.6.(ii) the sequence $X_{1,n} - X_n = x_n(X_0') - x_n(X_0)$ has at most rate κ. By the "design feature" (3.5) of a_{2,J_n} the second component of $V_{2,n}$ equals $a_{J_n}(Z_0) = Z_n$ which is also the second component of $V_{0,n}$. Moreover, $V_{2,n}$ has a.s. upcircle rates κ, κ and therefore in particular a.s. upcircle rates κ, 0. Thus Lemma 5.9 in the case $v_n = V_{2,n}$ and $v_n' = V_{0,n}$ states that $V_{0,n}$ has a.s. upcircle rates κ, 0. ■

Remark 5.18 Let us comment on the above proof:

(i) A comparison with Section 1.4 shows that all five stochastic processes $V_{t,n}$ are barycentric processes. For $t = 1, 2$ the definitions of $V_{t,n}$ and $R_{t,n}$ are identical with Definitions (1.17) and (1.20).

(ii) It would have sufficed to show that $V_{t,n}$ has a.s. upcircle rates κ, 0 for $t = 1, -1, -2, 2$. However, we find the following byproduct of our proof interesting for its own sake: There is a barycentric process $V_{2,n}$ with second component Z_n such that $V_{2,n}$ has a.s. upcircle rates κ, κ. In particular, the angles $D(V_{2,n})$ converge a.s. (and asymptotically exponentially fast) to some random variable D_∞ taking values in $[0, \infty)$.

6 Triangles

This section discusses the objects from Sections 2.1, 2.3, 2.4. For the reader's convenience, we depict again the figures from Sections 1.1, 2.1, 2.4. We always assume that $z \in \overline{\mathbb{H}}_0$.

6.1 Geometric setting

In Section 1.1 we defined for any child $\Delta := \Delta_z^{(l,k)}$ of z the outer vertex of Δ as the vertex of Δ that is also a vertex of Δ_z, and we defined the outer edge of Δ as the edge that is one half of an edge (namely the k-edge) of Δ_z, see Figure 6.

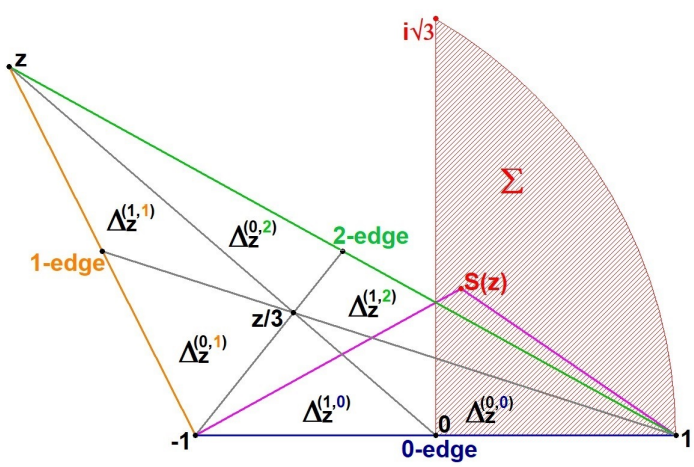

Figure 6: (Copy of Figure 1). A triangle Δ_z, its shape $S(z) \in \Sigma$ and its children $\Delta_z^{(l,k)}$. Their second index k is depicted in the same colour as the corresponding k-edge of z.

Let us now restate Definition 2.1 (on page 12).

Definition 6.1 Let $\Delta := \Delta_z^j$ be some child of Δ_z. If the outer edge of Δ has length 0, then we set $A_j(z) := \infty$. Otherwise, there is a unique similarity transformation that maps the outer edge of Δ to the line segment between -1 and 1, the outer vertex of Δ to 1 and the inner vertex of Δ, namely the barycentre $\frac{1}{3}z$ of Δ_z, to some new point $z' \in \mathbb{H}_0$. We set

$A_j(z) := z'$. This defines a function $A_j : \mathbb{H}_0 \to \overline{\mathbb{H}}_0$.

Figure 7 compares the above definition to the child functions.

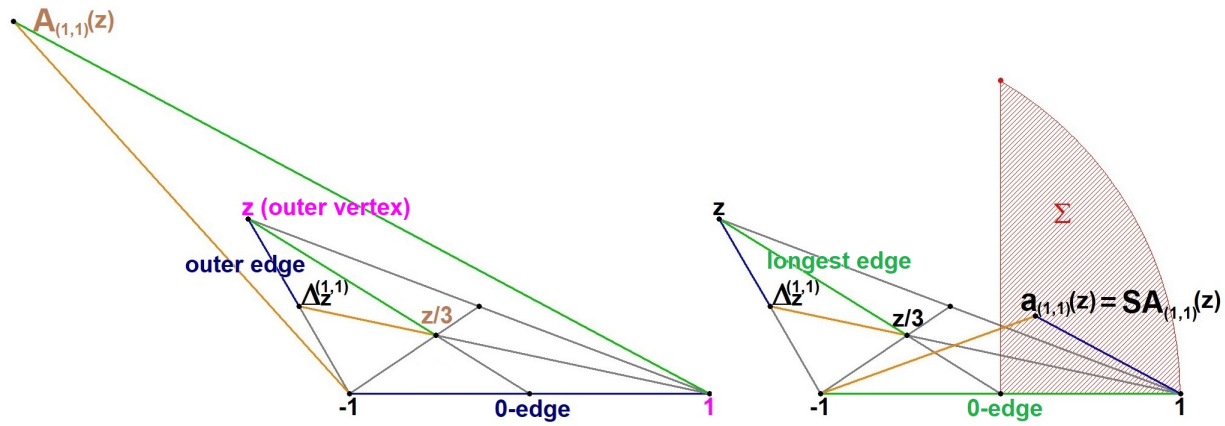

Figure 7: (Copy of Figure 2). Comparison of $A_{(1,1,)}(z)$ and $a_{(1,1,)}(z)$. Corresponding edges in similar triangles have the same colour.

In Section 1.1 we defined points $z, z' \in \mathbb{H}_0$ to be similar, if their associated triangles Δ_z, $\Delta_{z'}$ are similar, and we introduced the shape function $S : \mathbb{H}_0 \to \Sigma$ that maps any point z to the unique point in Σ which is similar to z. Now we extend the notions of similarity and associated triangles to all points in $\overline{\mathbb{H}}_0$. Moreover, using the same symbol, we extend the shape function to a function $S : \overline{\mathbb{H}}_0 \to \Sigma$ by restating Definition 2.2 in part (iii) of the following definition.

Definition 6.2 We define:

(i) $\Delta_\infty := \Delta_1$,

(ii) ∞ and $z \in \overline{\mathbb{H}}_0$ are *similar*, if $z \in \{-1, 1, \infty\}$,

(iii) $S(\infty) := 1 \in \Sigma$.

Remark 6.3 Let us comment on the above extended definition:

(i) For any $z, z' \in \overline{\mathbb{H}}_0$ it is still true that z and z' are similar, if and only if Δ_z and $\Delta_{z'}$ are similar.

(ii) For any $z \in \overline{\mathbb{H}}_0$ it is still true that $S(z)$ is the unique point in Σ which is similar to z.

Let us restate Lemma 2.3 (on page 13) which we have proven already.

Lemma 6.4 Any $z \in \mathbb{H}_0$ satisfies

$$a_j(z) = SA_j(z). \tag{6.1}$$

6.2 Explicit formula for the barycentric subdivision

Let us restate Definition 2.10 (on page 16).

Definition 6.5 We define:

(i) the *rotation function* $h \in \mathcal{F}$ by

$$h(z) := \frac{z-3}{z+1} = 1 - \frac{4}{z+1},$$

(ii) the six functions $S_j \in \mathcal{G}$ by

$$S_{(l,k)} := m^l h^k, \quad (l,k) \in \mathcal{J},$$

(iii) the *rescaling function* $\alpha \in \mathcal{F}$ by

$$\alpha(z) := \tfrac{2}{3}z - 1.$$

Remark 6.6 Let $z \in \overline{\mathbb{H}}_0$. The first and the second representation of $h(z)$ in Definition 6.5.(i) show the first and the last equation in the following line:

$$h^2(z) = \frac{-z-3}{z+1} = -1 + \frac{4}{z-1} = h^{-1}(z). \tag{6.2}$$

Hence $h^3 = \mathrm{id}$.

Lemma 6.7 The rotation function h has the following geometric meaning:

(i) Let $z \in \mathbb{H}_0 \setminus \{-1\}$. Then $h(z) \neq \infty$ and there is an orientation preserving similarity transformation that maps the k-edge of z to the $(k-1)$-edge of $h(z)$.

(ii) Let $z \in \mathbb{H}_0 \setminus \{1\}$. Then $h^2(z) \neq \infty$ and there is an orientation preserving similarity transformation that maps the k-edge of z to the $(k+1)$-edge of $h^2(z)$.

Proof. Let us define the complex affine linear transformations

$$\varphi_{1,z} : w \mapsto \frac{-2w + z - 1}{z+1}, \quad \varphi_{2,z} : w \mapsto \frac{2w - z - 1}{z-1}.$$

We see that the orientation preserving similarity transformation $\varphi_{1,z}$ maps the vertices of (the associated triangle of) z to the vertices of $h(z)$ in the following way:

$$z \mapsto -1, \quad -1 \mapsto 1, \quad 1 \mapsto \frac{z-3}{z+1} = h(z). \tag{6.3}$$

Hence $\varphi_{1,z}$ maps the k-edge of z to the $(k-1)$-edge of $h(z)$. The analogous statement holds for $\varphi_{2,z}$. ∎

Lemma 6.8 The rotation function h and the mirror function m have the following properties:

$$h^3 = \mathrm{id}, \quad m^2 = \mathrm{id}, \quad h^k m = m h^{-k}. \tag{6.4}$$

This means that $(\{S_j : j \in \mathcal{J}\}, \circ)$ is a Dihedral group which we denote by $\mathcal{S}$. The map $j \mapsto S_j$ is a group isomorphism between $(\mathcal{J}, \cdot)$ and $(\mathcal{S}, \circ)$, where the operation $\cdot$ is defined by

$$(l', k') \cdot (l, k) := (l' + l, (-1)^l k' + k). \tag{6.5}$$

The identity element of $\mathcal{J}$ is $(0,0)$ and the formula for the inverse is $(l,k)^{-1} = (l, (-1)^{l+1}k)$.

Proof. By Remark 6.6 it is clear that $m, h, h^2 \neq \mathrm{id}$ and that (6.4) holds. Hence the six (pairwise different) functions S_j form the Dihedral group $\mathcal{S}$. Since $\phi : \mathcal{J} \to D_6 : j \mapsto S_j$ is bijective, it suffices to show that $\phi(j)\phi(j') = \phi(j \cdot j')$ in order to conclude that $\mathcal{J}$ is a group and ϕ a group isomorphism. Let $j = (l,k)$ and $j' = (l', k')$. By (iii) we have $h^{k'} m^l = m^l h^{k''}$ for $k'' := (-1)^l k'$. Hence $\phi(j')\phi(j) = \phi(l' + l, k'' + k) = \phi(j \cdot j')$. The statements about the identity element and the inverse are clear. ∎

Proposition 6.9 For any $z \in \mathbb{H}_0$ the explicit formula for the barycentric subdivision is

$$A_j(z) = \alpha S_j(z).$$

Proof. For $z \in \{-1, 1\}$ we verify the above formula by hand. Now we assume $z \notin \{-1, 1, \infty\}$. Due to Lemma 6.7 we know the following: For $k' \in \{0, 1, 2\}$ there is a similarity transformation $\varphi_{k',z}$ (in the case $k' = 0$ we set $\varphi_{0,z} := \mathrm{id}$) which maps the k-edge of z to the $(k-k')$-edge of $h^{k'}(z)$. It follows with elementary geometry that

$$\varphi_{k,z} \quad \text{maps} \quad \Delta := \Delta_z^{(l,k)} \quad \text{to} \quad \Delta' := \Delta_{h^k(z)}^{(l,0)}$$

and the outer edge of Δ to the outer edge of Δ' and the outer vertex of Δ to the outer vertex of Δ'. The example case $(l, k) = (1, 2)$ is illustrated in Figure 8.

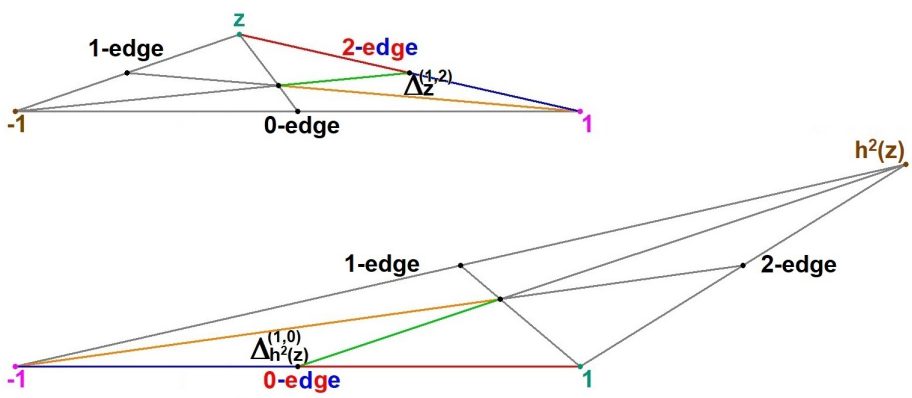

Figure 8: Triangles Δ_z and $\Delta_{h^2(z)}$ with their similar children $\Delta_z^{(1,2)}$ and $\Delta_{h^2(z)}^{(1,0)}$. Corresponding edges have the same colour. The outer edges of both children are blue and the outer vertices are magenta.

For $z' := S_{(l,k)}(z)$ we notice that

$$m^l \quad \text{maps} \quad \Delta' \quad \text{to} \quad \Delta'' := \Delta_{z'}^{(0,0)}$$

and the outer edge to the outer edge and the outer vertex to the outer vertex. The similarity transformation $\varphi_\alpha : w \mapsto 2w - 1$ maps 0, 1, $\frac{1}{3}z'$ to -1, 1, $\alpha(z')$ (in this order). Thus

$$\varphi_\alpha \quad \text{maps} \quad \Delta'' \quad \text{to} \quad \Delta_{\alpha(z')}$$

and the outer edge to the 0-edge and the outer vertex to 1. Combining the three steps shows that there is a similarity transformation that maps Δ to $\Delta_{\alpha(z')}$ and the outer edge to the 0-edge and the outer vertex to 1. It follows by definition of A_j that $A_{(l,k)}(z) = \alpha(z') = \alpha S_{(l,k)}(z)$. ∎

As in Definition 2.13, we can extend A_j and a_j to the following functions:

$$A_j = \alpha S_j \in \mathcal{G}, \quad a_j = SA_j : \overline{\mathbb{H}}_0 \to \Sigma. \tag{6.6}$$

6.3 Shape sets

Let us restate Definition 2.14 (on page 17).

Definition 6.10 We define the *shape sets*

$$\Sigma_{0,k} := \{z : \text{ the } k\text{-edge is a longest edge and the } (k-1)\text{-edge is a shortest edge of } z\},$$

$$\Sigma_{1,k} := \{z : \text{ the } k\text{-edge is a longest edge and the } (k+1)\text{-edge is a shortest edge of } z\}.$$

We include ∞ in $\Sigma_{(0,1)}$ and $\Sigma_{(1,2)}$, but not in the other shape sets.

Figure 9 illustrates the above definition.

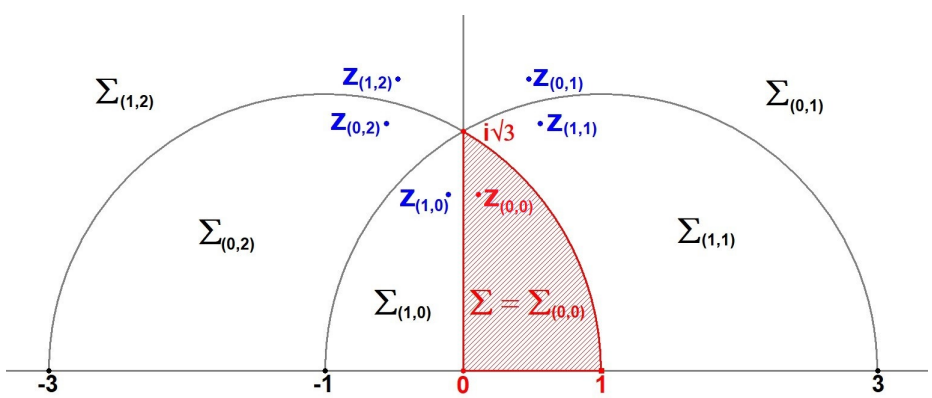

Figure 9: (Copy of Figure 3). The shape sets Σ_j and elements $z_j \in \Sigma_j$ with $z_{(0,0)} = S(z_j) = S_j(z_j)$.

In the following lemma we assume that $z \in \overline{\mathbb{H}}_0$.

Lemma 6.11 The functions S_j and the shape sets Σ_j have the following properties:

(i) The set $\{S_j(z) : j \in \mathcal{J}\}$ consists of those points that are similar to z.

(ii) The invariance $SS_j = S$ holds.

(iii) $\Sigma_j = \{z : S_j(z) \in \Sigma\} = \{z : S(z) = S_j(z)\}$.

(iv) $S_j(\Sigma_{j'}) = \Sigma_{j''}$ with $j'' := j' \cdot j^{-1}$.

(v) $S(x) = \min_j |S_j(x)|$ for any $x \in \overline{\mathbb{R}}$.

Proof. (i): Let T denote the above set $\{S_j(z) : j \in \mathcal{J}\}$. If $z \in \{1, -1, \infty\}$, then $T = \{-1, 1, \infty\}$ consists of those points that are similar to z, see Definition 6.2. Suppose $z \notin \{-1, 1, \infty\}$. We see by Lemma 6.7 that $H := \{z, h(z), h^2(z)\}$ consists of those points that are similar to z with the same orientation as z. Since $T = H \cup m(H)$ we can conclude (i).

(ii): For any $z \in \overline{\mathbb{H}}_0$ we know by Remark 6.3 that $S(z)$ is the unique point in Σ that is similar to z. Together with (i) it follows (ii).

(iii): In the second equation the inclusion "$\supseteq$" is trivial and "$\subseteq$" is a consequence of (ii) and the fact that $S : \overline{\mathbb{H}}_0 \to \Sigma$ is constant on Σ. For the first equation we show that

$$z \in \Sigma_{(l,k)} \quad \Leftrightarrow \quad h^{k'}(z) \in \Sigma_{(l,k-k')} \tag{6.7}$$

because then, the special case $k' = k$ and the fact $m^l(\Sigma_{(l,0)}) = \Sigma_{(0,0)}$ show that

$$z \in \Sigma_{(l,k)} \quad \Leftrightarrow \quad S_{(l,k)}(z) = m^l h^k(z) \in \Sigma_{(0,0)} = \Sigma. \tag{6.8}$$

For the proof of (6.7) we may assume w.l.o.g. that $k' = 1$ ($k' = 0$ is trivial and for $k' = 2$ we apply the result for $k' = 1$ twice). We also assume $l = 0$ (the proof for $l = 1$ is similar). So we need to show that

$$z \in \Sigma_{(0,k)} \quad \Leftrightarrow \quad h(z) \in \Sigma_{(0,k-1)}. \tag{6.9}$$

The shape set $\Sigma_{(0,k)}$ contains $z = 1, \infty, -1$ if and only if $k = 0, 1, 2$ (in this order), see Figure 9. For $z = 1, \infty, -1$ we have $h(z) = -1, 1, \infty$ and the shape set $\Sigma_{(0,k-1)}$ contains $h(z) = -1, 1, \infty$ if and only if $k - 1 = 2, 0, 1$. This shows the equivalence (6.9) in the case $z \in \{-1, 1, \infty\}$. Now suppose that $z \notin \{-1, 1, \infty\}$. By definition, $\Sigma_{(0,k)}$ contains z if and only if the k-edge is a longest and the $(k-1)$-edge is a shortest edge of z, and $\Sigma_{(0,k-1)}$ contains $h(z)$ if and only if the $(k-1)$-edge is a longest and the $(k-2)$-edge is a shortest edge of $h(z)$. According to Lemma 6.7 there exists a similarity transformation that maps the k-edge of z to the $(k-1)$-edge of $h(z)$. This shows the equivalence (6.9).

(iv): By Lemma 6.8 we know that $S_{j'} = S_{j''\cdot j} = S_{j''}S_j$. Applying (iii), (ii) and again (iii) yields

$$z \in \Sigma_{j'} \quad \Leftrightarrow \quad SS_j(z) = S_{j'}(z) = S_{j''}S_j(z) \quad \Leftrightarrow \quad S_j(z) \in \Sigma_{j''}.$$

Since S_j is bijective, it follows $S_j(\Sigma_{j'}) = \Sigma_{j''}$.

(v): As a consequence of Lemma 6.8 and the bijectivity of the function $\mathcal{J} \to \mathcal{J} : j \mapsto j \cdot j'$ we see that the set of the six compositions $S_j S_{j'}$ equals the set of the six functions $S_{j'}$. So we may assume w.l.o.g. that $x \in \Sigma$ (or equivalently $S(x) = x$), because $S(x) = S_{j'}(x)$ for some index $j' \in \mathcal{J}$.

For $j \in \{(0,0), (1,0)\} =: \mathcal{J}^*$ we have $x = |S_j(x)|$. Suppose $j \in \mathcal{J} \setminus \mathcal{J}^*$. It follows that $j^{-1} \in \mathcal{J} \setminus \mathcal{J}^*$. Thus $S_{j^{-1}} \cap \mathbb{R}$ is one of the intervals $[-\infty, -3]$, $[-3, -1]$, $[1, 3]$, $[3, \infty]$, see Figure 9. Since $x \in \Sigma = \Sigma_{(0,0)}$ and due to part (iv) we have $S_j(x) \in \Sigma_{j^{-1}}$ and therefore $x \leq 1 \leq |S_j(x)|$. ∎

7 Up-to-sign derivatives

7.1 Basic properties

In this subsection we always assume that x, y are real variables with $x \neq y$. We remind ourselves of Definition 3.2 which defines up-to-sign derivatives and the difference quotient $|b'(x,y)| = |b(x) - b(y)|/|x - y|$.

Definition 7.1 Let $D_b \subseteq \mathbb{R}$ and $b : D_b \to \mathbb{R}$ be an up-to-sign differentiable function. We set $Q_b(x,x) := 0$ and

$$Q_b(x,y) := \frac{|b'(x,y)|^2}{|b'(x)| \cdot |b'(y)|}.$$

Remark 7.2 Let $g \in \mathcal{G}$ and $b : D_b \to \mathbb{R}$ be the restriction of g to $D_b := \mathbb{R} \setminus \{g^{-1}(\infty)\}$. Lemma 2.9.(ii) states thate the square of the difference quotient $(g(x) - g(y))/(x - y)$ equals $|g'(x)| \cdot |g'(y)|$. This means that $Q_b(x,y)$ is constantly 1.

Lemma 7.3 Let $b : D_b \to D_a$ and $a : D_a \to \mathbb{R}$ be functions with $D_a, D_b \subseteq \mathbb{R}$. Suppose that a and b are up-to-sign differentiable. Their composition ab satisfies the following "chain rules" for any $x, y \in D_b$:

(i) ab is up-to-sign differentiable and $|(ab)'(x)| = |a'b(x)| \cdot |b'(x)|$.

(ii) $|(ab)'(x,y)| = |a'(b(x), b(y))| \cdot |b'(x,y)|$.

(iii) $Q_{ab}(x,y) = Q_a(b(x), b(y)) \cdot Q_b(x,y)$.

Proof. (i): There are only finitely many $x \in D_b$ for which b is not differentiable in x or a not differentiable in $b(x)$. For any other $x \in D_b$ we have $|(ab)'(x)| = |a'b(x)| \cdot |b'(x)|$ according to the ordinary chain rule. The function $|a'b| \cdot |b'|$ is continuous and is therefore the up-to-sign derivative of ab.

(ii) is clear for $b(x) \neq b(y)$, as well as for $b(x) = b(y)$ (in which case both sides of (ii) vanish).

(iii) follows from (i) and (ii). ∎

7.2 Up-to-sign derivative of the shape function

In this subsection we always assume that x, y are real variables with $x \neq y$.

Definition 7.4 We define the *fixpoint polynomial* $\widehat{h} : \overline{\mathbb{R}} \to [3, \infty]$ by

$$\widehat{h}(x) := x^2 + 3.$$

Remark 7.5 Let us explain the reason for the name "fixpoint polynomial" (even though we do not need this in the following): The formula $h(z) = (z-3)/(z+1)$ shows that the fixpoints of h are exactly the two roots of the complex polynomial $z^2 + 3$. These roots are $\pm i\sqrt{3}$. Notice that the associated triangle of the point $i\sqrt{3}$ is equilateral, so it is also geometrically clear that this point is a fixpoint of h.

Lemma 7.6 For any j and $x \in \overline{\mathbb{R}}$ we can express $|S_j'(x)|$ in terms of $\widehat{h}$ via

$$|S_j'(x)| \cdot \widehat{h}(x) = \widehat{h} S_j(x).$$

Proof. Let $(l, k) := j$. We have that $S_j = m^l h^k$ and $|S_j'| = (h^k)'$. Since $\widehat{h}$ is even, the right-hand side in the above formula is $\widehat{h} h^k(x)$. So we need to show that

$$(h^k)'(x) \cdot \widehat{h}(x) = \widehat{h} h^k(x). \tag{7.1}$$

The case $k = 0$ is trivial. For $k \in \{1, 2\}$ we easily verify the formula by plugging in the formulas $h(x) = (x-3)/(x+1)$ and $h'(x) = 4/(x+1)^2$, or $h^2(x) = -(x+3)/(x-1)$ and $(h^2)'(x) = 4/(x-1)^2$, see Remark 6.6. ∎

Definition 7.7 We define the *non-isosceles set* Σ^* as the set of all points $x \in \mathbb{R}$ that lie in exactly one of the shape sets Σ_j.

Notice that $\Sigma^* = \mathbb{R} \setminus \{-3, -1, 0, 1, 3\}$ contains exactly those points $x \in \mathbb{R}$ that are not isosceles.

Proposition 7.8 Let us view the shape function as a function $S : \mathbb{R} \to [0, 1]$. Then S is Lipschitz-continuous and up-to-sign differentiable. Any $x, y \in \mathbb{R}$ satisfy:

(i) $|S'(x)| = \widehat{h}S(x)/\widehat{h}(x) = |S_j'(x)|$, if $x \in \Sigma_j$,

(ii) $|S'(x)| = \min_j |S_j'(x)| \in (0, 1]$,

(iii) $Q_S(x, y) \leq 1$.

Proof. We set

$$\beta(x) := \frac{\widehat{h}S(x)}{\widehat{h}(x)} \in (0, \infty), \quad Q(x, y) := \frac{|S'(x, y)|^2}{\beta(x)\beta(y)}. \tag{7.2}$$

The proof is structured in the following steps:

- Step 1 explains why it suffices to show that $Q(x, y) \leq 1$ for any x, y.
- Step 2 shows the invariance $Q(x, y) = Q(S_j(x), S_j(y))$.
- Step 3 justifies why we may assume w.l.o.g. that $y \in \{0, 1\}$ and $x \notin \Sigma \cap \mathbb{R} = [0, 1]$.
- Step 4 considers the cases $y = 0$ and $y = 1$.

Step 1: Due to Lemma 7.6 the function β equals $|S_j'|$ on $\Sigma_j \cap \mathbb{R}$. Any $x \in \Sigma^* = \mathbb{R} \setminus \{-3, -1, 0, 1, 3\}$ lies in the interior of one shape set Σ_j, so β is differentiable in each such x with

$$\beta(x) = |S_j'(x)|. \tag{7.3}$$

Due to Lemma 7.6 and $S(x) = \min_j |S_j(x)|$, see Lemma 6.11.(v), we know that $\beta(x) = \min_j |S_j'(x)|$ is at most 1. So $|S'(x, y)|^2 \leq Q(x, y)$. Suppose we can show that

$$Q(x, y) \leq 1. \tag{7.4}$$

Then S is Lipschitz-continuous. Thus β is continuous and is therefore the up-to-sign derivative of S. Hence $Q_S = Q$.

Step 2: First, we show the following analogue of the chain rule from Lemma 7.3.(i):

$$\beta(x) = \beta S_j(x) \cdot |S_j'(x)|. \tag{7.5}$$

This follows from a combination of Lemma 7.6, and the invariance $SS_j = S$, see Lemma 6.11.(ii).

Using again the invariance $SS_j = S$ we conclude that $Q(x,y) = Q(S_j(x), S_j(y)) \cdot Q_{S_j}(x,y)$. Due to Remark 7.2 we obtain the invariance

$$Q(x,y) = Q(S_j(x), S_j(y)). \tag{7.6}$$

In the following two steps we prove (7.4) by a series of "w.l.o.g."-assumptions.

Step 3: Because of the invariance (7.6) we may assume that $y \in \Sigma$, so $S(y) = y$ and $\beta(y) = 1$. By definition of Q, β and by $\widehat{h}(x) = x^2 + 3$ we obtain the formula

$$Q(x,y) = |S'(x,y)|^2 \cdot \frac{x^2+3}{S(x)^2+3}. \tag{7.7}$$

If x also lies in Σ, then $Q(x,y) = 1$. We assume $x \notin \Sigma$. This implies that the function

$$\varphi : [0,1] \to \mathbb{R} : t \mapsto \frac{S(x)-t}{x-t} = \frac{S(x)-x}{x-t} + 1 \tag{7.8}$$

is monotone. Thus $|\varphi(y)| \leq \max\{|\varphi(0)|, |\varphi(1)|\}$. Since $Q(x,t) = \varphi(t)^2 \beta(x)^{-1}$ for any $t \in [0,1]$ it follows $Q(x,y) \leq \max\{Q(x,0), Q(x,1)\}$. So we may assume that $y \in \{0,1\}$.

Step 4: Case $y = 0$: The invariance $Q(m(x),0) = Q(x,0)$, see (7.6), justifies the assumption $x \geq 0$. Together with $x \notin \Sigma$ it follows that $x > 1 \geq S(x) =: \widetilde{x} \geq 0$. Formula (7.7) yields

$$Q(x,0) = \frac{\widetilde{x}^2}{x^2} \cdot \frac{x^2+3}{\widetilde{x}^2+3} = \frac{1+3x^{-2}}{1+3\widetilde{x}^{-2}} \leq 1. \tag{7.9}$$

Case $y = 1$: The representation $mh(z) + 1 = 4/(z+1)$ shows that the function mh maps $\mathbb{R} \setminus [-3,1] = \{t \in \mathbb{R} : |t+1| > 2\}$ to $(-3,1) = \{t \in \mathbb{R} : |t+1| < 2\}$. Since mh has the fixpoint 1, we have the invariance $Q(mh(x),1) = Q(x,1)$, see (7.6). So we may assume $x \in [-3,1]$. Together with $x \notin \Sigma$ it follows that $-3 \leq x < 0 \leq s(x) =: \widetilde{x} \leq 1$. Formula (7.7) yields

$$Q(x,1) = \frac{(1-\widetilde{x})^2}{(1-x)^2} \cdot \frac{x^2+3}{\widetilde{x}^2+3} = \frac{\varphi(x)}{\varphi(\widetilde{x})}, \quad \varphi(t) := \frac{t^2+3}{(1-t)^2}. \tag{7.10}$$

In the special case $\widetilde{x} = 1$ we have $Q(x,1) = 0$. Suppose $\widetilde{x} \neq 1$. Due to $\varphi'(t) = 2(t+3)/(1-t)^3$ the function φ is increasing on $[-3,1)$. This shows that $\varphi(x) \leq \varphi(\widetilde{x})$, so $Q(x,1) \leq 1$. ∎

7.3 Up-to-sign derivative of the child functions

In this subsection we always assume that x, y are real variables with $x \neq y$.

Lemma 7.9 Let $g \in \mathcal{G}$. We define $b : \mathbb{R} \to [0,1]$ by $b(x) := Sg(x)$. Then b is Lipschitz-continuous and up-to-sign differentiable. Any $x, y \in D_b$ satisfy:

(i) $|b'(x)| = |(S_j g)'(x)|$, if $g(x) \in \Sigma_j$,

(ii) $|b'(x)| = \min_j |(S_j g)'(x)| \in (0, \infty)$,

(iii) $Q_b(x, y) \leq 1$.

Proof. Let $g = m^l f$ with $f \in \mathcal{F}$. We use the standard notation $f(z) = (c_1 z + c_2)/(c_3 z + c_4)$ with coefficients $c_i \in \mathbb{R}$ such that $c := c_1 c_4 - c_2 c_3 > 0$. Let us define a continuous functions $\alpha, \beta : \mathbb{R} \to (0, \infty)$ by

$$\alpha(x) := \frac{c}{(c_1 x + c_2)^2 + 3(c_3 x + c_4)^2}, \quad \beta(x) := \widehat{h} Sg(x) \cdot \alpha(x). \tag{7.11}$$

Let $D_g := \mathbb{R} \setminus \{g^{-1}(\infty)\}$. For any $x \in D_g$ we have $|g'(x)| = c/(c_3 x + c_4)^2$ and therefore $\alpha(x) = |g'(x)|/\widehat{h} g(x)$. Together with the formula $|S'g(x)| = \widehat{h} Sg(x)/\widehat{h} g(x)$, see Proposition 7.8.(i), it follows that

$$\beta(x) = |S'g(x)| \cdot |g'(x)|, \quad x \in D_g. \tag{7.12}$$

According to the chain rule from Lemma 7.3.(i), the right-hand side of the above formula is the up-to-sign derivative of the restriction $b : D_g \to [0,1]$. Hence β is the up-to-sign derivative $|b'|$ of $b : \mathbb{R} \to [0,1]$.

By using formula (7.12) and the same chain rule from Lemma 7.3.(i) again, we see that parts (i) and (ii) from Proposition 7.8 imply the respective part of this lemma.

For (iii) let $\widetilde{g} : D_g \to \mathbb{R}$ denote the restriction of g. The chain rule from Lemma 7.3.(iii) states that

$$Q_b(x, y) = Q_{S\widetilde{g}}(x, y) = Q_S(g(x), g(y)) \cdot Q_{\widetilde{g}}(x, y), \quad x, y \in D_g. \tag{7.13}$$

According to Proposition 7.8.(iii) and Remark 7.2 we have $Q_S \leq 1$ and $Q_{\widetilde{g}} = 1$. Since $Q_b(x, y)$ is continuous in x and y it follows that $Q_b(x, y) \leq 1$ for any $x, y \in \mathbb{R}$. ∎

Let us restate and prove Proposition 3.3 (on page 21).

Proposition 7.10 We restrict the child functions to functions $a_j : [0,1] \to [0,1]$. Let $I \in \mathcal{J}^n$ be a multiindex with $n \in \mathbb{N}_0$. Then the composition a_I is Lipschitz-continuous and has an up-to-sign derivative $|a_I'| : [0,1] \to (0,\infty)$. Moreover, this up-to-sign derivative satisfies:

(i) the chain rule $|a_I'(x)| = |a_i'(a_J(x))| \cdot |a_J'(x)|$, if we write $I = (J,i)$ with $J \in \mathcal{I}^{n-1}$,

(ii) the "Cauchy-Schwarz inequality" $|a_I'(x,y)| \leq |a_I'(x)|^{1/2}|a_I'(y)|^{1/2}$.

In the "$n = 1$" case $I = j \in \mathcal{J}$ we have:

(iii) $|a_j'(x)| = |S_{j'}A_j(x)|$, if $a_j(x) = S_{j'}A_j(x)$,

(iv) $|a_j'(x)| = \min_{j'} |(S_{j'}A_j)'(x)|$.

Proof. (i) holds by the chain rule Lemma 7.3.(i).

(ii): Lemma 7.9, in the special case $g = A_j \in \mathcal{G}$, states that a_j has an up-to-sign derivative $|a_j'| : [0,1] \to (0,\infty)$ and we have the bound $Q_{a_j} \leq 1$. It follows with the chain rule from Lemma 7.3.(i) that the composition a_I has an up-to-sign derivative $|a_I'| : [0,1] \to (0,\infty)$, and it follows with the chain rule in part (iii) of the same lemma that $Q_{a_I} \leq 1$ which is equivalent to (ii).

(iii) follows from Lemma 7.9.(i) and the equivalence of the statements $a_j(x) = S_{j'}A_j(x)$ and $A_j(x) \in \Sigma_{j'}$.

(iv) follows from Lemma 7.9.(ii). ∎

Remark 7.11 In the above proposition we could as well view the child functions as functions $a_j : \mathbb{R} \to \mathbb{R}$ and the same results hold.

8 Generalized child functions

8.1 Exceptional set

In Definition 7.7 we introduced the non-isosceles set Σ^* consisting of all points $x \in \mathbb{R}$ that lie in exactly one of the shape sets Σ_j and noticed that $\Sigma^* = \mathbb{R} \setminus \{-3, -1, 0, 1, 3\}$ contains exactly those points $x \in \mathbb{R}$ that are not isosceles. We say that z' is a descendant of z if $\Delta_{z'}$ is similar to one of the descendants of Δ_z in the iterated barycentric subdivision. In particular, z is a descendant of itself.

Definition 8.1 We define

(i) the *exceptional set* E as the set of all points $x \in \mathbb{R}$ that have an isosceles descendant,

(ii) $D := (\mathbb{R} \setminus E) \times \mathbb{H}$,

Let us restate Lemma 3.7 (on page 23) and augment it with the following part (ii).

Lemma 8.2 The sets E and D have the following properties:

(i) E is countable.

(ii) If $v \in D$, then $A_j(v) \in D$.

(iii) If $A_j(x) \in E$, then $x \in E$.

Proof. (i): It is geometrically clear that x' is a descendant of $x \in \mathbb{R}$ in generation $n \in \mathbb{N}_0$ if and only if x' is similar to $x'' := A_I(x)$ for a certain multiindex $I \in \mathcal{J}^n$. Note that x' is isosceles if and only if x'' is isosceles, i.e. $x'' \in \mathbb{R} \setminus \Sigma^* = \{-3, -1, 0, 1, 3\}$. The bijectivity of A_I and the following representation of E imply that E is countable:

$$E = \bigcup_{n \in \mathbb{N}_0} \bigcup_{I \in \mathcal{J}^n} (A_I)^{-1} (\mathbb{R} \setminus \Sigma^*). \tag{8.1}$$

For (iii) we only need to notice that if the descendant $A_j(x)$ of x has an isosceles descendant, then x has an isosceles descendant.

(ii): follows from (iii). ∎

8.2 Construction of the generalized shape functions

For the rest of this section we consider a fixed component index $c \in \{1, 2\}$. In Lemma 6.8 (on page 51) we saw that the map $j \mapsto S_j$ is a group isomorphism between the Dihedral groups $(\mathcal{J}, \cdot)$ and $(\{S_j : j \in \mathcal{J}\}, \circ)$, where the operation $\cdot$ is defined by

$$(l', k') \cdot (l, k) := (l' + l, (-1)^l k' + k). \tag{8.2}$$

The identity element of $\mathcal{J}$ is $(0, 0)$ and the formula for the inverse is $(l, k)^{-1} = (l, (-1)^{l+1} k)$.

Proposition 8.3 There is a function $S^c : D \to D$ such that any $v \in D$ satisfies the following two conditions:

(i) $S^c(v) = S_j(v)$ for some j such that $v_c \in \Sigma_j$.

(ii) $S^c S_j(v) = S^c(v)$ for any j.

Corollary 8.4 Any $v \in D$ satisfies the following two conditions:

(i) The multiset of the six evaluations $S_j S^c(v)$ equals the multiset of the six evaluations $S_j(v)$.

(ii) $S^{c'} S^c(v) = S^{c'}(v)$ for any $c, c' \in \{1, 2\}$.

Proof (of the corollary). We have $w := S^c(v) = S_{j'}(v)$ for some index j'.

(i): The map $S_j \mapsto S_j S_{j'}$ is a permutation on the group $\{S_j : j \in \mathcal{J}\}$.

(ii) follows from $S^{c'} S_{j'} = S^{c'}$. ■

Proof (of the proposition). This very technical proof may be skipped at first reading. The set $\mathbb{R} \setminus E$ of all points $x \in \mathbb{R}$ that have only non-isosceles descendants is a subset of the non-isosceles set Σ^* which is a subset of $\mathbb{R} \setminus \{-1, 1\}$. The last set contains all points $x \in \mathbb{R}$ such that all six evaluations $A_j(x)$ are finite.

We set $D^* := \Sigma^* \times \mathbb{H} \supseteq D$. Due to the similarity of x and $S_j(x)$ it is geometrically clear that x lies in $\mathbb{R} \setminus E$ if and only if $S_j(x)$ does. Hence v lies in D if and only if $S_j(v)$ does. As

a consequence, it suffices to show the following in order to prove the proposition: There is a function $S^c : D^* \to D^*$ such that any $v \in D^*$ satisfies (i) and (ii).

In the following we always assume $v \in D^*$. We show that there is a function $J^c : D^* \to \mathcal{J}$ with the following two properties:

(i)$_J$ $J^c(v) = j$ for some j such that $v_c \in \Sigma_j$,

(ii)$_J$ $J^c S_j(v) \cdot j = J^c(v)$ for any j.

We then define

$$S^c(v) := S_{J^c(v)}(v). \tag{8.3}$$

We see that (i)$_J$ trivially implies (i) and (i) implies that S^c actually maps D^* to D^* because of the obvious invariance $S_j(\Sigma^*) = \Sigma^*$. We quickly verify that (ii)$_J$ implies (ii):

$$S^c S_j(v) = S_{J^c S_j(v)} S_j(v) = S_{J^c(v)}(v) = S^c(v). \tag{8.4}$$

The case $c = 1$ is easy: By definition of D^* the first component v_1 lies in exactly one shape set Σ_{j_1} and we set $J^1(v) := j_1$. In that case the first component $S_j(v_1)$ of $S_j(v)$ lies in $\Sigma_{j''}$ with $j'' := j_1 \cdot j^{-1}$. Then $J^c S_j(v) = j''$, thus $J^c S_j(v) \cdot j = j_1 = J^c(v)$.

From now on let $c = 2$. We compute that

$$(l', k') \cdot (l, k)^{-1} = (l' + l, (-1)^l (k' - k)). \tag{8.5}$$

In the following, we will use this formula twice.

Let us write $M_l := \bigcup_k \Sigma_{(l,k)}$ and $H_k := \bigcup_l \Sigma_{(l,k)}$. We always assume that $j = (l, k)$. We show that there are functions $L^2, K^2 : D^* \to \mathcal{J}$ with the following properties:

(i)$_L$ $L^2(v) = l$ for some l such that $v_2 \in M_l$,

(i)$_K$ $K^2(v) = k$ for some k such that $v_2 \in H_k$,

(ii)$_L$ $L^2 S_j(v) = L^2(v) + l$,

(ii)$_K$ $K^2 S_j(v) = (-1)^l (K^2(v) - k)$.

We then define

$$J^2(v) := (L^2(v), K^2(v)). \tag{8.6}$$

Since $\Sigma_{l,k} = M_l \cap H_k$ we see that (i)$_L$, (i)$_K$ imply (i)$_J$. Because of (8.5) we also see that (ii)$_L$, (ii)$_K$ imply $J^2 S_j(v) = J^2(v) \cdot j^{-1}$, which is equivalent to (ii)$_J$.

We set $M_l^1 := M_l \setminus M_{l+1}$ and $M^2 := M_0 \cap M_1$, as well as $H_k^1 := H_k \setminus (H_{k+1} \cup H_{k+2})$ and $H_k^2 := H_k \cap H_{k+1} \setminus H_{k+2}$ and $H^3 := H_0 \cap H_1 \cap H_2$. The superscripts of these sets indicate for each element in how many of the sets M_0, M_1 or H_0, H_1, H_2 this element lies. By definition of D^* the first component v_1 lies in exactly one of the sets M_l. We define

$$L^2(v) := \begin{cases} l_2, & v_2 \in M_{l_2}^1 \\ l_1, & v_2 \in M^2,\ v_1 \in M_{l_1} \end{cases} \tag{8.7}$$

$$K^2(v) := \begin{cases} k_2, & v_2 \in H_{k_2}^1 \\ k_2 + l_1, & v_2 \in H_{k_2}^2,\ v_1 \in M_{l_1} \\ k_1, & v_2 \in H^3,\ v_1 \in H_{k_1} \end{cases} \tag{8.8}$$

In the definition of $K^2(v)$ we treat $l_1 \in \{0, 1\}$ as an element of $\mathbb{Z}_3$ and not, as usually, as an element of $\mathbb{Z}_2$.

The properties (i)$_L$ and (i)$_K$ are obviously fulfilled. Before we prove (ii)$_L$ and (ii)$_K$ we investigate the image sets of M_l and H_k under S_j. Proposition 6.11.(iv) states that S_j maps $\Sigma_{(l',k')}$ to $\Sigma_{j''}$ with $j'' := (l', k') \cdot j^{-1}$. By (8.5) we have $j'' = (l' + l, (-1)^l (k' - k))$. It follows for $M_{l'} = \bigcup_{k'} \Sigma_{(l',k')}$ and $H_{k'} = \bigcup_{l'} \Sigma_{(l',k')}$ that

$$S_j(M_{l'}) = M_{l'+l}, \quad S_j(H_{k'}) = H_{(-1)^l (k'-k)}. \tag{8.9}$$

Morreover, since S_j is bijective, we also know that S_j maps the complement of $M_{l'}$ or $H_{k'}$ to the complement of $M_{l'+l}$ or $H_{(-1)^l (k'-k)}$.

(ii)$_L$: If $v_2 \in M_{l_2}^1$, then $S_j(v_2) \in M_{l_2+l}^1$ and therefore $L^2 S_j(v) = l_2 + l = L^2(v) + l$. If $v_2 \in M^2$ and $v_1 \in M_{l_1}$, then $S_j(v_2) \in M^2$ and $S_j(v_1) \in M_{l_1+l}$, hence $L^2 S_j(v) = l_1 + l = L^2(v) + l$.

(ii)$_K$: If $v_2 \in H^1_{k_2}$, then $S_j(v_2) \in H^1_{(-1)^l(k_2-k)}$ and therefore $K^2 S_j(v) = (-1)^l(K^2(v) - k)$. If $v_2 \in H_3$, then $S_j(v_2) \in H^3$ and $S_j(v_1) \in H_{(-1)^l(k_1-k)}$, hence $K^2 S_j(v) = (-1)^l(K^2(v) - k)$. The remaining case is $v_2 \in H^2_{k_2}$ and $v_1 \in M_{l_1}$. For $l = 0, 1$ the point $S_j(v_2)$ lies in $H^2_{k_2-k}$ or $H^2_{-(k_2+1-k)}$ (in this order). Both cases are captured by the formula

$$S_j(v_2) \in H^2_{(-1)^l(k_2-k)-l}. \tag{8.10}$$

We also know that $S_j(v_1) \in M_{l'}$, where the number $l' \in \{0, 1\}$ is equal to $l_1 + l$ modulo 2. So we can write $l' = (-1)^l l_1 + l$. Combining both formulas shows

$$K^2 S_j(v) = (-1)^l(k_2 - k) - l + l' = (-1)^l(k_2 + l_1 - k) = (-1)^l(K^2(v) - k). \tag{8.11}$$

This finishes the proof. ∎

8.3 Definition of the generalized child functions

Let us restate Definition 3.10 (on page 24).

Definition 8.5 For $c \in \{1, 2\}$ we define the *generalized child function*

$$a_{c,j} := S^c A_j : D \to D.$$

Remark 8.6 Let $\pi_c(v_1, v_2) := v_c$ denote the projection onto the cth component. As already seen in (3.1) and (3.5), Proposition 8.3.(i) and the above definition ensure the "design features"

$$\pi_c S^c = S\pi_c : D \to \Sigma, \tag{8.12}$$

$$\pi_c a_{c,j} = a_j \pi_c : D \to \Sigma. \tag{8.13}$$

8.4 Equality of distributions

In this subsection we always assume that $n \in \mathbb{N}$. Let us restate Proposition 5.15 (on page 45).

Proposition 8.7 Let $v \in D$. The two stochastic processes $a_{c,J_n}(v)$ and $S^c A_{J_n}(v)$ have the same distribution in the Borel space of sequences in D.

Proof. Let $w := S^c(v)$. As already mentioned before in the heuristic argument of Proposition 5.15, the key is the following consequence of Corollary 8.4.(i):

$$\{A_j(v) : j \in \mathcal{J}\} = \{A_j(w) : j \in \mathcal{J}\}. \tag{8.14}$$

We set

$$B_n(v) := (A_{J_1}(v), \dots, A_{J_n}(v)), \quad b_{c,n} := \Big(a_{c,J_1}(v), \dots, a_{c,J_n}(v)\Big). \tag{8.15}$$

For any tuple $\widetilde{v} = (v^1, \dots, v^n) \in D^n$ of n pairs v^t we define $S^c(\widetilde{v}) = (S^c(v^1), \dots, S^c(v^n))$ component-wise. For any n we need to show that $S^c(B_n(v))$ and $b_{c,n}(v)$ have the same distribution.

We study all possible compositions of the six functions A_j and of the six functions a_j^c. Let $I = (i_1, i_2, \dots)$ be a sequence of indices $i_n \in \mathcal{J}$ and let $\mathcal{J}^{\mathbb{N}}$ denote the set of all such sequences I. For any $I \in \mathcal{J}^{\mathbb{N}}$ we define the multiindex $I_n := (i_1, \dots, i_n)$. We also define the "family trees"

$$T_n(v) := \Big\{\Big(A_{I_1}(v), \dots, A_{I_n}(v)\Big) : I \in \mathcal{J}^{\mathbb{N}}\Big\}, \tag{8.16}$$

$$\tau_{c,n}(v) := \Big\{\Big(a_{c,I_1}(v), \dots, a_{c,I_n}(v)\Big) : I \in \mathcal{J}^{\mathbb{N}}\Big\}. \tag{8.17}$$

So $T_n(v)$ or $\tau_{c,n}(v)$ is the set of all tupels $(v^1, \dots, v^n) \in D^n$ such that for each $t = 2, \dots, n$ there exists $j \in \mathcal{J}$ such that v^t is $A_j(v^{t-1})$ or $a_{c,j}(v^{t-1})$, resp. Since our random indices $j_1, j_2, \dots$ are independent and uniform on $\mathcal{J}$, the random vector $J_n = (j_1, \dots, j_n)$ is uniform on $\mathcal{J}^n$. Hence $S^c(B_n(v))$ is uniform on $S^c(T_n(v))$ and $b_{c,n}(v)$ is uniform on $\tau_{c,n}(v)$. So we need to show that

$$S^c(T_n(v)) = \tau_{c,n}(v). \tag{8.18}$$

Let $w := S^c(v)$. By (8.14) the set of the evaluations $A_j(w)$ equals the set of the evaluations $A_j(v)$. We obtain for fixed $I \in \mathcal{J}^{\mathbb{N}}$ that

$$\Big\{\Big(v', A_{I_1}(v'), \dots, A_{I_{n-1}}(v')\Big) : v' = A_j(w), j \in \mathcal{J}\Big\} \tag{8.19}$$

$$= \Big\{\Big(v', A_{I_1}(v'), \dots, A_{I_{n-1}}(v')\Big) : v' = A_j(v), j \in \mathcal{J}\Big\}. \tag{8.20}$$

Taking the union over all $I \in \mathcal{I}^{\mathbb{N}}$ shows

$$T_n(w) = T_n(v). \tag{8.21}$$

Now we prove (8.18) by induction over n. The case $n = 1$ is trivial. The induction step "$n-1 \mapsto n$" relies on the definition $a_{c,n} = S^c A_j$:

$$\begin{aligned} \tau_{c,n}(v) &= \bigcup_{j \in \mathcal{J}} \{a_{c,j}(v)\} \times \tau_{c,n-1}(a_{c,j}(v)) \\ S^c(T_n(v)) = S^c\left(\bigcup_{j \in \mathcal{J}} \{A_j(v)\} \times T_{n-1}(A_j(v))\right) &= \bigcup_{j \in \mathcal{J}} \{a_{c,j}(v)\} \times S^c(T_{n-1}(A_j(v))). \end{aligned}$$

We only need to show for $v' := A_j(v)$ and $w := a_{c,j}(v) = S^c(v')$ that $\tau_{c,n-1}(w) = S^c(T_{n-1}(v'))$. This is clear by the induction hypothesis $\tau_{c,n-1}(w) = S^c(T_{n-1}(w))$ and by $T_{n-1}(w) = T_{n-1}(v')$, see (8.21). ∎

9 Upcircles

In this section we always assume that $(x, r, d), (x', r', d') \in \mathbb{R} \times (0, \infty) \times \mathbb{R}$ and $z \in \mathbb{H}$. For the reader's convenience, we depict again the figures from Sections 2.5 and 5.4.

9.1 Upcircle coordinates

Let us restate Definition 2.16 (on page 18).

Definition 9.1 We define the

(i) *upcircle* $U(x, r) := \{z : |z - (x + ir)| = r\} \subseteq \mathbb{H}_0$,

(ii) *radius function* $R(x, z) := |z - x|^2/(2\Im z) \in (0, \infty)$,

(iii) *angle function* $D(x, z) := (\Re z - x)/\Im z \in \mathbb{R}$,

(iv) *unit upcircle function* $\gamma(d) := 2/(d - i) \in \mathbb{H}$,

(v) *upcircle coordinates function* $u(x, z) := (x, R(x, z), D(x, z))$.

Figure 10 illustrates the above definition.

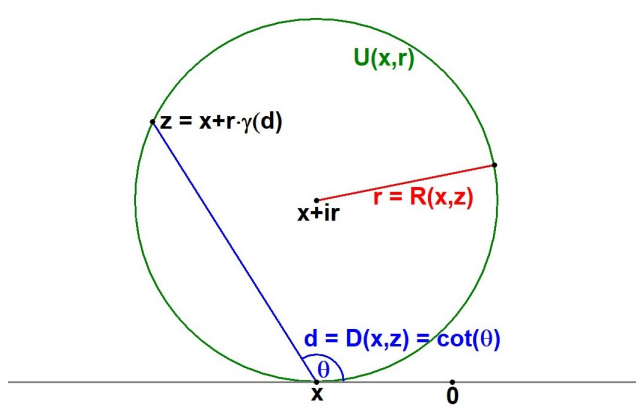

Figure 10: (Copy of Figure 4). Upcircle $U(x, r)$.

Lemma 9.2 Let $z := x + r \cdot \gamma(d)$ and $x' \in \mathbb{R}$. For $c := d^2 + 1$ and $y := (x - x')/(2r)$ we have the formulas

$$\frac{R(x', z)}{r} = cy^2 + 2dy + 1, \quad D(x', z) - d = cy.$$

Proof. Since $R(x', z) = R(0, z - x')$ and $D(x', z) = D(0, z - x')$ we may assume w.l.o.g. that $x' = 0$, i.e. $2ry = x$. Since $z = 2r(y + \frac{1}{2}\gamma(d))$ and $\frac{1}{2}\gamma(d) = 1/(d - i) = (d + i)/c$ we have

$$\Re z = 2r\left(y + \frac{d}{c}\right), \quad \Im z = \frac{2r}{c}. \tag{9.1}$$

This implies $D(0, z) = \Re z/\Im z = cy + d$. Together with $R(0, z) = |z|^2/(2\Im z)$ it follows that

$$\frac{R(0, z)}{r} \cdot c = \frac{|z|^2}{\Im z} \cdot \frac{c}{2r} = \frac{|z|^2}{(\Im z)^2} = \left(\frac{\Re z}{\Im z}\right)^2 + 1 \tag{9.2}$$

$$= c^2y^2 + 2cdy + d^2 + 1 = c \cdot \Big(cy^2 + 2dy + 1\Big). \tag{9.3}$$

Dividing by $c > 0$ yields the formula for $R(0, z)/r$. ∎

Let us restate and prove Proposition 2.18 (on page 19).

Proposition 9.3 The objects from Definition 9.1 have the following properties:

(i) $\gamma : \mathbb{R} \to U(0, 1) \setminus \{0\}$ is bijective.

(ii) u is bijective with inverse $u^{-1}(x, r, d) = (x, x + r \cdot \gamma(d)) \in \{x\} \times U(x, r)$.

Proof. We compute

$$R(0, z) \cdot \gamma(D(0, z)) = \frac{|z|^2}{2\Im z} \cdot \frac{2}{\frac{\Re z}{\Im z} - i} = \frac{|z|^2}{\overline{z}} = z. \tag{9.4}$$

(i): The image $\gamma(d)$ obviously lies in $\mathbb{H}$ and therefore cannot be 0. The following computation shows that $\gamma(d) \in U(0, 1)$:

$$|\gamma(d) - i| = \left|\frac{2 - i(d - i)}{d - i}\right| = \frac{|1 - id|}{|d - i|} = 1. \tag{9.5}$$

It is clear that γ is injective. For the surjectivity choose $z \in U(0, 1) \setminus \{0\}$. Due to this section's underlying assumption $z \in \mathbb{H}$ this is equivalent to $|z - i| = 1$. We can rewrite this as

$$0 = (z - i)(\overline{z} + i) - 1 = |z|^2 + i(z - \overline{z}) = |z|^2 - 2\Im z. \tag{9.6}$$

Hence $R(0,z)=1$, so by (9.4) we obtain $\gamma(D(0,z))=z$.

(ii): Let $\widetilde{u}(x,r,d) := (x, x+r\cdot\gamma(d))$. We verify both identities $\widetilde{u}u = \mathrm{id}$ and $u\widetilde{u} = \mathrm{id}$:

- We have $R(x,z)\cdot\gamma(D(x,z)) = R(0,z-x)\cdot\gamma(D(0,z-x)) = z-x$, see (9.4). Hence $\widetilde{u}u(x,z) = (x,z)$.
- Lemma 9.2 in the special case $x = x'$ states $R(x,z) = r$ and $D(x,z) = d$ for $z = x + r\cdot\gamma(d)$. Thus $u\widetilde{u}(x,r,d) = (x,r,d)$.

This shows that u^{-1} exists and that $u^{-1} = \widetilde{u}$. The second component lies indeed on the upcircle $U(x,r)$ because $\gamma(d)\in U(0,1)$ according to (i). ∎

9.2 Möbius transformations in upcircle coordinates

Let us restate Definition 2.21.

Definition 9.4 For $g\in\mathcal{G}$ we define $g^\sharp : \mathbb{R}\to[-\infty,\infty]\setminus\{0\}$ by

$$g^\sharp(x) := \frac{2}{x-g^{-1}(\infty)}.$$

Remark 9.5 If $g\in\mathcal{G}$ is bounded on some set $B\subseteq\mathbb{R}$, then $g^\sharp$ is also bounded on B.

Let us restate and prove Proposition 2.23 (on page 20).

Proposition 9.6 Let $g\in\mathcal{G}$ and $(x,z)\in\mathbb{R}\times\mathbb{H}$ such that $g(x)\neq\infty$ (which is equivalent to $|g'(x)|\neq\infty$). Suppose we know the upcircle coordinates $(x,r,d) := u(x,z)$.

(i) The upcircle coordinates $(x',r',d') := ug(x,z)$ have the formulas

$$x' = g(x), \quad r' = |g'(x)|\cdot r, \quad |d'| = \left|d + g^\sharp(x)\cdot r\right|.$$

(ii) g preserves upcircles in the following way:

$$g(U(x,r)) = U(x',r').$$

Proof. (i): We first show the following two special cases:

- Case 1: If $g = m$ then $(x', r', d') = (-x, r, -d)$.
- Case 2: If $g = f \in \mathcal{F}$ then

$$x' = f(x), \quad r' = f'(x) \cdot r, \quad d' = d + f^\sharp(x) \cdot r. \tag{9.7}$$

Then the general case $g = m^l f$ follows by combining both special cases and using that $|g'(x)| = f'(x)$ and $g^{-1}(\infty) = f^{-1}(\infty)$.

Case 1 only consists of the verification

$$r' = R(m(x), m(z)) = \frac{|m(z-x)|^2}{2\Im z} = R(x, z) = r, \tag{9.8}$$

$$d' = D(m(x), m(z)) = \frac{-\Re z - (-x)}{\Im z} = -D(x, z) = -d. \tag{9.9}$$

Case 2: We set

$$\widetilde{r} := f'(x) \cdot r, \quad \widetilde{d} := d + g^\sharp(x) \cdot r. \tag{9.10}$$

We show that $f(z) = f(x) + \widetilde{r} \cdot \gamma(\widetilde{d})$. The formula for u^{-1}, see Proposition 9.3.(ii), then implies that $r' = R(f(x), f(z)) = \widetilde{r}$ and $d' = D(f(x), f(z)) = \widetilde{d}$. Similarly, we also have $z = x + r \cdot \gamma(d)$.

Let us use our standard notation for functions $f \in \mathcal{F}$, namely $f(z) = (c_1 z + c_2)/(c_3 z + c_4)$ with coefficients $c_i \in \mathbb{R}$ such that $c_1 c_4 - c_2 c_3 > 0$. Notice that the following computations are even true in the affine linear case $c_3 = 0$ which is equivalent to $f^{-1}(\infty) = \infty$. We have by Definition 2.5 (on page 14) and by the subsequent Remark 2.6.(iv) that

$$\frac{f(z) - f(x)}{z - x} = f'(x, z) = f'(x) \cdot \frac{c_3 x + c_4}{c_3 z + c_4}. \tag{9.11}$$

Together with $z = x + r \cdot \gamma(d)$ it follows that

$$\frac{f(z) - f(x)}{\widetilde{r}} = \frac{z - x}{r} \cdot \frac{c_3 x + c_4}{c_3 z + c_4} = \gamma(d) \cdot \frac{c_3 x + c_4}{c_3 x + c_3 r \gamma(d) + c_4} = \left(\frac{1}{\gamma(d)} + \frac{c_3 r}{c_3 x + c_4} \right)^{-1} \tag{9.12}$$

$$= 2 \cdot \left(\frac{2}{\gamma(d)} + \frac{2r}{x - f^{-1}(\infty)} \right)^{-1} = 2 \cdot \Big((d - i) + (\widetilde{d} - d) \Big)^{-1} = \gamma(\widetilde{d}). \tag{9.13}$$

Hence $f(z) = f(x) + \widetilde{r} \cdot \gamma(\widetilde{d})$.

(ii): Notice that $(x', r', d') = ugu^{-1}(x, r, d)$. In the following, we fix x and r and consider d a free variable. Let us define x' and r' by the first two formulas from part (i) which do not depend on d. The proof of (i) shows that there is a sign $\sigma \in \{-1, 1\}$ such that any d satisfies

$$ugu^{-1}(x, r, d) = \Big(x', r', \sigma \cdot (d + g^\sharp(x) \cdot r)\Big). \tag{9.14}$$

Thus

$$\Big\{gu^{-1}(x, r, d) : d \in \mathbb{R}\Big\} = \Big\{u^{-1}(x', r', d') : d' \in \mathbb{R}\Big\}. \tag{9.15}$$

As a consequence of both parts of Proposition 9.3 we see that the right-hand side is $\{x'\} \times (U(x', r') \setminus \{x'\})$ and the left-hand side is the image under g of the set $\{x\} \times (U(x, r) \setminus \{x\})$. This implies that $U(x', r') = g(U(x, r))$. ∎

9.3 Generalized child functions in upcircle coordinates

Let us consider the generalized child function $a_{c,j} : D \to D$ in the case $c = 1$.

Definition 9.7 We define $a_j^\sharp : \mathbb{R} \setminus E \to \mathbb{R}$ by

$$a_j^\sharp(x) := (S_{j'} A_j)^\sharp(x), \quad \text{if} \quad a_j(x) = S_{j'} A_j(x).$$

Remark 9.8 We notice the similarity to Proposition 7.10.(iii) which states for $x \in \mathbb{R}$ (see Remark 7.11) that

$$|a_j'(x)| = |(S_{j'} A_j)'(x)|, \quad \text{if} \quad a_j(x) = S_{j'} A_j(x). \tag{9.16}$$

Lemma 9.9 The function $a_j^\sharp$ has the following properties:

(i) $a_j^\sharp$ is well-defined and bounded.

(ii) Given the upcircle coordinates $(x, r, d) := u(x, z)$, we have the following formula for the upcircle coordinates $(x', r', d') := ua_{1,j}(x, z)$:

$$x' = a_j(x), \quad r' = |a_j'(x)| \cdot r, \quad |d'| = |d + a_j^\sharp(x) \cdot r|.$$

Proof. (i): For $x \in \mathbb{R} \setminus E$ we have $A_j(x) \in \mathbb{R} \setminus E \subseteq \Sigma^*$ according to Lemma 8.2.(ii). This means that there is a unique index j' such that $A_j(x) \in \Sigma_{j'}$ or equivalently $a_j(x) = S_{j'}A_j(x)$. For the boundedness, we fix j' and show that $a_j^\sharp$ is bounded on $B := \{x \in \mathbb{R}\setminus E : A_j(x) \in \Sigma_{j'}\}$. Let $g := S_{j'}A_j \in \mathcal{G}$. For any $x \in B$ we have $a_j^\sharp(x) = g^\sharp(x)$ and $g(x) \in \Sigma$. Hence g is bounded on B. It follows with Remark 9.5 that $a_j^\sharp$ is bounded on B.

(ii): By definition of S^1 we have $a_{1,j}(x,z) = S_{j'}A_j(x)$, where j' is chosen such that $A_j(x) \in \Sigma_{j'}$ or equivalently $a_j(x) = S_{j'}A_j(x)$. Thus $a_j^\sharp(x) = (S_{j'}A_j)^\sharp(x)$ and $|a_j'(x)| = |(S_{j'}A_j)'(x)|$. Now, Proposition 9.6 shows the three formulas in (ii). ∎

9.4 Upcircles intersecting two shape sets

Let us restate and prove Proposition 5.10 (on page 42). The case $c = 2$ is illustrated in Figure 11. We obtain an illustration of the case $c = 1$ by interchanging the labels x, r, z with x', r', z'.

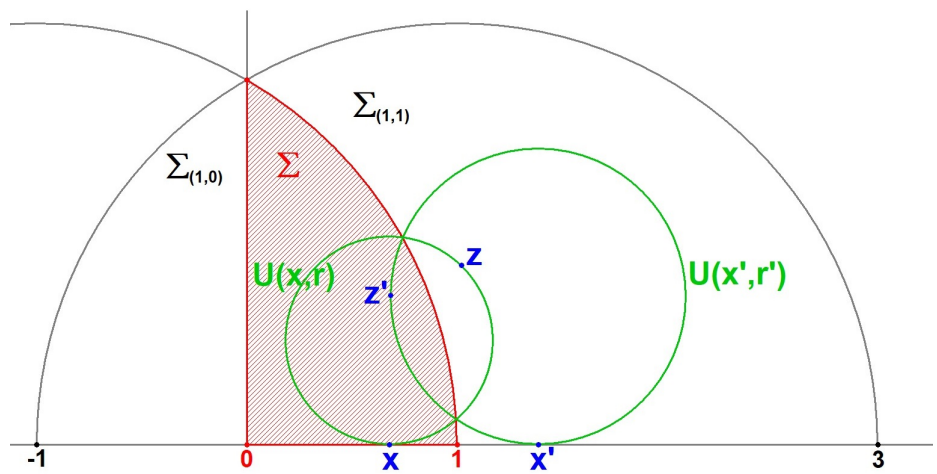

Figure 11: (Copy of Figure 5). Upcircles $U(x,r)$ and $U(x',r') = S_{(1,1)}(U(x,r))$ with elements z and $z' = S_{(1,1)}(z)$.

Proposition 9.10 Let $(x,z) \in D$ and $(x',z') := S^c(x,z)$. If $c = 2$ we assume $x \in \Sigma$, and if $c = 1$ we assume $z \in \Sigma$. We define the upcircle coordinates $(x,r,d) := u(x,z)$ and

$(x', r', d') := u(x', z')$. Suppose r is small enough such that $U(x, r)$ intersects at most two shape sets. Then the following bounds hold:

$$|x - x'| \leq 4r, \quad \tfrac{1}{4}r \leq r' \leq 4r, \quad \Big||d| - |d'|\Big| \leq 2r. \tag{9.17}$$

Proof. The points x and z lie on the upcircle $U(x, r)$ which intersects Σ and at most one other shape set Σ_j. If such j exists, then j is either $(1, 0)$ or $(1, 1)$. Figure 11 illustrates the case $j = (1, 1)$.

$c = 2$: The case $z \in \Sigma$ is trivial because then $(x', z') = (x, z)$. Assume $z \notin \Sigma$, thus Σ_j is the only shape set that contains z. It follows by construction of S^2 that $(x', z') = S^2(x, z) = S_j(x, z)$. Together with Lemma 9.11 (which is stated right after this proof) it follows (9.17).

$c = 1$: The case $x \in \Sigma$ is trivial because then $(x', z') = (x, z)$. Assume $x \notin \Sigma$, thus Σ_j is the only shape set that contains x. It follows by construction of S^1 that $(x', z') = S^1(x, z) = S_j(x, z)$. Lemma 6.11.(iv) states that $S_j(\Sigma_{j'}) = \Sigma_{j''}$ with $j'' = j' \cdot j^{-1}$. Hence $x' \in \Sigma_{j \cdot j^{-1}} = \Sigma$ and $z' \in \Sigma_{j^{-1}}$. In both cases $j = (1, 0)$ and $j = (1, 1)$ we easily check $j \cdot j = (0, 0)$. Thus $j^{-1} = j$ and $S_j^{-1} = S_j$. We conclude that $z' \in \Sigma_j$ and $(x, z) = S_j(x', z')$.

The points x' and z' lie on the upcircle $U(x', r')$ which therefore intersects Σ and Σ_j. Now, we show that $U(x', r')$ does not intersect any other shape set. Assume that $U(x', r')$ contains an element $w \in \Sigma_{j'}$ with $j' \in \mathcal{J} \setminus \{(0, 0), j\}$. Since $U(x', r') = S_j(U(x, r))$, see Proposition 9.6.(ii), it follows that $U(x, r)$ contains $S_j^{-1}(w) = S_{j^{-1}}(w) \in \Sigma_{j''}$ with $j'' = j' \cdot j \in \mathcal{J} \setminus \{(0, 0), j\}$ in contradiction to our assumption that $U(x, r)$ intersects at most two shape sets.

Together with Lemma 9.11 (where we interchange x, r, d, z with x', r', d', z') we obtain the following bounds:

$$|x - x'| \leq 4r', \quad r' \leq r \leq 4r', \quad \Big||d| - |d'|\Big| \leq 2r'. \tag{9.18}$$

It follows that (9.17) holds. ∎

Figure 11 illustrates the case $j = (1,1)$ of the following lemma.

Lemma 9.11 Let $(x,z) \in D$ and $(x',z') := S_j(x,z)$, where j is either $(1,0)$ or $(1,1)$. We define the upcircle coordinates $(x,r,d) := u(x,z)$ and $(x',r',d') := u(x',z')$. Suppose that $x \in \Sigma$ and that $U(x,r)$ intersects Σ_j (not necessarily in z). Then the following bounds hold:

$$|x - x'| \leq 4r, \quad r \leq r' \leq 4r, \quad \Big||d| - |d'|\Big| \leq 2r.$$

Proof. Σ is the intersection of the first quadrant Q and the disk D around -1 of radius 2. The center of the upcircle $U(x,r)$ is $x + ir$.

Case 1: $j = (1,0)$. The parallel to the real line passing through $x + ir$ intersects $U(x,r)$ in some point w outside (or on the boundary) of Q, i.e. $\Re w \leq 0$. Since $w = x + ir - r$ it follows that $x \leq r$. Proposition 9.6 states that $x' = -x$ and $r' = r$ and $|d'| = |d|$. Hence $|x - x'| = 2x \leq 2r$.

Case 2: $j = (1,1)$. The line passing through the centers -1 and $x + ir$ intersects $U(x,r)$ in some point w outside (or on the boundary) of D, i.e. $|w - (-1)| \geq 2$. See Figure 12.

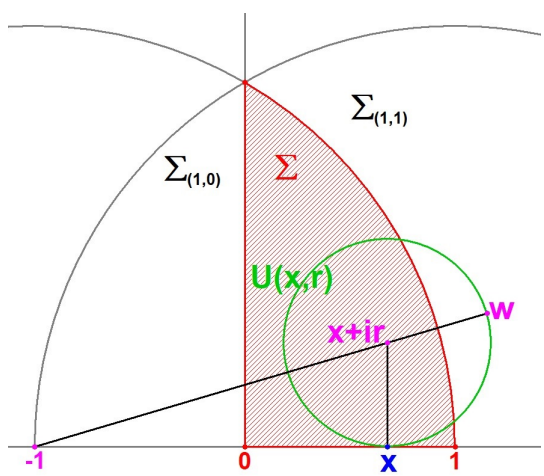

Figure 12: Upcircle $U(x,r)$ intersecting $\Sigma_{(1,1)}$.

Since

$$|w-(-1)| = |w-(x+ir)| + |(x+ir)-(-1)| = r + |x+1+ir| \tag{9.19}$$

it follows that $2-r \leq |x+1+ir|$ and therefore $4-4r \leq (x+1)^2$. Proposition 9.6 states that

$$x' = -1 + \frac{4}{x+1}, \quad r' = \frac{4}{(x+1)^2} \cdot r, \quad |d'| = \left| d + \frac{2r}{x+1} \right|. \tag{9.20}$$

In the following, we use $x \in [0,1]$ several times:

$$|x'-x| = x'-x = \frac{4-(x+1)^2}{x+1} \leq \frac{4r}{x+1} \leq 4r, \tag{9.21}$$

$$\frac{r'}{r} = \frac{4}{(x+1)^2} \in [1,4], \tag{9.22}$$

$$\big||d|-|d'|\big| \leq \frac{2r}{x+1} \leq 2r. \tag{9.23}$$

■

10 Ergodicity and the law of large numbers

10.1 First setting

In this subsection we assume that x, y are real variables with $x \neq y$ and we impose Setting 4.1 (on page 28) which we first restate.

Setting 10.1 Let $d \in \mathbb{N}$ and C be a compact subset of $(\mathbb{R}^d, |\cdot|)$, where $|\cdot|$ is some norm. We fix a finite index set $\mathcal{I}$ and always assume $i \in \mathcal{I}$. Let $(f_i)_i$ be a family of Lipschitz-continuous functions $f_i : C \to C$. Consider the Markov kernel M from C to C given by

$$M(x, \cdot) := \frac{1}{|\mathcal{I}|} \sum_i \delta_{f_i(x)}, \quad x \in C. \tag{10.1}$$

Let $i_1, i_2, \ldots$ be a sequence of independent random indices such that each i_n is uniformly distributed on $\mathcal{I}$. For $n \in \mathbb{N}_0$ we define the random multiindex $I_n := (i_1, \ldots, i_n)$ which is uniformly distributed on $\mathcal{I}^n$. Notice for any $x \in C$ that

$$U_n^x := f_{I_n}(x), \quad n \in \mathbb{N}_0$$

is a Markov chain with $U_0^x = x$ and kernel M because $U_n^x = f_{i_n}(U_{n-1}^x)$.

Let us restate Definitions 4.4 and 4.9 (on pages 29 and 31) which use the notation $|b'(x, y)| = |b(x) - b(y)|/|x - y|$, see Definition 3.2.

Definition 10.2 Let us define the following notions of contractivity on average:

(i) $(f_i)_i$ *contracts on average in step* $N \in \mathbb{N}$, if

$$\sup_{x,y \in C} \mathbb{E} \ln |f'_{I_N}(x, y)| < 0,$$

(ii) $(f_i)_i$ *strongly contracts on average in step* $N \in \mathbb{N}$, if there exists $\epsilon > 0$ such that

$$\sup_{x,y \in C} \mathbb{E} \ln \Big(\max\{|f'_{I_N}(x, y)|, \epsilon\} \Big) < 0,$$

(iii) $(f_i)_i$ (*strongly*) *contracts on average*, if it (strongly) contracts on average in step 1.

Definition 10.3 We define:

(i) M is *ergodic*, if it has a unique invariant probability measure μ which is also attracting, i.e. $\mu M = \mu$ and νM^n converges weakly to μ for any probability measure ν on C.

(ii) For any $x \in C$ we define the Markov chain

$$U_n^x := f_{I_n}(x), \quad n \in \mathbb{N}_0,$$

which starts in $U_0^x = x$ and has kernel M.

(iii) M satisfies the *law of large numbers*, if M is ergodic with invariant measure μ and for any continuous function $\phi : C \to \mathbb{R}$ and $x \in C$ we have a.s.

$$\lim_n \frac{1}{n+1} \sum_{t=0}^{n} \phi(U_t^x) = \int \phi \, d\mu.$$

Let us restate and prove Lemma 4.12 (on page 32).

Lemma 10.4 Let $N \in \mathbb{N}$. The kernel $M^N = M \dots M$ has the formula

$$M^N(x, \cdot) = \frac{1}{|\mathcal{I}^N|} \sum_{I \in \mathcal{I}^N} \delta_{f_I(x)}, \quad x \in C, \tag{10.2}$$

and satisfies the following properties:

(i) If M^N is ergodic, then M is ergodic.

(ii) If M^N satisfies the law of large numbers, then M satisfies the law of large numbers.

Proof. (i): The kernel M^N has a unique invariant probability measure μ and μ is also attracting. We have $(\mu M)M^N = (\mu M^N)M = \mu M$ and by the uniqueness of μ it follows $\mu M = \mu$, so μ is also invariant for M. The uniqueness of the invariant measure for M is clear because any invariant measure for M is also invariant for M^N. To see that μ is attracting for M let ν be a probability measure on C. We decompose the sequence $\nu_n := \nu M^n$ into N subsequences $\nu_{n,r} := \nu_{Nn+r}$ with $r \in \{0, \dots, N-1\}$. Each subsequence $\nu_{n,r} = (\nu M^r)(M^N)^n$ converges weakly to μ, thus ν_n converges weakly to μ.

(ii): For fixed $x \in C$ we decompose U_n^x into N subsequences $U_{n,r}^x := U_{Nn+r}^x$ with $r \in \{0, \dots, N-1\}$ and show the law of large numbers for each of them. For each r and $n \in \mathbb{N}$ the random multiindex $I_{n,r} := (i_{N(n-1)+r+1}, \dots, i_{Nn+r})$ is uniformly distributed on $\mathcal{I}^N$. Moreover, the random multiindex I_{Nn+r} is the concatenation of $I_{n,r}$ and $I_{N(n-1)+r}$. This shows the independence of $I_r, I_{1,r}, I_{2,r}, \dots$ and it also shows that

$$U_{0,r}^x = f_{I_r}(x), \quad U_{n,r}^x = f_{I_{n,r}}(U_{n-1,r}^x). \tag{10.3}$$

Hence $U_{n,r}^x$ is a Markov chain with kernel M^N. The law of large numbers for M^N shows for any (non-random) multiindex $I^* \in \mathcal{I}^r$ that

$$\mathbb{P}\left(\lim_n \frac{1}{n+1} \sum_{t=0}^n \phi(U_{t,r}^x) = \int \phi \, d\mu \;\middle|\; I_r = I^* \right) = 1. \tag{10.4}$$

It follows for any (non-random) sequence t_n in $\mathbb{N}_0$ with $\lim_n t_n = \infty$ that a.s.

$$\lim_n \frac{1}{t_n+1} \sum_{t=0}^{t_n} \phi(U_{t,r}^x) = \int \phi \, d\mu. \tag{10.5}$$

Let $t_{n,r}$ be the largest $t \in \mathbb{N}_0$ such that $Nt + r \leq n$. Then

$$\sum_{t=0}^n \phi(U_t^x) = \sum_{r=0}^{N-1} V_{n,r}^x, \quad V_{n,r} := \sum_{t=0}^{t_{n,r}} \phi(U_{t,r}^x). \tag{10.6}$$

Since $\lim_n (t_{n,r}+1)/(n+1) = \frac{1}{N}$ and in particular $\lim_n t_{n,r} = \infty$ we know that a.s.

$$\lim_n \frac{1}{n+1} V_{n,r}^x = \frac{1}{N} \int \phi \, d\mu. \tag{10.7}$$

This immediately implies that a.s.

$$\lim_n \frac{1}{n+1} \sum_{t=0}^n \phi(U_t^x) = \int \phi \, d\mu. \tag{10.8}$$

■

Let us restate Proposition 4.13 (on page 33) which we have proven already to be a consequence of Lemma 4.12 (or its restated version Lemma 10.4).

Proposition 10.5 If $(f_i)_i$ contracts on average in some step N, then M satisfies the law of large numbers.

10.2 Extended setting

In this subsection we impose Setting 4.22 (on page 36) which we first restate.

Setting 10.6 We extend Setting 4.1 (restated in Setting 10.1) in the following way:

(i) Let $(f_i')_i$ be a family of Lipschitz-continuous functions $f_i' : C \to C'$, where $C' \subseteq \mathbb{R}$ is compact. Thus $\widetilde{C} := C \times C' \subseteq \mathbb{R}^{d+1}$ is compact, too.

(ii) For any $n \in \mathbb{N}$ and multiindices $I = (i, J) \in \mathcal{I}^n$ with $J \in \mathcal{I}^{n-1}$ let us define $F_I : C \to \widetilde{C}$ by

$$F_I := (f_I, f_i' f_J). \tag{10.9}$$

(iii) In Definition 10.3.(ii) we defined for any $x \in C$ the Markov chain $U_n^x = f_{I_n}(x)$ with state space C. Now we define for any $x \in C$ a stochastic process $\widetilde{U}_n^x$ with state space $\widetilde{C}$ by

$$\widetilde{U}_n^x := F_{I_n}(x), \quad n \in \mathbb{N}_0. \tag{10.10}$$

Remark 10.7 The proof of the following proposition shows that $\widetilde{U}_n^x$ is also a Markov chain.

Let us restate and prove Proposition 4.25 (on page 37), the modified law of large numbers.

Proposition 10.8 Suppose that $(f_i)_i$ strongly contracts on average in some step N. Let μ denote the invariant measure of M, see Proposition 4.13. For any continuous function $\phi : \widetilde{C} \to \mathbb{R}$ and $x \in C$ we have a.s.

$$\lim_n \frac{1}{n} \sum_{t=1}^{n} \phi(\widetilde{U}_t^x) = \frac{1}{|\mathcal{I}|} \sum_i \int \phi F_i \, d\mu. \tag{10.11}$$

Proof. By Definition 10.2.(ii), there exists $\epsilon > 0$ such that

$$\sup_{x,y \in C} \mathbb{E} \ln \Big(\max\{|f_{I_N}'(x,y)|, \epsilon\} \Big) < 0. \tag{10.12}$$

In particular, $(f_i)_i$ contracts on average in step N. Proposition 10.5 ensures that M satisfies the law of large numbers. Let μ denote the invariant measure of M.

Let us define $\widetilde{f}_i : \widetilde{C} \to \widetilde{C}$ and a Markov kernel $\widetilde{M}$ from $\widetilde{C}$ to $\widetilde{C}$ by

$$\widetilde{f}_i(x, x') := (f_i(x), f_i'(x)) \tag{10.13}$$

$$\widetilde{M}(v, \cdot) := \frac{1}{|\mathcal{I}|} \sum_i \delta_{\widetilde{f}_i(v)}. \tag{10.14}$$

Notice that $\widetilde{f}_i(x, x')$ does not depend on x'.

The proof is structured in the following way:

- Step 1 shows $\widetilde{f}_I(x) = F_I(x, x')$ for $(x, x') \in \widetilde{C}$ and $I \in \mathcal{I}^n$ with $n \in \mathbb{N}$. This implies the analogue of $U_n^x = f_{I_n}(x)$, namely $\widetilde{U}_n^x = \widetilde{f}_{I_n}(x, x')$. Hence $\widetilde{U}_n^x = \widetilde{f}_{i_n}(\widetilde{U}_{n-1}^x)$ is a Markov chain with kernel $\widetilde{M}$.

- Step 2 justifies why we may assume w.l.o.g. that ϵ is a Lipschitz-constant of $g := f_{i_N}' f_{I_{N-1}}$ which is the second component of $F_{I_N} = (f_{I_N}, g)$.

- Step 3 verifies that $\widetilde{M}$ satisfies the law of large numbers.

- Step 4 shows that the invariant measure $\widetilde{\mu}$ of $\widetilde{M}$ is the arithmetic mean of the distributions of $F_i : C \to \widetilde{C}$ w.r.t μ. Thus any integral $\int \phi \, d\widetilde{\mu}$ can be written as $|\mathcal{I}|^{-1} \sum_i \int \phi F_i \, d\mu$.

For any continuous function $\phi : \widetilde{C} \to \mathbb{R}$ and $x \in C$ the combination of Steps 1, 3, 4 leads to the a.s. convergence

$$\lim_n \frac{1}{n} \sum_{t=1}^n \phi(\widetilde{U}_t^x) = \int \phi \, d\widetilde{\mu} = \frac{1}{|\mathcal{I}|} \sum_i \int \phi F_i \, d\mu. \tag{10.15}$$

The very technical proofs of Steps 1 and 2 may be skipped at first reading.

Step 1: Let $I^n = (i^1, \ldots, i^n) \in \mathcal{I}^n$. We verify

$$\widetilde{f}_{I^n}(x, x') = F_{I^n}(x) \tag{10.16}$$

with the following illustration of the composition $\widetilde{f}_{I^n}(x,x') = \widetilde{f}_{i^n} \dots \widetilde{f}_{i^1}(x,x')$, where the number $t = 1,\dots,n$ over an arrow signifies the application of the function $\widetilde{f}_{i^t} : (x,x') \mapsto (f_{i^t}(x), f'_{i^t}(x))$:

$$(x,x') \overset{1}{\mapsto} \Big(f_{i^1}(x), f'_{i^1}(x)\Big) \overset{2}{\mapsto} \Big(f_{I^2}(x), f'_{i^2} f_{I^1}(x)\Big) \overset{3}{\mapsto} \dots \overset{n}{\mapsto} \Big(f_{I^n}(x), f'_{i^n} f_{I^{n-1}}(x)\Big). \quad (10.17)$$

This informal explanation is more intuitive than the formal induction proof: The case $n = 1$ holds by definition and for "$n-1 \mapsto n$" we apply the induction hypothesis to $(i^2,\dots,i^{n-1})$.

Step 2: Let us assume that the proposition holds under the additional assumption that ϵ is a Lipschitz-constant of g. Now we show the general case.

Since all functions f_i and f'_i are Lipschitz-continuous and $\mathcal{I}$ is finite, there are constants $L, L' \in (0,\infty)$ such that L is a Lipschitz-constant for each function f_i and L' is a Lipschitz-constant for each function f'_i. Hence $L'L^{N-1}$ is a Lipschitz-constant of g. Using the rescaling factor $\lambda := \epsilon/(L'L^{N-1}) > 0$ we set

$$f'_{i,\star} := \lambda f'_i : C \to \lambda C' =: C'_\star \quad (10.18)$$

and $\widetilde{C}_\star := C \times C'_\star$. We define $F_{I,\star}$, $\widetilde{U}^x_{n,\star}$, $g_\star$ analogously to F_I, $\widetilde{U}^x_n$, g, where the family $(f'_{i,\star})_i$ replaces $(f'_i)_i$. Then ϵ is a Lipschitz-constant of $g_\star = \lambda g$ and the continuous function $\phi_\star : \widetilde{C}_\star \to \mathbb{R} : (x,x') \mapsto \phi(x, \frac{1}{\lambda} \cdot x')$ fulfills $\phi_\star(\widetilde{U}^x_{n,\star}) = \phi(\widetilde{U}^x_n)$ for any $x \in C$.

Step 3: We show that $\widetilde{M}$ satisfies the law of large numbers. According to Proposition 10.5 we only need to show that $(\widetilde{f}_i)_i$ contracts on average in step N. Let $(x,x'),(y,y')$ be different pairs in $\widetilde{C}$. Step 1 and the function g (with Lipschitz-constant ϵ) from Step 2 allow us to write

$$\widetilde{f}_{I_N}(x,x') - \widetilde{f}_{I_N}(y,y') = F_{I_N}(x) - F_{I_N}(y) = \Big(f_{I_N}(x) - f_{I_N}(y),\ g(x) - g(y)\Big). \quad (10.19)$$

We equip $\mathbb{R}^{d+1}$ with the norm $|(x,x')|_\infty := \max\{|x|,|x'|\}$, where $x \in \mathbb{R}^d$, $x' \in \mathbb{R}$. It follows that

$$|\widetilde{f}_{I_N}(x,x') - \widetilde{f}_{I_N}(y,y')|_\infty \le \max\Big\{|f_{I_N}(x) - f_{I_N}(y)|,\ \epsilon \cdot |x-y|\Big\}. \quad (10.20)$$

In this inequality both sides vanish for $x = y$. Suppose $x \neq y$. Using $|(x,x') - (y,y')|_\infty \ge |x-y|$ leads to

$$\Big|\widetilde{f}'_{I_N}\Big((x,x'),(y,y')\Big)\Big| \le \max\Big\{|f'_{I_N}(x,y)|,\ \epsilon\Big\}. \quad (10.21)$$

Together with our assumption (10.12) it follows that $(\widetilde{f}_i)_i$ contracts on average in step N.

Step 4: Let $\widetilde{\mu}$ denote the invariant measure of $\widetilde{M}$ and μ_i denote the distribution of $F_i : C \to \widetilde{C}$ w.r.t. μ. We show that

$$\widetilde{\mu} = \frac{1}{|\mathcal{I}|} \sum_i \mu_i.$$

Since $\widetilde{\mu} = \widetilde{M}\widetilde{\mu}$, any Borel set $\widetilde{B} \subseteq \widetilde{C}$ satisfies

$$\widetilde{\mu}(\widetilde{B}) = \int \mathbf{1}_{\widetilde{B}} \, d\widetilde{\mu} = \int \int \mathbf{1}_{\widetilde{B}} \, d\widetilde{M}(v, \cdot) \, \widetilde{\mu}(dv) = \frac{1}{|\mathcal{I}|} \sum_i \int \mathbf{1}_{\widetilde{B}}(\widetilde{f}_i(v)) \widetilde{\mu}(dv). \tag{10.22}$$

Let μ^* be the probability measure on C defined by $\mu^*(B) := \widetilde{\mu}(B \times C')$ and μ_i^* be the distribution of $F_i : C \to \widetilde{C}$ w.r.t. μ^*. For $v = (x, x')$ we have $\widetilde{f}_i(v) = F_i(x)$ and therefore

$$\int \mathbf{1}_{\widetilde{B}}(\widetilde{f}_i(v)) \widetilde{\mu}(dv) = \int \mathbf{1}_{\widetilde{B}}(F_i(x)) \widetilde{\mu}(d(x, x')) = \widetilde{\mu}(\{F_i \in \widetilde{B}\} \times C') \tag{10.23}$$

$$= \mu^*(\{F_i \in \widetilde{B}\}) = \mu_i^*(\widetilde{B}). \tag{10.24}$$

Combining both results yields

$$\widetilde{\mu}(\widetilde{B}) = \frac{1}{|\mathcal{I}|} \sum_i \mu_i^*(\widetilde{B}). \tag{10.25}$$

It remains to show that $\mu^* = \mu$ because this implies that $\mu_i^* = \mu_i$. Let $B \subseteq C$ be a Borel set. If we apply formula (10.25) in the special case $\widetilde{B} := B \times C'$, then the left-hand side is $\mu^*(B)$ and for the right-hand side we notice that

$$\mu_i^*(\widetilde{B}) = \mu^*\Big(\Big\{x \in C : F_i(x) \in B \times C'\Big\}\Big) = \mu^*(\{f_i \in B\}). \tag{10.26}$$

Thus

$$\mu^*(B) = \frac{1}{|\mathcal{I}|} \sum_i \mu^*(\{f_i \in B\}) = \int \int \mathbf{1}_B \, dM(x, \cdot) \, \mu^*(dx) = (\mu^* M)(B). \tag{10.27}$$

So μ^* is invariant for M and must therefore be equal to μ. ∎

11 Shape integrand

11.1 Definition and connection to contractivity on average

We remind ourselves of Definition 7.4 (on page 57) of the fixpoint polynomial $\widehat{h} : \overline{\mathbb{R}} \to [3, \infty]$ and the subsequent Lemma 7.6 which states for any j and $x \in \overline{\mathbb{R}}$ that

$$|S_j'(x)| \cdot \widehat{h}(x) = \widehat{h} S_j(x). \tag{11.1}$$

Definition 11.1 We define functions $\psi, \Psi, \Psi_n : \overline{\mathbb{R}} \to \mathbb{R}$ with $n \in \mathbb{N}_0$ by

$$\psi(x) := \ln\left(\frac{\frac{2}{3} \cdot \widehat{h}(x)}{\widehat{h}\alpha(x)}\right), \quad \Psi(x) := \tfrac{1}{6} \sum_j \psi S_j(x), \quad \Psi_n(x) := \mathbb{E}\Psi x_n(x),$$

where $\psi(\infty)$ is defined as $\lim_{|x|\to\infty} \psi(x) = \ln \frac{3}{2}$. We call Ψ the *shape integrand* and Ψ_n the *shape integrand of order n*.

Let us restate parts of Definition 3.14 (on page 26).

Definition 11.2 Let $n \in \mathbb{N}_0$ and $J_n := (j_1, \ldots, j_n)$ be a random multiindex. We define (random) functions x_n, r_n and a (non-random) function Φ_n:

(i) $x_n : [0,1] \to [0,1] : \; x \mapsto a_{J_n}(x)$,

(ii) $r_n : [0,1] \to (0,\infty) : \; x \mapsto |x_n'(x)|$,

(iii) $\Phi_n : [0,1] \to \mathbb{R} : \; x \mapsto \mathbb{E} \ln r_n(x)$.

Note that $x_0(x) = x$ and $r_0(x) = 1$ and $\Phi_0(x) = 0$.

We remind ourselves of the chain rule $|(ab)'(x)| = |a'b(x)| \cdot |b'(x)|$ in Lemma 7.3.(i) (on page 56) for up-to-sign differentiable functions a, b. Let us restate and prove Proposition 4.15 (on page 34).

Proposition 11.3 For any $n \in \mathbb{N}_0$ and $x \in [0,1]$ we have the following approximation of $\Phi_n(x)$ by the shape integrands of orders up to $n-1$:

$$\left|\Phi_n(x) - \sum_{t=0}^{n-1} \Psi_t(x)\right| \leq \ln \tfrac{4}{3}.$$

Proof. For $n \in \mathbb{N}_0$ we set $H_n(x) := \mathbb{E} \ln \widehat{h} x_n(x)$. For $x \in [0,1]$ we notice that $\widehat{h}(x) = x^2 + 3 \in [3,4]$ and therefore $H_n(x) \in [\ln 3, \ln 4]$. We are done if we can show that

$$\Phi_n(x) - \sum_{t=0}^{n-1} \Psi_t(x) = H_n(x) - H_0(x). \tag{11.2}$$

Let $n \in \mathbb{N}$. Conditioning on the last index j_n and applying the chain rule yield

$$\Phi_n(x) = \tfrac{1}{6} \sum_j \mathbb{E} \ln |(a_j x_{n-1})'(x)| = \mathbb{E} \tfrac{1}{6} \sum_j \ln |a_j' x_{n-1}(x)| + \mathbb{E} \ln r_{n-1}(x) \tag{11.3}$$

$$= \mathbb{E} \Phi_1 x_{n-1}(x) + \Phi_{n-1}(x). \tag{11.4}$$

Similarly, we see that $H_n(x) = \mathbb{E} \frac{1}{6} \sum_j \ln \widehat{h} a_j x_{n-1}(x) = \mathbb{E} H_1(x_{n-1}(x))$. We remind ourselves of $A_j = \alpha S_j$ and $a_j = S A_j$. By (11.1) we have

$$|S' A_j(x)| = \frac{\widehat{h} a_j(x)}{\widehat{h} A_j(x)}, \quad |S_j'(x)| = \frac{\widehat{h} S_j(x)}{\widehat{h}(x)}. \tag{11.5}$$

Together with the chain rule $|a_j'(x)| = |S' A_j(x)| \cdot \frac{2}{3} \cdot |S_j'(x)| = \frac{2}{3} \cdot \widehat{h} S_j(x) / \widehat{h} \alpha S_j(x) \cdot \widehat{h} a_j(x) / \widehat{h}(x)$ it follows that

$$\ln |a_j'(x)| = \psi S_j(x) + \ln \widehat{h} a_j(x) - \ln \widehat{h}(x). \tag{11.6}$$

Interpreting j as a random index uniformly distributed on $\mathcal{J}$ and taking the expectation leads to $\Phi_1(x) = \Psi(x) + H_1(x) - \ln \widehat{h}(x)$. Plugging in $x_{n-1}(x)$ for x, taking the expectation and applying (11.3), (11.4) leads to

$$\Phi_n(x) - \Phi_{n-1}(x) = \mathbb{E} \Phi_1 x_{n-1}(x) = \Psi_{n-1}(x) + H_n(x) - H_{n-1}(x). \tag{11.7}$$

Writing $\Phi_n(x) = \Phi_n(x) - \Phi_0(x)$ and $H_n(x) - H_0(x)$ as a telescopic sum yields (11.2). ∎

11.2 Extrema of the shape integrand

Lemma 11.4 The shape integrand is invariant under the application of any shape function, that is

$$\Psi S_j(x) = \Psi(x), \quad x \in \overline{\mathbb{R}}.$$

Proof. The lemma is an immediate consequence of $\{S_{j'}S_j : j' \in \mathcal{J}\} = \{S_{j'} : j' \in \mathcal{J}\}$, see Lemma 6.8. ∎

Proposition 11.5 The shape integrand has the following extrema:

$$\max_{x\in[0,1]} \Psi(x) = \Psi(0) = -\chi_{\min}, \qquad \min_{x\in[0,1]} \Psi(x) = \Psi(1) = -\chi_{\max}$$

Proof. A simple computation verifies

$$-\Psi(0) = \tfrac{1}{3}\ln\tfrac{3}{2} = \chi_{\min} < \chi_{\max} = \tfrac{1}{3}\ln\tfrac{91}{2} - \ln 3 = -\Psi(1). \tag{11.8}$$

We show that Ψ is decreasing on $[0,1]$. Since Ψ is differentiable on $[0,1]$ it suffices to show that Ψ has no critical points in $(0,1)$.

In the following we view the six functions S_j as functions $\overline{\mathbb{R}} \to \overline{\mathbb{R}}$. Lemma 11.4 states the invariance $\Psi S_j = \Psi$ and implies $\Psi' S_j(x) \cdot S_j'(x) = \Psi'(x)$ for $x \in \mathbb{R}$. Notice that $S_j'(x) \neq 0$. So if x is a critical point of Ψ, then $S_j(x)$ is critical, too.

Step 1: We show that $-3, -1, 0, 1, 3$ are critical points of Ψ. Since $-3 = S_{(0,1)}(0)$ and $3 = S_{(0,2)}(0)$ and $-1 = S_{(0,1)}(1)$ it suffices to show that 0 and 1 are critical points of Ψ. The invariance $\Psi m = \Psi$ means that Ψ is even and therefore Ψ' is odd. Notice that $h(0) = -h^2(0)$ and $h(1) = -1$. Thus

$$\Psi'(h(0)) + \Psi'(h^2(0)) = 0, \quad \Psi'(h(1)) + \Psi'(1) = 0. \tag{11.9}$$

We multiply the first equation by $h'(0) = (h^2)'(0)$ and the second one by $h'(1) = 1$. Then the chain rule applied to Ψh^k and the invariance $\Psi h^k = \Psi$ yield

$$2\Psi'(0) = 0, \quad 2\Psi'(1) = 0. \tag{11.10}$$

Step 2: We show that Ψ has at most 11 critical points. Let $Q(x) := \frac{9}{4}\widehat{h}\alpha(x) = x^2 - 3x + 9$ and $P(x) := x^2 - 4x - 3$. We compute

$$\psi'(x) = \frac{2x}{\widehat{h}(x)} - \frac{3\alpha(x)}{Q(x)} = \frac{-3P(x)}{\widehat{h}(x)Q(x)}. \tag{11.11}$$

Let us define the even function $\psi_\pm(x) := \psi(x) + \psi(-x)$, the even polynomial $Q_0(x) := Q(x) \cdot Q(-x)$ of degree 4 and the odd polynomial $P_0(x) := P(x)Q(-x) - P(-x)Q(x)$. Since $\widehat{h}$ is even we have

$$\psi'_\pm(x) \cdot \widehat{h}(x) = \psi'(x) \cdot \widehat{h}(x) - \psi'(-x) \cdot \widehat{h}(-x) = \frac{P(x)}{Q(x)} - \frac{P(-x)}{Q(-x)} = \frac{P_0(x)}{Q_0(x)}. \tag{11.12}$$

We have $Q_0(x) = (x^2+9)^2 - (3x)^2$ and $P(x)Q(-x) = c_4x^4 - x^3 + c_2x^2 - 45x + c_0$ for certain (irrelevant) coefficients c_i. Hence

$$Q_0(x) = x^4 + 3^2x^2 + 3^4, \quad P_0(x) = -3x^3 - 5 \cdot 3^3x. \tag{11.13}$$

Since $6\Psi(x) = \sum_k \psi_\pm h^k(x)$ we can write

$$6\Psi'(x) \cdot \widehat{h}(x) = \sum_k \varphi_k(x), \quad \varphi_k(x) := (\psi_\pm h^k)'(x) \cdot \widehat{h}(x). \tag{11.14}$$

Lemma 7.6 (more exactly equation (7.1) in its proof) states that $(h^k)'(x) \cdot \widehat{h}(x) = \widehat{h}h^k(x)$. Thus

$$\varphi_k(x) = \psi'_\pm h^k(x) \cdot \widehat{h}h^k(x) = \frac{P_0h^k(x)}{Q_0h^k(x)}. \tag{11.15}$$

For $k \in \{-1, 1\}$ the formulas $h(x) = (x-3)/(x+1)$ and $h^{-1}(x) = -(x+3)/(x-1)$ show that

$$Q_k(x) := Q_0h^k(x) \cdot (x+k)^4, \quad P_k(x) := P_0h^k(x) \cdot (x+k)^4 \tag{11.16}$$

are polynomials of degree 4. We compute

$$Q_1(x) = (x-3)^4 + 3^2(x-3)^2(x+1)^2 + 3^4(x+1)^4 \tag{11.17}$$

$$= 91x^4 + 276x^3 + 58 \cdot 3^2x^2 + 4 \cdot 3^4x + 3^5, \tag{11.18}$$

$$P_1(x) = -3(x-3)^3(x+1) - 5 \cdot 3^3(x-3)(x+1)^3 \tag{11.19}$$

$$= -138x^4 + 24x^3 + 28 \cdot 3^2x^2 + 40 \cdot 3^2x - 2 \cdot 3^5. \tag{11.20}$$

Since $h^{-1} = mhm$ and Q_0 is even and P_0 is odd we see that

$$Q_{-1}(x) = Q_1(-x), \quad P_{-1}(x) = -P_1(-x). \tag{11.21}$$

By (11.14) we can write

$$6\Psi'(x) \cdot \widehat{h}(x) = \sum_k \frac{P_k(x)}{Q_k(x)} = \frac{P^*(x)}{Q^*(x)} \tag{11.22}$$

with the polynomials

$$Q^*(x) := \prod_k Q_k(x), \quad Q_k^*(x) := \frac{Q^*(x)}{Q_k(x)}, \quad P^*(x) := \sum_k P_k(x) Q_k^*(x). \tag{11.23}$$

Notice that Q^* has degree 12 and Q_k^* has degree 8 and P^* has at most degree 12. We see by (11.21) that Q_0^* is even, so $Q^* = Q_0^* \cdot Q_0$ is also even. It follows that $P^* = 6\Psi' \cdot \widehat{h} \cdot Q^*$ is odd because Ψ' is odd and $\widehat{h}$ is even. So the degree of P^* is at most 11. Notice that any critical point of Ψ is a root of P^*.

Step 3: We show that $-3, -1, 0, 1, 3$ are the only critical points of Ψ. Assume that Ψ has a critical point $y \in \mathbb{R} \setminus \{-3, -1, 0, 1, 3\} = \Sigma^*$. As mentioned before Step 1 this implies that each point $S_j(y)$ is also critical. These six critical points $S_j(y)$ are pairwise different and each of them lies in Σ^*. So there exist at least 11 critical points. We saw in Step 2 that P^* has at most degree 11. Hence P^* has degree 11 and 11 simple roots. Exactly 5 of the roots are positive, because 0 is a root and P^* is odd. So P^* changes the sign exactly 5 times on $(0, \infty)$. Now we show that

1. $P^*(x)$ converges to $+\infty$ for $x \to +\infty$,
2. $(P^*)'(0) > 0$.

Since P^* changes the sign an odd number of times on $(0, \infty)$, part 1 implies that P^* is negative on some interval $(0, \epsilon)$ in contradiction to part 2 and $P^*(0) = 0$.

It suffices to show that both the x^{11} and the x^1 coefficient of P^* are positive. We compute

$$Q_0^*(x) = 91^2 x^8 + c_6 x^6 + c_4 x^4 + c_2 x^2 + 3^{10}, \tag{11.24}$$

$$Q_{-1}^*(x) = 91x^8 + 276x^7 + c_6' x^6 + c_5' x^5 + c_4' x^4 + c_3' x^3 + c_2' x^2 + 4 \cdot 3^8 x + 3^9 \tag{11.25}$$

for certain (irrelevant) coefficients c_i, c_i'. Notice that $Q_1^*(x) = Q_{-1}^*(-x)$ due to (11.21). Simple computations show that the x^{11} coefficient of $P_0 \cdot Q_0^*$ is $-3 \cdot 91^2$ and the x^1 coefficient is $-5 \cdot 3^{13}$. The x^{11} coefficient of $P_1 \cdot Q_1^*$ is $40\,272$ and the x^1 coefficient is $112 \cdot 3^{11}$. As a consequence of $P_{-1}(x) Q_{-1}^*(x) = -P_1(-x) Q_1^*(-x)$, see (11.21), the x^{11} and x^1 coefficients of $P_{-1} Q_{-1}^*$ and $P_1 Q_1^*$ coincide. In combination, we see that the x^{11} coefficient of P^* is $-3 \cdot 91^2 + 2 \cdot 40\,272 > 0$ and the x^1 coefficient is $-5 \cdot 3^{13} + 2 \cdot 112 \cdot 3^{11} > 0$. ∎

12 Conclusion and conjecture

12.1 Conclusion

Consider a Markov chain of triangles each of which is uniformly chosen amongst the children of the previous triangle. The sequence of the shapes of these triangles is also a Markov chain and we call it a shape chain.

This thesis provides a new description of shape chains by expressing them in upcircle coordinates. Simply speaking, instead of starting with a single point (which represents its associated triangle), we can start with a whole upcircle. Applying the same (random) Möbius transformations as for a certain shape chain results in a Markov chain of upcircles.

The new description of shape chains in upcircle coordinates allows for the application of our modified version of the law of large numbers. The original version was proved in [6] where it was not used for the proof of the exponentially fast "flattening" of the non-flat shape chain and where a technical condition, namely contractivity on average in step 2, was verified via numerical computations. We show the contractivity on average in some step N without numerical computations by reducing the problem to the analysis of a one-dimensional function, the shape integrand.

A byproduct of our analysis of the shape integrand are the bounds $\chi_{\min}$ and $\chi_{\max}$ for the universal constant $\chi := -\frac{1}{6}\sum_j \int \ln|a_j'|\, d\mu > 0$, where $|a_j'|$ is the up-to-sign derivative of the child function a_j and μ is the invariant measure of the flat shape chain.

Our second theorem considers the same coupling (X_n, Z_n) of a flat and a non-flat shape chain as [6] and presents the new result that the radius R_n of the unique upcircle containing X_n and Z_n decays (asymptotically) with the exact exponential rate χ from above. From here, we concluded that $Z_n - X_n$ and $\Im Z_n = \mathrm{flat}(\Delta^{(n)})$ decay exactly with rate χ, too.

The latter result about $\Im Z_n$ shows that the (upper) Lyapunov constant of the associated dynamical system (investigated in [1], [9], [10]) is $\frac{1}{2}\chi$ and therefore has the bounds $\frac{1}{2}\chi_{\min}$ and $\frac{1}{2}\chi_{\max}$ which are sharper than the previously known bounds.

12.2 Conjecture

Theorem 1.2 in [6] states that the largest angle of $\Delta^{(n)}$ converges to π in probability. One can show that the a.s. convergence follows from our **conjecture**:

$$\lim_n \frac{1}{n} \ln(1 - X_n) = 0 \quad \text{a.s.} \tag{12.1}$$

It suffices to prove that the limit $\lim_n \frac{1}{n} \ln(1 - X_n)$ exists a.s. because this limit can neither be positive (that would mean that X_n is unbounded), nor negative (because then the limit set of X_n would a.s. be $\{1\}$ instead of $[0, 1]$ as stated in Theorem 1.3 in [6]).

Let $\lambda \in [0, 2]$. For a flat triangle Δ we set $\text{flat}_\lambda := 0$. For any other triangle Δ we define

$$\text{flat}_\lambda(\Delta) := \frac{(\text{shortest height in } \Delta)^{2-\lambda} \cdot (\text{longest height in } \Delta)^\lambda}{\text{area of } \Delta}. \tag{12.2}$$

In particular $\text{flat}_0 = \text{flat}$. We notice that $\text{flat}_\lambda(\Delta)$ is non-decreasing in λ. According to the formula for the area of a triangle, a shortest height is orthogonal to a longest edge and a longest height is orthogonal to a shortest edge. Hence

$$\text{flat}_1(\Delta) = 2 \cdot \frac{\text{shortest height in } \Delta}{\text{shortest edge in } \Delta},$$
$$\text{flat}_2(\Delta) = 2 \cdot \frac{\text{longest height in } \Delta}{\text{shortest edge in } \Delta}.$$

With this representation of flat_1 we can easily verify that $\frac{1}{2}\text{flat}_1(\Delta)$ is the sine of the second-largest angle of Δ (or largest angle in case of several largest angles). We set $\text{flat}_\lambda(z) := \text{flat}_\lambda(\Delta_z)$ and say that a sequence of triangles or points λ-*flattens* if their flat_λ-values converge to 0. So a sequence of triangles 1-flattens if and only if the largest angle converges to π. One can show that the conjecture (12.1) does not only imply that Z_n 1-flattens, but even that $\text{flat}_\lambda(Z_n)$ decays exactly with the same rate χ for each $\lambda \in [0, 2]$.

The following two examples show that 2-flattening is strictly stronger than 1-flattening and 1-flattening is strictly stronger than 0-flattening. The sequence $n + i$ does not 2-flatten, but it 1-flattens. The sequence of isosceles points $-1 + 2\exp(i/n)$ does not 1-flatten, but it 0-flattens.

If we could show that $\phi : [0, 1] \to \mathbb{R} : x \mapsto \ln(1 - x)$ is integrable w.r.t. μ and the law of large numbers (1.13) still holds, then our conjecture (12.1) would follow.

References

[1] I. Bárány, A.F. Beardon, and T.K. Carne. Barycentric subdivision of triangles and semigroups of Möbius maps. *Mathematika,* **43**:165-171, (1996).

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

[3] M. Begue, D.J. Kelleher, A. Nelson, H. Panzo, R. Pellico, and A. Teplyaev. Random walks on barycentric subdivisions and the Strichartz hexacarpet. *Exp. Math.,* **21**(4):402-417, (2012).

[4] S. Butler, and R. Graham. (2010) Iterated Triangle Partitions. In: G.O.H. Katona, A. Schrijver, T. Szőnyi, and G. Sági (eds). Fete of Combinatorics and Computer Science. *Bolyai Soc. Math. Stud.,* **20**:23-42, (2010).

[5] B. Casselman. The joy of barycentric subdivision. *Amer. Math. Soc. Feature column,* (electronic), (June, 2017).

[6] P. Diaconis, and L. Miclo. On Barycentric Subdivision. *Combin. Probab. Comput.,* **20**(2):213-237, (2011).

[7] P. Diaconis, and L. Miclo. On barycentric subdivision, with simulations, (2010).

[8] H. Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.,* **108**:377-428 (1963).

[9] B. Hough. Tessellation of a triangle by repeated barycentric subdivision. *Elect. Commun. in Probab.,* **14**:270-277, (2009).

[10] A. Wilkinson. What are Lyapunov exponents, and why are they interesting? *Amer. Math. Soc. Bulletin,* **54**:79-105 (electronic), (2016).

Appendix

A Asymptotic exponential rates

Lemma A.1 Let w_n, w'_n be sequences in $\overline{\mathbb{C}}$ and κ,κ' be real constants.

(i) w_n has exactly rate κ if and only if w_n has at most and at least rate κ.

(ii) w_n has at most (or at least) rate κ, if for any $\kappa' > \kappa$ (or $\kappa' < \kappa$) there is some $n_0 \in \mathbb{N}$ such that any $n \geq n_0$ satisfies

$$|w_n| \leq e^{\kappa' n} \quad \text{or} \quad |w_n| \geq e^{\kappa' n}, \quad \text{resp.}$$

(iii) w_n has at most rate κ if and only if w_n^{-1} has at least rate $-\kappa$.

(iv) If w_n has at most rate κ and $p \in (0, \infty)$, then w_n^p has at most rate $p \cdot \kappa$. The same is true with "at most" replaced by "at least" or "exactly".

(v) If w_n has at most rate $\kappa < 0$, then w_n converges to 0. If w_n has at least some rate $\kappa > 0$, then $|w_n|$ converges to ∞.

(vi) If w_n is bounded, then w_n has at most rate 0. If w_n is bounded away from 0, then w_n has at least rate 0. In particular, if w_n converges to a non-zero complex number, then w_n has exactly rate 0.

(vii) If w_n has at most rate κ and w'_n has at most rate κ', then $w_n \cdot w'_n$ has at most rate $\kappa + \kappa'$. The same is true with "at most" replaced by "at least" or "exactly".

(viii) If w_n has at most rate κ and w'_n has at most rate κ', then $w_n + w'_n$ has at most rate $\max\{\kappa, \kappa'\}$.

Proof. The parts (i), (ii), (iv), (vi), (vii) are clear and the parts (iii), (v) follow from (ii).

(viii): For any $a, b \in \mathbb{C}$ we know $|a + b| \leq 2 \max\{|a|, |b|\}$ and therefore

$$\ln |a + b| \leq \ln 2 + \max\{\ln |a|, \ln |b|\}. \tag{A.1}$$

We plug in $a = w_n$ and $b = w'_n$, multiply by $\frac{1}{n}$, take $\limsup_n$ and use that the limit superior of the sum or the maximum of two real sequences is at most the sum or the maximum of the limits inferior of those two sequences. This shows (viii). ∎