Comment. Math. Helv. 98 (2023), 41–134
DOI 10.4171/CMH/548

Convergence of spherical averages for actions of Fuchsian groups

Alexander I. Bufetov, Alexey Klimenko, and Caroline Series

Abstract. We prove pointwise convergence of spherical averages for a measure-preserving action of a Fuchsian group. The proof is based on a new variant of the Bowen–Series symbolic coding for Fuchsian groups that, developing a method introduced by Wroten, simultaneously encodes all possible shortest paths representing a given group element. The resulting coding is self-inverse, giving a reversible Markov chain to which methods previously introduced by the first author for the case of free groups may be applied.

1. Introduction

1.1. Formulation of the main result

Let G be a finitely generated group with a symmetric set of generators G_0. For $g \in G$, denote by $|g|$ the length of the shortest word in G_0 representing g. Let $S(n)$ be the sphere of radius n in G:

$$S(n) = \{g \in G : |g| = n\}.$$

Suppose that G acts on a probability space (X, μ) by measure-preserving transformations T_g, $g \in G$. For a function $f \in L^1(X, \mu)$, consider spherical averages

$$\mathbf{S}_n(f) = \frac{1}{\#S(n)} \sum_{g \in S(n)} f \circ T_g. \tag{1}$$

The main result of this paper, Theorem A below, gives the almost sure convergence of spherical averages for measure-preserving actions of Fuchsian groups and for $f \in L \log L(X, \mu)$, that is, whenever

$$\int |f| \log^+ |f| d\mu < \infty,$$

where $\log^+ |f| = \max(0, \log |f|)$.

2020 Mathematics Subject Classification. Primary 22D40; Secondary 20H10, 37A30.
Keywords. Ergodic theorem, pointwise convergence, Fuchsian group, spherical averages.

Let G be a Fuchsian group and let $\mathcal{R}$ be a fundamental domain for G. Assume that the sides of $\mathcal{R}$ are paired by a set of elements $G_0 \subset G$. As is well known, G_0 is a symmetric set of generators for G. The images of $\mathcal{R}$ under the action of G induce a tessellation $\mathbf{T}_{\mathcal{R}} = \{g\mathcal{R} : g \in G\}$ of the hyperbolic disk $\mathbb{D}$. Following [10], we say that $\mathcal{R}$ has *even corners* if the geodesic extension of every side of $\mathcal{R}$ is entirely contained in $\mathbf{T}_{\mathcal{R}}$, more precisely in the union of boundaries of all domains $g\mathcal{R} \in \mathbf{T}_{\mathcal{R}}$.

Let $v \in \mathbb{D}$ be a vertex of $\mathbf{T}_{\mathcal{R}}$. If $\mathcal{R}$ has even corners, then the boundary of $\mathbf{T}_{\mathcal{R}}$ in a small neighborhood of v consists of n geodesic segments intersecting at v and dividing our neighborhood into $2n$ sectors. Write $n = n(v)$ and let $N(\mathcal{R})$ denote the number of sides of $\mathcal{R}$ inside $\mathbb{D}$. We need the following assumption on $\mathcal{R}$.

Assumption 1.1. *The assumption is in two parts.*

(i) *$\mathcal{R}$ has even corners.*

(ii) *One of the following conditions holds for $\mathcal{R}$:*

- $N(\mathcal{R}) \geq 5$,
- $N(\mathcal{R}) = 4$ *and either $\mathcal{R}$ is non-compact or $\mathcal{R}$ is compact and does not have two opposite vertices v, v' such that $n(v) = n(v') = 2$,*
- $N(\mathcal{R}) = 3$ *and $\mathcal{R}$ is non-compact.*

Our main result is the following:

Theorem A. *Let G be a non-elementary Fuchsian group G and let $\mathcal{R}$ be its fundamental domain with side-pairing transformations G_0 and satisfying Assumption 1.1. Let G act on a Lebesgue probability space (X, μ) by measure-preserving transformations. Denote by $\mathcal{I}_{G_0^2}$ the sigma-algebra of sets invariant under all maps $T_{g_1 g_2}$, $g_1, g_2 \in G_0$. Then, for any function $f \in L \log L(X, \mu)$, as $n \to \infty$, we have*

$$\mathbf{S}_{2n}(f) \to \mathsf{E}(f|\mathcal{I}_{G_0^2}) \quad \textit{almost surely and in } L^1.$$

The condition that $\mathcal{R}$ has even corners is not as restrictive as it appears. In fact, it is clear that our result only depends on the generators G_0 and the coding, and not on the precise geometry of $\mathcal{R}$. Thus Theorem A extends immediately to any presentation of a Fuchsian group for which one can find deformed group G' which has a fundamental domain $\mathcal{R}'$ with the same pattern of sides and side-pairings and even corners, see [10, 20] and [48] for a detailed discussion. The need to restrict to spheres of even radius can be seen by considering the action of the free group F_2 on the two-element set $\{0, 1\}$ in which both generators of F_2 act by interchanging the elements, in which case the value of $\mathbf{S}_n(f)$ depends on the parity of n.

We note that the conditions of Assumption 1.1 are not quite identical with those in [3,46] and elsewhere, the main difference being the weaker restriction if $N(\mathcal{R}) = 4$.

In fact, all results of those papers should apply under these somewhat weaker assumptions.

The Cesàro convergence of the averages $\mathbf{S}_{2n}(f)$ is proven by Bufetov and Series in [20] using the Bowen–Series Markovian coding [10] (see also [3, 46, 48]) in order to reduce the statement to the ergodic theorem for Markov operators, cf. [12, 13]. The Bowen–Series coding allows one to assign states to group generators in a suitable product representing an arbitrary group element as a shortest word in such a way that the admissible sequences of states form a Markov chain. This gives rise to a Markov operator as described in [20]. However the proof in [14], which establishes convergence of the spherical averages themselves for free groups, does not extend in any obvious way. This is because the argument of [14] relies on a symmetry condition for the coding, namely that the coding is reversible or self-inverse, which allows one to relate the Markov operator generated by the coding to its adjoint. The Bowen–Series coding of [10] fails to be symmetric in this sense.

The main construction of this paper is a new self-inverse coding for Fuchsian groups, which allows us to adapt the proof in [14] to this new case.

This new coding is constructed using a variant of that introduced by Matthew Wroten [50], see also a related idea in [23] and [47]. Wroten's idea is to encode all possible representations of a group element as a shortest word simultaneously. This involves assigning states to all possible ways of building up shortest words step by step. The set of states together with allowed transitions defines a Markov chain with the property that the transition rules, that is the set of all admissible paths, can be inverted. From this we construct an associated Markov operator with the required symmetry condition on its adjoint, and then derive a suitably modified version of the convergence theorem in [14].

It would be interesting to obtain a similar coding for a more general hyperbolic groups. In particular, it is not clear to us how to invert paths in the classical Cannon–Gromov coding [22, 31].

To explain the ideas in a bit more detail, let us briefly describe Wroten's approach in our setting. Every shortest word in the Fuchsian group G corresponds to a shortest path in the Cayley graph of G relative to the given generators G_0. This graph is embedded in $\mathbb{D}$ by sending $g \in G$ to $gO \in \mathbb{D}$, where O is some fixed base point in int $\mathcal{R}$. Vertices gO, hO are joined by an edge if and only if $g^{-1}h \in G_0$. If β is a shortest path in the Cayley graph, we refer to the sequence of domains traversed by the edges of β also as a shortest path. If $g \in G$ then the *thickened path* $[g]$ associated to g is by definition the collection of all those $h\mathcal{R}, h \in G$ which are traversed by some shortest path from $\mathcal{R}$ to $g\mathcal{R}$. Every domain $h\mathcal{R} \in [g]$ is endowed with an *index*, which equals the distance in the Cayley graph from $\mathcal{R}$ to $h\mathcal{R}$. The set of all domains with index k we will refer to as a *level* of $[g]$ and denote by $[g]_k$.

The coding works as follows. We will define a space of states $\Xi = \{X_1, \ldots, X_k\}$ and a $\Xi \times \Xi$ transition matrix $\Pi = (\Pi_{ij})$ such that $\Pi_{ij} = 1$ if transition from X_i to X_j is possible and $\Pi_{ij} = 0$ otherwise. There is a subset $\Xi_S \subset \Xi$ of start states, and another subset $\Xi_F \subset \Xi$ of end states.

The states in Ξ represent how $[g]_k$ and $[g]_{k+1}$ are attached to each other. It turns out that every $[g]_k$ contains at most two fundamental domains and the domains from $[g]_{k+1}$ are glued to the ones from $[g]_k$ across one, two or three sides, see Figure 6. We endow this geometrical configuration with some additional data to obtain a Markov chain generating thickened paths; in particular, the data records the generators needed to carry out the gluing. Then we define a transition matrix Π and subsets Ξ_S and Ξ_F of Ξ and prove that thickened paths from $\mathcal{R}$ to $g\mathcal{R}$ with $|g| = n$ are in one-to-one correspondence with admissible sequences of length n starting in Ξ_S and ending in Ξ_F. The required reversibility or self-inverse symmetry condition follows since inverting a thickened path yields a thickened path and the coding preserves this symmetry.

In terms of the associated Markov operators, this symmetry property can be expressed as follows. We introduce two maps $\gamma, \omega \colon \Xi \to G$, closely related to the attaching maps between $[g]_k$ and $[g]_{k+1}$, see Section 7. These maps satisfy certain relations, see Lemma 7.1. Following [14], we then construct Markov operators P and U on $L^1(X \times \Xi)$, which as a consequence of these relations satisfy

$$P^* = UPU, \quad U^* = U^{-1} = U.$$

Then we can apply the Alternierende Verfahren method in a manner similar to [14].

For this application we need an inequality between P^n and $(P^*)^k P^k$, which is the basis for the maximal inequality in the Alternierende Verfahren scheme. For free groups this inequality was

$$cUP^{2n-1}\varphi \leq (P^*)^n P^n \varphi$$

for any non-negative φ, see [14]. In the present case, the inequality becomes more complicated, both because the index on the right-hand side may vary slightly and also because there are a small number of possible sequences for which the required geometrical statements fail. To correct this, terms on the right-hand side of the inequality have to be summed over a small bounded interval of indices near n, and the inequality also contains an error term $A_n\varphi$, see (14) in Section 8 below. The proof of the geometrical statement associated to the proof of this inequality, Lemma 8.12, is one of the most technically complicated parts of the paper.

A short announcement of the results of this paper with a more detailed outline of the coding can be found in [19]. The purpose of the present paper is to provide detailed proofs and the reader may well find it helpful to look at [19] first before becoming involved in the details explained here.

1.2. Organization of the paper

The paper is organized as follows. In the next section we give some notation and preliminaries regarding Fuchsian groups and their fundamental domains. In particular, we show in Lemma 2.1 that under Assumption 1.1 three geodesic lines in the boundary of the tessellation $\mathbf{T}_{\mathcal{R}}$ cannot form a triangle.

Section 3 deals with the local structure of the thickened paths. Namely, we show that each thickened path is split into *bottles* by levels (*bottlenecks*) which contain only one copy of the fundamental domain, and the structure of each bottle is then described by Lemma 3.9. The local rules from this description give rise to the construction of the Markov chain in Section 4. The main result of this section, Theorem 4.10, shows that every thickened path can be produced by this Markov chain, and conversely that every path defined by this chain is indeed a thickened path. This result should be of independent interest and may have applications elsewhere.

In Section 5 we present some techniques for cutting and joining thickened paths which we use in Section 6 to show that the Markov chain is strongly connected and aperiodic or, in other words, its adjacency matrix Π has a power with all elements positive. The same techniques are also used in Section 8.

Section 7 shows that these properties of the Markov chain allow us to construct its Parry measure and then to relate the spherical averages for our group to powers of a Markov operator P associated to this coding. We also show that the symmetry of the coding yields a relation between P and its adjoint P^*.

Section 8 concludes the proof of the main theorem. To do this, we first formulate the new general theorem on pointwise convergence of powers of a Markov operator, Theorem 8.6. Most of Section 8 is then devoted to checking that the conditions necessary for this theorem apply in our case, including the most complicated one, that involving an inequality between the operator and its adjoint, as discussed above. This is proved in Section 8.3 using techniques from Section 5. The proof of Theorem A assuming Theorem 8.6 is concluded in Section 8.4.

Finally, in Section 9 we give the proof of the new general result, Theorem 8.6 on pointwise convergence for Markov operators. As discussed above, the argument here follows that in [14] and is based on Rota's "Alternierende Verfahren" scheme.

We remark that many of the proofs, especially in Sections 5 and 8.3, may seem rather long and complicated; this is partly because of the generality in which we are working. In many cases the situation with $N(\mathcal{R}) \geq 5$ simplifies considerably; on the other hand the cases $N(\mathcal{R}) = 3, 4$ simplify in different ways and $N(\mathcal{R}) = 3$ encompasses in particular the modular group $SL(2, \mathbb{Z})$. In almost all cases (to be precise, everywhere except in case (4) of Proposition 6.4), our proofs depend only on the geometry of R and not on analysing the particular pattern of side pairings.

1.3. Historical remarks

For two rotations of a sphere, convergence of spherical averages was established by Arnold and Krylov [1], and a general mean ergodic theorem for actions of free groups was proved by Guivarc'h [32].

The first general pointwise ergodic theorem for convolution averages on a countable group is due to Oseledets [42] who relied on the martingale convergence theorem. The first general pointwise ergodic theorems for free semigroups and groups were given by Grigorchuk in 1986 [28], where the main result is Cesàro convergence of spherical averages for measure-preserving actions of a free semigroup and group. Convergence of the actual spherical averages for free groups was established by Nevo [37] for functions in L^2, and Nevo and Stein [39] for functions in L^p, $p > 1$ using spectral theory methods. Nevo, Stein, and Margulis [36, 40] considered ball averages for actions of connected semisimple Lie group with finite center and no non-trivial compact factors and showed that these ball averages converge almost everywhere and in L^p, $p > 1$. Note that, as shown by Tao [49], whose argument is inspired by Ornstein's counterexample [41], pointwise convergence of spherical averages for functions in L^1 does not hold even for actions of free groups.

The method of Markov operators in the proof of ergodic theorems for actions of free semigroups and groups was suggested by Grigorchuk [29, 30], Thouvenot (oral communication), and in [12]. In [14] pointwise convergence is proved for Markovian spherical averages under the additional assumption that the Markov chain be reversible. The key step in [14] is the triviality of the tail sigma-algebra for the corresponding Markov operator; this is proved using Rota's "Alternierende Verfahren" [45], that is to say, martingale convergence. Another result in this direction was obtained in [5]; it states the mean convergence for analogues of spherical averages for an arbitrary Markov chain satisfying very mild conditions. It is not known whether similar result holds for pointwise convergence.

The study of Markovian averages is motivated by the problem of ergodic theorems for general countable groups, specifically, for groups admitting a Markovian coding such as Gromov hyperbolic groups [31] (see e.g. Ghys–de la Harpe [25] for a detailed discussion of the Markovian coding for Gromov hyperbolic groups). The first results on convergence of spherical averages for Gromov hyperbolic groups, obtained under strong exponential mixing assumptions on the action, are due to Fujiwara and Nevo [24]. For actions of hyperbolic groups on finite spaces, an ergodic theorem was obtained by Bowen in [4].

Cesàro convergence of spherical averages for all measure-preserving actions of Markov semigroups, and, in particular, Gromov hyperbolic groups, was established in [15, 16]; earlier partial results were obtained in [11, 13]. In the special case of hyperbolic groups a shorter proof of this theorem, using the method of Calegari and Fujiwara [21], was later given by Pollicott and Sharp [43]. Using the method of amenable equivalence relations, Bowen and Nevo [6–9] established ergodic theorems for "spherical shells" in Gromov hyperbolic groups. For further background see the surveys [17, 27, 38].

2. Definitions and notation

Let G be a finitely generated non-elementary Fuchsian group acting on the hyperbolic disk $\mathbb{D}$ with fundamental domain $\mathcal{R}$, which we take to be closed. We suppose $\mathcal{R}$ to be a finite-sided convex polygon with vertices contained in $\overline{\mathbb{D}} = \mathbb{D} \cup \partial\mathbb{D}$, such that the interior angle at each vertex is strictly less than π. By a *side* of $\mathcal{R}$ we mean the closure in $\mathbb{D}$ of the geodesic arc joining a pair of adjacent vertices. We allow the infinite area case in which some adjacent vertices on $\partial\mathbb{D}$ are joined by an arc contained in $\partial\mathbb{D}$; we do not count these arcs as sides of $\mathcal{R}$. Further we usually mean by *vertices* of $\mathcal{R}$ only vertices inside $\mathbb{D}$. Sometimes it is convenient to count as vertices also the side ends that belong to $\partial\mathbb{D}$, these instances will be specified explicitly. Two sides are called *adjacent* if they share a common vertex lying in $\mathbb{D}$. We refer to each image $g\mathcal{R}$ of $\mathcal{R}$ by an element $g \in G$ as either a *fundamental domain* or, for brevity, a *domain*.

We assume that the sides of $\mathcal{R}$ are paired; that is, for each side s of $\mathcal{R}$ there is a (unique) element $e \in G$ such that $e(s)$ is also a side of $\mathcal{R}$ and the domains $\mathcal{R}$ and $e(\mathcal{R})$ are adjacent along $e(s)$. Notice that this includes the possibility that $e(s) = s$, in which case e is elliptic of order 2 and the side s contains the fixed point of e in its interior. The condition that the vertex angle be strictly less than π excludes the possibility that the fixed point of e is counted as a vertex of $\mathcal{R}$. Since the element pairs the side to itself the possibility of more than one elliptic fixed point on one side is excluded, for the existence of two such points implies the existence of infinitely many contained in the one side. Note also that the treatment of order two elliptic fixed points in [3] and elsewhere is slightly different.

We denote by $\partial\mathcal{R}$ the union of the sides of $\mathcal{R}$, in other words, $\partial\mathcal{R}$ is the part of the boundary of $\mathcal{R}$ inside the disk $\mathbb{D}$. Each side of $\partial\mathcal{R}$ is assigned two labels, one interior to $\mathcal{R}$ and one exterior, in such a way that the interior and exterior labels are mutually inverse elements of G. We label the side $s \subset \partial R$ interior to $\mathcal{R}$ by e if e carries s to another side $e(s)$ of $\mathcal{R}$, while we label the same side exterior to $\mathcal{R}$ by e^{-1}, see Figure 1. With this convention, $\mathcal{R}$ and $e^{-1}(\mathcal{R})$ are adjacent along the side s whose *interior* label is e, while the side $e(s)$ has interior label e^{-1}.

Let G_0 denote the set of group elements which label sides of $\mathcal{R}$. The labelling extends to a G-invariant labelling of all sides of the tessellation $\mathbf{T}_{\mathcal{R}}$ of $\mathbb{D}$ by images of $\mathcal{R}$, where by a *side* of $\mathbf{T}_{\mathcal{R}}$, we mean a side of $g\mathcal{R}$ for some $g \in G$. The conventions have been chosen in such a way that if two domains $g\mathcal{R}, h\mathcal{R}$ are adjacent along a

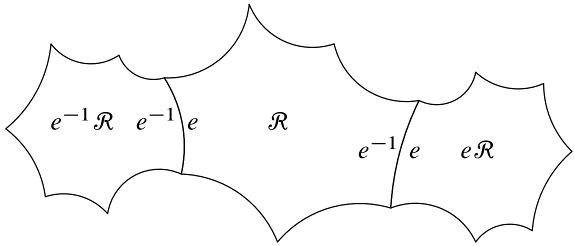


Figure 1. Labelling the sides of the fundamental domain $\mathcal{R}$. Note that the label e is interior to $\mathcal{R}$ on the side of $\mathcal{R}$ adjacent to the domain $e^{-1}\mathcal{R}$.

common side s, then $h^{-1}g \in G_0$ and the label on s interior to $g\mathcal{R}$ is $h^{-1}g$, while that on the side interior to $h\mathcal{R}$ is $g^{-1}h$. Suppose that O is a fixed basepoint in $\mathcal{R}$ and that γ is an oriented path in $\mathbb{D}$ from O to gO, $g \in G$, which avoids all vertices of $\mathbf{T}_{\mathcal{R}}$, and which passes through in order adjacent domains

$$\mathcal{R} = g_0\mathcal{R}, g_1\mathcal{R}, \ldots, g_n\mathcal{R} = g\mathcal{R}.$$

Then the labels of the sides crossed by γ, read in such a way that if γ crosses from $g_{i-1}\mathcal{R}$ into $g_i\mathcal{R}$ we read off the label $e_i = g_{i-1}^{-1}g_i$ of the common side interior to $g_i\mathcal{R}$, are in order $e_1, e_2, \ldots, e_n$ so that $g = e_1e_2\ldots e_n$. This proves the well-known fact that G_0 generates G; see, for example, [2].

As explained in the introduction, the fundamental domain $\mathcal{R}$ is said to have *even corners* if for each side s of $\mathcal{R}$, the complete geodesic in $\mathbb{D}$ which extends s is contained in the sides of $\mathbf{T}_{\mathcal{R}}$. This condition is satisfied for example, by the regular $4g$-gon of interior angle $\pi/2g$ whose sides can be paired with the standard generating set

$$\left\{a_i, b_i,\ i = 1, \ldots, g \ \Big|\ \prod_{i=1}^{g}[a_i, b_i]\right\}$$

to form a surface of genus g. It is also satisfied by the modular group $SL(2,\mathbb{Z})$ with the classical fundamental domain $\{z : |\Re z| < 1/2, |z| > 1\}$ in the upper half plane. For further discussion on the even corners condition, see the references in the introduction.

Note that under the even corners condition there exists a “chequered coloring” of the domains in $\mathbf{T}_{\mathcal{R}}$ (or elements of G): one can color each domain either in black or white in such a way that each side of $\mathbf{T}_{\mathcal{R}}$ separates domains of different color.

We will frequently consider the union (or the collection) of all $2n(v)$ domains in $\mathbf{T}_{\mathcal{R}}$ adjacent to a vertex v. We call this the *flower* at v and denote it by $\mathcal{F}_v$ and refer to the individual domains in $\mathcal{F}_v$ as *petals*, while the sides between its petals we call its *radii*. Note that $\mathcal{F}_v$ is a convex polygon. Indeed, it is a star domain (where this use of domain is not to be confused with our usual convention that a domain is a copy of the fundamental region $\mathcal{R}$) with respect to v and the internal angle at any vertex u on its boundary contains either one or two sectors; since $n(u) \geq 2$, this angle does not exceed π. Moreover, the angle π may occur only at the common vertex w of two petals of the flower, and in this case $n(w) = 2$.

Let us also denote the geodesic line passing through a side s or a pair of vertices u, v in $\mathbf{T}_{\mathcal{R}}$ as $\ell(s)$ or $\ell(uv)$.

We start with some properties of the tessellation $\mathbf{T}_{\mathcal{R}}$ which are consequences of Assumption 1.1.

Lemma 2.1. *Under Assumption 1.1, there are no vertices a, b, c of $\mathbf{T}_{\mathcal{R}}$ such that the lines $\ell(ab)$, $\ell(bc)$, and $\ell(ca)$ belong to $\partial\mathbf{T}_{\mathcal{R}}$.*

Proof. Assume the contrary: there exists a triangle $\Delta = abc$ in $\partial\mathbf{T}_{\mathcal{R}}$. Note that Δ cannot be a fundamental domain since Assumption 1.1 excludes compact triangular domains. Therefore, on $\partial\Delta$ there is a point p belonging to at least two fundamental domains in Δ. Then there is a ray α in $\partial\mathbf{T}_{\mathcal{R}}$ that starts at p and goes inside Δ. The ray α cuts Δ into two regions, at least one of them being triangular. Choose this region as a new triangle Δ', which also violates the statement of the lemma.

This process can be repeated indefinitely, and each iteration decreases the number of fundamental domains inside the triangle. However, this number is finite since the area of the triangle is finite, and we arrive at a contradiction. ■

The next proposition was stated under slightly stronger assumptions in [10, Lemma 2.2] in the case in which P is a fundamental domain. We will use it for P equal to either a fundamental domain or a flower; see Corollary 2.3 below.

Lemma 2.2. *Suppose that Assumption 1.1 holds for $\mathbf{T}_{\mathcal{R}}$ and consider a convex polygon P with sides lying in $\partial\mathbf{T}_{\mathcal{R}}$. Take any two different lines ℓ_1, ℓ_2 from $\partial\mathbf{T}_{\mathcal{R}}$ that intersect ∂P but not* int P. *Then either ℓ_1 and ℓ_2 do not intersect or they intersect at a vertex of P.*

Proof. Assume the contrary: $\ell_1 \cap \ell_2 = p \notin \partial P$; see Figure 2. The line ℓ_j meets ∂P either in a side s_j or a vertex v_j. In the former case let v_j be the end of s_j closest to p. Let $u_0 u_1 \dots u_k$ ($u_0 = v_1$, $u_k = v_2$) be the vertices, in order, of the segment of ∂P between v_1 and v_2 that lies inside the triangle $\Delta = v_1 v_2 p$, and let γ be the piecewise geodesic joining these points. Note that the s_j's are not included in γ.

Among all pairs (ℓ_1, ℓ_2) violating the statement of the lemma choose the one for which the number k of sides of γ is minimal.

If $k = 1$, Lemma 2.1 for the triangle $v_1 v_2 p$ in $\partial\mathbf{T}_{\mathcal{R}}$ yields a contradiction. Otherwise, consider the ray $\alpha \subset \ell(u_0 u_1)$ that is the continuation of the side $u_0 u_1$ past the point u_1. This ray enters the region D bounded by the curve γ and the segments $v_1 p$, $v_2 p$, since the inner angle of D at u_1 is more that π. This ray must exit D at

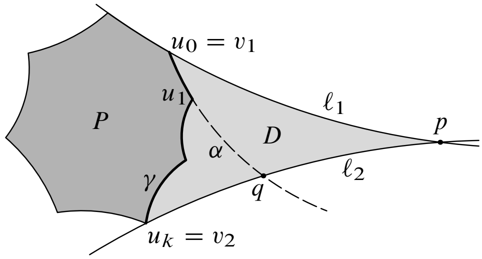


Figure 2. The proof of Lemma 2.2.

some point q, which cannot belong to γ (otherwise P is not convex) or to $v_1 p$ (since $\ell(u_0u_1)$ and ℓ_1 intersect only at v_1). Therefore, q belongs to $v_2 p$, and $\ell_1' = \ell(u_0u_1)$, $\ell_2' = \ell_2$ is a pair of lines also violating the statement of the lemma but for which $\gamma' = \gamma \setminus u_0u_1$ has $k-1$ sides. ∎

Corollary 2.3. *Suppose that Assumption 1.1 holds for $\mathbf{T}_{\mathcal{R}}$. Consider any two different sides s_1, s_2 of fundamental domains lying on the boundary of the flower $\mathcal{F}_v$. Then either $\ell(s_1)$ and $\ell(s_2)$ coincide, or they intersect at one or other end of either s_1 or s_2, or they do not intersect.*

Proof. We need some care since *one* side of $\mathcal{F}_v$ *as a convex polygon* can contain *two* sides of $\mathbf{T}_{\mathcal{R}}$: if a common vertex $w \neq v$ of two domains in $\mathcal{F}_v$ has $n(w) = 2$, then two sides on the boundary of $\mathcal{F}_v$ adjacent to w lie on the same geodesic and thus they form one side of $\mathcal{F}_v$ as a polygon. Let us refer to such side of the polygon $\mathcal{F}_v$ as a *compound* side.

In view of Lemma 2.2, we only need to rule out the case in which s_1 and s_2 are contained in two adjacent compound sides of the polygon $\mathcal{F}_v$, separated by an intervening side or sides. We will show that compound sides cannot be adjacent. Indeed, assume that both u_1u_2 and u_2u_3 are compound, where the u_i are vertices of petals of $\mathcal{F}_v$. Thus u_1u_2 and u_2u_3 contain vertices w_1 and w_2 in their interiors. Since each vw_j is a side common to two petals of $\mathcal{F}_v$, it follows that either $vw_1u_2w_2$ or vw_1u_2 is a fundamental domain. We see that either $N(\mathcal{R}) = 4$ and $n(w_{1,2}) = 2$, or $N(\mathcal{R}) = 3$ and $\mathcal{R}$ is compact, both of which cases are excluded by Assumption 1.1. ∎

3. Structure of thickened paths

As explained in the introduction, a thickened path between two domains $\mathcal{A}$ and $\mathcal{B}$ in $\mathbf{T}_{\mathcal{R}}$ is the union of all translates of $\mathcal{R}$ crossed by any possible shortest paths between $\mathcal{A}$ and $\mathcal{B}$. In this section we describe the detailed structure of thickened paths. The results will be used in Section 4 to construct a Markov coding that generates all possible thickened paths: we will show there that the features discussed are also *sufficient* conditions for a union of fundamental domains to be a thickened path.

From now on, we suppose Assumption 1.1 holds for G and $\mathcal{R}$ and do not mention this explicitly in the statements.

3.1. Thickened paths

Let $\mathcal{A}$, $\mathcal{B}$ be two fundamental domains from $\mathbf{T}_{\mathcal{R}}$. A *path* from $\mathcal{A}$ to $\mathcal{B}$ is a sequence $\underline{\mathcal{R}} = (\mathcal{R}_0 = \mathcal{A}, \mathcal{R}_1, \dots, \mathcal{R}_n = \mathcal{B})$ of domains from $\mathbf{T}_{\mathcal{R}}$, so that $\mathcal{R}_i$ and $\mathcal{R}_{i+1}$ have a common side for all $i = 0, \dots, n-1$. The number n here is the *length* of the path $\underline{\mathcal{R}}$. Equivalently, if $\mathcal{R}_i = g_i \mathcal{R}$ then $\underline{\mathcal{R}}$ is a path if and only if $\underline{g} = (g_0, \dots, g_n)$ is a path from $a = g_0$ to $b = g_n$ in the Cayley graph of G with respect to G_0.

The path $\underline{\mathcal{R}}$ is *shortest* if its length is minimal among all paths with the same ends. The *distance* $\operatorname{dist}(\mathcal{A}, \mathcal{B})$ between $\mathcal{A}$ and $\mathcal{B}$ is the length of a shortest path between them.

The *thickened path* from $\mathcal{A}$ to $\mathcal{B}$ is the collection of all domains in $\mathbf{T}_{\mathcal{R}}$ that belong to some shortest path from $\mathcal{A}$ to $\mathcal{B}$. This thickened path $\underline{\mathcal{S}}$ is decomposed into *levels* $\mathcal{S}_k$, $k = 0, 1, \dots, n = \operatorname{dist}(\mathcal{A}, \mathcal{B})$. Namely, a domain $\mathcal{C} \in \underline{\mathcal{S}}$ belongs to the level $\mathcal{S}_k$ if $\operatorname{dist}(\mathcal{A}, \mathcal{C}) = k$, and therefore

$$\operatorname{dist}(\mathcal{C}, \mathcal{B}) = n - k.$$

We also observe that two domains $\mathcal{C}, \mathcal{C}'$ in $\underline{\mathcal{S}} = (\mathcal{S}_0, \dots, \mathcal{S}_n)$ can have a common side only if their levels k, k' differ by one. Indeed, if $k < k' - 1$, then

$$\begin{aligned} n = \operatorname{dist}(\mathcal{A}, \mathcal{B}) &\le \operatorname{dist}(\mathcal{A}, \mathcal{C}) + \operatorname{dist}(\mathcal{C}, \mathcal{C}') + \operatorname{dist}(\mathcal{C}', \mathcal{B}) \\ &= k + 1 + (n - k') < n. \end{aligned}$$

The case $k = k'$ is impossible: the cycle $\mathcal{A}\mathcal{C}\mathcal{C}'\mathcal{A}$ of odd length $2k + 1$ contradicts the “chequered coloring” of $\mathbf{T}_{\mathcal{R}}$ (see the discussion of the even corners condition in Section 2).

3.2. Convexity of thickened paths

In this subsection we prove that the thickened path between two domains $\mathcal{A}$ and $\mathcal{B}$ is the smallest convex union of domains containing them. We begin by describing an alternative method of finding the distance between two fundamental domains.

Let us say that a geodesic γ in $\partial\mathbf{T}_{\mathcal{R}}$ *separates* domains $\mathcal{A}$ and $\mathcal{B}$ if $\mathcal{A}$ and $\mathcal{B}$ lie in different half-planes with respect to γ and denote the set of geodesics separating $\mathcal{A}$ and $\mathcal{B}$ by $\mathbf{S}_{\mathcal{A},\mathcal{B}}$. In particular, if $\mathcal{A}$ and $\mathcal{B}$ share a common side s, we have $\mathbf{S}_{\mathcal{A},\mathcal{B}} = \{\ell(s)\}$.

Now consider any path $\underline{\mathcal{R}} = (\mathcal{R}_0, \dots, \mathcal{R}_n)$ from $\mathcal{A}$ to $\mathcal{B}$ and let s_i be the common side of $\mathcal{R}_i$ and $\mathcal{R}_{i+1}$. Then every geodesic $\gamma \in \mathbf{S}_{\mathcal{A},\mathcal{B}}$ appears at least once among the $\ell(s_i)$, $i = 0, \dots, n-1$. (Note that we are not at this point assuming the $\ell(s_i)$ are distinct, see below.) Indeed, otherwise for every i the domains $\mathcal{R}_i$ and $\mathcal{R}_{i+1}$ lie in the same half-plane with respect to γ, so by transitivity $\mathcal{A} = \mathcal{R}_0$ and $\mathcal{B} = \mathcal{R}_n$ also lie in the same half-plane. This means that $\operatorname{dist}(\mathcal{A}, \mathcal{B}) \geq \#\mathbf{S}_{\mathcal{A},\mathcal{B}}$. Let us show that this inequality is indeed an equality.

Lemma 3.1. *The distance between two fundamental domains $\mathcal{A}$ and $\mathcal{B}$ equals the number of geodesics in $\mathbf{S}_{\mathcal{A},\mathcal{B}}$.*

Proof. From the consideration above we see that it is sufficient to construct a path of length $m = \#\mathbf{S}_{\mathcal{A},\mathcal{B}}$ from $\mathcal{A}$ to $\mathcal{B}$. To do so choose points $a \in \operatorname{int} \mathcal{A}$ and $b \in \operatorname{int} \mathcal{B}$ so that the geodesic segment $I = ab$ does not pass through any vertex of $\mathbf{T}_{\mathcal{R}}$. Then the geodesic lines from $\partial \mathbf{T}_{\mathcal{R}}$ crossed by I are exactly the geodesics separating a and b, or, equivalently, $\mathcal{A}$ and $\mathcal{B}$. The points of intersection of I with these m lines from $\mathbf{S}_{\mathcal{A},\mathcal{B}}$ are different and by convexity I cannot enter any domain twice. Hence, I traverses $(m+1)$ domains $\mathcal{R}_0 = \mathcal{A}, \mathcal{R}_1, \dots, \mathcal{R}_m = \mathcal{B}$, and for any $i = 0, \dots, m-1$ the domains $\mathcal{R}_i$ and $\mathcal{R}_{i+1}$ have a common side. ∎

Remark 3.2. Let us say that a path $\underline{\mathcal{R}} = (\mathcal{R}_0 = \mathcal{A}, \dots, \mathcal{R}_n = \mathcal{B})$ *crosses a line* γ if $\gamma = \ell(s_i)$ for some i, where $s_i = \mathcal{R}_i \cap \mathcal{R}_{i+1}$. Then if γ is the shortest path between $\mathcal{A}$ and $\mathcal{B}$, it crosses every line γ from $\mathbf{S}_{\mathcal{A},\mathcal{B}}$ exactly once (i.e. $\gamma = \ell(s_i)$ for only one i) and does not cross any other lines.

Proof. We saw above that every line from $\mathbf{S}_{\mathcal{A},\mathcal{B}}$ appears in the sequence $\{\ell(s_i)\}_{i=0}^{n-1}$ *at least* once. If some line from $\mathbf{S}_{\mathcal{A},\mathcal{B}}$ appears twice in this sequence, or if any line outside of $\mathbf{S}_{\mathcal{A},\mathcal{B}}$ appears there, we have $n > \#\mathbf{S}_{\mathcal{A},\mathcal{B}}$. On the other hand, for the shortest path we have $n = \operatorname{dist}(\mathcal{A}, \mathcal{B}) = \#\mathbf{S}_{\mathcal{A},\mathcal{B}}$ by Lemma 3.1. ∎

The following proposition describes a thickened path as a convex set.

Proposition 3.3. *Let $\mathcal{A}$ and $\mathcal{B}$ be two fundamental domains in $\mathbf{T}_{\mathcal{R}}$. The thickened path from $\mathcal{A}$ to $\mathcal{B}$ is the minimal convex union of fundamental domains that contains both $\mathcal{A}$ and $\mathcal{B}$. If $\mathcal{R}$ is compact, then the boundary of this thickened path contains at least one side of $\mathcal{A}$ and one of $\mathcal{B}$.*

Proof. We divide the proof into steps.

Step 1. Denote by $\mathbf{NS}_{\mathcal{A},\mathcal{B}}$ the set of all lines $\ell \subset \partial \mathbf{T}_{\mathcal{R}}$ that do not separate $\mathcal{A}$ and $\mathcal{B}$. For every $\ell \in \mathbf{NS}_{\mathcal{A},\mathcal{B}}$ consider the half-plane H_ℓ bounded by ℓ and containing both $\mathcal{A}$

and $\mathcal{B}$. Set

$$\mathcal{G} = \bigcap_{\ell \in \mathbf{NS}_{\mathcal{A},\mathcal{B}}} H_\ell.$$

We claim that $\mathcal{G}$ is the thickened path from $\mathcal{A}$ to $\mathcal{B}$. By the previous corollary, no shortest path from $\mathcal{A}$ to $\mathcal{B}$ intersects any line $\ell \in \mathbf{NS}_{\mathcal{A},\mathcal{B}}$. Therefore, the thickened path from $\mathcal{A}$ to $\mathcal{B}$ is contained in $\mathcal{G}$.

On the other hand, consider any fundamental domain $\mathcal{C} \subset \mathcal{G}$. As in the proof of Lemma 3.1, choose generic points $a \in \mathcal{A}$, $b \in B$, $c \in C$ so that the segments ac and cb do not contain vertices. The sequences

$$\mathcal{R}_0 = \mathcal{A}, \mathcal{R}_1, \ldots, \mathcal{R}_k = \mathcal{C}, \quad \mathcal{R}_k = \mathcal{C}, \mathcal{R}_{k+1}, \ldots, \mathcal{R}_{k+l} = \mathcal{B}$$

of fundamental domains traversed respectively by ac and cb are shortest paths from $\mathcal{A}$ to $\mathcal{C}$ and $\mathcal{C}$ to $\mathcal{B}$, respectively. Since $\mathcal{G}$ is convex, the segments ac and cb lie in $\mathcal{G}$ and hence so do all the $\mathcal{R}_j$'s.

It remains to show that $\mathcal{R}_0, \mathcal{R}_1, \ldots, \mathcal{R}_{k+l}$ is a shortest path. Let ℓ_j be the geodesic line separating $\mathcal{R}_j$ and $\mathcal{R}_{j+1}$. Then the lines $(\ell_j)_{j=0}^{k+l-1}$ are the consecutive geodesics intersected by the path acb. Every geodesic γ in $\partial \mathbf{T}_{\mathcal{R}}$ intersects either none or two sides of the triangle abc. If γ intersects the sides ac and cb, then γ does not separate a and b. But then $c \in \mathcal{C}$ does not belong to H_γ, and hence to $\mathcal{G}$. On the other hand, γ intersects ab if and only if it belongs to $\mathbf{S}_{\mathcal{A},\mathcal{B}}$, thus each line from $\mathbf{S}_{\mathcal{A},\mathcal{B}}$ crosses the curve acb exactly once. Therefore,

$$k + l = \#\mathbf{S}_{\mathcal{A},\mathcal{B}} = \operatorname{dist}(\mathcal{A}, \mathcal{B}).$$

Step 2. Assume that $\mathcal{G}' \subset \mathcal{G}$ is a smaller convex union of fundamental domains containing $\mathcal{A}$ and $\mathcal{B}$. Note that $\mathcal{G}$ (hence, $\mathcal{G}'$) contains only finite number of fundamental domains, since there are only finitely many domains at distance not more than $\operatorname{dist}(\mathcal{A}, \mathcal{B})$ from $\mathcal{A}$. Thus $\mathcal{G}'$ is a convex polygon and the supporting half-planes of its sides are H_{ℓ_k}, $k = 1, \ldots, K$, for some lines $\ell_k \in \mathbf{NS}_{\mathcal{A},\mathcal{B}}$. Therefore,

$$\mathcal{G}' = \bigcap_{k=1}^{K} H_{\ell_k} \supset \bigcap_{\ell \in \mathbf{NS}_{\mathcal{A},\mathcal{B}}} H_\ell = \mathcal{G}.$$

Step 3. To prove the final statement, choose generic points $a \in \mathcal{A}$, $b \in \mathcal{B}$ so that the line $\ell(ab)$ does not contain any vertices of $\mathbf{T}_{\mathcal{R}}$. Denote by α the ray in $\ell(ab)$ starting from a in the direction away from b. Let a' be the first intersection of α with $\partial \mathbf{T}_{\mathcal{R}}$ and let $\ell' \ni a'$ be the corresponding geodesic line from $\partial \mathbf{T}_{\mathcal{R}}$; a' exists since $\mathcal{A}$ is compact. Thus $\ell' \in \mathbf{NS}_{\mathcal{A},\mathcal{B}}$ as it does not cross the segment ab. Therefore, no point on α after a' belongs to $H_{\ell'}$, and hence to $\mathcal{G}$, and no point before a' is separated from a by any H_ℓ. Thus $\alpha \cap \mathcal{G} = aa'$, so $a' \in \partial \mathcal{G}$ and the side of $\mathcal{A}$ that contains a' belongs to $\partial \mathcal{G}$. ■

3.3. Levels of thickened paths

Each level in a thickened path is a union of domains. We now show that each level can contain at most two domains.

Let $\underline{\mathcal{S}} = (\mathcal{S}_0 = \mathcal{A}, \mathcal{S}_1, \dots, \mathcal{S}_n = \mathcal{B})$ be the thickened path from $\mathcal{A}$ to $\mathcal{B}$. Consider its closure $\mathrm{clos}_{\overline{\mathbb{D}}}\,\underline{\mathcal{S}}$ and its boundary $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}}$ in $\overline{\mathbb{D}}$. They are homeomorphic respectively to a closed disk and a circle.

Consider the intersection $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}} \cap \partial_{\overline{\mathbb{D}}}\mathcal{A}$. It is nonempty: if $\mathcal{R}$ is compact, this is stated in Proposition 3.3, otherwise any point of $\partial_{\overline{\mathbb{D}}}\mathcal{A} \cap \partial\mathbb{D}$ belongs to $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}}$. Moreover, it is connected. Indeed, take any p and p' lying on different sides in this intersection and consider a segment $J \subset \mathcal{A}$ connecting p and p'. Then J separates $\mathrm{clos}_{\overline{\mathbb{D}}}\,\underline{\mathcal{S}}$ into two connected components. If both components contain fundamental domains (besides parts of $\mathcal{A}$), choose any $\mathcal{C} \neq \mathcal{A}, \mathcal{B}$ such that $\mathcal{B}$ and $\mathcal{C}$ lie in different components. Then the path from $\mathcal{C}$ to $\mathcal{B}$ inside $\mathcal{G}$ must cross J and hence $\mathcal{A}$, so $\mathcal{C}$ cannot belong to a shortest path from $\mathcal{A}$ to $\mathcal{B}$. Therefore, one connected component contains only a part of $\mathcal{A}$. But then the boundary of this connected component apart from J is a segment of $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}} \cap \partial_{\overline{\mathbb{D}}}\mathcal{A}$ that connects p and p'.

We conclude that $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}} \setminus (\partial_{\overline{\mathbb{D}}}\mathcal{A} \cup \partial_{\overline{\mathbb{D}}}\mathcal{B})$ consists of two arcs, which we call the *left* and *right* boundaries of $\underline{\mathcal{S}}$ and denote $\partial_{L,R}\underline{\mathcal{S}}$. Namely, going clockwise around $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}}$ we pass through an arc of $\partial_{\overline{\mathbb{D}}}\mathcal{A}$, then $\partial_L\underline{\mathcal{S}}$, then an arc of $\partial_{\overline{\mathbb{D}}}\mathcal{B}$, and $\partial_R\underline{\mathcal{S}}$. Both $\partial_{L,R}\underline{\mathcal{S}}$ are oriented from $\partial_{\overline{\mathbb{D}}}\mathcal{A}$ to $\partial_{\overline{\mathbb{D}}}\mathcal{B}$. Sometimes we will use the same notation $\partial_{L,R}\underline{\mathcal{S}}$ for the parts of these boundaries that lie inside $\mathbb{D}$.

Proposition 3.4. *Let $\underline{\mathcal{S}} = (\mathcal{S}_0 = \mathcal{A}, \mathcal{S}_1, \dots, \mathcal{S}_n = \mathcal{B})$ be a thickened path from $\mathcal{A}$ to $\mathcal{B}$. Consider the sequence of adjacent domains $\mathcal{L}_0 = \mathcal{A}, \mathcal{L}_1, \dots, \mathcal{L}_m = \mathcal{B}$ which meet $\partial_L\underline{\mathcal{S}}$ in a point or side and the similar sequence of domains $\mathcal{R}_0 = \mathcal{A}, \mathcal{R}_1, \dots, \mathcal{R}_{m'} = \mathcal{B}$ which meet $\partial_R\underline{\mathcal{S}}$. Then:*

(1) *both these sequences are shortest paths from $\mathcal{A}$ to $\mathcal{B}$, hence $m = m' = n$;*

(2) *every domain in $\underline{\mathcal{S}}$ belongs to one of these two sequences, hence $\mathcal{S}_j = \{\mathcal{L}_j, \mathcal{R}_j\}$ for every $j = 0, \dots, n$; it is possible that $\mathcal{L}_j = \mathcal{R}_j$.*

Proof. We consider each point in turn.

(1) Consider the side s_j between $\mathcal{L}_j$ and $\mathcal{L}_{j+1}$. Then $\ell(s_j)$ passes inside $\underline{\mathcal{S}}$ and hence separates $\mathcal{A}$ and $\mathcal{B}$ by Proposition 3.3. Moreover, every geodesic ℓ in $\mathbf{T}_{\mathcal{R}}$ separating $\mathcal{A}$ and $\mathcal{B}$ intersects $\underline{\mathcal{S}}$ in a segment $I(\ell)$ with one end on $\partial_L\underline{\mathcal{S}}$ and another one on $\partial_R\underline{\mathcal{S}}$. Thus if $s_j \subset I(\ell)$ then s_j is adjacent to the end of $I(\ell)$ lying on $\partial_L\underline{\mathcal{S}}$. Therefore, each of the n geodesics separating $\mathcal{A}$ and $\mathcal{B}$ produces exactly one such s_j, so $m = n$.

(2) This is [3, Lemma 2.7]. We give a proof for the sake of completeness.

Assume that the fundamental domain $\mathcal{C}$ lies strictly inside $\underline{\mathcal{S}}$. Then $\mathcal{C}$ is compact. Let s_i, $i = 1, \dots, N(\mathcal{R})$ be the consecutive sides of $\mathcal{C}$ and let H_i be the half-plane bounded by $\ell(s_i)$ that does not contain $\mathcal{C}$. By Lemma 2.2 the lines $\ell(s_j)$ and $\ell(s_k)$ intersect only for adjacent sides s_j and s_k. Therefore,

$$H_j \cap H_k \neq \varnothing$$

also only for adjacent s_j and s_k. This means that at most two lines $\ell(s_j)$ separate $\mathcal{A}$ from $\mathcal{C}$ and at most two separate $\mathcal{B}$ from $\mathcal{C}$. Hence there are at least $N(\mathcal{R}) - 4$ lines of the form $\ell(s_j)$ that do not separate $\mathcal{A}$ and $\mathcal{B}$. In the case $N(\mathcal{R}) \geq 5$ we arrive at the contradiction with Step 1 in Proposition 3.3: a line not separating $\mathcal{A}$ and $\mathcal{B}$ cannot enter the interior of $\underline{\mathcal{S}}$.

If $N(\mathcal{R}) = 4$, the only remaining case (up to renumbering the s_j) is that

$$\mathcal{A} \subset H_1 \cap H_2, \quad \mathcal{B} \subset H_3 \cap H_4.$$

Then $p = s_2 \cap s_3$, $q = s_4 \cap s_1$ are opposite vertices of $\mathcal{C}$. Assumption 1.1 states that $n(p) > 2$ or $n(q) > 2$. If, say, $n(p) > 2$, there is a line ℓ^* from $\partial \mathbf{T}_{\mathcal{R}}$ that intersects $\partial\mathcal{C}$ only at p. Denote the half-plane bounded by ℓ^* and not containing $\mathcal{C}$ by H^*. Lemma 2.1 yields that ℓ^* does not intersect $\ell(s_1)$ and $\ell(s_4)$, thus

$$H_1 \cap H^* = H_4 \cap H^* = \varnothing.$$

Therefore, again using Step 1 in the proof of Proposition 3.3, both $\mathcal{A}$ and $\mathcal{B}$ lie outside H^*, so we arrive at the same contradiction for ℓ^*. ∎

3.4. Bottles and bottlenecks

Continuing our consideration of thickened paths, let $\underline{\mathcal{S}} = (\mathcal{S}_0 = \mathcal{A}, \mathcal{S}_1, \dots, \mathcal{S}_n = \mathcal{B})$ be the thickened path between fundamental domains $\mathcal{A}$ and $\mathcal{B}$. We call a level $\mathcal{S}_j \in \underline{\mathcal{S}}$ a *bottleneck* if it contains only one fundamental domain $\mathcal{C}$ say, so that, with the notation of Proposition 3.4, $\mathcal{L}_j = \mathcal{R}_j = \mathcal{C}$. Thus all shortest paths from $\mathcal{S}_0 = \mathcal{A}$ to $\mathcal{S}_n = \mathcal{B}$ must pass through $\mathcal{C}$. The bottlenecks of a thickened path $\underline{\mathcal{S}}$ divide it into *bottles*. Namely, if $\mathcal{S}_r$ and $\mathcal{S}_s$ are bottlenecks and $\mathcal{S}_k$, $r < k < s$ are not, then

$$\underline{\mathcal{S}}_r^s = \bigcup_{k=r}^{s} \mathcal{S}_k$$

is a *bottle* (so each bottle has two bottlenecks, one at each end). The bottle is *trivial* if $s = r + 1$.

We now focus on the structure of one bottle. Note that if in a thickened path $\underline{\mathcal{S}}$ the levels $\mathcal{S}_j = \{\mathcal{A}'\}$ and $\mathcal{S}_k = \{\mathcal{B}'\}$ contain one domain each, then $(\mathcal{S}_i)_{i=j}^k$ is the thickened path from $\mathcal{A}'$ to $\mathcal{B}'$. Therefore, every bottle is a thickened path between its two bottlenecks, and we assume for the rest of this section that $\underline{\mathcal{S}} = (\mathcal{S}_0, \mathcal{S}_1, \dots, \mathcal{S}_n)$ is a non-trivial bottle: $\mathcal{S}_j$ contains two domains for every $j = 1, \dots, n-1$ and one domain for $j = 0$ and $j = n$.

Denote by s_j^L (respectively, s_j^R) the common side of $\mathcal{L}_j$ and $\mathcal{L}_{j+1}$ (respectively, $\mathcal{R}_j$ and $\mathcal{R}_{j+1}$) and consider the set

$$\mathbb{A} = \bigcup_{i=0}^{n} \big(\mathrm{int}(\mathcal{L}_i) \cup \mathrm{int}(\mathcal{R}_i)\big) \cup \bigcup_{i=0}^{n-1} \big(\mathrm{int}(s_i^L) \cup \mathrm{int}(s_i^R)\big).$$

This set is homeomorphic to an annulus and can be retracted to the boundary of the closed disk $\mathrm{clos}_{\overline{\mathbb{D}}}(\underline{\mathcal{S}})$. Now let

$$\mathcal{T} = \mathrm{int}\,\underline{\mathcal{S}} \setminus \mathbb{A}.$$

This set is a union of all vertices and sides of $\mathbf{T}_{\mathcal{R}}$ that lie inside $\underline{\mathcal{S}}$ and have no points in common with $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}}$; note that $\mathcal{T}$ cannot contain fundamental domains by Proposition 3.4. Thus the set $\mathcal{T}$ is closed, connected and simply connected. In other words, the graph $\mathcal{T}$ is a tree. We call $\mathcal{T}$ the *core* of the bottle.

Proposition 3.5. *The core $\mathcal{T}$ of a non-trivial bottle is a linear graph, that is, its vertices can be enumerated as $v_0, v_1, \dots, v_r$ in such a way that for every $j = 1, \dots, r$ the vertices v_{j-1} and v_j are adjacent and its edges are sides $v_{j-1}v_j$ for $j = 1, \dots, r$.*

Proof. (We are grateful to the referee for suggesting the following rephrasing of our original proof.)

A tree is a linear graph if and only if it has not more than two *leaves*, a leaf being a vertex with only one adjacent edge. Let $\mathcal{S}$ be a thickened path between $\mathcal{A}$ and $\mathcal{B}$, assumed to be a bottle with core $\mathcal{T}$. We want to check that $\mathcal{T}$ cannot have more than two leaves.

If v is a vertex of $\mathcal{T}$, then the flower $\mathcal{F}_v$ cannot be contained in $\bigcup_{i=0}^{n} \mathcal{L}_i$ or $\bigcup_{i=0}^{n} \mathcal{R}_i$. Indeed, if $\mathcal{F}_v$ contains two fundamental domains of least index, we are done; otherwise, there is a unique fundamental domain $\mathcal{C}$ of least index i and the two fundamental domains adjacent to $\mathcal{C}$ in $\mathcal{F}_v$ have same index, $i+1$, therefore one of them is $\mathcal{L}_{i+1}$ and the other is $\mathcal{R}_{i+1}$.

Suppose that v is a leaf of $\mathcal{T}$. If v is the only vertex in $\mathcal{T}$, we are done. Otherwise, let s be the only edge adjacent to v. Walking through fundamental domains, going cyclically around v, one encounters $\mathcal{L}_i$'s and $\mathcal{R}_i$'s. There are only two ways of going from the left side to the right side and vice-versa: either crossing an edge of $\mathcal{T}$, or passing through a bottleneck; since s is the only edge that can be crossed, the flower $\mathcal{F}_v$ contains a bottleneck.

Moreover, if a bottleneck is contained in a flower contained in the bottle, then the center of the flower must be the intersection of the bottleneck with its two adjacent fundamental domains in the bottle.

In summary, the flower around any leaf of $\mathcal{T}$ contains a bottleneck and a bottleneck is contained in a unique flower; therefore, there are at most two leaves. ∎

In reading what follows, the reader may find it helpful to refer to Figure 3.

Since each of $\mathcal{F}_{v_0}$ and $\mathcal{F}_{v_r}$ contains a bottleneck, inverting the sequence (v_j) if necessary, we will assume for the rest of the section that $\mathcal{S}_0 \subset \mathcal{F}_{v_0}$, $\mathcal{S}_n \subset \mathcal{F}_{v_r}$. Each flower $\mathcal{F}_{v_j}$, $j = 1, \ldots, r-1$, is split by the edges $v_{j-1}v_j$ and $v_j v_{j+1}$ into two sectors, one containing domains from $\underline{\mathcal{L}}$ and the other from $\underline{\mathcal{R}}$. For $j = 0$ or $j = r$ similar sectors overlap only in the respective bottlenecks. Thus the "left" sector of $\mathcal{F}_{v_0}$ consists of the petals $\mathcal{L}_0, \ldots, \mathcal{L}_{i_0}$ in counterclockwise order. The part of the boundary of $\mathcal{L}_{i_0}$ lying inside the bottle is the union of the side $s^L_{i_0-1}$, which goes from $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}}$ to v_0, a segment $v_0 v_1 \ldots v_t \subset \mathcal{T}$, $t \geq 1$, and the side $s^L_{i_0}$, which goes from v_t to $\partial_{\overline{\mathbb{D}}}\underline{\mathcal{S}}$. Then there are some domains $\mathcal{L}_{i_0}, \ldots, \mathcal{L}_{i_t}$ that are consecutive petals in the "left" sector at v_t, and so on. Denoting $i_1 = \cdots = i_{t-1} = i_0$, etc., we arrive at the following statement.

Proposition 3.6. *Let a bottle $\underline{\mathcal{S}} = (\mathcal{S}_0, \ldots, \mathcal{S}_n)$ have a core $\mathcal{T} = v_0 \ldots v_r$ oriented in such a way that $\mathcal{S}_0 \in \mathcal{F}_{v_0}$, $\mathcal{S}_n \in \mathcal{F}_{v_r}$. Then there exist numbers*

$$i_{-1} = 0 \leq i_0 \leq i_1 \leq \cdots \leq i_{r-1} \leq n = i_{r+1}$$

such that the sequence $\mathcal{L}_{i_{s-1}}, \ldots, \mathcal{L}_{i_s}$ enumerates in clockwise order the consecutive petals of $\mathcal{F}_{v_s}$ from $v_{s-1}v_s$ to $v_s v_{s+1}$, where $v_{-1}v_0 = s^R_0$, $v_r v_{r+1} = s^R_{n-1}$.

The same is true for the path $\underline{\mathcal{R}}$ and a sequence (i'_j). Here the petals are enumerated in counterclockwise order and one should assume $v_{-1}v_0 = s^L_0$, $v_r v_{r+1} = s^L_{n-1}$.

The next two lemmas establish some properties of the sequences (i_j) and (i'_j).

Lemma 3.7. *If v_j is incident to the domains $\mathcal{L}_i, \ldots, \mathcal{L}_{i+k-1}$ or $\mathcal{R}_i, \ldots, \mathcal{R}_{i+k-1}$, then $k \leq n(v_j) + 1$.*

Proof. If this is not the case, then the part of the path $\underline{\mathcal{L}}$ between $\mathcal{L}_i$ and $\mathcal{L}_{i+k-1}$ can be replaced by a shorter one which goes around the other side of v_j, and similarly for $\underline{\mathcal{R}}$. This would contradict Proposition 3.4 (1). ∎

Let l'_j and r'_j be the number of domains in $\underline{\mathcal{L}}$ and $\underline{\mathcal{R}}$, respectively, that are incident to v_j so that $l'_j \leq n(v_j) + 1$, and define $l_j = l'_j - n(v_j)$, $r_j = r'_j - n(v_j)$.

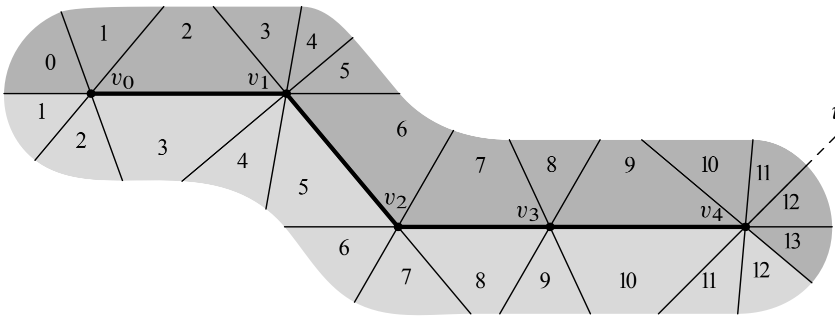


Figure 3. A typical bottle, in this case with levels from 0 to 13. The bold piecewise geodesic joining $v_0, \ldots, v_4$ is its core and the shaded regions indicate the left hand path $\underline{\mathcal{L}}$. With the notation of Lemma 3.8 one has $(l_0, r_0) = (0, 1)$; $(l_1, r_1) = (1, -1)$; $(l_2, r_2) = (-1, 1)$; $(l_3, r_3) = (0, 0)$; $(l_4, r_4) = (1, 0)$. Note that a continuation of the core to a vertex $v_5 = u$ as indicated by the dotted line, with two consecutive right turns, would give instead $(l_4, r_4) = (-1, 1)$ resulting in the inadmissible string $(r_2, r_3, r_4) = (1, 0, 1)$.

Lemma 3.8. *Let $\mathcal{T}$ be the core of a bottle, that is, the chain of sides between vertices $v_0, \ldots, v_r$.*

(1) *We have $l_j, r_j \leq 1$ for all $j = 0, \ldots, r$.*

(2) *If $0 < j < r$ then $l_j + r_j = 0$, hence $(l_j, r_j) \in \{(-1, 1), (0, 0), (1, -1)\}$. If $j = 0$ or $j = r$, where $r > 0$, then $l_j + r_j = 1$, hence $(l_j, r_j) \in \{(0, 1), (1, 0)\}$. If $j = 0 = r$ then $l_j + r_j = 2$, hence $(l_j, r_j) = (1, 1)$.*

(3) *The sequences $(l_j)_{j=0}^r$ and $(r_j)_{j=0}^r$ cannot contain a segment of the form $1, 0, \ldots, 0, 1$ (with $t \geq 0$ zeroes between the two 1's).*

(4) *The side $v_k v_{k+1}$ separates either $\mathcal{L}_s$ and $\mathcal{R}_{s+1}$ if the last $(l_j, r_j) \neq (0, 0)$ with $j \leq k$ has $r_j = 1$, or $\mathcal{L}_{s+1}$ and $\mathcal{R}_s$ if this (l_j, r_j) has $l_j = 1$.*

(5) *For any s the domains $\mathcal{L}_s$ and $\mathcal{R}_s$ of the same level share a common vertex $v_j \in \mathcal{T}$.*

This lemma controls the shape of the core of a bottle. Items (1) and (2) together say that the segments of the core of a bottle turn through at most one sector (petal) at each vertex v_j. Item (3) implies that successive bends alternate between turns to the right and turns to the left. The lemma is illustrated in Figure 3, which shows a typical bottle.

Proof. We consider each point in turn.

(1) This follows from the previous lemma.

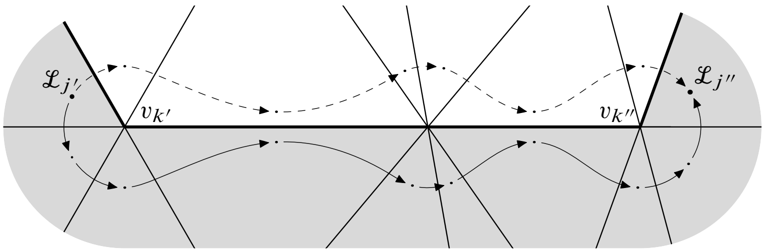


Figure 4. The proof of Lemma 3.8. The solid lines represent the subpath $(\mathcal{L}_{j'}, \ldots, \mathcal{L}_{j''})$ and the dashed ones indicate a shorter alternative.

(2) All $2n(v)$ domains around v are counted either in l'_j or in r'_j. The only domains that are counted both in l'_j and r'_j are $\mathcal{L}_0 = \mathcal{R}_0$ and $\mathcal{L}_n = \mathcal{R}_n$. As we have seen above, these domains are incident only to v_0 and v_r respectively.

(3) Assume that $l_{k'} = l_{k''} = 1$, $l_i = 0$ for $k' < i < k''$. Then the edges $v_{k'}v_{k'+1}, \ldots, v_{k''-1}v_{k''}$ form a geodesic segment γ. Let $\mathcal{L}_{j'}$ be the first domain in $\underline{\mathcal{L}}$ adjacent to $v_{k'}$ and $\mathcal{L}_{j''}$ be the last one adjacent to $v_{k''}$. The condition implies that the path from $\mathcal{L}_{j'}$ to $\mathcal{L}_{j''}$ crosses γ both at $v_{k'}$ and $v_{k''}$. But this is impossible by Remark 3.2; see Figure 4.

(4) We prove this by induction on k. The statement clearly holds for $k = 0$. Suppose the edge $v_{k-1}v_k$ separates $\mathcal{L}_s$ and $\mathcal{R}_{s'}$, while $v_k v_{k+1}$ separates $\mathcal{L}_t$ and $\mathcal{R}_{t'}$. If $(l_k, r_k) = (0, 0)$, then

$$t = s + n(v_k) - 1, \quad t' = s' + n(v_k) - 1,$$

hence $t - t' = s - s'$, so the statement for $k - 1$ implies it for k.

Similarly, if $(l_k, r_k) = (1, -1)$ then the previous non-zero (l_j, r_j) should have $r_j = 1$ by item (3), hence $s' = s + 1$ by the induction assumption. Also,

$$t = s + n(v), \quad t' = s' + n(v) - 2,$$

thus $t = t' + 1$ and we have proved the statement for k.

(5) Define $m_k \in \mathbb{Z} + 1/2$ so that the edge $v_k v_{k+1}$ separates domains from the levels $m_k - 1/2$ and $m_k + 1/2$, (thus, for example, in Figure 3, $m_2 = 7.5$, since $v_2 v_3$ separates $\mathcal{L}_7$ and $\mathcal{R}_8$), also set

$$m_{-1} = -1/2, \quad m_{r+1} = n + 1/2.$$

Then from item (4) and Proposition 3.6 one can see that (m_j) is an increasing sequence and the flower $\mathcal{F}_{v_k}$ contains all domains in the levels $\mathcal{S}_s$ with $m_{k-1} < s < m_k$. ∎

Besides describing the core of the bottle, Lemma 3.8 also allows us to describe how adjacent levels in the bottle are attached to one other. The reader may find it helpful to refer again to Figure 3, see also Remark 3.10.

Lemma 3.9. *Let $\underline{\mathcal{S}} = (\mathcal{S}_0, \mathcal{S}_1, \dots, \mathcal{S}_n)$ be a bottle.*

Every level $\mathcal{S}_k$, $k = 1, \dots, n-1$ contains two domains with a common vertex v. The flower $\mathcal{F}_v$ is split by these two domains into two sectors, each containing an odd number of petals. The sector bounded by the two radii of $\mathcal{F}_v$ which make up $\mathcal{S}_{k-1} \cap \mathcal{S}_k$ we call the "past" sector and that bounded by the radii making up $\mathcal{S}_k \cap \mathcal{S}_{k+1}$ the "future" sector.

(1) *If the future sector for $\mathcal{S}_k$ contains at least three petals, then $\mathcal{S}_{k+1}$ consists of the two petals in this sector adjacent to $\mathcal{S}_k$.*

(2) *If the future sector for $\mathcal{S}_k$ contains only one petal, there are the following possibilities:*

  (a) *$k + 1 = n$ and $\mathcal{S}_n$ is the only domain in the future sector; or*

  (b) *let the boundary of the future sector be $v_L v v_R$, where the segment $v v_L$ is the next one after $v v_R$ when going counterclockwise around v. Then $\mathcal{S}_{k+1}$ contains either the two domains from $\mathcal{F}_{v_L}$ adjacent to $\mathcal{S}_k$ (the "left" subcase), or the two domains from $\mathcal{F}_{v_R}$ adjacent to $\mathcal{S}_k$ (the "right" subcase).*

Proof. Assume that $\mathcal{S}_k$ belongs to $\mathcal{F}_{v_j}$. We have seen that the future sector of $\mathcal{F}_{v_j}$ contains s domains (in clockwise order): $\mathcal{L}_{k+1}, \dots, \mathcal{L}_{k+p}, \mathcal{R}_{k+q}, \dots, \mathcal{R}_{k+1}$, where either $\mathcal{L}_{k+p} = \mathcal{R}_{k+q} = \mathcal{S}_n$ or $\mathcal{L}_{k+p}$ and $\mathcal{R}_{k+q}$ are adjacent via a segment $v_j v_{j+1}$. In the former case

$$p = q, \quad s = p + q - 1,$$

and in the latter case

$$p = q \pm 1, \quad s = p + q$$

by item (4) of Lemma 3.8. Thus in both cases s is odd, as is the number $2n(v_j) - s - 2$ of petals in the past sector.

If $s \geq 3$, we have $p, q \geq 1$, so $\mathcal{L}_{k+1}$ and $\mathcal{R}_{k+1}$ belong to the future sector and are adjacent to its radii; this is case 1. Now assume $s = 1$ and let $\mathcal{A}$ be the single domain in the future sector. If $\mathcal{A}$ is the only domain in $\mathcal{S}_{k+1}$, it is a bottleneck, so $k + 1 = n$ (case (2a)). Otherwise $\mathcal{A}$ is either $\mathcal{R}_{k+1}$ or $\mathcal{L}_{k+1}$. Note that $v_j v_{j+1}$ is the common side of the last domains in $\underline{\mathcal{L}}$ and $\underline{\mathcal{R}}$ belonging to $\mathcal{F}_{v_j}$, so either $\mathcal{A} = \mathcal{R}_{k+1}$ and $v_{j+1} = v_L$ (the "left" subcase) or $\mathcal{A} = \mathcal{L}_{k+1}$ and $v_{j+1} = v_R$ (the "right" one). It remains to consider the flower $\mathcal{F}_{v_{j+1}}$: say, in the left subcase it contains $\mathcal{L}_k$ and $\mathcal{R}_{k+1}$ adjacent via $v_j v_{j+1}$, and $\mathcal{L}_{k+1}$ is the other petal adjacent to $\mathcal{L}_k$. ∎

Remark 3.10. Case (2b) corresponds to the transition from the domains around v_j to the ones around v_{j+1}. If transitions $v_j \to v_{j+1}$ and $v_{j+1} \to v_{j+2}$ belong to the same ("left" or "right") subcase, then item (4) of Lemma 3.8 implies that $\mathcal{F}_{v_{j+1}}$ contains the same number of domains from $\underline{\mathcal{L}}$ and $\underline{\mathcal{R}}$, so the segment $v_j v_{j+1} v_{j+2}$ of the core curve is straight. Similarly, if the transition $v_j \to v_{j+1}$ belongs to the "left" subcase and $v_{j+1} \to v_{j+2}$ to the "right" one, the core curve $v_j v_{j+1} v_{j+2}$ bends to the right through one petal. For example, in Figure 3, the transitions $v_0 \to v_1$ ($k = 2$), $v_2 \to v_3$ ($k = 7$), $v_3 \to v_4$ ($k = 9$) are in case (2b) left, and $v_1 \to v_2$ ($k = 5$) is in the case (2b) right, hence the core curve bends to the right at v_1, to the left at v_2, and remains straight at v_3.

4. The Markov coding

As was explained in the introduction, states of our Markov chain should describe how the "past" level $\mathcal{S}_- = \mathcal{S}_k$ of a thickened path is attached to its "future" level $\mathcal{S}_+ = \mathcal{S}_{k+1}$. More specifically, a state of the Markov chain should describe the arrangement of $\mathcal{S}_-$ and $\mathcal{S}_+$ up to the G-action. The set $\widehat{\Xi}$ of possible arrangements is listed in Definition 4.2. However, to construct the actual states Ξ of our Markov chain we have to endow these arrangements with some additional data; this is done in Section 4.1. In Section 4.2 we list the admissible transitions between states, thus defining the transition matrix Π. In Section 4.3 we prove the important result that there is a bijective correspondence between admissible sequences and thickened paths. Finally, Section 4.4 presents a time-reversing involution on the set of states.

4.1. States of the Markov chain

4.1.1. Types of adjacency. As we have seen in Lemma 3.9, the adjacency graph for the domains in $\mathcal{S}_- \cup \mathcal{S}_+$ has one of the following types as illustrated in Figure 6 below:

A. $\#\mathcal{S}_- = 1$, $\#\mathcal{S}_+ = 1$, and the graph contains the only possible edge from $\mathcal{S}_-$ to $\mathcal{S}_+$. This corresponds to a trivial bottle.

B. $\#\mathcal{S}_- = 1$, $\#\mathcal{S}_+ = 2$, and the graph contains both edges from $\mathcal{S}_-$ to $\mathcal{S}_+$. This state starts a bottle.

C. $\#\mathcal{S}_- = 2$, $\#\mathcal{S}_+ = 2$, and the edges join the left domain in $\mathcal{S}_-$ to the left domain in $\mathcal{S}_+$ and the right domain in $\mathcal{S}_-$ to the right domain in $\mathcal{S}_+$ (this is the case from item (1) of Lemma 3.9).

D. $\#\mathcal{S}_- = 2$, $\#\mathcal{S}_+ = 1$, and the graph contains both edges from $\mathcal{S}_-$ to $\mathcal{S}_+$. This state ends a bottle (the case (2a) of the lemma).

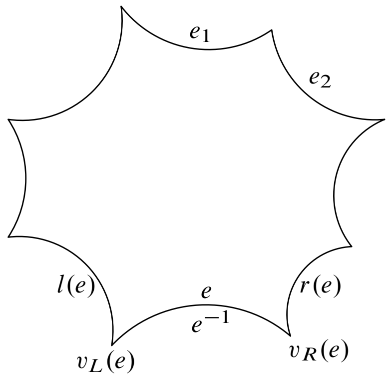


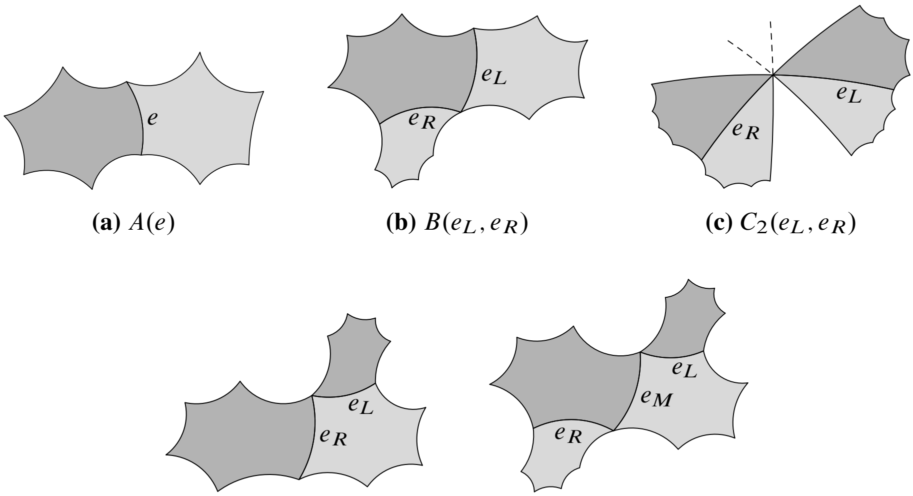
Figure 5. The definitions of vertices $v_L(e)$, $v_R(e)$ and labels $l(e)$, $r(e)$. Labels $e_1, e_2 \in G_0$ are adjacent.

E. $\#\mathcal{S}_- = 2$, $\#\mathcal{S}_+ = 2$, and the graph contains three edges, the two described for type C, and one more. This type is subdivided into the type E_L, where the third edge goes from the left domain in $\mathcal{S}_-$ to the right domain in $\mathcal{S}_+$, and the type E_R, where it goes from the right domain in $\mathcal{S}_-$ to the left domain in $\mathcal{S}_+$. The states of type E correspond to the transitions from one flower to the next one inside a bottle (the "left" and "right" subcases in case (2b) of the lemma).

4.1.2. Labelling. The notation for each state of our Markov chain includes the type $A \dots E$ of the state from the list above and the label(s) corresponding to generators on the sides separating $\mathcal{S}_-$ and $\mathcal{S}_+$. More precisely, these separating sides form a polygonal curve, which is co-oriented from $\mathcal{S}_-$ to $\mathcal{S}_+$, and thus oriented from left to right when looking from $\mathcal{S}_-$ to $\mathcal{S}_+$. The notation for a state is found by recording in order from left to right the labels on the $\mathcal{S}_+$-side of the separating sides; see Definition 4.2 below.

Importantly, the same orientation on this separating curve allows us to define "left" and "right" domains in each of $\mathcal{S}_\pm$, namely, $\mathcal{L}_\pm$ (respectively, $\mathcal{R}_\pm$) is the only domain in $\mathcal{S}_\pm$ that borders the leftmost (respectively, the rightmost) edge of the separating curve. As before, if $\mathcal{S}_\pm$ contains only one domain we have $\mathcal{L}_\pm = \mathcal{R}_\pm$.

Clearly, the labels which appear in the notation of the state must satisfy some restrictions. To express these we introduce some notation regarding vertices, sides, and labels as shown in Figure 5. For any $e \in G_0$, consider the side s_e of $\mathcal{R}$ so that its label inside $\mathcal{R}$ is e. We co-orient this side from the outside to the inside of $\mathcal{R}$, and the corresponding orientation of s_e allows us to define its *left vertex* $v_L(e)$ and the *right vertex* $v_R(e)$. Note that $v_L(e)$ or $v_R(e)$ is undefined if the corresponding end of s_e lies on $\partial\mathbb{D}$. The same notation $v_{L,R}(s)$ will be used for the ends of a co-oriented side s of the tessellation $\mathbf{T}_{\mathcal{R}}$.



(a) $A(e)$ **(b)** $B(e_L, e_R)$ **(c)** $C_2(e_L, e_R)$

(d) $D(e_L, e_R)$ **(e)** $E_R(e_L, e_M, e_R)$

Figure 6. Configurations for states of the Markov coding. The domains in $\mathcal{S}_-$ and $\mathcal{S}_+$ are indicated respectively by the dark and the light shades of gray.

Definition 4.1. The labels e_1 and e_2 are called *adjacent* if the sides of $\mathcal{R}$ with these *outgoing* labels have a common vertex, i.e. either $e_1 = e_2$, or $v_L(e_1^{-1}) = v_R(e_2^{-1})$, or vice versa, see Figure 5.

We now define maps l and r on the set of labels as shown in Figure 5. Informally speaking, we do the following: for $e \in G_0$ we go around $v_L(s_e)$ in the counterclockwise direction, then the next side we cross after s_e has the label $l(e)$ outside $\mathcal{R}$. Similarly, going clockwise around $v_R(s_e)$ we obtain $r(e)$. Formally we define $l(e)$ and $r(e)$ as the labels such that

$$v_R(l(e)^{-1}) = v_L(e), \quad v_L(r(e)^{-1}) = v_R(e).$$

Note that $l(e)$ or $r(e)$ is undefined if the corresponding end of s_e lies on $\partial\mathbb{D}$.

4.1.3. The possible arrangements $\widehat{\Xi}$. The following definition specifies the set $\widehat{\Xi}$ of all possible arrangements of $\mathcal{S}_-$ and $\mathcal{S}_+$ up to the action of G. Later, we will refine this in order to list the actual states Ξ of the Markov chain.

Definition 4.2. The set $\widehat{\Xi}$ consists of the following elements (see Figure 6):

$A(e)$: $\#\mathcal{S}_- = \#\mathcal{S}_+ = 1$, and e is the label on the $\mathcal{S}_+$-side of the common side of $\mathcal{S}_-$ and $\mathcal{S}_+$.

$B(e_L, e_R)$: $\#\mathcal{S}_- = 1, \#\mathcal{S}_+ = 2$, e_L and e_R are $\mathcal{S}_+$-labels on the common sides of $\mathcal{S}_-$ with the left and the right domains in $\mathcal{S}_+$ respectively. Since these sides of $\mathcal{S}_-$ are adjacent, we have that

$$v_L(e_L^{-1}) = v_R(e_R^{-1}).$$

$C_k(e_L, e_R)$: $\#\mathcal{S}_- = \#\mathcal{S}_+ = 2$, and all four domains in $\mathcal{S}_\pm$ share a common vertex v. The label e_L (respectively, e_R) is the $\mathcal{S}_+$-label on the common side of the left (respectively, right) domains in $\mathcal{S}_-$ and $\mathcal{S}_+$, and the sector of the flower at v between these two sides that contains $\mathcal{S}_-$ consists of $2k+1$ petals. Denoting

$$n(e_L, e_R) = n(v) = n(v_R(e_L)) = n(v_L(e_R)),$$

then $1 \le k \le n(e_L, e_R) - 2$, and we have

$$l^{2k+1}(e_L^{-1}) = e_R.$$

$D(e_L, e_R)$: $\#\mathcal{S}_- = 2, \#\mathcal{S}_+ = 1$, e_L and e_R are $\mathcal{S}_+$-labels on the common sides of the left and the right domain in $\mathcal{S}_-$ with the domain $\mathcal{S}_+$. The adjacency condition gives

$$v_R(e_L) = v_L(e_R).$$

$E_{L,R}(e_L, e_M, e_R)$: $\#\mathcal{S}_- = \#\mathcal{S}_+ = 2$. The four domains in $\mathcal{S}_-$ and $\mathcal{S}_+$ do not have a common vertex, and there are three sides separating them. The state E_L represents the case when these sides form an N-shaped line, that is, the left past domain borders both future domains via sides with the $\mathcal{S}_+$-labels e_L and e_M, and the right past domain borders only the right future domain via the side with the label e_R. Thus we have

$$v_L(e_L^{-1}) = v_R(e_M^{-1}), \quad v_R(e_M) = v_L(e_R).$$

The state E_R is the same with left and right inverted: the boundary is И-shaped, and

$$v_R(e_L) = v_L(e_M), \quad v_L(e_M^{-1}) = v_R(e_R^{-1}).$$

4.1.4. Refining the arrangements. It is clear that every configuration of adjacent levels in a thickened path belongs to the set $\widehat{\Xi}$. On the other hand, the set of all possible sequences of configurations cannot be generated by a Markov chain. For example, for a vertex v with $n(v) \ge 3$ it is allowed that $\mathcal{S}_i, \mathcal{S}_{i+1}, \mathcal{S}_{i+2}$ are consecutive petals around v, say, in the counterclockwise direction. Then if e is the label on the future side of $\mathcal{S}_i \cap \mathcal{S}_{i+1}$, the label on the future side of $\mathcal{S}_{i+1} \cap \mathcal{S}_{i+2}$ is $l(e)$, and we have that the transition $A(e) \to A(l(e))$ is admissible. On the other hand, a long sequence

$$A(e) \to A(l(e)) \to A(l(l(e))) \to \cdots$$

is not admissible, since the respective sets $\mathcal{S}_i$ are still the consecutive petals around v, and a thickened path cannot have v on its boundary and contains more than $n(v)$ petals around v.

To solve this problem we endow the states of type A with some additional information based on the following statement.

Proposition 4.3. *Let $\underline{\mathcal{S}}$ be a thickened path. Suppose a vertex $v \in \partial\underline{\mathcal{S}}$ belongs to the boundary of $\mathcal{S}_k$ for $k = i, \dots, j+1$, where $j > i$. Then either*

(1) *$j = i + 2$ and both pairs $(\mathcal{S}_i, \mathcal{S}_{i+1})$, $(\mathcal{S}_{i+1}, \mathcal{S}_{i+2})$ represent E-states (see Figure 7), or*

(2) *for all $k = i+1, \dots, j-1$ the pair $(\mathcal{S}_k, \mathcal{S}_{k+1})$ represents a state of type A, for $k = i$ it represents a state of type A or D, and for $k = j$ it represents a state of type A or B, depending on whether $\mathcal{S}_i$ or $\mathcal{S}_{j+1}$ respectively contains two domains. Moreover, $j - i \leq n(v)$.*

Proof. Suppose first that a vertex $v \in \mathbb{D}$ belongs to three consecutive levels $\mathcal{S}_k$, $\mathcal{S}_{k+1}$, $\mathcal{S}_{k+2}$ of the thickened path, and $\#\mathcal{S}_{k+1} = 2$. If, say, v belongs to $\partial_L\underline{\mathcal{S}}$, then $\partial_L\mathcal{S}_{k+1}$ consists of the vertex v only. Then $\mathcal{R}$ must be compact and hence $N(\mathcal{R}) \geq 4$.

The two domains in $\mathcal{S}_{k+1}$ must be petals of $\mathcal{F}_u$ for some vertex $u \neq v$. The left domain $\mathcal{L}_{k+1}$ in the level $\mathcal{S}_{k+1}$ must meet $\mathcal{S}_{k+2}$ along at least two sides, one emanating from v and one from u. Moreover, these sides, being common sides of one domain in two levels, must themselves be adjacent, hence meet in a vertex w say. By the same argument, $\mathcal{L}_{k+1}$ meets $\mathcal{S}_k$ along two adjacent sides which meet in a vertex w' say. Hence, $N(\mathcal{R}) = 4$, and each of $\mathcal{L}_{k+1} \cap \mathcal{S}_k$ and $\mathcal{L}_{k+1} \cap \mathcal{S}_{k+2}$ contains two sides, see Figure 7. We deduce that the states representing the pairs $(\mathcal{S}_k, \mathcal{S}_{k+1})$, $(\mathcal{S}_{k+1}, \mathcal{S}_{k+2})$ are of types E_R and E_L respectively. In particular, this means that v cannot belong to four consecutive levels of the thickened path, and we see that were are in case (1) of the proposition.

It remains to consider the case when $\#\mathcal{S}_k = 1$ for all $k = i+1, \dots, j-1$. If $\mathcal{S}_i$ contains two domains and $\mathcal{S}_{i+1}$ contains one, then $(\mathcal{S}_i, \mathcal{S}_{i+1})$ must be of type D; otherwise $\mathcal{S}_i$ contains one domain and $(\mathcal{S}_i, \mathcal{S}_{i+1})$ must be of type A, with a similar argument for the transition $(\mathcal{S}_j, \mathcal{S}_{j+1})$, which implies case (2) of the proposition. ∎

Remark 4.4. Assumption 1.1 yields that in the first case in this proposition we have $n(v) \geq 3$. Indeed, $N(\mathcal{R}) = 4$, $\mathcal{R}$ is compact, and for the common vertex u of the two domains in $\mathcal{S}_{k+1}$ we have $n(u) = 2$; see Figure 7.

Let $(\mathcal{S}_k, \mathcal{S}_{k+1})$ form a configuration $A(e)$ and s_k be the common side of $\mathcal{S}_k$ and $\mathcal{S}_{k+1}$. Then one can define four numbers $i_{\pm,L}$ and $i_{\pm,R}$ as follows: $i_{-,\alpha}$ (resp., $i_{+,\alpha}$), $\alpha \in \{L, R\}$, is the number of $m \leq k$ (resp., $m \geq k+1$) such that $\mathcal{S}_m$ contains $v_\alpha(s_k)$. If the vertex $v_\alpha(s_k)$ is not defined, we set $i_{\pm,\alpha} = 1$.

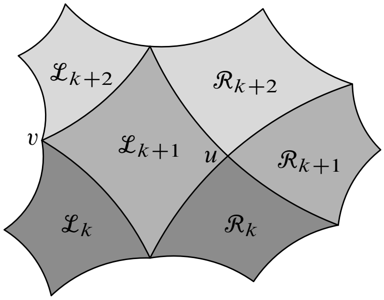


Figure 7. Two consecutive E states, see Proposition 4.3.

Note that it is not possible to have $i_{-,L} > 1$ and $i_{-,R} > 1$ simultaneously: these conditions mean that both $\partial_L \mathcal{S}_k$ and $\partial_R \mathcal{S}_k$ each consist of a vertex only, say v and v'. Indeed since the sides of $\mathcal{S}_k$ adjacent to v and v' must both be in common with $\mathcal{S}_{k+1}$, and since by assumption the transition is of type A, this would mean that vv' is the common side s of $\mathcal{S}_k, \mathcal{S}_{k+1}$. But the sides t, t' of $\mathcal{S}_k$ adjacent to s must also both be adjacent to $\mathcal{S}_{k-1}$. Now $\ell(t), \ell(t')$ cannot meet, otherwise we have a triangle in $\mathbf{T}_{\mathcal{R}}$, so they separate $\mathcal{S}_{k-1}$ into two disconnected components which is impossible. The same argument applies to $i_{+,L}$ and $i_{+,R}$.

The convexity of $\underline{\mathcal{S}}$ at $v_\alpha(s_k)$, $\alpha = L, R$, implies that $i_{-,\alpha} + i_{+,\alpha} \le n(v_\alpha(s_k))$. It follows that the configuration $A(e)$ can be subdivided as follows (see Figure 8):

$A_0(e)$: all four $i_{\pm,L/R}$ equal one.

$A_L[i_-, i_+](e)$: here $i_{-,L} = i_-$, $i_{+,L} = i_+$, $i_{-,R} = i_{+,R} = 1$, and the indices $i_\pm$ should satisfy

$$3 \le i_- + i_+ \le n(v_L(e)).$$

$A_R[i_-, i_+](e)$: symmetric to the previous case; here

$$3 \le i_- + i_+ \le n(v_R(e)).$$

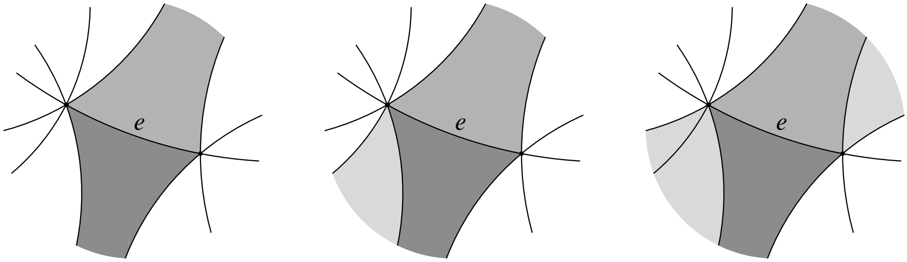
$A_{LR}[i_-, i_+](e)$: here $i_{-,L} = i_-$, $i_{+,R} = i_+$, and $i_{+,L} = i_{-,R} = 1$. The conditions on the indices $i_\pm$ are

$$2 \le i_- \le n(v_L(e)) - 1, \quad 2 \le i_+ \le n(v_R(e)) - 1.$$

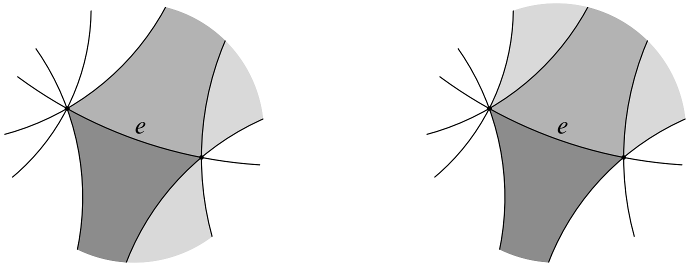
$A_{RL}[i_-, i_+](e)$: symmetric to the previous case; here

$$2 \le i_- \le n(v_R(e)) - 1, \quad 2 \le i_+ \le n(v_L(e)) - 1.$$

Remark 4.5. If $N(\mathcal{R}) = 3$ and $\mathcal{R}$ has a compact side (as is the case for example for the classical fundamental domain for the group $SL(2, \mathbb{Z})$), some of these states may



(a) $A_0(e)$ **(b)** $A_L[2,1](e)$ **(c)** $A_{LR}[3,2](e)$



(d) Impossible "$A_R[2,2](e)$": $i_- + i_+ > n(v_R(e))$, thus convexity in $v_R(e)$ fails.

(e) Impossible subtype: both $i_{+,L}$ and $i_{+,R}$ are greater than 1.

Figure 8. Possible (a–c) and impossible (d–e) subtypes for type A states. The dark and medium gray domains are respectively the past and the future domain for the current state, light gray domains are other domains from the thickened path. In the figure, $n(v_L(e)) = 4$, $n(v_R(e)) = 3$.

be absent. Namely, let s be the only compact side of $\mathcal{R}$ and g be its label outside of $\mathcal{R}$. If $(\mathcal{S}_k, \mathcal{S}_{k+1})$ has the form $A(g)$, then $\mathcal{S}_{k+2}$ must contain at least one of the domains adjacent to the sides of $\mathcal{S}_{k+1}$, hence either $i_{+,L}$ or $i_{+,R}$ is greater than one so there are no states $A_0(g)$. Similarly, either $i_{-,L} > 1$ or $i_{-,R} > 1$. This case needs special consideration in several statements below, and we usually refer to it as "the special case from Remark 4.5".

Note that even in this case the list of $A_{\dots}(g)$-states is not completely empty. Indeed, since s is the only compact side, it must be paired to itself: $g = g^{-1}$. Therefore, the ends of s are swapped by the action of g, hence $n(v_L(g)) = n(v_R(g)) = n$. Let α and β be the angles of $\mathcal{R}$ at the ends of s. Consider the flower around a vertex $v \in \mathbb{D}$. Note that the sides incident to v are alternately compact and non-compact, and the angles between these sides are alternately α and β. Therefore, $n\alpha + n\beta = 2\pi$. On the other hand, the sum of angles in the hyperbolic triangle $\mathcal{R}$ is $\alpha + \beta < \pi$. Consequently, $n \geq 3$, and for example, the state $A_{LR}[2,2](g)$ is allowed.

4.1.5. The states of the Markov chain. Finally, we are able to list the states Ξ.

Definition 4.6. The *set of states* Ξ of our Markov chain is the set of all states of types B, C, D, E from the set $\widehat{\Xi}$ and of all subtypes of type A states enumerated in the previous list. We denote the projection from Ξ to $\widehat{\Xi}$ by π.

Finally, let us define sets $\Xi_S, \Xi_F \subset \Xi$ as follows:

$$\begin{aligned}\Xi_S &= \{A_0(e), A_L[1,i_+](e), A_R[1,i_+](e), B(e_L,e_R)\},\\ \Xi_F &= \{A_0(e), A_L[i_-,1](e), A_R[i_-,1](e), D(e_L,e_R)\},\end{aligned}$$

where the parameters $i_\pm$, e, e_L, e_R admit all possible values. In the special case from Remark 4.5 these definitions are amended as follows: if $g = g^{-1}$ is the label on the compact side of $\mathcal{R}$, we include $A_{LR}[2,i_+](g), A_{RL}[2,i_+](g)$ in Ξ_S and $A_{LR}[i_-,2](g), A_{RL}[i_-,2](g)$ in Ξ_F in place of the A states listed above.

4.2. The admissible transitions

Definition 4.7. The set of *admissible transitions* in our Markov coding is enumerated in the following list. We denote by Π the corresponding $\Xi \times \Xi$ adjacency matrix and write $j \to j'$ if the transition from j to j' is admissible (thus $\Pi_{jj'} = 1$). (Recall that the adjacency of labels was described in Definition 4.1; in particular, every label is adjacent to itself.)

- $$A_0(e) \to \begin{cases} A_0(e') & \text{for } e' \text{ non-adjacent to } e^{-1},\\ A_L[1,i_+](e') & \text{for } e' \text{ non-adjacent to } e^{-1}, \text{ any admissible } i_+,\\ A_R[1,i_+](e') & \text{for } e' \text{ non-adjacent to } e^{-1}, \text{ any admissible } i_+,\\ B(e_L,e_R) & \text{for any } e_L, e_R \text{ non-adjacent to } e^{-1}.\end{cases}$$
- If $i_+ > 1$ then
$$A_L[i_-,i_+](e) \to \begin{cases} A_L[i_- + 1, i_+ - 1](l(e)), & \\ A_{LR}[i_- + 1, j_+](l(e)), & \left|\begin{array}{l}\text{if } i_+ = 2,\\ \text{for any admissible } j_+,\end{array}\right.\\ B(l(e), r(l(e))^{-1}), & \text{if } i_+ = 2.\end{cases}$$
- $A_L[i_-,1](e) \to$ (the same cases as for $A_0(e)$).
- $A_{RL}[i_-,i_+](e) \to$ (the same cases as for $A_L[1,i_+](e)$).
- The transitions for the A_R- and A_{LR}-states are similar with the exchange of left and right.
- $B(e_L,e_R) \to C_1(r(e_L), l(e_R))$ if $n(e_L,e_R) \geq 3$, if $n(e_L,e_R) = 2$ the transitions for $B(e_L,e_R)$ are the same as for $C_{n(e_L,e_R)-2}(e_L,e_R)$ below.
- $C_i(e_L,e_R) \to C_{i+1}(r(e_L), l(e_R))$, for $i < n(e_L,e_R) - 2$,

- $C_{n(e_L,e_R)-2}(e_L,e_R) \to \begin{cases} D(r(e_L), l(e_R)), \\ E_L(r(r(e_L)^{-1}), r(e_L), l(e_R)), \\ E_R(r(e_L), l(e_R), l(l(e_R)^{-1})). \end{cases}$
- $D(e_L,e_R) \to \begin{cases} A_0(e') & \text{for } e' \text{ non-adjacent to } e_L^{-1}, e_R^{-1}, \\ \left.\begin{matrix} A_L[1,i_+](e') \\ A_R[1,i_+](e') \end{matrix}\right\} & \left| \begin{matrix} \text{for } e' \text{ non-adjacent to } e_L^{-1}, e_R^{-1}, \\ \text{any admissible } i_+, \end{matrix} \right. \\ B(e'_L, e'_R) & \left| \begin{matrix} \text{for } e'_L, e'_R \text{ either not adjacent to } e_L^{-1}, e_R^{-1}, \\ \text{or adjacent via a vertex } v \text{ with } n(v) > 2, \\ \text{or adjacent via two such vertices} \\ \text{(the latter is possible for } N(\mathcal{R}) = 4 \text{ only),} \end{matrix} \right. \\ \left.\begin{matrix} A_L[2,i_+](l(e_L)) \\ A_{LR}[2,i_+](l(e_L)) \end{matrix}\right\} & \left| \begin{matrix} \text{if } n(v_L(e_L)) > 2, \\ \text{for any admissible } i_+, \end{matrix} \right. \\ \left.\begin{matrix} A_R[2,i_+](r(e_R)) \\ A_{RL}[2,i_+](r(e_R)) \end{matrix}\right\} & \left| \begin{matrix} \text{if } n(v_R(e_R)) > 2, \\ \text{for any admissible } i_+. \end{matrix} \right. \end{cases}$
- $E_L(e_L, e_M, e_R)$ has the same set of transitions as $B(e_L, e_M)$.
- $E_R(e_L, e_M, e_R)$ has the same set of transitions as $B(e_M, e_R)$.

Remark 4.8. The reader may check that the transition $j \to j'$ is admissible if and only if it satisfies the following three sets of restrictions. First, there should be three unions $\mathcal{S}_-$, $\mathcal{S}_+$, $\mathcal{S}_{++}$ of fundamental domains such that $(\mathcal{S}_-, \mathcal{S}_+)$ represents $\pi(j)$, $(\mathcal{S}_+, \mathcal{S}_{++})$ represents $\pi(j')$, and $\mathcal{S}_-$ and $\mathcal{S}_{++}$ have no common domains. Second, the indices of A-states should be compatible: for example, if sides $s = \mathcal{S}_- \cap \mathcal{S}_+$ and $s' = \mathcal{S}_+ \cap \mathcal{S}_{++}$ have the same left end, then the left indices for j and j' should satisfy $i'_{\pm,L} = i_{\pm,L} \mp 1$. Finally, the boundary of a thickened path should be convex at any of its vertices v, that is, no more than $n(v)$ domains should belong to $\underline{\mathcal{S}}$. This is mostly guaranteed by the inequalities for the indices $i_\pm$ given in the list of subtypes of A-states above. Part 1 of the proof of Theorem 4.10 discusses this in more detail.

4.3. Correspondence with thickened paths

We now come to the important result that admissible sequences of states do indeed correspond to thickened paths. Precisely, define

$$\mathbf{P}_{N-1}^{S\to F} = \{(j_0, \dots, j_{N-1}) \subset \Xi^N : j_0 \in \Xi_S,\ j_{N-1} \in \Xi_F, \\ \Pi_{j_k, j_{k+1}} = 1 \text{ for } k = 0, \dots, N-2\}. \quad (2)$$

We will show that $\mathbf{P}_{N-1}^{S\to F}$ is in $1:1$-correspondence with the set of the thickened paths of length N.

Definition 4.9. Say that a sequence $\underline{j}$ of states *generates* a sequence of domains $\underline{\mathcal{S}}$ if for each k the pair $(\mathcal{S}_k, \mathcal{S}_{k+1})$ represents the configuration $\pi(j_k)$.

Theorem 4.10. *Let $\underline{\mathcal{S}} = (\mathcal{S}_0, \dots, \mathcal{S}_N)$ be a thickened path starting at $\mathcal{R}$. Then there exists a unique sequence of states $\underline{j} \in \mathbf{P}_{N-1}^{S\to F}$ which generates the sequence $\underline{\mathcal{S}}$. Moreover, this mapping of thickened paths of length N starting in $\mathcal{R}$ to the set $\mathbf{P}_{N-1}^{S\to F}$ is a bijection.*

Proof. The proof is split into two parts. It is rather straightforward to see that every thickened path $\underline{\mathcal{S}}$ is generated by a unique admissible sequence $\underline{j}$: this follows from the local properties of the thickened path, mostly from its convexity at boundary vertices. Thus the map $\underline{\mathcal{S}} \mapsto \underline{j}$ is well-defined and also injective: if for two thickened paths $\underline{\mathcal{S}}$ and $\underline{\mathcal{S}}'$ we have $\mathcal{S}_i = \mathcal{S}_i'$ for $i = 0, \dots, m-1$ and $\mathcal{S}_m \neq \mathcal{S}_m'$ then for the generating sequences we have $\pi(j_{m-1}) \neq \pi(j_{m-1}')$. The second part of the proof, the surjectivity of this map, is more difficult: we need to deduce a *global* property of the sequence $\underline{\mathcal{S}}$ from its local behavior given by Definition 4.7. The precise global property is given by Claim 4.11, informally it says that the levels of $\underline{\mathcal{S}}$ cannot first "leave" a vertex v and then "return" back to it.

Part 1. Every thickened path $\underline{\mathcal{S}} = (\mathcal{S}_0, \dots, \mathcal{S}_N)$ can be generated by a unique sequence $\underline{j} \in \mathbf{P}_{N-1}^{S\to F}$.

Step 1. By hypothesis, each pair $(\mathcal{S}_k, \mathcal{S}_{k+1})$ in a thickened path $\underline{\mathcal{S}}$ represents a unique configuration $\hat{j}_k \in \hat{\Xi}$. Further, for every configuration of type A one can recover the indices $i_{\pm,L/R}$ as described above, thus arriving at the states j_k with $\pi(j_k) = \hat{j}_k$. Note that if $\hat{j}_0 = A(e)$ then the state j_0 has $i_{-,L} = i_{-,R} = 1$, so $j_0 \in \Xi_S$. In the special case from Remark 4.5 we need to amend these indices as follows: if $\pi(j_0) = A(g)$, where g is the label on the compact side, then either $i_{+,L} \geq 2$ or $i_{+,R} \geq 2$. In the former case we then set $i_{-,L} = 1$, $i_{-,R} = 2$ and in the latter we set $i_{-,L} = 2$, $i_{-,R} = 1$; this corresponds to the addition of the "virtual domain" $\mathcal{S}_{-1}$ to our thickened path. Note that the state $A_{RL}[2, i_+](g)$ has the same set of allowed transitions as the non-existent "state $A_L[1, i_+](g)$".

Now we have to check that all transitions $j_k \to j_{k+1}$ are admissible. There are three types of restrictions on the pair of states (j_k, j_{k+1}) in the list of Definition 4.7.

First, there are restrictions on the configurations $\hat{j}_k, \hat{j}_{k+1}, \pi(j_k) = \hat{j}_k$. For example, if $\hat{j}_k = C_l(e_L, e_R)$, then $\mathcal{S}_+ = \mathcal{S}_{k+1}$ is a pair of petals meeting at a vertex v with $2n(v) - 2l - 1$ petals in the "future" sector in $\mathcal{F}_v$. Therefore, if $l < n(v) - 2$ by Lemma 3.9 we see that $\mathcal{S}_{++} = \mathcal{S}_{k+2}$ is the pair of petals in the "future" sector adjacent to $\mathcal{S}_+$, hence

$$\hat{j}_{k+1} = C_{l+1}(e_L', e_R')$$

with $e_L' = r(e_L)$, $e_R' = l(e_R)$.

Second, there are restrictions on the indices $i_\pm$ of A-states. If j_k is an A-state with $i_{+,L} > 1$ then $\mathcal{S}_{k+2}, \dots, \mathcal{S}_{k+i_{+,L}+1}$ should contain the consecutive petals around the vertex $v_L(s_k)$ going from $\mathcal{S}_{k+1}$ in the counterclockwise direction. Hence, by Proposition 4.3 the sequence $(j_{k+1}, \dots, j_{k+i_{+,L}})$ is either of type $(A, \dots, A)$ or $(A, \dots, A, B)$. Therefore, if $i_{+,L} > 2$ then $\hat{j}_{k+1} = A(l(e))$, and for its indices $i'_{\pm,L}$, we have

$$i'_{-,L} = i_{-,L} + 1, \quad i'_{+,L} = i_{+,L} - 1.$$

Similarly, if $i_{+,L} = 2$, we have either $\hat{j}_{k+1} = B(l(e), r(l(e))^{-1})$ or $\hat{j}_{k+1} = A(l(e))$ with the same relations for $i'_{\pm,L}$. In the latter case $i'_{+,L} = 1$ so we have two subcases: either

$$i'_{+,R} = 1 \quad \text{or} \quad i'_{+,R} > 1,$$

which correspond in Definition 4.7 to the transitions to A_L- and A_{LR}-states respectively.

Finally, there are restrictions related to the convexity of $\partial \underline{\mathcal{S}}$ (see Proposition 3.3) which we need to check for the boundary vertices v that are incident to at least three levels in $\underline{\mathcal{S}}$. These cases are enumerated in Proposition 4.3. In the cases when the corresponding sequence of states contains A-states, the convexity is guaranteed by the inequalities on the indices $i_\pm$ for these states, so we need to consider only the cases when (j_k, j_{k+1}) have types (E_L, E_R), (E_R, E_L), and (D, B). In the first two cases the convexity at v holds, see Remark 4.4. The remaining case (D, B) is specially mentioned in Definition 4.7: if v is a common vertex of $\mathcal{S}_{k-1} \cap \mathcal{S}_k$ and $\mathcal{S}_k \cap \mathcal{S}_{k+1}$, we require that $n(v) > 2$.

Step 2. It remains to prove that the above-constructed sequence $\underline{j}$ is the only one in $\mathbf{P}_{N-1}^{S \to F}$ that generates $\underline{\mathcal{S}}$. Namely, we have to check that the indices $i_{\pm,L/R}$ for A-states cannot be chosen in a different way. One can see that the "past" indices $i_{-,L/R}$ for the state j_k are uniquely defined by the configurations $\pi(j_{k-1}), \pi(j_k)$ and by the past indices for the state j_{k-1} (assuming j_{k-1} has type A). The past indices for $j_0 \in \Xi_S$ are $i_{-,L/R}(j_0) = 1$, hence one can successively find these indices for all successive states $j_1, \dots, j_{N-1}$. Similarly, the "future" indices $i_{+,L/R}(j_k)$ are successively found starting from the end of the sequence:

$$i_{+,L/R}(j_{N-1}) = 1.$$

The special case from Remark 4.5 again needs separate consideration if $\hat{j}_0 = A(g)$. Here $\hat{j}_1 = A(e)$, where e is a label on a non-compact side of $\mathcal{R}$, and either $e = l(g)$ or $e = r(g)$. In the former case, say, this yields $i_{+,L}(j_0) \geq 2$, thus $j_0 \in \Xi_S$ implies $j_0 = A_{RL}[2, i_+](g)$ with some i_+. Therefore, we find the past indices for j_0 and can now proceed as in the general case.

Part 2. Every sequence $\underline{j} \in \mathbf{P}_{N-1}^{S \to F}$ generates a thickened path.

The proof is in several steps. In Step 1 we reformulate what needs to be proved as Claim 4.11 and Steps 2–5 are devoted to proving this claim.

Step 1. We begin by constructing the $\mathcal{S}_k$'s inductively, starting with $\mathcal{S}_0 = \mathcal{R}$. To define $\mathcal{S}_{k+1}$ we take a configuration $(\mathcal{S}_-, \mathcal{S}_+)$ representing $\pi(j_k)$, choose h such that $h\mathcal{S}_- = \mathcal{S}_k$ and define $\mathcal{S}_{k+1} = h\mathcal{S}_+$. The choice of such an h is possible for $k = 0$ since all states from Ξ_S have only one "past" domain, and for $k \geq 1$ this is possible since if $j \to j'$ is admissible, and $(\mathcal{S}_-, \mathcal{S}_+)$, $(\mathcal{S}'_-, \mathcal{S}'_+)$ are configurations representing $\pi(j)$ and $\pi(j')$, respectively, then $\mathcal{S}_+$ and $\mathcal{S}'_-$ can be translated to each other by an element of G: $\mathcal{S}'_- = h\mathcal{S}_+$. Moreover, as described in Section 4.1.2, one can define left domains $\mathcal{L}_\pm$ and right domains $\mathcal{R}_\pm$ in $\mathcal{S}_\pm$, as well as $\mathcal{L}'_\pm, \mathcal{R}'_\pm$ in $\mathcal{S}'_\pm$, and these definitions agree: $\mathcal{L}'_- = h\mathcal{L}_+$, $\mathcal{R}'_- = h\mathcal{R}_+$. Thus we have defined all levels $\mathcal{S}_k$ and the domains $\mathcal{L}_k, \mathcal{R}_k \subset \mathcal{S}_k$ for all $k = 0, \dots, N$. Observe also that $\mathcal{S}_N$ contains only one domain since $j_{N-1} \in \Xi_F$.

We need to show that

$$\bigcup \underline{\mathcal{S}} = \bigcup_{k=0}^{N} \mathcal{S}_k$$

is the thickened path from $\mathcal{S}_0$ to $\mathcal{S}_N$ and that the $\mathcal{S}_k$'s are its levels. This is obtained from the following statement, which will be established below.

Claim 4.11. *The following assertions hold.*

(1) *The intersection $\mathcal{S}_l \cap \mathcal{S}_m$ with $l \neq m$ contains no fundamental domains, and contains sides only if $|l - m| = 1$.*

(2) *The set $\bigcup \underline{\mathcal{S}}$ is convex.*

To prove the result given Claim 4.11, let $\underline{\mathcal{T}}$ be the actual thickened path from $\mathcal{S}_0$ to $\mathcal{S}_N$. The second item in the claim together with Proposition 3.3 gives

$$\bigcup \underline{\mathcal{T}} \subset \bigcup \underline{\mathcal{S}}.$$

Now consider some shortest path $\underline{\mathcal{A}} = (\mathcal{A}_0 = \mathcal{S}_0, \mathcal{A}_1, \dots, \mathcal{A}_M = \mathcal{S}_N)$ from $\mathcal{S}_0$ to $\mathcal{S}_N$. Then,

$$\mathcal{A}_m \subset \bigcup \underline{\mathcal{T}} \subset \bigcup \underline{\mathcal{S}}.$$

Define d_m by $\mathcal{A}_m \subset \mathcal{S}_{d_m}$ so that in particular $d_M = N$. The first item of the claim implies that $|d_m - d_{m-1}| = 1$, hence $d_M \leq M$. On the other hand, the two length N paths $\underline{\mathcal{L}} = (\mathcal{L}_0, \dots, \mathcal{L}_N)$ and $\underline{\mathcal{R}} = (\mathcal{R}_0, \dots, \mathcal{R}_N)$ connect $\mathcal{S}_0$ and $\mathcal{S}_N$, and hence $M \leq N$. Therefore, $M = N$ and $\underline{\mathcal{L}}, \underline{\mathcal{R}}$ are shortest paths connecting $\mathcal{S}_0$ and $\mathcal{S}_N$. Thus $\underline{\mathcal{L}}, \underline{\mathcal{R}}$ are contained in $\bigcup \underline{\mathcal{T}}$. But

$$\bigcup \underline{\mathcal{S}} = \underline{\mathcal{L}} \cup \underline{\mathcal{R}},$$

and hence $\bigcup \underline{\mathcal{S}} \subset \bigcup \underline{\mathcal{T}}$.

We now turn to the proof of Claim 4.11. To do this, we will construct a nested sequence of convex regions $\mathcal{D}_k \supset \mathcal{D}_{k+1}$ whose intersection is $\bigcup \underline{\mathcal{S}}$, with some further properties listed in Claim 4.12. Step 2 sets up notation and Step 3 proves that $\bigcup \underline{\mathcal{S}}$ is locally convex. Step 4 states Claim 4.12 and proves that this is sufficient to deduce Claim 4.11. Lastly, Step 5 proves Claim 4.12.

Step 2. We begin with some notation. From here to the end of the proof we consider all domains, curves, vertices, etc. to lie in the closure $\overline{\mathbb{D}}$ of the hyperbolic plane.

Let $s_k = \mathcal{S}_k \cap \mathcal{S}_{k+1}$ be the curve separating $\mathcal{S}_k$ and $\mathcal{S}_{k+1}$. This curve is the union of no more than three sides of domains in $\mathbf{T}_{\mathcal{R}}$, and, as one can see from the adjacency matrix, the curves s_{k-1} and s_k contain no common sides of $\partial \mathcal{S}_k$. Moreover, $\partial \mathcal{S}_k \setminus (s_{k-1} \cup s_k)$ is a union of two curves, one joining the left ends of s_{k-1} and s_k, and the other joining the right ends; it is possible that one or both of these curves consists of a single vertex and no sides, see the discussion following Remark 4.4. We denote these curves by $\partial_L \mathcal{S}_k$ and $\partial_R \mathcal{S}_k$ respectively. For $k = 0, N$ there is only one s_j to remove, so we define

$$\partial_O \mathcal{S}_0 = \partial \mathcal{S}_0 \setminus s_0, \quad \partial_O \mathcal{S}_N = \partial \mathcal{S}_N \setminus s_{N-1}.$$

For $k = 1, \dots, N-1$, we also denote $\partial_O \mathcal{S}_k = \partial_L \mathcal{S}_k \cup \partial_R \mathcal{S}_k$.

Next, let us orient the curves $\partial_{\dots} \mathcal{S}_k$ in such a way that $\mathcal{S}_k$ lies locally to the left of the curve when moving in positive direction. Then all these curves can be joined into one closed oriented curve in the following order, so that the end of each curve coincides with the beginning of the next:

$$\begin{aligned} &\dots, \partial_O \mathcal{S}_N, \partial_L \mathcal{S}_{N-1}, \dots, \partial_L \mathcal{S}_{k+1}, \\ &\qquad \underbrace{\partial_L \mathcal{S}_k, \dots, \partial_L \mathcal{S}_1, \partial_O \mathcal{S}_0, \partial_R \mathcal{S}_1, \dots, \partial_R \mathcal{S}_k}_{\partial_O \mathcal{S}_0^k}, \partial_R \mathcal{S}_{k+1}, \dots, \partial_R \mathcal{S}_{N-1}, \partial_O \mathcal{S}_N, \dots \end{aligned}$$

Denote the part of this curve bracketed in the formula by $\partial_O \mathcal{S}_0^k$.

Step 3. Since our object is to show that the region $\bigcup \underline{\mathcal{S}}$ is convex, we start by checking local convexity of the closed oriented curve above, that is, that for a vertex v of the boundary there are no more than $n(v) - 2$ consecutive boundary curves $\partial_\alpha \mathcal{S}_k$ which consist solely of the vertex v and no sides.

First of all, let us find all transitions $j_{k-1} \to j_k$ such that $\partial_L \mathcal{S}_k$ contains only a vertex. This is an analogue to Proposition 4.3 but for sequences of levels generated by admissible steps in our Markov coding instead of by thickened paths. In all steps of the arguments which follow it is understood that there is a similar statement with the indices L, R interchanged.

Assume first that $\mathcal{S}_k$ contains only one domain. Then the outgoing indices on the sides of the curve $s_k \subset \partial \mathcal{S}_k$ are the indices from the state j_k, while the outgoing indices on the sides of $s_{k-1} \subset \partial \mathcal{S}_k$ are the indices from j_{k-1} inverted (see Section 4.4 below for a precise definition). It follows that in all "non-adjacent" cases in the list in the Definition 4.7, the curves s_{k-1} and s_k have no common vertices, as for example in the transition $A_0(e) \to A_0(e')$ with e non-adjacent to e'. The remaining cases are

$$(A_L, A_{RL}, \text{ or } D) \to (A_L, A_{LR}, \text{ or } B).$$

Now suppose that $\mathcal{S}_k$ contains two domains $\mathcal{L}_k$ and $\mathcal{R}_k$. Referring to the configurations in Figure 6, we see that the sections of the curves s_{k-1} and s_k contained in $\partial\mathcal{L}_k$ meet in a vertex of $\mathcal{L}_k$ forming a continuous segment σ in $\partial\mathcal{L}_k$. Moreover, $\partial_L \mathcal{S}_k$ is the complement of σ in the boundary $\partial\mathcal{L}_k$ of $\mathcal{L}_k$. However, if $\partial_L \mathcal{S}_k$ consists of a single vertex v, then the sides of $\partial\mathcal{L}_k$ adjacent to v are contained in $s_{k-1} \cup s_k$ so that $\partial_L \mathcal{S}_k \subset s_{k-1} \cup s_k$, from which it follows that $\mathcal{R}$ is compact (note that none of the interior vertices of s_k lie on $\partial\mathbb{D}$) and hence that $N(\mathcal{R}) \geq 4$. On the other hand, each of s_{k-1} and s_k contains at most two sides of $\mathcal{L}_k$, so they can cover $\partial\mathcal{L}_k$ entirely only if $N(\mathcal{R}) = 4$ and they contain exactly two sides each. Thus the transition $(j_{k-1} \to j_k)$ must be $(E_R \to E_L)$ as shown in Figure 7.

Now consider a segment $\{k, \dots, k+m-1\}$, with

$$v = \partial_L \mathcal{S}_k = \partial_L \mathcal{S}_{k+1} = \cdots = \partial_L \mathcal{S}_{k+m-1},$$

and the corresponding sequence of states $j_{k-1} \to \cdots \to j_{k+m-1}$. We are going to prove that $m \leq n(v) - 2$.

If there are no A_L-states in this sequence, then we are in one of the cases above, in particular $m = 1$. Thus we need to check that $n(v) > 2$. The case $E_R \to E_L$ is shown in Figure 7 (with the levels numbered $k, k+1, k+2$ there instead of $k-1, k, k+1$ as here). The common vertex u of $\mathcal{L}_k$ and $\mathcal{R}_k$ has $n(u) = 2$, so for the opposite vertex v of L_k we have $n(v) \geq 3$. For a $D \to B$ transition, Definition 4.7 explicitly specifies that $n(v) > 2$. Finally, if $j_{k-1} = A_{RL}[i_-, 2](e)$ or $j_k = A_{LR}[2, i_+](e)$, we have

$$n(v) = n(v_L(e)) \geq 3$$

from the definition of the $A_{\dots}$-states above.

Now assume that there are A_L-states in the sequence. Then one can see that its i_--indices grow by one with each step to the right in the sequence $j_{k-1} \to \cdots \to j_{k+m-1}$ as long as we remain in an A_L-state. Moreover, if j_{k-1} is of type A_{RL} or D then $j_k = A_L[2, i_+]$. Therefore, in any case for which j_{k-1+s} is of type A_L, it has $i_- \geq s+1$. Similarly, the i_+-index of A_L-states grows by one with each step from right to left, so we have $i_+ \geq m - s + 1$. Therefore, considering any $j_{k-1+s} = A_L[i_-, i_+]$, we have

$$m + 2 \leq i_- + i_+ \leq n(v_L(e)) = n(v),$$

and the statement is proven.

Step 4. We now construct the sets $\mathcal{D}_k$ referred to above. Let $\mathcal{H}_k$ be the collection of all half-planes H such that ∂H contains a side of $\partial_O \mathcal{S}_k$ and H contains $\mathcal{S}_k$. Denote

$$\mathcal{H}_0^k = \bigcup_{j=0}^{k} \mathcal{H}_j, \quad \mathcal{D}_k = \bigcap_{H \in \mathcal{H}_k^0} H.$$

By construction, $\mathcal{D}_k$ is a union of fundamental domains and is convex, moreover clearly $\mathcal{D}_k \subset \mathcal{D}_{k-1}$. We want to show that

$$\bigcap \mathcal{D}_k = \bigcup \underline{\mathcal{S}}.$$

From the definition of the D_k, we will use the local convexity of the curve going around $\underline{\mathcal{S}}$ to see that each D_k contains $\bigcup \underline{\mathcal{S}}$, and further, that D_{k+1} can obtained from D_k by removing half-planes until we touch all segments of $\partial_L \mathcal{S}_{k+1} \cup \partial_R \mathcal{S}_{k+1}$. In more detail:

Claim 4.12. *For $k = 0, \dots, N-1$ we have the following (see Figure 9):*

(i) *The curve s_k lies in the interior of $\mathcal{D}_k$, joins two points on its boundary, and divides $\mathcal{D}_k$ into two parts.*

(ii) *One of these parts is the union of all $\mathcal{S}_j$ with $j = 0, \dots, k$. We denote this part by $\mathcal{D}_k^-$ and the other part by $\mathcal{D}_k^+$.*

(iii) *$\partial \mathcal{D}_k^- = s_k \cup \partial_O \mathcal{S}_0^k$, while $\partial \mathcal{D}_k^+$ consists of s_k and two rays $\gamma_{k,L}$, $\gamma_{k,R}$ that are continuations of the first and the last sides in $\partial_O \mathcal{S}_0^k$ beyond the ends of s_k. These rays do not intersect inside $\mathbb{D}$.*

(iv) *$\mathcal{D}_k \subset \mathcal{D}_{k-1}$, or, more precisely, $\mathcal{S}_k \cup \mathcal{D}_k^+ \subset \mathcal{D}_{k-1}^+$. Finally, for $k = N$, we have $\mathcal{D}_N = \bigcup \underline{\mathcal{S}}$.*

To deduce Claim 4.11 from this statement note that if $l = m + b$, $b > 0$, then

$$\mathcal{S}_{m+b} \subset \mathcal{D}_{m+b-1}^+ \subset \mathcal{D}_{m+b-2}^+ \subset \cdots \subset \mathcal{D}_m^+,$$

hence $\mathcal{S}_l$ has no domain in common with $\mathcal{S}_m \subset \mathcal{D}_m^-$. Moreover, if $\mathcal{S}_m \cap \mathcal{S}_{m+b}$ contains a common side s, then $s \subset \mathcal{D}_m^+ \cap \mathcal{D}_m^- = s_m$, and by construction one of the two domains adjacent to s_m belongs to $\mathcal{S}_m$ and the other does to $\mathcal{S}_{m+1}$. This proves the first statement in Claim 4.11. The second is immediate from $\mathcal{D}_N = \bigcup \underline{\mathcal{S}}$.

Step 5. Finally, let us verify Claim 4.12 by induction on k. The base, $k = 0$, is clear.

Let us assume that this claim holds for some k and check it for $k+1$. Since s_k lies in the interior of $\mathcal{D}_k$, the points of $\mathcal{S}_{k+1}$ that are close to s_k lie inside $\mathcal{D}_k^+$, and since $\mathcal{D}_k^+$ is the union of fundamental domains, $\mathcal{S}_{k+1} \subset \mathcal{D}_k^+$.

To construct $\mathcal{D}_{k+1}$, we need to add to the intersection $\bigcap_{H \in \mathcal{H}_k^0} H$ defining $\mathcal{D}_k$ the half-planes $H \in \mathcal{H}_{k+1}$. Assume first that $k + 1 < N$.

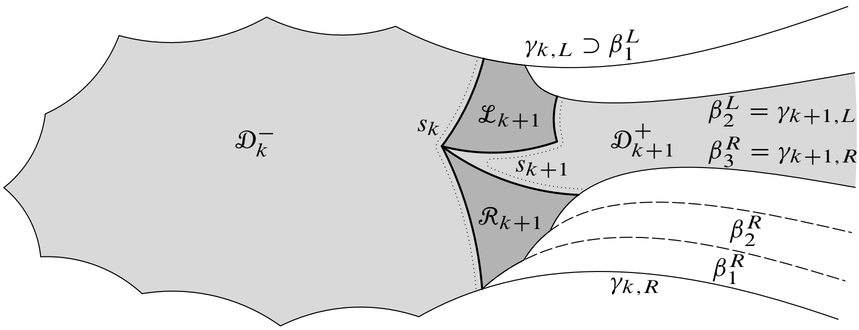


Figure 9. Illustrating the proof of Claim 4.12.

If $\partial_\alpha \mathcal{S}_{k+1}$ for $\alpha = L, R$ consists of a single vertex, there are no new half-planes to be added on $\partial_\alpha \mathcal{S}_{k+1}$. Otherwise, $\partial_\alpha \mathcal{S}_{k+1}$ consists of sides joining a sequence of vertices $u_0^\alpha u_1^\alpha \dots u_{m_\alpha}^\alpha$. Let H_j^α be the half-plane in $\mathcal{H}_{k+1}$ with $\partial H_j^\alpha = \ell(u_{j-1}^\alpha u_j^\alpha)$, and let $\beta_j^\alpha \subset \ell(u_{j-1}^\alpha u_j^\alpha)$ be the ray starting from u_{j-1}^α and passing through u_j^α; see Figure 9.

Add the half-planes H_j^α to the intersection one by one with j increasing. For $j = 1$ the line $\ell(u_0^\alpha u_1^\alpha)$ may contain the ray $\gamma_{k,\alpha}$, so the intersection is not changed. Otherwise the line $\ell(u_{j-1}^\alpha u_j^\alpha)$ crosses the boundary of the current intersection at u_{j-1}^α and by the local convexity from Step 3 it enters this intersection in the direction of β_j^α. Therefore, by adding H_j^α we cut the intersection along β_j^α and remove the part that does not contain $\mathcal{S}_{k+1}$. The removed regions are shown in white in Figure 9.

Note that β_j^L and $\beta_{j'}^R$ do not intersect for any j, j'. This follows from Lemma 2.2 applied to the domain $P = \mathcal{S}_{k+1}$ if $\mathcal{S}_{k+1}$ contains only one fundamental domain, and to $P = \mathcal{F}_u$ if both domains in $\mathcal{S}_{k+1}$ share a vertex u. Therefore, by adding H_j^α, the intersection is cut along the whole of β_j^α, not just its initial segment.

Let us now check that s_{k+1} lies in the interior of $\mathcal{D}_{k+1}$. Let $v_{j,L}, v_{j,R}$ be the left and right ends of s_j and let $v_{j,L}u_{j,L}$ and $u_{j,R}v_{j,R}$ be the sides of s_j adjacent to these ends. We will show that $v_{k+1,L}u_{k+1,L}$ lies inside $\mathcal{D}_{k+1}$.

There are two cases. First, if $v := v_{k+1,L}$ coincides with $v_{k,L}$ then using Step 3 again, for some $r \le n(v)$ the domains $\mathcal{L}_{k+1-r}, \dots, \mathcal{L}_{k+2}$ are consecutive petals in the flower $\mathcal{F}_v$ while $\mathcal{L}_{k+1-r}$ is not in $\mathcal{F}_v$, and the ray $\gamma_{k+1,L} = \gamma_{k,L}$ contains the side $wv \subset \partial_L \mathcal{S}_{k+1-r}$. But then the angle $wvu_{k+1,L}$ contains at most $n(v) - 1$ petals, so $vu_{k+1,L}$ is not the continuation of wv and hence is not contained in $\gamma_{k+1,L}$.

Similarly, if $v \neq v_{k,L}$, then $\gamma_{k+1,L}$ is the continuation of the side $wv = u_{m_L-1}^L u_{m_L}^L$ of $\partial_L \mathcal{S}_{k+1}$ adjacent to v, and the angle $wvu_{k+1,L}$ contains only one sector, namely, $\mathcal{L}_{k+1}$. Thus again, $vu_{k+1,L}$ lies inside $\mathcal{D}_{k+1}$.

This proves item (i) of Claim 4.12 if s_{k+1} contains one or two sides. If it contains three sides: $s_{k+1} = v_{k+1,L} u_{k+1,L} u_{k+1,R} v_{k+1,R}$, it remains to rule out the possibility that $u_{k+1,L} \in \gamma_{k+1,R}$. But in this case the triangle $u_{k+1,L} u_{k+1,R} v_{k+1,R}$ has all of its sides lying in $\partial \mathbf{T}_{\mathcal{R}}$, and this is impossible by Lemma 2.1. Therefore, $u_{k+1,L}$ and $u_{k+1,R}$ lie inside $\mathcal{D}_{k+1}$ and item (i) is fully established.

Further, by construction $\mathcal{D}_{k+1}^- \setminus \mathcal{D}_k^-$ is bounded by

$$\partial_O \mathcal{S}_{k+1} \cup s_k \cup s_{k+1} = \partial \mathcal{S}_{k+1},$$

hence $\mathcal{D}_{k+1}^- \setminus \mathcal{D}_k^- = \mathcal{S}_{k+1}$, and item (ii) holds. Items (iii) and (iv) for $k+1$ are also now clear.

In the case $k+1 = N$ we likewise consecutively cut $\mathcal{D}_{N-1}$ along the rays through the segments of $\partial_O \mathcal{S}_N$; on the last step we cut along the last segment of this boundary. Therefore, $\mathcal{D}_N \setminus \mathcal{D}_{N-1}^-$ is bounded by

$$s_{N-1} \cup \partial_O \mathcal{S}_N = \partial \mathcal{S}_N,$$

hence $\mathcal{D}_N \setminus \mathcal{D}_{N-1}^- = \mathcal{S}_N$, verifying the final statement of Claim 4.12. ∎

4.4. The time-reversing involution

The Markov coding defined above has the following property: the Markov chain with time reversed, that is, the Markov chain with the matrix Π^T, is the same as the initial one with the states renamed. This is possible precisely because the thickened path between domains $\mathcal{B}$ and $\mathcal{A}$ is exactly the thickened path from $\mathcal{A}$ to $\mathcal{B}$ read in the opposite direction. Thus we obtain an involution which inverts arrangements and states; informally it swaps the past and the future domains for each state. Precisely, we define the involution $\iota\colon \Xi \to \Xi$ by:

$$A_0(e) \leftrightarrow A_0(e^{-1}), \qquad A_\alpha[i_-, i_+](e) \leftrightarrow A_\alpha[i_+, i_-](e^{-1}) \quad (\alpha = LR, RL),$$
$$A_L[i_-, i_+](e) \leftrightarrow A_R[i_+, i_-](e^{-1}), \qquad B(e_L, e_R) \leftrightarrow D(e_R^{-1}, e_L^{-1}),$$
$$C_k(e_L, e_R) \leftrightarrow C_{n(e_L,e_R)-k-1}(e_R^{-1}, e_L^{-1}),$$
$$E_\alpha(e_L, e_M, e_R) \leftrightarrow E_\alpha(e_R^{-1}, e_M^{-1}, e_L^{-1}) \quad (\alpha = L, R).$$

Proposition 4.13. *The involution ι maps the Markov chain with the adjacency matrix Π to the same chain with reversed time, that is, $\Pi_{\iota(j)\iota(k)} = \Pi_{kj}$. Also, $\iota(\Xi_S) = \Xi_F$ and vice versa.*

Proof. This follows directly from the definitions. ∎

5. Operations with thickened paths

In this section we develop some techniques for manipulating thickened paths which will be used in Section 6 to establish strong connectivity and aperiodicity of the Markov chain. The same techniques will also be used in Section 8.3 to verify the convergence conditions for Theorem A.

5.1. Adjusting the labels of states

Recall from Definition 4.9 that a sequence $\underline{j}$ of states generates a sequence of domains $\underline{\mathcal{S}}$ if for each k the pair $(\mathcal{S}_k, \mathcal{S}_{k+1})$ represents the configuration $\pi(j_k)$. In certain circumstances, we will need to adjust the sequence $\underline{j}$ while leaving the sequence it generates, and hence the sequence $\pi(j_k)$, unchanged. Such an adjustment is achieved by the following technical lemma. It will be crucial later to note that the required changes do not propagate beyond a definite bounded distance which depends only on the tessellation $\mathbf{T}_{\mathcal{R}}$.

As we have noted in the proof of Theorem 4.10, the sequence $\pi(\underline{j})$ is uniquely determined by $\underline{\mathcal{S}}$, while the indices $i_{-,L/R}$ of any A-state j_k are defined uniquely from the corresponding indices for the state j_{k-1}, and the indices $i_{+,L/R}$ are similarly defined in the backwards direction.

Lemma 5.1. *Suppose that the admissible sequence $\underline{j} = (j_0, \dots, j_{m-1})$ generates a sequence of domains $\underline{\mathcal{S}} = (\mathcal{S}_0, \dots, \mathcal{S}_m)$. Assume that $\pi(j_{m-1}) = A(e)$. Let j'_{m-1} be any A-state such that $\pi(j'_{m-1}) = A(e)$ and $i_{-,\alpha}(j_{m-1}) = i_{-,\alpha}(j'_{m-1})$ for $\alpha = L, R$. Then there exists an admissible sequence $\underline{j}' = (j'_0, \dots, j'_{m-1})$ generating the same sequence $\underline{\mathcal{S}}$. Moreover, if $s_i = \mathcal{S}_i \cap \mathcal{S}_{i+1}$, then $j_i = j'_i$ whenever either j_i is not of type A or s_i and s_{m-1} have no common points. In particular, $j_i = j'_i$ for $i \leq m - n_0 + 1$, where $n_0 = \max\{n(v) : v$ is a vertex of $\mathbf{T}_{\mathcal{R}}\}$. (While the lemma is trivial if $\mathcal{R}$ has no vertices inside $\mathbb{D}$, we set $n_0 = 2$ in this case. This will be used below.)*

We remark that the same result holds *mutatis mutandi* adjusting states forwards from j_0.

Proof. First assume that $i_{-,L}(j_{m-1}) = i_{-,R}(j_{m-1}) = 1$. Then the states j_{m-1} and j'_{m-1} belong to the set

$$\{A_0(e), A_L[1, i_+](e), A_R[1, i_+](e)\}.$$

One can see from Definition 4.7 that all these states have the same set of allowed preceding states, so the sequence $\underline{j}' = (j_0, \dots, j_{m-2}, j'_{m-1})$ is admissible.

Now assume that, say, $i_{-,L}(j_{m-1}) = k > 1$. Then the states j_{m-1} and j'_{m-1} belong to the set

$$\{A_L[k, i_+](e), A_{LR}[k, i_+](e)\}.$$

Let $l = i_{+,L}(j_{m-1})$, i.e.

$$\begin{aligned} l &= i_+ \quad \text{if } j_{m-1} = A_L[k, i_+](e), \text{ and} \\ l &= 1 \quad \text{if } j_{m-1} = A_{LR}[k, i_+](e); \end{aligned}$$

let l' be defined in the same way for j'_{m-1}. Then the suffix $(j_{m-k}, \dots, j_{m-1})$ of the sequence $\underline{j}$ has the following form:

$$\left.\begin{matrix} A_L[1, l+k-1](e_1) \\ A_{RL}[i_-, l+k-1](e_1) \\ D(e_1, \widetilde{e}_1) \end{matrix}\right\}, A_L[2, l+k-2](e_2), \dots,$$

$$A_L[k-1, l+1](e_{k-1}), \begin{cases} A_L[k, i_+](e) \\ A_{LR}[k, i_+](e). \end{cases} \tag{3}$$

Here $e_{i+1} = l(e_i)$ for $i = 1, \dots, k-1$, where $e_k = e$. Formally speaking, it is possible that the whole sequence $\underline{j}$ is only a suffix of the sequence in (3); the proof for this case is the same.

Define the sequence $\underline{j}'$ as follows: $j'_i = j_i$ for $i = 0, \dots, m-k-1$, and for $i = m-k, \dots, m-2$ define j'_i by the formula (3) with l' in place of l (that is, $j'_{m-k} = j_{m-k}$ if $j_{m-k} = D(e_1, \widetilde{e}_1)$). Observe that these states are allowed: it is clear that all indices are positive and

$$\begin{aligned} i_{-,L}(j'_{m-i}) + i_{+,L}(j'_{m-i}) &= (k+1-i) + (l'+i-1) \\ &= k + l' = i_{-,L}(j'_{m-1}) + i_{+,L}(j'_{m-1}) \le n(v_L(e)), \end{aligned} \tag{4}$$

and since s_{m-i} and s_{m-1} have the same left end, we may replace e by e_{k+1-i} in the right-hand side of this formula. Also it is clear that $\underline{j}'$ is an admissible sequence.

To prove the last statement observe that if $n_0 = 2$, then $\pi(\hat{\jmath}) = A(e)$ implies $\hat{\jmath} = A_0(e)$, hence $\underline{j} = \underline{j}'$, while for $n_0 \ge 3$ we have the following two cases. If $l = l'$, then $j'_i = j_i$ for all $i \le m-2$; otherwise, $\max(l, l') \ge 2$, so (4) for $\underline{j}$ and $\underline{j}'$ yields $k \le n_0 - 2$, and $j'_i = j_i$ for $i \le m-k-1$. Hence, in all cases

$$j'_i = j_i \quad \text{for } i \le m - n_0 + 1. \qquad \blacksquare$$

5.2. Narrowing

The next lemma shows how a sequence $\underline{\mathcal{S}}$ can be "narrowed" by reducing its final level from two domains to one. This will be useful, for example, when $\underline{\mathcal{S}} = (\mathcal{S}_0, \dots, \mathcal{S}_N)$ is a thickened path between its ends and we need to find a thickened path $\underline{\mathcal{S}}'$ between $\mathcal{S}_0$ and some intermediate domain $\mathcal{L}_k$ for which $\mathcal{S}_k$ contains two domains.

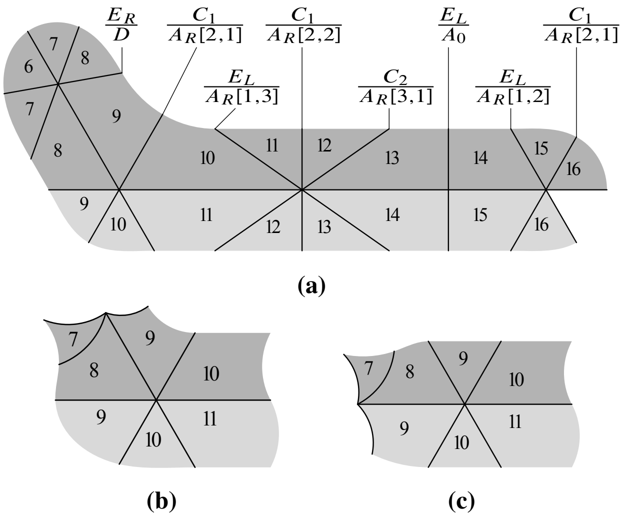


Figure 10. The proof of Lemma 5.2. The original sequence $\underline{\mathcal{S}}$ contains all shaded domains, while $\underline{\mathcal{S}}'$ contains only those shaded dark gray. Numbers inside the domains indicate their levels. The states j_i and j_i' are shown above the curves s_i. In all figures $k = 8$.

One way to deal with this situation is to consider the thickened path as a minimal convex union of fundamental domains. Then one can see that the desired thickened path is a subset of $(\mathcal{S}_0, \ldots, \mathcal{S}_k)$ obtained from $\bigcup_{j=0}^{k} \mathcal{S}_j$ by cutting along a line $\ell \subset \partial \mathbf{T}_{\mathcal{R}}$ incident to a common vertex of $\mathcal{L}_k$ and R_k as shown in Figure 10.

However, since below we are mostly interested in the corresponding sequences of states, from now on we will consider not only thickened paths but any sequence $(\mathcal{S}_0, \ldots, \mathcal{S}_N)$ of domains generated by admissible sequences of states $(j_0, \ldots, j_{N-1})$, in other words, we drop the conditions that $j_0 \in \Xi_S$, $j_{N-1} \in \Xi_F$.

Suppose the sequence $\underline{\mathcal{S}} = (\mathcal{S}_0, \ldots, \mathcal{S}_m)$ is generated by an admissible sequence of states $\underline{j} = (j_0, \ldots, j_{m-1})$. By the *directed adjacency graph* associated to $\underline{\mathcal{S}}$ we mean the graph whose vertices are the domains in $\underline{\mathcal{S}}$ with a directed edges going from a domain $\mathcal{A} \in \mathcal{S}_k$ to $\mathcal{B} \in \mathcal{S}_{k+1}$ whenever $\mathcal{A}$ and $\mathcal{B}$ share a common side.

Lemma 5.2. *Suppose that the sequence $\underline{\mathcal{S}} = (\mathcal{S}_0, \ldots, \mathcal{S}_m)$ is generated by an admissible sequence of states $\underline{j} = (j_0, \ldots, j_{m-1})$. Assume that $\mathcal{S}_m$ contains two domains $\mathcal{L}_m, \mathcal{R}_m$ and let $\mathcal{S}_m'$ be one of them, for definiteness $\mathcal{L}_m$. Then there is a narrowed sequence $\underline{\mathcal{S}}'$ from $\mathcal{S}_0$ to $\mathcal{S}_m'$ which can be described as follows.*

(1) *Let $\mathcal{S}_l'$ be the set of domains $\mathcal{A} \subset \mathcal{S}_l$ such that there exists a path in the directed graph from $\mathcal{A}$ to $\mathcal{S}_m'$. Let k be the maximal number such that there is an edge $\mathcal{R}_k \to \mathcal{L}_{k+1}$; we set $k = -1$ if there is no such edge. Then $\mathcal{S}_l' = \mathcal{S}_l$ for $l \leq k$, and $\mathcal{S}_l = \mathcal{L}_l$ for $l \geq k + 1$.*

j_l	j'_l
$B(e_L, e_R)$	$A_R[1, n(e_L, e_R) - 1](e_L)$ or $A_{LR}[i_-, n(e_L, e_R) - 1](e_L)$
$E_R(e_L, e_M, e_R)$	$D(e_L, e_M)$
$C_i(e_L, e_R)$	$A_R[i + 1, n(e_L, e_R) - i - 1](e_L)$
$E_L(e_L, e_M, e_R)$	$A_R[1, n(e_L, e_M) - 1](e_L)$

Table 1

(2) *The sequence* $\underline{\mathcal{S}}'$ *can be generated by an admissible sequence*

$$\underline{j}' = (j'_0, \ldots, j'_{m-1}),$$

where one can assume that $j_i = j'_i$ *if* $\pi(j_i) = \pi(j'_i)$ *and either* j_i *is not of type* A *or* s_i *has no common vertex with* s_k*. In particular,* $j_0 = j'_0$ *if* $\mathcal{S}'_{n_0-1} = \mathcal{S}_{n_0-1}$*, where* n_0 *is defined in Lemma 5.1.*

Proof. We consider each point in turn.

(1) This is illustrated in Figure 10. Let us go backward from $l = m$ to $l = 0$. Every path from $\mathcal{A} \in \mathcal{S}'_l$ to $\mathcal{S}'_m$ should pass through $\mathcal{B} \in \mathcal{S}'_{l+1}$. For $l \geq k + 1$ there is only one edge, $\mathcal{L}_l \to \mathcal{L}_{l+1}$ that ends in $\mathcal{S}'_{l+1} = \mathcal{L}_{l+1}$, hence $\mathcal{S}'_l = \mathcal{L}_l$. Similarly, for $l \leq k$, we obtain $\mathcal{S}'_l = \mathcal{S}_l$.

(2) Assume that $\mathcal{S}'_m = \mathcal{L}_m$ and let k be the same as in the first statement. Then, following Lemma 3.8 (4) and still referring to Figure 10, the state j_k is of type either B or E_R (if $k \geq 0$), and the states $j_{k+1}, \ldots, j_{m-1}$ are of types C and E_L.

Let us define the states j'_l for $l = k, \ldots, m - 1$ as in Table 1 (see Figure 10 (a)).

In the first line there are several options, we will choose one of them later. Note also that if the table suggests $A_R[1, 1](e)$, it should be replaced by $A_0(e)$. It is clear by the construction that for all $l \geq k$ the state j'_l is well-defined, $\pi(j'_l)$ is represented by the pair $(\mathcal{S}'_l, \mathcal{S}'_{l+1})$ and the transition $j'_l \to j'_{l+1}$ is allowed.

If $k = -1$, this ends the proof; otherwise we have to define j'_l for $l = 0, \ldots, k - 1$, as well as to choose j'_k from the set given above in such a way that the transitions $j'_0 \to \cdots \to j'_k$ are allowed. By default we set $j'_l = j_l$, although sometimes a correction is needed: this will be stated explicitly and only occurs in the final paragraph of the proof.

If $j_k = E_R(e_L, e_M, e_R)$, then either

$$n(e_L, e_M) > 2 \quad \text{and} \quad j_{k-1} = C_{n(e_L, e_M)-2}(\hat{e}_L, \hat{e}_R),$$

or

$$n(e_L, e_M) = 2 \quad \text{and} \quad j_{k-1} \in \{B(\hat{e}_L, \hat{e}_R), E_R(\tilde{e}, \hat{e}_L, \hat{e}_R), E_L(\hat{e}_L, \hat{e}_R, \tilde{e})\}.$$

In all these cases $e_L = r(\hat{e}_L), e_M = l(\hat{e}_R)$, and the transition $j_{k-1} \to j'_k = D(e_L, e_M)$ is allowed.

If $j_k = B(e_L, e_R)$, there are several cases. As above, let $s_j = \mathcal{S}_j \cap \mathcal{S}_{j+1}$.

First of all, suppose that s_k and s_{k-1} have no common vertices, i.e. j_{k-1} equals $A_0(\hat{e})$, $A_L[i_-, 1](\hat{e})$, $A_R[i_-, 1](\hat{e})$, or $D(\hat{e}_L, \hat{e}_R)$ with e_L, e_R non-adjacent to $\hat{e}^{-1}$ (or $\hat{e}_L^{-1}, \hat{e}_R^{-1}$). Then we set

$$j'_k = A_R[1, n(e_L, e_R) - 1](e_L),$$

so the transition $j_{k-1} \to j'_k$ is allowed.

Now assume that the left ends of s_k and s_{k-1} coincide (see Figure 10 (b)), i.e. j_{k-1} equals $A_L[i_-, 2](\hat{e})$, $A_{RL}[i_-, 2](\hat{e})$ with $e_L = l(\hat{e})$, or j_{k-1} equals $D(\hat{e}_L, \hat{e}_R)$ with $e_L = l(\hat{e}_L)$. In the first of these three subcases we set

$$j'_k = A_{LR}[i_- + 1, n(e_L, e_R) - 1](e_L),$$

and in the remaining two we set $j'_k = A_{LR}[2, n(e_L, e_R) - 1](e_L)$.

It remains to consider the case when the right (but not left) ends of s_k and s_{k-1} coincide (see Figure 10 (c), hence s_{k-1} and $s'_k = \mathcal{S}'_k \cap \mathcal{S}'_{k+1}$ have no common vertices. Then either j_{k-1} equals $A_R[i_-, 2](\hat{e})$ or $A_{LR}[i_-, 2](\hat{e})$ with $e_R = r(\hat{e})$, or it equals $D(\hat{e}_L, \hat{e}_R)$ with $e_R = r(\hat{e}_R)$. Let

$$j'_k = A_R[1, n(e_L, e_R) - 1](e_L),$$

so in the last subcase the transition $j_{k-1} \to j'_k$ is admissible. In the first two subcases apply Lemma 5.1 to define the sequence $(j'_0, \dots, j'_{k-1})$: if $j_{k-1} = A_R[i_-, 2](\hat{e})$, let $j'_{k-1} = A_R[i_-, 1](\hat{e})$, and if $j_{k-1} = A_{LR}[i_-, 2](\hat{e})$, let $j'_{k-1} = A_L[i_-, 1](\hat{e})$.

The last part of the statement, which estimates the common part of $\underline{j}$ and $\underline{j}'$, follows from the corresponding part in Lemma 5.1: $\mathcal{S}'_{n_0-1} = \mathcal{S}_{n_0-1}$ yields $k \geq n_0 - 1$, so when we apply that lemma with $m = k \geq n_0 - 1$, we can have $j_i \neq j'_i$ only for $i > m - n_0 + 1 \geq 0$. ∎

5.3. Joining

The next lemma deals with how to join two sequences of domains, $\underline{\mathcal{S}}^- = (\mathcal{S}_{-m}, \dots, \mathcal{S}_0)$ and $\underline{\mathcal{S}}^+ = (\mathcal{S}_0, \dots, \mathcal{S}_n)$, which have a common end $\mathcal{S}_0$ containing only one domain. The result of such a join is not necessarily a thickened path between its ends, since the union

$$\mathcal{U} = \bigcup \underline{\mathcal{S}}^- \cup \bigcup \underline{\mathcal{S}}^+$$

may fail to be convex at the vertices of $\mathcal{S}_0$.

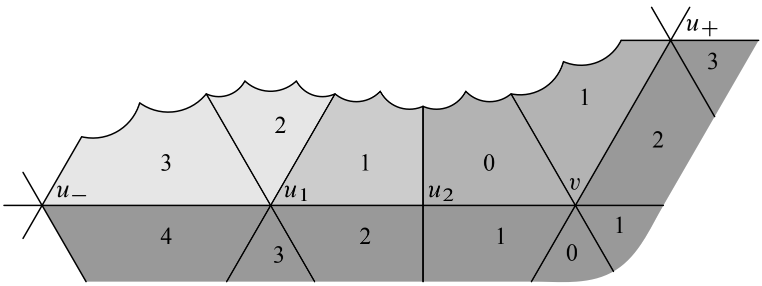


Figure 11. Joining two paths at a minimally concave vertex; see Definition 5.3. The union of the two thickened paths shown in dark gray is minimally non-convex at the vertex v. The lighter colors show the consecutive flowers we add to the path: we first add the flower $\mathcal{F}_v$, so the adjacent vertex u_2 on γ becomes minimally concave, then we add $\mathcal{F}_{u_2}$, making $u_1 = w_-$ minimally concave, and then $\mathcal{F}_{u_1}$. The vertices $u_\pm$ are non-concave since they are incident to at most $n(u_\pm) - 1$ domains in $\mathcal{U}$ and the addition of the flowers increases this number by one. For details, see Lemma 5.4 (3).

If the union is convex, we show that it can be generated by an admissible sequence of states. If convexity fails at some vertex v, there are two cases: either v is incident to $n(v) + 1$ domains in the union, or to more than $n(v) + 1$. In the former case we will show that the union can be enlarged to a thickened path, while in the latter case the union cannot be so enlarged: the part of $\mathcal{U}$ between the first and the last domain adjacent to v can be shortcut by a path that goes "the other way around v". A possible enlargement in the first case is illustrated in Figure 11.

Consider the case in which convexity fails because v is incident to $n(v) + 1$ domains in the union $\mathcal{U}$. Then one can add to $\mathcal{U}$ the remaining domains in the flower $\mathcal{F}_v$, and these domains can be assigned levels in such a way that the levels of adjacent domains differ by one, see Figure 11. However, this adds one domain adjacent to a vertex u next to v in $\partial\mathcal{U}$, so if $\partial\mathcal{U}$ had a straight angle at u, now u is adjacent to $n(u) + 1$ domains in $\mathcal{U}' = \mathcal{U} \cup \mathcal{F}_v$. Thus we need to add $\mathcal{F}_u$, and so on until we arrive to the ends $u_\pm$ of the maximal geodesic segments in $\partial\mathcal{U}$ starting from v in both directions. Since the vertices $u_\pm$ are incident to less than $n(u_\pm)$ domains in $\mathcal{U}$, the addition of one more domain does not destroy convexity at these points. It can be shown that the resulting union of fundamental domains can be generated by an admissible sequence of states, and hence is a thickened path between its ends.

In fact, the whole analysis of thickened paths presented in this paper can be performed using this "convexification" technique, i.e. the successive addition of flowers to vertices of a collection of domains, where the convexity failed; one can find a detailed exposition of this approach in the first version of the preprint [18]. Rather than doing this, however, we keep track of states using Lemma 5.4 below.

Definition 5.3. Suppose given sequences of domains $\underline{\mathcal{S}}^- = (\mathcal{S}_{-m}, \dots, \mathcal{S}_0)$ and $\underline{\mathcal{S}}^+ = (\mathcal{S}_0, \dots, \mathcal{S}_n)$ such that $\mathcal{S}_0$ contains only one domain, and as usual denote $s_l = \mathcal{S}_l \cap \mathcal{S}_{l+1}$. We call the common vertex v of s_{-1} and s_0 *concave*, if it is incident to more than $n(v)$ domains in the sequence $\underline{\mathcal{S}} = (\mathcal{S}_{-m}, \dots, \mathcal{S}_n)$. It is *minimally concave* it is incident to exactly $n(v) + 1$ domains in $\underline{\mathcal{S}} = (\mathcal{S}_{-m}, \dots, \mathcal{S}_n)$.

Lemma 5.4. *We have the following cases:*

(1) *Let sequences of levels $\underline{\mathcal{S}}^- = (\mathcal{S}_{-m}, \dots, \mathcal{S}_0)$ and $\underline{\mathcal{S}}^+ = (\mathcal{S}_0, \dots, \mathcal{S}_n)$ be generated by sequences of states $\underline{j}^- = (j_{-m}, \dots, j_{-1})$ and $\underline{j}^+ = (j_0, \dots, j_{n-1})$. Assume that $\mathcal{S}_0$ contains only one domain and that the curves s_{-1} and s_0 have no common sides. Then at most one vertex in $s_{-1} \cap s_0$ is concave.*

(2a) *Assume that there are no concave common vertices. Then the sequence $\underline{\mathcal{S}} = (\mathcal{S}_{-m}, \dots, \mathcal{S}_0, \dots, \mathcal{S}_n)$ can be generated by a sequence of states $\underline{\hat{j}}$.*

(2b) *(See Figure 11.) Assume that there is a minimally concave common vertex v. Let u_-v and vu_+ be the maximal geodesic segments in $\partial\underline{\mathcal{S}}^- \setminus \partial\mathcal{S}_0$ and $\partial\underline{\mathcal{S}}^+ \setminus \partial\mathcal{S}_0$ respectively. Let $w_\pm$ be the internal vertices of the curve u_-vu_+ that are closest to $u_\pm$; it is possible that one or both of $w_\pm$ coincide with v. Define the curve $\gamma = w_-vw_+$. Then there exists a sequence $\underline{\hat{\mathcal{S}}} = (\hat{\mathcal{S}}_{-m}, \dots, \hat{\mathcal{S}}_n)$ generated by an admissible sequence of states $\underline{\hat{j}}$ such that*

(i) $\mathcal{S}_i \subset \hat{\mathcal{S}}_i$ *for all i;*

(ii) $\hat{\mathcal{S}}_i = \mathcal{S}_i$ *if $\mathcal{S}_i$ has no common vertex with γ;*

(iii) $\bigcup \underline{\hat{\mathcal{S}}}$ *is the union of $\bigcup \underline{\mathcal{S}}$ and of all flowers $\mathcal{F}_w$, where w is a vertex of γ;*

(3) *Let $\delta = s_{-1} \cup s_0$ in the case from* (2a) *and $\delta = u_-vu_+$ in the case from* (2b)*. Then one can assume that $\hat{j}_t = j_t$ if $\pi(\hat{j}_t) = \pi(j_t)$ and either j_t is not of type A or s_t has no common points with δ. In particular, $j_{n-1} = \hat{j}_{n-1}$ if $n \geq n_0 - 1$ and (in the case from* (2b)*) $\mathcal{S}_{n-n_0+1} = \hat{\mathcal{S}}_{n-n_0+1}$, where n_0 is defined in Lemma 5.1.*

Proof. We consider each case in turn.

(1) Assume the contrary: both ends v_L, v_R of the curves s_{-1} and s_0 coincide. Then $\mathcal{S}_0$ is compact, so $N(\mathcal{R}) = 4$, and j_{-1} is of type D and j_0 is of type B. Therefore, both v_L and v_R are incident only to $\mathcal{S}_{-1}, \mathcal{S}_0, \mathcal{S}_1$, but at least one of v_L, v_R has $n(v) \geq 3$ by Assumption 1.1.

(2a) Since $\mathcal{S}_0$ contains one domain, the state j_{-1} is of type D or A, and the state j_0 is of type A or B.

First suppose that s_{-1} and s_0 have no common vertices. Then if j_{-1} is of type A, apply Lemma 5.1 for $\hat{j}_{-1}$ with $i_{+,L/R}(\hat{j}_{-1}) = 1$, otherwise let $\underline{\hat{j}}^- = \underline{j}^-$. Similarly, if j_0 is of type A, apply the analogue of Lemma 5.1 with time inverted for the state $\hat{j}_0$ with $i_{-,L/R}(\hat{j}_0) = 1$, otherwise let $\underline{\hat{j}}^- = \underline{j}^-$.

The case when s_{-1} and s_0 share both ends is trivial: here j_{-1} is of type D, j_0 is of type B, and the convexity in the common vertices $v_{L,R}$ means that $n(v_{L,R}) \geq 3$, so the transition $j_{-1} \to j_0$ is admissible.

Finally, suppose that s_{-1} and s_0 share their left end only: $s_{-1} \cap s_0 = v_L$. Let n_+ (respectively, n_-) be the number of levels $\mathcal{S}_k$ with $k \geq 0$ (respectively, $k \leq 0$) that are adjacent to v_L. By our assumption,

$$n_- + n_+ - 1 \leq n(v_L), \quad n_\pm \geq 2.$$

Let us construct the sequence $\underline{\hat{\jmath}}^- = (\hat{\jmath}_{-m}, \dots, \hat{\jmath}_{-1})$ as follows. If j_{-1} is of type D we set $\underline{\hat{\jmath}}^- = \underline{j}^-$; otherwise the construction is performed in two steps. First we have to ensure that

$$i_{-,L}(\hat{\jmath}_{-1}) = n_- - 1.$$

This is in fact true for j_{-1} unless all domains in $\underline{\mathcal{S}}^-$ are consecutive petals in the same flower. Namely, let us say that the state j_l (with $l < 0$) is "correct", if either it is not of type A, or it is of type A and the index $i_{-,L}(j_l)$ coincides with the number of domains in $(\mathcal{S}_{-m}, \dots, \mathcal{S}_l)$ that are incident to the left end of s_l. One can check that if the transition $j_l \to j_{l+1}$ is not of type $A \to A$, or if it is of type $A \to A$ and $\partial_L \mathcal{S}_{l+1}$ contains at least one side, then j_{l+1} is correct. Also, if $j_l \to j_{l+1}$ is of type $A \to A$ *and* the state j_l is correct, then j_{l+1} is correct.

Thus the state j_{-1} is correct unless the states $j_{-m}, \dots, j_{-1}$ are of type A and the curves $s_{-m}, \dots, s_{-1}$ are segments with the same left end v_L. If j_{-1} is correct, let $\underline{\tilde{\jmath}}^- = \underline{j}^-$, otherwise use Lemma 5.1 to obtain the sequence $\underline{\tilde{\jmath}}^-$ with $i_{-,L}(\tilde{\jmath}_{-m}) = 1$; note that since

$$i_{-,L}(\tilde{\jmath}_{-m}) \leq i_{-,L}(j_{-m}),$$

the state $\tilde{\jmath}_{-m}$ is allowed. Now $\tilde{\jmath}_{-m}$, and hence $\tilde{\jmath}_{-1}$ are correct.

Apply Lemma 5.1 again to transform $\underline{\tilde{\jmath}}^-$ into $\underline{\hat{\jmath}}^-$ with $i_{+,L}(\hat{\jmath}_{-1}) = n_+$ and $i_{+,R}(\hat{\jmath}_{-1}) = 1$. Note that the state $\hat{\jmath}_{-1}$ is allowed, since

$$i_{-,L}(\hat{\jmath}_{-1}) = i_{-,L}(\tilde{\jmath}_{-1}) = n_- - 1$$

due to its "correctness", so

$$i_{-,L}(\hat{\jmath}_{-1}) + i_{+,L}(\hat{\jmath}_{-1}) = n_- + n_+ - 1 \leq n(v_L).$$

Likewise, we construct a sequence $\underline{\hat{\jmath}}^+ = (\hat{\jmath}_0, \dots, \hat{\jmath}_{n-1})$ with $i_{-,L}(\hat{\jmath}_0) = n_-$, $i_{+,L}(\hat{\jmath}_0) = n_+ - 1$ that generates the same domains $\underline{\mathcal{S}}^+$. It remains to check that the transition $\hat{\jmath}_{-1} \to \hat{\jmath}_0$ is admissible. Clearly it belongs to one of the four types,

$$D \to B, \quad D \to A, \quad A \to B, \quad A \to A.$$

In the first case the argument from part 1 above works, while the next two cases correspond to the transitions

$$D(e_L, e_R) \to A_{L,LR}[2, \dots](l(e_L)), \quad A_{L,RL}[\dots, 2](e) \to B(l(e), \dots),$$

which are admissible. Finally, the case $A \to A$ splits into four subcases:

$$\left.\begin{matrix} A_L[i_-, i_+](e) \\ A_{RL}[i_-, i_+](e) \end{matrix}\right\} \to \left\{\begin{matrix} A_L[i'_-, i'_+](l(e)) \\ A_{LR}[i'_-, i'_+](l(e)) \end{matrix}\right. .$$

Definition 4.7 reduces the existence of such a transition to the subcase $A_L \to A_L$, replacing $A_{RL}[i_-, i_+](e)$ by $A_L[1, i_+](e)$ and $A_{LR}[i'_-, i'_+](l(e))$ by $A_L[i'_-, 1](l(e))$. In the resulting transition

$$A_L[\tilde{\imath}_-, \tilde{\imath}_+](e) \to A_L[\tilde{\imath}'_-, \tilde{\imath}'_+](l(e)) \tag{5}$$

one can express the indices in terms of $n_\pm$, so that in all four subcases the transition (5) is of the form $A_L[n_- - 1, n_+](e) \to A_L[n_-, n_+ - 1](l(e))$, which is allowed.

(2b) Let k be such that $\partial_L \mathcal{S}_{-k} \supset u_- w_-$; see Figure 11. Let us show that the states $j_{-k+1}, \dots, j_{-1}$ are of type A, while j_{-k} is of type A or D. Indeed, let $u_0 = u_-, u_1, \dots, u_p = v$ be the consecutive vertices on $u_- v$, thus $u_- w_- = u_0 u_1$. Then there are

$$-k = i_0 < i_1 < \dots < i_{p-1} \le -1$$

such that $\partial_L \mathcal{S}_{i_t} = u_t u_{t+1}$ for $t = 0, \dots, p-1$, and for $t = 0, \dots, p-1$ and $i_t < i < i_{t+1}$, we have $\partial_L \mathcal{S}_i = u_{t+1}$, here $i_p = 0$. By Proposition 4.3, each segment $\mathbf{J}_t = (j_{i_t}, \dots, j_{i_{t+1}-1})$ has the form

$$(E_R, E_L) \quad \text{or} \quad (A \text{ or } D, A, \dots, A, A \text{ or } B).$$

Since $\mathcal{S}_0$ contains one domain, the segment $\mathbf{J}_{p-1}$ has the form $(A \text{ or } D, A, \dots, A)$. Assume that $p > 1$ and $\mathbf{J}_{p-1}$ starts with D. Then $\mathbf{J}_{p-2}$ ends with B or E_L, hence $\mathcal{L}_{i_{p-1}}$ intersects each of $s_{i_{p-1}-1}$ and $s_{i_{p-1}}$ in one side and these sides meet in a vertex of $\mathcal{L}_{i_{p-1}}$. Thus if $\partial_L \mathcal{S}_{i_{p-1}} = u_{p-1} u_p$, the domain $\mathcal{L}_{i_{p-1}}$ would be a compact triangle, which is not allowed. Hence, $\mathbf{J}_{p-1}$ has the form $(A, \dots, A)$ and $\mathbf{J}_{p-2}$ has the form $(A \text{ or } D, \dots, A)$. Repeating this argument p times, we obtain the desired statement.

Moreover, let e_l be the internal label on $u_{l-1} u_l$ and let $n_l = n(u_l)$. Then one can see that the sequence $j_{-k}, \dots, j_{-1}$ has the following structure:

$$\left.\begin{matrix} A_L[1, n_1 - 1](l(e_1)) \\ A_{RL}[i_-, n_1 - 1](l(e_1)) \\ D(l(e_1), r(l(e_1))^{-1}) \end{matrix}\right\}, A_L[2, n_1 - 2](l^2(e_1)), \dots, A_L[n_1 - 1, 1](l^{n_1 - 1}(e_1)), \therefore,$$

$$A_L[1,n_t-1](l(e_t)), A_L[2,n_t-2](l^2(e_t)), \dots, A_L[n_t-1,1](l^{n_t-1}(e_1)), \because\because,$$
$$A_L[1,m_- -1](l(e_p)), \dots, A_L[n_- -2, m_- - n_-](l^{n_- -2}(e_p)),$$
$$\begin{cases} A_L[n_- -1, m_- - n_- + 1](l^{n_- -1}(e_p)) \\ A_{LR}[n_- -1, i_+](l^{n_- -1}(e_p))^\dagger \end{cases},$$

where n_- is the number of levels in $\underline{\mathcal{S}}^-$ adjacent to v, and the number m_- can be chosen arbitrarily satisfying the inequality $n_- \le m_- \le n(v)$; the case marked with † is allowed only if $m_- = n_-$. Here every segment separated by "$\because\because$" is one of the $\mathbf{J}_t$'s. (Note that here and below we somewhat abuse the notation for states: thus for example, $A_{LR}[1,i_+](e)$ means $A_R[1,i_+](e)$, $A_L[1,1](e)$ means $A_0(e)$, etc.)

Similarly, let K be such that $\partial_L \mathcal{S}_{K+1} \supset w_+u_+$, let $U_0 = v, U_1, \dots, U_q = u_+$ be the consecutive vertices on vu_+, let E_l be the external label on $U_{l-1}U_l$ and let $N_l = n(U_{l-1})$. Then the sequence $(j_0, \dots, j_K)$ has the form

$$\left.\begin{matrix} A_L[m_+ - n_+ + 1, n_+ - 1](l^{1-n_+}(E_1)) \\ A_{RL}[i_+, n_+ - 1](l^{1-n_+}(E_1))^\dagger \end{matrix}\right\},$$
$$A_L[m_+ - n_+ + 2, n_+ - 2](l^{2-n_+}(E_1)), \dots, A_L[m_+ - 1, 1](l^{-1}(E_1)), \because\because,$$
$$A_L[1, N_t - 1](l^{1-N_t}(E_t)), A_L[2, N_t - 2](l^{2-N_t}(E_t)), \dots,$$
$$A_L[N_t - 1, 1](l^{-1}(E_t)), \because\because, A_L[N_q - 1, 1](l^{1-N_q}(E_q)), \dots,$$
$$A_L[N_q - 2, 2](l^{-2}(E_q)), \begin{cases} A_L[N_q - 1, 1](l^{-1}(E_q)) \\ A_{LR}[N_q - 1, i_+](l^{-1}(E_q)) \\ B(l^{-1}(E_q), r(l^{-1}(E_q)^{-1})) \end{cases},$$

where n_+ is the number of levels in $\underline{\mathcal{S}}^+$ adjacent to v, the number m_+ should satisfy $n_+ \le m_+ \le n(v)$, and the case marked with † is allowed only if $m_+ = n_+$.

Let us now define the states $\hat{\jmath}_{-k}, \dots, \hat{\jmath}_K$; see Figure 12. In some sense, this is opposite to the transformation in the proof of Lemma 5.2. For $l = -k, \dots, -1$ we use the following substitutions:

$$\left.\begin{matrix} A_L[1,\dots](l(e)) \\ A_{RL}[\dots](l(e)) \end{matrix}\right\} \mapsto \begin{cases} B(e^{-1}, l(e)) & \text{if } l = -k, \\ E_R(l(e^{-1})^{-1}, e^{-1}, l(e)) & \text{otherwise,} \end{cases}$$
$$A_{L,LR}[i,\dots](l^i(e)) \mapsto C_{i-1}(r^{i-1}(e^{-1}), l^i(e)), \quad i \ge 2,$$
$$D(l(e), \tilde{e}) \mapsto E_L(e^{-1}, l(e), \tilde{e}),$$

while for $l = 0, \dots, K$, we use

$$\left.\begin{matrix} A_L[\dots,1](l^{-1}(f)) \\ A_{LR}[\dots](l^{-1}(f)) \end{matrix}\right\} \mapsto \begin{cases} D(f^{-1}, l^{-1}(f)) & \text{if } l = K, \\ E_L(r(f), f^{-1}, l^{-1}(f)) & \text{otherwise,} \end{cases}$$

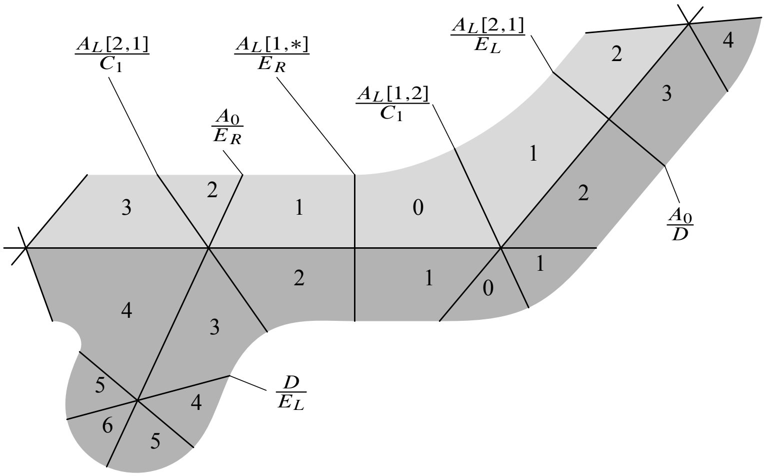


Figure 12. The proof of Lemma 5.4. The original sequence $\underline{S}^- \cup \underline{S}^+$ contains only dark gray domains, while $\widehat{\underline{S}}$ contains both light and dark gray ones. Numbers inside the domains indicate their levels. The states j_i and $\hat{j}_i$ are shown above or below the curves s_i. A The asterisk in j_{-1} represents 1 or 2. It is also possible that $j_{-1} = A_R[1, i_+]$ and/or $j_0 = A_{RL}[i_-, 2]$.

$$A_{L,RL}[\dots, i](l^{-i}(f)) \mapsto C_{n(v_L(f))-i}(r^{-i+1}(f^{-1}), l^{-i}(f)), \quad i \geq 2,$$
$$B(l^{-1}(f), \tilde{e}) \mapsto E_R(f^{-1}, l^{-1}(f), \tilde{e}).$$

It is straightforward to check all transitions within $(\hat{j}_{-k}, \dots, \hat{j}_{-1})$ and $(\hat{j}_0, \dots, \hat{j}_{K-1})$ are admissible. Let us verify that $\hat{j}_{-1} \to \hat{j}_0$ is also admissible.

Indeed, if $n_- > 2$, we have

$$\hat{j}_{-1} = C_{n_--2}(r^{n_--2}(e_1^{-1}), l^{n_--1}(e_1)),$$

and if $n_- = 2$, this state is of type B or E_R and has the same allowed next states as "the state $C_0(e_1, l(e_1))$". Similarly,

$$\hat{j}_0 = C_{n(v)-n_++1}(r^{2-n_+}(f_1^{-1}), l^{1-n_+}(f_1));$$

for $n_+ = 2$, this formula means that $\hat{j}_0$ has the same allowed previous states as this "state $C_{n(v)-1}(\dots)$".

By the assumptions of the lemma, v is adjacent to $n(v) + 1$ domains in $\underline{S}$, hence

$$n(v) = n_- + n_+ - 2, \quad r^{n(v)-1}(e_1^{-1}) = f_1^{-1}, \quad l^{n(v)+1}(e_1) = f_1.$$

Indeed, going around v from $u_{p-1}v$ to vu_1 in an anti-clockwise direction, one crosses all these $n(v)+1$ domains in $\underline{S}$, while going clockwise one passes through the remaining $n(v)-1$ domains in $\mathcal{F}_v$. Therefore, the transition $\widehat{\jmath}_{-1} \to \widehat{\jmath}_0$ takes the form

$$C_s(e_L, e_R) \to C_{s+1}(l(e_L), r(e_R)),$$

which is admissible.

It remains to define the states $\widehat{\jmath}_l$ for $l < -k$ and $l > K$. To do so, we join the sequences

$$\underline{\mathcal{S}}^{--} = (\mathcal{S}_{-m}, \dots, \mathcal{S}_{-k}), \quad \underline{\mathcal{S}}^{++} = (\mathcal{S}_{K+1}, \dots, \mathcal{S}_n)$$

with the respective sequences of states

$$\underline{j}^{--} = (j_{-m}, \dots, j_{-k-1}), \quad \underline{j}^{++} = (j_{K+1}, \dots, j_{n-1})$$

to the constructed sequences

$$\underline{\widehat{\mathcal{S}}}^{\diamond} = (\widehat{\mathcal{S}}_{-k}, \dots, \widehat{\mathcal{S}}_{K+1}), \quad \underline{\widehat{\jmath}}^{\diamond} = (\widehat{\jmath}_{-k}, \dots, \widehat{\jmath}_K).$$

If $\mathcal{S}_{-k}$ contains two domains, then j_{-k} is of type D, $\widehat{\jmath}_{-k}$ is of type E_L, and they have the same set of admissible previous states, thus the joining is just a concatenation.

Otherwise let us show that the joining can be done by item (2)a of this lemma. Compare the conditions there for the joining of $\underline{\mathcal{S}}^{--}$ and $\underline{\widehat{\mathcal{S}}}^{\diamond}$ and for that of $\underline{\mathcal{S}}^{--}$ and $(\mathcal{S}_{-k}, \dots, \mathcal{S}_0)$, which is clearly admissible. Since $\widehat{s}_{-k} = s_{-k} \cup u_0u_1$ and $u_0u_1 \subset \partial \underline{\mathcal{S}}^-$, the curves s_{-k-1} and $\widehat{s}_{-k}$ have no common sides. If s_{-k-1} and s_{-k} share their right end w, no domains adjacent to w are added in $\underline{\widehat{\mathcal{S}}}^{\diamond}$, so w is convex for both joinings. Now consider u_0. By definition the angle of $\partial_L \underline{\mathcal{S}}^-$ at u_0 is less than π, hence it is adjacent to less than $n(u_0)$ domains in the first joining. The only domain adjacent to u_0 that is added in $\underline{\widehat{\mathcal{S}}}^{\diamond}$ is the one adjacent to $\mathcal{S}_{-k}$ via the side u_0u_1. Therefore, u_0 is adjacent to at most $n(u_0)$ domains in the second joining. Note also that since $\widehat{\jmath}_{-k}$ is of type B, the sequence $\underline{\widehat{\jmath}}^{\diamond}$ stays the same after this joining.

Therefore, both sequences can be joined to $\underline{\widehat{\mathcal{S}}}^{\diamond}$ at the same time. This operation can change only those states in $\underline{j}^{--}$ and $\underline{j}^{++}$ where the corresponding levels are adjacent to an end of $\widehat{s}_{-k}$ or $\widehat{s}_K$. Moreover, this end should be adjacent to different numbers of domains in $\underline{\widehat{\mathcal{S}}}^{\diamond}$ and in $(\mathcal{S}_{-k}, \dots, \mathcal{S}_{K+1})$, hence it is u_- or u_+.

(3) This statement is again a direct consequence of the estimates in Lemma 5.1. In the case from (2a) we replace j_0 by $\widehat{\jmath}_0$ representing the same configuration and adjust the next states by that lemma. Hence we can get $j_t \neq \widehat{\jmath}_t$ only for $t \le n_0 - 3$, and not for $t = n-1 \ge n_0 - 2$. Similarly, in the case (2b) the equality $\mathcal{S}_{n-n_0+1} = \widehat{\mathcal{S}}_{n-n_0+1}$ means that $K < n - n_0 + 1$, hence the modification of $\underline{j}^{++} = (j_{K+1}, \dots, j_{n-1})$ reaches to at most $j_{(K+1)+n_0-3}$, and $(K+1)+n_0-3 < n-1$, whence again $\widehat{j}_{n-1} = j_{n-1}$. ■

The following corollary describes a precise sense in which the operations of narrowing and joining are mutually inverse. It is needed in the proof of Lemma 8.12 which is crucial for verifying Assumption 8.4 for convergence in Section 8.1.

Corollary 5.5. *Let $\widehat{\underline{\mathcal{S}}}$ be constructed from $\underline{\mathcal{S}}^-$ and $\underline{\mathcal{S}}^+$ as in Lemma 5.4. Then applying Lemma 5.2 to the sequence $\widehat{\underline{\mathcal{S}}}^- = (\widehat{\mathcal{S}}_i)_{i=-m}^0$ or $\widehat{\underline{\mathcal{S}}}^+ = (\widehat{\mathcal{S}}_i)_{i=0}^n$ and the domain $\widehat{\mathcal{S}}_0' = \mathcal{S}_0$, we arrive at the original sequences $\underline{\mathcal{S}}^\pm$.*

Proof. If there are no concave vertices in $\underline{\mathcal{S}}$, then $\widehat{\underline{\mathcal{S}}} = \underline{\mathcal{S}}$, $\widehat{\mathcal{S}}_0$ contains only one domain, hence no domains are removed when applying Lemma 5.2. Let us now assume that the vertex v, which is the common left end of s_0 and s_{-1}, is minimally concave. Using the notation of Lemma 5.4, one can see that any side in u_-v separates $\widehat{\mathcal{L}}_{l+1}$ and $\widehat{\mathcal{R}}_l = \mathcal{S}_l$ for some $l < 0$, while sides in vu_+ separate $\widehat{\mathcal{R}}_l = \mathcal{S}_l$ and $\widehat{\mathcal{L}}_{l-1}$ with $l > 0$. Hence it is impossible to connect the domain $\mathcal{B}$ in $\widehat{\underline{\mathcal{S}}}$ to $\widehat{\mathcal{R}}_0 = \mathcal{S}_0$ by a path of adjacent domains in $\widehat{\underline{\mathcal{S}}}$ with monotonic indices if and only if $\mathcal{B}$ is one of the domains $\widehat{\mathcal{L}}_{-k+1}, \dots, \widehat{\mathcal{L}}_K$, which comprise $\bigcup \widehat{\underline{\mathcal{S}}} \setminus \bigcup \underline{\mathcal{S}}$. Therefore, these are the domains removed when we apply Lemma 5.2. ∎

6. Strong connectivity and aperiodicity

In this section we will establish strong connectivity and aperiodicity of the Markov chain (Ξ, Π) constructed in Section 4. Together these properties yield the existence of $N > 0$ such that all entries of the matrix Π^N are positive, where Π is the adjacency matrix as defined in Section 4.2.

Definition 6.1. Let X and M be respectively the set of states and the adjacency matrix of a topological Markov chain.

(1) (X, M) is *strongly connected* if for any $x, y \in X$ there exists a sequence

$$z_0 = x, z_1, \dots, z_k = y$$

such that $z_j \to z_{j+1}$ is an admissible transition for all $j = 0, 1, \dots, k-1$.

(2) (X, M) is *aperiodic* if there does not exist $c > 1$, together with a map $\tau\colon X \to \mathbb{Z}/c\mathbb{Z}$, such that for every admissible transition $x \to y$, one has

$$\tau(y) = \tau(x) + 1.$$

We remark that the terms “Markov chain” and “topological Markov chain” as used here are interchangeable, both meaning a “subshift of finite type”.

6.1. Head and tail paths

We start by constructing some special short admissible sequences, of length depending only on the geometry of $\mathbf{T}_{\mathcal{R}}$, which will be useful in enabling us to pass from one state to another. Since these sequences can be attached either to the beginning or end of an admissible sequence, we call them "head" and "tail" sequences, whence the notation, $\underline{\mathcal{H}}^e$ and $\underline{\mathcal{T}}^e$.

Proposition 6.2. *There exists $M \le 4$, depending only on $\mathbf{T}_{\mathcal{R}}$, such that for every $e \in G_0$ there is an admissible sequence of states*

$$\underline{i}^e = (i_0^e \to \cdots \to i_{M-1}^e = \mathbf{t}(e))$$

with the following properties. First, $i_0^e = A_0(e)$ unless $\mathcal{R}$ belongs to the special case of Remark 4.5 and e is the label on the compact side, in which case i_0^e can be chosen to be either $A_{LR}[2,2](e)$ or $A_{RL}[2,2](e)$. Let $\underline{\mathcal{T}}^e = (\mathcal{T}_0^e = \mathcal{R}, \mathcal{T}_1^e, \dots, \mathcal{T}_M^e)$ be the sequence of levels generated by $\underline{i}^e$ and let γ_L, γ_R be the maximal geodesic segments of the curve $\partial\underline{\mathcal{T}}^e \setminus \partial\mathcal{T}_0^e$ adjacent respectively to the left or right end of $s_0 = \mathcal{T}_0^e \cap \mathcal{T}_1^e$ provided that this end is in $\mathbb{D}$, otherwise let $\gamma_{L,R}$ be empty. Then $\gamma_{L,R}$ have no common points with $s_{M-1} = \mathcal{T}_{M-1}^e \cap \mathcal{T}_M^e$.

Similarly, there exists a sequence of states

$$\underline{j}^e = (\mathbf{h}(e) = j_{-M}^e \to \cdots \to j_{-1}^e),$$

so that if $\underline{j}^e$ generates a sequence of levels $\underline{\mathcal{H}}^e = (\mathcal{H}_{-M}^e, \mathcal{H}_{-M+1}^e, \dots, \mathcal{H}_0^e = \mathcal{R})$, then the maximal geodesic segments of the curve $\partial\underline{\mathcal{H}}^e \setminus \partial\mathcal{H}_0^e$ starting at its ends lying inside $\mathbb{D}$ have no common points with $\mathcal{H}_{-M}^e \cap \mathcal{H}_{-M+1}^e$, and $j_{-1}^e = A_0(e)$ except for the label on the compact side in the special case of Remark 4.5, where j_{-1}^e is any of

$$A_{LR}[2,2](e), \quad A_{RL}[2,2](e).$$

Proof. The proposition and proof is illustrated in Figure 13. We will construct the paths $\underline{\mathcal{T}}^e$ in such a way that all states i_k^e are of type A.

If $N(\mathcal{R}) \ge 5$, take $M = 3$ and consider a path $\underline{\mathcal{T}}^e$ such that $\partial_L\mathcal{T}_1^e$ and $\partial_R\mathcal{T}_2^e$ contain at least two sides each, see Figure 13 (a). Then the geodesic segments $\gamma_{L,R}$ cannot pass through the internal vertices of $\partial_L\mathcal{T}_1^e$ and $\partial_R\mathcal{T}_2^e$, hence they cannot touch $\mathcal{T}_3^e$.

If $N(\mathcal{R}) = 4$ and $\mathcal{R}$ is non-compact, one can construct $\underline{\mathcal{T}}^e$ as shown in Figure 13 (b) or (c).

Suppose $N(\mathcal{R}) = 4$ and $\mathcal{R}$ is compact and has no opposite vertices with $n(v) = 2$. Let s_1 be the side of $\mathcal{T}_1^e = e\mathcal{R}$ opposite to s_0, and choose $\mathcal{T}_2^e$ to be the domain on the other side of s_1. Then there are the following three cases. If both ends of s_1 have $n(\dots) \ge 3$, we use the path shown in Figure 13 (d). Alternatively, if both ends of s_0

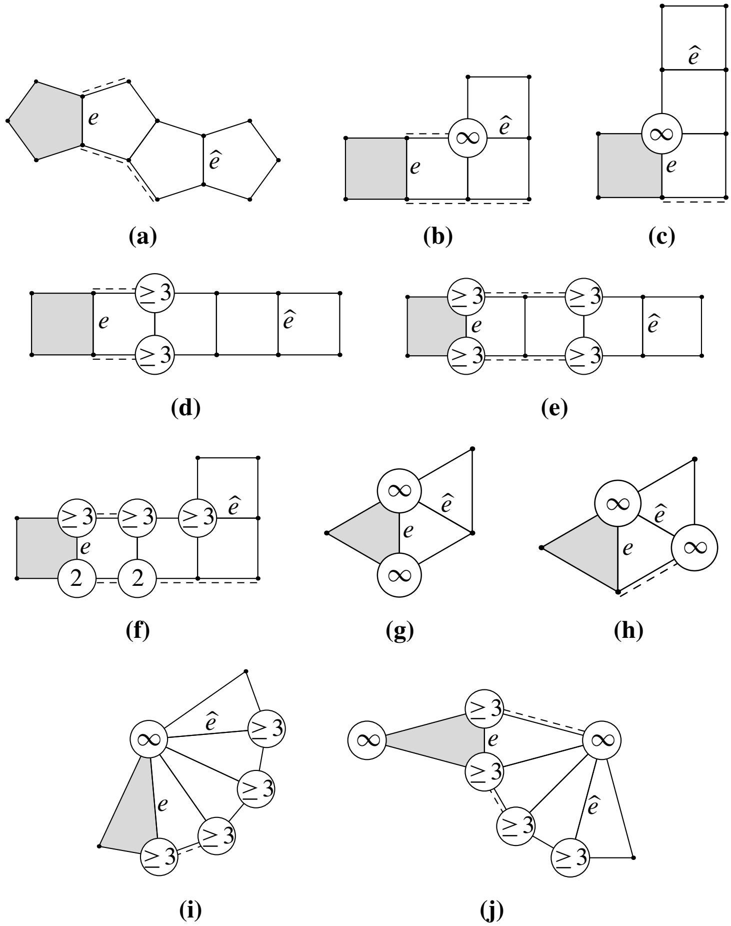


Figure 13. "Tail" paths $\underline{\mathcal{T}}^e$ from Proposition 6.2. The domains $\mathcal{T}_0^e$ are shaded gray. Dashed lines show the maximal possible extent of the segments $\gamma_{L,R}$. Numbers in circles indicate $n(\dots)$ for the vertex, ∞ meaning that the vertex lies on $\partial\mathbb{D}$. Note that $\mathbf{t}(e) = A_0(\hat{e})$ in all cases except (f), in which case $\mathbf{t}(e) = A_L[2,1](\hat{e})$.

have $n(\dots) \geq 3$, the same holds for both ends of s_2, the side of $\mathcal{T}_2^e$ opposite to s_1, and we use the path from Figure 13 (e). The remaining case is when both s_0 and s_1 have ends with $n(\dots) = 2$. Then these ends lie on the same (say, right) boundary, and the left ends of s_0, s_1, and s_2 all have $n(\dots) \geq 3$. Then we construct our path as shown in Figure 13 (f).

If $N(\mathcal{R}) = 3$ and all of the sides are non-compact, we can use the paths shown in Figure 13 (g),(h).

Finally, in the special case from Remark 4.5 the paths $\underline{\mathcal{T}}^e$ are defined as shown in Figure 13 (i),(j). If e is the label on the compact side then as in Figure 13 (j), we have

$$i_0^e = A_{LR}[2,2](g),$$

while the analogous symmetrical path gives the sequence with $i_0^e = A_{RL}[2,2](g)$.

As for the second statement concerning the head paths $\underline{\mathcal{H}}^e$, one can set

$$j_{-k}^e = \iota(i_{k-1}^{e^{-1}}), \quad \mathcal{H}_{-k}^e = \mathcal{T}_k^{e^{-1}}, \qquad k = 0, \dots, M.$$

In particular, we have $\mathbf{h}(e) = \iota(\mathbf{t}(e^{-1}))$. ∎

The next proposition describes two important combinations of the "head" and "tail" sequences, which will be directly used to prove connectivity and aperiodicity.

Proposition 6.3. *There are two cases to consider.*

(1) *Take any $e, \widehat{e}$ such that $\widehat{e} \neq e^{-1}$ and let $\underline{\mathcal{H}}^e$, $\underline{\mathcal{T}}^e$ be the paths from Proposition 6.2. Then Lemma 5.4 can be applied to join $\underline{\mathcal{H}}^{\widehat{e}}$ and $\underline{\mathcal{T}}^e$. The resulting path is generated by an admissible sequence $\underline{k}$ of states with $k_{-M} = \mathbf{h}(\widehat{e})$, $k_{M-1} = \mathbf{t}(e)$.*

(2) *Take any e, $\widehat{e}$ such that $\widehat{e} \neq e^{-1}$ and denote $\mathcal{S}_0 = \mathcal{R}$, $\mathcal{S}_1 = e\mathcal{R}$. At most one vertex w of $\mathcal{S}_1$ is shared with $\mathcal{H}_{-1}^{\widehat{e}}$. Choose any side of $\mathcal{S}_1$ that is not incident to w and let $\widetilde{e}$ be the label on this side outside $\mathcal{S}_1$. Denote $\mathcal{S}_{l+1} = e\mathcal{T}_l^{\widetilde{e}}$ for $l = 0, \dots, M$. Then one can apply Lemma 5.4 twice to join the three sequences,*

$$\underline{\mathcal{S}}^- = \underline{\mathcal{H}}^{\widehat{e}}, \quad \underline{\mathcal{S}}^\diamond = (\mathcal{S}_l)_{l=0}^1, \quad \underline{\mathcal{S}}^+ = (\mathcal{S}_l)_{l=1}^{M+1}.$$

The sequence of states $\underline{k}$ corresponding to this joining starts with $\mathbf{h}(\widehat{e})$ and ends with $\mathbf{t}(\widetilde{e})$.

Proof. We consider each case in turn.

(1) We use the notation of Lemma 5.4. The sides

$$s_{-1} = \mathcal{S}_{-1} \cap \mathcal{S}_0 \quad \text{and} \quad s_0 = \mathcal{S}_0 \cap \mathcal{S}_1$$

do not coincide since $\widehat{e} \neq e^{-1}$. Assume that they have a common vertex v. In the non-special case, v is incident to only two domains, $\mathcal{S}_{-1}$ and $\mathcal{S}_0$, in $\underline{\mathcal{H}}^{\widehat{e}}$ since $i_{-1}^{\widehat{e}} = A_0(\widehat{e})$. Similarly, v is incident to only two domains in $\underline{\mathcal{T}}^e$. Therefore, v is incident only to $\mathcal{S}_{-1}$, $\mathcal{S}_0$, $\mathcal{S}_1$ in $\underline{\mathcal{S}}$, hence v is either convex or minimally concave. In the special case from Remark 4.5 at most one of the sides s_{-1}, s_0 is compact, hence by the same argument v is adjacent to at most four domains in $\underline{\mathcal{S}}$, and we use that $n(v) \geq 3$. The last part of the statement follows directly from Proposition 6.2.

(2) The paths $\underline{\mathcal{S}}^-$ and $\underline{\mathcal{S}}^+$ are generated by the sequences $\underline{j}^{\hat{e}}$ and $(i^{\tilde{e}}_{k-1})^M_{k=1}$ respectively, while $\underline{\mathcal{S}}^\diamond$ is generated by any state of the form $A_{\dots}(e)$. As in the first part of the proof, we can apply Lemma 5.4 to the paths $\underline{\mathcal{S}}^-$ and $\underline{\mathcal{S}}^\diamond$ and obtain the path $\underline{\widetilde{\mathcal{S}}}$ and the sequence of states $(\tilde{j}_l)^0_{l=-M}$ such that

$$s_{-M} = \mathcal{S}_{-M} \cap \mathcal{S}_{-M+1}$$

have no common points with the domains added in this step and $\tilde{j}_{-M} = \mathbf{h}(\hat{e})$.

Let us now consider two cases. Assume first that no domains are added at the first step. Then $\tilde{s}_0 = s_0$ and $s_1 = \mathcal{S}_1 \cap \mathcal{S}_2$ are different sides of $\mathcal{S}_1$, and if they have a common vertex u, then $u \neq w$, hence u is incident only to $\mathcal{S}_0$ and $\mathcal{S}_1$ in $\underline{\widetilde{\mathcal{S}}}$, that is, u is either convex or minimally concave in $\underline{\widetilde{\mathcal{S}}} \cup \underline{\mathcal{S}}^+$. Therefore, we can apply Lemma 5.4 to $\underline{\widetilde{\mathcal{S}}}$ and $\underline{\mathcal{S}}^+$ and obtain a path $\underline{\widehat{\mathcal{S}}}$ and corresponding sequence $\underline{k}$ of states. As above,

$$k_M = i^{\tilde{e}}_{M-1} = \mathbf{t}(\tilde{e}).$$

To show that $k_{-M} = \tilde{j}_{-M}$, we need to check that the maximal geodesic segment γ in $\partial\underline{\widetilde{\mathcal{S}}} \setminus s_1$ starting at u does not reach s_{-M}. Clearly, if γ does not go past an end z of s_0, it does not reach s_{-M}. On the other hand, if γ goes past z, then it ends at the same point as the maximal geodesic segment of $\partial\underline{\mathcal{H}}^{\hat{e}} \setminus s_0$ starting from z, hence γ does not reach s_{-M} in this case also.

Now assume that we have added some domains in the first step. This means that the vertex w exists, the curve $\tilde{s}_0 = \widetilde{\mathcal{S}}_0 \cap \mathcal{S}_1$ consists of the two sides of $\mathcal{S}_1$ that are adjacent to w, and $\widetilde{\mathcal{S}}_0 = \mathcal{S}_0 \cup \mathcal{A}$ consists of two domains adjacent to these sides. Therefore, the curves $\tilde{s}_0$ and s_1 have no common sides, and if they have a common vertex $\tilde{w}$, then $\tilde{w}$ is the end of $\tilde{s}_0$, hence it is adjacent only to $\widetilde{\mathcal{S}}_0$ and $\widetilde{\mathcal{S}}_1$ in $\underline{\widetilde{\mathcal{S}}}$. Therefore, we may apply Lemma 5.4 and get $k_M = \mathbf{t}(\tilde{e})$ as in the previous case. To show that $k_{-M} = \mathbf{h}(\hat{e})$, it remains to check that the domains added in this step have no common points with s_{-M}. Indeed, no domains are added to $\widetilde{\mathcal{S}}_l$ with $l \le 0$ since $\widetilde{\mathcal{S}}_0$ already contains two domains. On the other hand, if s_{-M} has a common point with $\widehat{\mathcal{S}}_l$ for $l \ge 1$, then it has a common point with

$$\widehat{\mathcal{S}}_0 = \widetilde{\mathcal{S}}_0 = \mathcal{S}_0 \cup \mathcal{A}.$$

The side s_{-M} has no common points with $\mathcal{S}_0$ by the construction of $\underline{\mathcal{H}}^{\hat{e}}$, and if s_{-M} has a common point with $\mathcal{A}$, we arrive at a contradiction with the first application of Lemma 5.4: the addition of $\mathcal{A}$ then changes the indices $i_{+,L/R}$ for the state corresponding to s_{-M}, i.e. $\tilde{j}_{-M} \neq j^{\hat{e}}_{-M}$. ∎

6.2. Connectivity and aperiodicity

Proposition 6.4. *The topological Markov chain* (Ξ, Π) *introduced in Definition 4.7 is strongly connected.*

Proof. The scheme of the proof is the following. We will consider several cases. In each case we choose the set $\Omega \subset \Xi$ with $\iota(\Omega) = \Omega$ and prove two properties:

(i) for every $j \in \Xi$ there exists a path along the arrows in the adjacency graph of the Markov chain from j to some state $k \in \Omega$;

(ii) for every $k_1, k_2 \in \Omega$ there exists a path from k_1 to k_2.

Let us write $j \rightsquigarrow k$ to indicate that there exists a path going from j to k. Observe that properties (i) and (ii) imply strong connectivity. Namely, from (i) we have that for any states $j, j' \in \Xi$ there exist $k, k' \in \Omega$ such that $j \rightsquigarrow k$ and $\iota(j') \rightsquigarrow k'$. Applying the involution ι to the second of these relations, we get $\iota(k') \rightsquigarrow j'$. Finally, (ii) yields $k \rightsquigarrow \iota(k')$, and we have

$$j \rightsquigarrow k \rightsquigarrow \iota(k') \rightsquigarrow j'.$$

Property (i) is proven in the same way in all cases. Namely, if $\mathcal{R}$ does not belong to the special case of Remark 4.5, let $\Omega = \{A_0(e) : e \in G_0\}$. First of all, we can reach an A-state from a state j via a series of states of the form $C \dots CDA$. Then we transform an A-state to a state with smaller index i_+, arriving eventually at a state k with $\pi(k) = A(\widehat{e})$ and $i_{+,L/R}(k) = 1$. Then the state k can be followed by a state $A_0(e)$, where e is any label not adjacent to $\widehat{e}^{-1}$.

In the special case of Remark 4.5 we set

$$\Omega = \{A_0(e) : s_e \text{ is not compact}\}. \tag{6}$$

The procedure given above brings us either to Ω or to an A-state with $i_+ = 2$, say,

$$\pi(k) = A(\widehat{e}), \quad i_{+,L}(k) = 2, \quad i_{+,R}(k) = 1.$$

If $s_{\widehat{e}}$ is compact, we have $k \to A_L[i_-, 1](\widetilde{e}) \to A_0(\widetilde{e})$, where $\widetilde{e} = l(\widehat{e})$ is the label on a non-compact side, and $i_- = i_{-,L}(k) + 1$. Similarly, if $s_{\widehat{e}}$ is non-compact, then $s_{\widetilde{e}}$ is compact, $k \to A_{LR}[i_-, 2](\widetilde{e})$, and we reduce this case to the previous one.

It remains to check property (ii).

(1) Assume $N(\mathcal{R}) \geq 5$. Let us construct the tail paths $\underline{\mathcal{T}}^e$ from Proposition 6.2 and shown in Figure 13 (a) in a uniform manner, namely, we choose the domains $\mathcal{T}^e_{2,3}$ in such a way that $\partial_R \mathcal{T}^e_1$ and $\partial_L \mathcal{T}^e_2$ contain one side each. Denote by $t(e)$ the label such that

$$\mathbf{t}(e) = A_0(t(e));$$

$t(e)$ is shown as $\widehat{e}$ in Figure 13 (a). Then $t \colon G_0 \to G_0$ is a bijection. Indeed, for any $e' \in G_0$, a path $\underline{\mathcal{T}} = (\mathcal{T}_i)_{i=0}^3$ such that $(\mathcal{T}_2, \mathcal{T}_3)$ represents the configuration $A(e')$ and $\partial_L \mathcal{T}_2$ and $\partial_R \mathcal{T}_1$ contain one side each is unique up to the group action. Hence, if we require $\mathcal{T}_0 = \mathcal{R}$, we have $\underline{\mathcal{T}} = \underline{\mathcal{T}}^e$ for some $e \in G_0$, thus $e' = t(e)$. Note also that for the map h defined by $\mathbf{h}(e) = A_0(h(e))$, we have

$$h(e) = (t(e^{-1}))^{-1}$$

by the construction of the paths $\underline{\mathcal{H}}^e$.

Now take any $f, \hat{f} \in G_0$ such that $\hat{f} \neq f^{-1}$ and denote

$$e = t^{-1}(f), \quad \hat{e} = h^{-1}(\hat{f}) = (t^{-1}(\hat{f}^{-1}))^{-1}.$$

Hence, $e \neq \hat{e}^{-1}$, and we may consider the path from the first part of Proposition 6.3 for these $e, \hat{e}$. This shows that $A_0(\hat{f}) \rightsquigarrow A_0(f)$ for any $f, \hat{f}$ such that $\hat{f} \neq f^{-1}$. Finally, to verify $A_0(f) \rightsquigarrow A_0(f^{-1})$ choose any $e \in G_0 \setminus \{f, f^{-1}\}$ and observe that

$$A_0(f) \rightsquigarrow A_0(e) \rightsquigarrow A_0(f^{-1}).$$

(2) Let us assume that $N(\mathcal{R}) = 4$ and $\mathcal{R}$ is compact, and define the bijection $\tau\colon G_0 \to G_0$ as follows. For $e \in G_0$ consider the side of $\mathcal{R}$ with the inside label e, then the outside label on the opposite side of $\mathcal{R}$ equals $\tau(e)$. Then $A_0(e) \to A_0(\tau(e))$. Choose m such that $\tau^m = \mathrm{id}$.

Assumption 1.1 implies that there is a label $f \in G_0$ such that $n(v_L(f)) \geq 3$ and $n(v_R(f)) \geq 3$. Then we have

$$\begin{aligned} A_0(f) \rightsquigarrow A_0(\tau^{m-1}(f)) &\to A_L[1,2](f) \\ &\to A_L[2,1](l(f)) \to A_0(\tau(l(f))) \rightsquigarrow A_0(l(f)). \end{aligned}$$

Similarly, $A_0(f) \rightsquigarrow A_0(r(f))$, thus $A_0(f) \rightsquigarrow A_0(g)$ for all $g \neq f^{-1}$. By the involution, we get $A_0(g) \rightsquigarrow A_0(f^{-1})$ for $g \neq f$. The inequalities $n(v_{L,R}(f)) \geq 3$ imply that $n(v_{R,L}(f^{-1})) \geq 3$, so by the same argument we have $A_0(f^{-1}) \rightsquigarrow A_0(g)$ for all $g \neq f$, $A_0(g) \rightsquigarrow A_0(f)$ for $g \neq f^{-1}$. Finally, $A_0(f) \leftrightsquigarrow A_0(f^{-1})$ since for $h \neq f, f^{-1}$, we have

$$A_0(f) \leftrightsquigarrow A_0(h) \leftrightsquigarrow A_0(f^{-1}),$$

hence property (ii) holds.

(3) The argument from the previous case works for a non-compact $\mathcal{R}$ with $N(\mathcal{R}) = 4$ if there exists a label $f \in G_0$ such that for both $\alpha = L, R$ we have that either $v_\alpha(f)$ is undefined or $n(v_\alpha(f)) \geq 3$. If, say, the left end of s_f lies on $\partial\mathbb{D}$ then one can simply write

$$A_0(f) \to A_0(l(f)).$$

(Here we use an extended definition of $l(f)$: the left end of s_f and right end of $s_{l(f)^{-1}}$ are either the same point of $\partial\mathbb{D}$ or they are joined by an arc from $\partial_{\overline{\mathbb{D}}}\mathcal{R} \cap \partial\mathbb{D}$.)

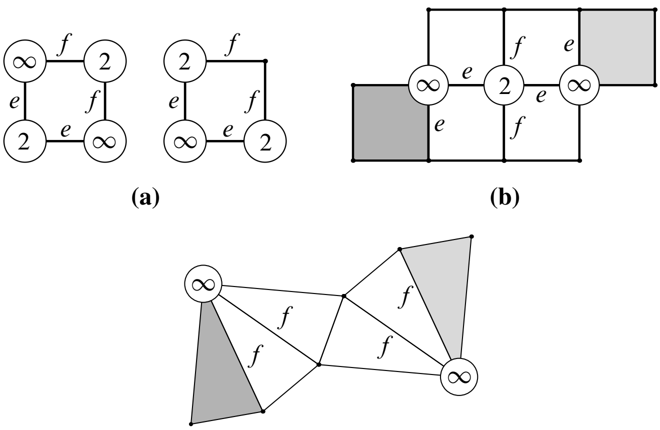


Figure 14. The proof of Proposition 6.4 (4) and (5). The notation is similar to that of Figure 13. The first and the last domains in the paths are shaded dark gray and light gray respectively.

(4) Consider now a non-compact $\mathcal{R}$ with $N(\mathcal{R}) = 4$ which is not covered by the previous case. Then $\mathcal{R}$ has either one vertex/arc on $\partial\mathbb{D}$ or two such vertices in opposite corners. It is easily checked that the only possible arrangements of side pairings are as shown in Figure 14 (a). In both cases τ has a pair of 2-cycles and we have

$$A_0(e) \to A_0(f) \to A_0(e), \quad A_0(e^{-1}) \to A_0(f^{-1}) \to A_0(e^{-1}).$$

In the left case we have $A_0(e) \to A_0(f^{-1})$, $A_0(e^{-1}) \to A_0(f)$, while in the right case the path shown in Figure 14 (b) yields $A_0(e) \rightsquigarrow A_0(e^{-1})$:

$$A_0(e) \to B(e, f) \to D(f^{-1}, e^{-1}) \to A_0(e^{-1}),$$

and similarly $A_0(e^{-1}) \rightsquigarrow A_0(e)$. Thus in both cases the cycles of τ are linked.

(5) Finally, consider the special case from Remark 4.5, where Ω is defined by (6). Let a side s_f be non-compact, say $g = r(f)$, where s_g is the compact side. Then property (ii), namely, that $A_0(f) \rightsquigarrow A_0(f^{-1})$, is shown in Figure 14 (c):

$$A_0(f) \to A_R[1,2](f) \to A_{RL}[2,2](g) \to A_L[2,1](f^{-1}) \to A_0(f^{-1}). \qquad \blacksquare$$

Proposition 6.5. *The topological Markov chain* (Ξ, Π) *defined in Definition 4.7 is aperiodic.*

Proof. Suppose that our Markov chain has a period c, that is, an index $\delta(i) \in \mathbb{Z}/c\mathbb{Z}$ can be assigned to every state $i \in \Xi$ in such a way that all allowed transitions $i \to j$ satisfy $\delta(j) = \delta(i) + 1$.

Take any e_1, e_2 and choose $\widehat{e} \neq e_1^{-1}, e_2^{-1}$. Then using the paths from the first part of Proposition 6.3, we have

$$\delta(\mathbf{t}(e_s)) = \delta(\mathbf{h}(\widehat{e})) + 2M - 1, \quad s = 1, 2,$$

whence $\delta(\mathbf{t}(e_1)) = \delta(\mathbf{t}(e_2))$. Therefore, $\delta(\mathbf{t}(e))$ is the same for all e, we denote it by $\delta_{\mathbf{t}}$. Similarly, $\delta(\mathbf{h}(e))$ equals the same number $\delta_{\mathbf{h}}$ for all e, and $\delta_{\mathbf{t}} = \delta_{\mathbf{h}} + 2M - 1$. On the other hand, any path from the second part of Proposition 6.3 yields $\delta_{\mathbf{t}} = \delta_{\mathbf{h}} + 2M$. Therefore,

$$2M - 1 \equiv 2M (\text{mod } c),$$

whence $c = 1$. ■

Corollary 6.6. *By Propositions 6.4 and 6.5 the Markov chain* (Ξ, Π) *is strongly connected and aperiodic, hence there exists* $N_0 > 0$ *such that all entries of the matrix* Π^{N_0} *are positive.*

Proof. That this result is implied by strongly connectivity and aperiodicity is well known; see, for example, [34]. ■

7. Spherical sums and Markov operator

In this section we express the spherical averages from equation (1) in terms of powers of a Markov operator (see Lemma 7.5), and obtain an identity relating this Markov operator with its adjoint; see Lemma 7.6.

7.1. Thickened paths and spheres in the group graph

Consider a state $k \in \Xi$ and let $(\mathcal{S}_-, \mathcal{S}_+)$ be a representation of the configuration $\pi(k)$. Let $\mathcal{L}_\pm, \mathcal{R}_\pm$ be the left and the right domains in $\mathcal{S}_\pm$; as usual, if the state has only one past (respectively, future) domain then $\mathcal{L}_- = \mathcal{R}_-$ (respectively, $\mathcal{L}_+ = \mathcal{R}_+$).

Define the maps $\gamma, \omega \colon \Xi \to G$ as follows: let $\mathcal{L}_- = h\mathcal{R}$, then

$$\mathcal{L}_+ = h\gamma(k)^{-1}\mathcal{R}, \quad \mathcal{R}_+ = h\omega(k)^{-1}\mathcal{R}. \tag{7}$$

Clearly, these definitions do not depend on the choice of a representation for k.

Lemma 7.1. *The maps defined above* γ *and* ω *satisfy the following identities:*

(1) $\omega(\iota(k)) = \omega(k)^{-1}$ *for any* $k \in \Xi$,

(2) $\gamma(k) = \omega(j)^{-1}\gamma(\iota(j))^{-1}\omega(k)$ *for any* $j, k \in \Xi$ *such that* $k \to j$ *is an admissible transition.*

Proof. We prove each identity separately.

(1) Consider a representation $(\mathcal{S}_-, \mathcal{S}_+)$ of the state k as above. Then $(\widetilde{\mathcal{S}}_-, \widetilde{\mathcal{S}}_+) = (\mathcal{S}_+, \mathcal{S}_-)$ is a representation of the state $\iota(k)$, and

$$\widetilde{\mathcal{L}}_\pm = \mathcal{R}_\mp, \quad \widetilde{\mathcal{R}}_\pm = \mathcal{L}_\mp.$$

Hence, if $\mathcal{L}_- = h\mathcal{R}$, $\mathcal{R}_+ = g\mathcal{R}$, then

$$\omega(k)^{-1} = h^{-1}g, \quad \omega(\iota(k))^{-1} = g^{-1}h.$$

(2) If $k \to j$ is admissible, one can consider the sets $\mathcal{S}_-$, $\mathcal{S}_+$, and $\mathcal{S}_{++}$ such that $(\mathcal{S}_-, \mathcal{S}_+)$ represents k and $(\mathcal{S}_+, \mathcal{S}_{++})$ represents j. Define $\mathcal{L}_\alpha, \mathcal{R}_\alpha$ as above for $\alpha \in \{-, +, ++\}$. Let $\mathcal{L}_- = h\mathcal{R}$. Then

$$\mathcal{L}_+ = h\gamma(k)^{-1}\mathcal{R}, \quad \mathcal{R}_+ = h\omega(k)^{-1}\mathcal{R}, \quad \mathcal{R}_{++} = h\gamma(k)^{-1}\omega(j)^{-1}\mathcal{R}.$$

On the other hand, $(\widetilde{\mathcal{S}}_-, \widetilde{\mathcal{S}}_+) = (\mathcal{S}_{++}, \mathcal{S}_+)$ represents the state $\iota(j)$, hence the formula for $\widetilde{\mathcal{L}}_- = \mathcal{R}_{++}$ yields that

$$\mathcal{R}_+ = \widetilde{\mathcal{L}}_+ = h\gamma(k)^{-1}\omega(j)^{-1}\gamma(\iota(j))^{-1}\mathcal{R}. \qquad \blacksquare$$

Recall that the set $\mathbf{P}_{n-1}^{S\to F}$ defined by equation (2) is the set of all admissible sequences of length n which begin and end in start and end states respectively. By Theorem 4.10, $\mathbf{P}_{n-1}^{S\to F}$ is in bijective correspondence with the set of thickened paths of length n and hence to the sphere of radius n in the graph of G. More precisely, we have the following lemma.

Lemma 7.2. *Consider the map* $\Phi\colon \mathbf{P}_{n-1}^{S\to F} \to G$, *where*

$$\Phi(j_0 \to \cdots \to j_{n-1}) = \omega(j_{n-1})\gamma(j_{n-2})\dots\gamma(j_0).$$

Then Φ *is a bijection of* $\mathbf{P}_{n-1}^{S\to F}$ *onto the set* $S_n(G) = \{g \in G : |g|_{G_0} = n\}$.

Remark 7.3. Note that for $j_{n-1} \in \Xi_F$ there is *only one* future fundamental domain, hence $\omega(j_{n-1}) = \gamma(j_{n-1})$. The reason for using ω rather than γ in the final step will become apparent later, see Lemma 7.5.

Proof. As observed above, sequences from $\mathbf{P}_{n-1}^{S\to F}$ bijectively correspond to thickened paths $\underline{\mathcal{S}}$ from $\mathcal{R}$ to $g\mathcal{R}$ with $g \in S_n(G)$. Take $\underline{j} \in \mathbf{P}_{n-1}^{S\to F}$. Let the sequence

$$\underline{\mathcal{S}} = (\mathcal{S}_0 = \mathcal{R}, \dots, \mathcal{S}_n)$$

be generated by $\underline{j}$ and let $\mathcal{L}_k, \mathcal{R}_k$ be the left and right domains in $\mathcal{S}_k$. Define $h_k \in G$ so that $\mathcal{L}_k = h_k\mathcal{R}$. Then

$$\begin{aligned} g &= h_{n-1}\omega(j_{n-1})^{-1} \\ &= h_{n-2}\gamma(j_{n-2})^{-1}\omega(j_{n-1})^{-1} \end{aligned}$$

$$= \cdots$$
$$= [\omega(j_{n-1})\gamma(j_{n-2}) \dots \gamma(j_0)]^{-1},$$

and it remains to use that $g \mapsto g^{-1}$ is a bijective map of the sphere $S_n(G)$. ∎

7.2. Parry measure

Let Π be the adjacency matrix of the topological Markov chain described in Definition 4.7. By Corollary 6.6 for some N_0 all elements of the matrix Π^{N_0} are positive. The Perron–Frobenius theorem then yields that the matrix Π has a unique (up to a scaling) eigenvector h with non-negative coordinates and that all coordinates of h are positive:

$$\sum_j \Pi_{ij} h_j = \lambda h_i \quad \text{and} \quad h_i > 0 \qquad \text{for all } i \in \Xi.$$

Moreover, the corresponding eigenvalue $\lambda > 0$ has multiplicity one and is larger than the absolute value of any other eigenvalue of Π. The eigenvalue λ is called the *Perron–Frobenius* (PF) eigenvalue and h is called the *right Perron–Frobenius* eigenvector. The matrix $\pi = (p_{ij})$ with entries

$$p_{ij} = \frac{h_j}{\lambda h_i} \Pi_{ij} \tag{8}$$

is stochastic and the corresponding Markov chain has the following property: the probability of an admissible sequence of transitions depends only on the initial and the final states in this sequence and the number of steps:

$$p_{i_0 i_1} \dots p_{i_{n-1} i_n} = \frac{h_{i_n}}{\lambda^n h_{i_0}} \Pi_{i_0 i_1} \dots \Pi_{i_{n-1} i_n} = \frac{h_{i_n}}{\lambda^n h_{i_0}}. \tag{9}$$

The Markov measure defined by the matrix $\pi = (p_{ij})$ is called the *Parry measure*. Its stationary distribution is

$$p_i = \alpha_i h_i, \tag{10}$$

where α is the left PF eigenvector of Π: $\alpha\Pi = \lambda\alpha$, normalized by $\alpha h = \sum_i \alpha_i h_i = 1$.

The time-reversing involution on the set of states implies certain symmetries for the Parry measure.

Proposition 7.4. *Suppose given an involution $\iota\colon \Xi \to \Xi$ such that $\Pi_{\iota(j)\iota(k)} = \Pi_{kj}$ for all $j, k \in \Xi$. Then the transition probability matrix (p_{ij}) and the stationary distribution (p_i) of the Parry measure corresponding to the matrix Π satisfy the following equations:*

$$p_{\iota(j)} = p_j, \quad p_{\iota(j)\iota(k)} = \frac{p_k p_{kj}}{p_j} \qquad \textit{for all } j, k \in \Xi.$$

Proof. Let J be the matrix for the substitution ι. Then $J = J^T = J^{-1}$, $J\Pi J = \Pi^T$. As above, let λ be the Perron–Frobenius (PF) eigenvalue for Π and let α and h be its left and right PF eigenvectors, normalized by $\alpha h = 1$. Then αJ is a left PF eigenvector for $J\Pi J = \Pi^T$, whence $(\alpha J)^T = J\alpha^T$ is a right PF eigenvector for Π. Therefore, $J\alpha^T$ is proportional to h: $\alpha_{\iota(k)} = ch_k$. Now

$$p_{\iota(j)} = \alpha_{\iota(j)} h_{\iota(j)} = ch_j \cdot \frac{1}{c}\alpha_j = p_j$$

and

$$p_{\iota(j)\iota(k)} = \frac{\Pi_{\iota(j)\iota(k)} h_{\iota(k)}}{\lambda h_{\iota(j)}} = \frac{\Pi_{kj} c^{-1}\alpha_k}{\lambda c^{-1}\alpha_j} = \frac{\Pi_{kj} h_j}{\lambda h_k}\frac{h_k \alpha_k}{h_j \alpha_j} = \frac{p_k p_{kj}}{p_j}. \qquad \blacksquare$$

7.3. The Markov operator

Recall that the group G acts on a Lebesgue probability space (X, μ) by measure-preserving maps T_g. We denote $T_g f := f \circ T_g^{-1}$ for any function $f \in L^p(X, \mu)$. Denote

$$\widetilde{\mathbf{S}}_n(f) = \sum_{|g|=n} T_g^{-1} f,$$

then

$$\mathbf{S}_n(f) = \frac{\widetilde{\mathbf{S}}_n(f)}{\widetilde{\mathbf{S}}_n(1)} = \frac{\sum_{|g|=n} T_g^{-1} f}{\#\{g : |g| = n\}},$$

where $\mathbf{S}_n(f)$ is defined by (1).

Consider the probability space $Y = \Xi \times X$ with the product measure $\nu = p \times \mu$. Here $p(\{i\}) = p_i$, where p_i is defined by (10). It is convenient to identify a function $\varphi \in L^1(Y, \nu)$ with a tuple of functions $(\varphi_i)_{i \in \Xi}$, where $\varphi_i(\,\cdot\,) = \varphi(i, \,\cdot\,)$, so that its L^p-norm is given by

$$\|\varphi\|_p^p = \sum_i p_i \int_X |\varphi_i|^p \, d\mu$$

for any $p \in [1, \infty)$ and $\|\varphi\|_\infty = \max_i \|\varphi_i\|_\infty$.

Define the following operators $P, U \colon L^1(Y, \nu) \to L^1(Y, \nu)$:

$$(P\varphi)_i = \sum_j p_{ij} T_{\gamma(i)}^{-1} \varphi_j, \quad (U\varphi)_j = T_{\omega(j)}^{-1} \varphi_{\iota(j)}. \tag{11}$$

Then P and U are measure-preserving Markov operators, meaning that both are unit norm contractions on every $L^p(Y, \nu)$, $p \in [1, \infty]$, and map the positive cone into itself.

Lemma 7.5. *For any function $f \in L^1(X, \mu)$ define a function $\varphi^{(f)} \in L^1(Y, \nu)$ by*

$$(\varphi^{(f)})_j = \begin{cases} \frac{1}{h_{\iota(j)}} f & j \in \Xi_S, \\ 0 & \textit{otherwise}. \end{cases}$$

Then

$$\widetilde{\mathbf{S}}_n(f) = \lambda^{n-1} \sum_{j \in \Xi_S} h_j (P^{n-1} U \varphi^{(f)})_j. \tag{12}$$

Proof. Indeed,

$$\widetilde{\mathbf{S}}_n(f) = \sum_{\substack{i_0 \in \Xi_S, i_{n-1} \in \Xi_F, \\ i_1, \ldots, i_{n-2} \in \Xi}} \Pi_{i_0 i_1} \ldots \Pi_{i_{n-2} i_{n-1}} T^{-1}_{\omega(i_{n-1})\gamma(i_{n-2})\ldots\gamma(i_0)} f$$

$$= \lambda^{n-1} \sum_{\substack{i_0 \in \Xi_S, i_{n-1} \in \Xi_F, \\ i_1, \ldots, i_{n-2} \in \Xi}} h_{i_0} p_{i_0 i_1} \ldots p_{i_{n-2} i_{n-1}} \frac{1}{h_{i_{n-1}}} T^{-1}_{\gamma(i_0)} \ldots T^{-1}_{\gamma(i_{n-2})} T^{-1}_{\omega(i_{n-1})} f$$

$$= \lambda^{n-1} \sum_{i_0 \in \Xi_S} h_{i_0}$$

$$\times \left(\sum_{i_1} p_{i_0 i_1} T^{-1}_{\gamma(i_0)} \left(\cdots \left(\sum_{i_{n-1}} p_{i_{n-2} i_{n-1}} T^{-1}_{\gamma(i_{n-2})} \underbrace{T^{-1}_{\omega(i_{n-1})} \left(\frac{\chi_{\Xi_F}(i_{n-1})}{h_{i_{n-1}}} f \right)}_{(U\varphi^{(f)})_{i_{n-1}}} \right) \cdots \right) \right)$$

$$= \lambda^{n-1} \sum_{i_0} h_{i_0} (P^{n-1} U \varphi^{(f)})_{i_0}.$$

■

7.4. The dual (adjoint) operator

Let us recall that for $\varphi, \psi \in L^2(Y, \nu)$ we have

$$\langle \varphi, \psi \rangle = \sum_{k \in \Xi} p_k \langle \varphi_k, \psi_k \rangle.$$

A short computation shows that if an operator Q has the form

$$(Q\varphi)_i = \sum_{j \in \Xi} p_{ij} T_{ij} \varphi_j,$$

then its dual satisfies

$$(Q^* \psi)_j = \sum_{k \in \Xi} \frac{p_k p_{kj}}{p_j} T^*_{kj} \psi_k.$$

Therefore, for P defined by (11) we have

$$(P^* \psi)_j = \sum_{k \in \Xi} \frac{p_k p_{kj}}{p_j} T_{\gamma(k)} \psi_k. \tag{13}$$

Lemma 7.6. *The Markov operators P and U defined by (11) satisfy the following identities:*

$$U = U^{-1} = U^*, \quad P^* = UPU.$$

Proof. These identities follow from Lemma 7.1 and Proposition 7.4. For example, let us prove the second:

$$\begin{aligned}(UPU\psi)_j &= T_{\omega(j)}^{-1}(PU\psi)_{\iota(j)} = T_{\omega(j)}^{-1}\sum_l p_{\iota(j),l}T_{\gamma(\iota(j))}^{-1}(U\psi)_l \\ &= \sum_l p_{\iota(j),l}T_{\omega(j)}^{-1}T_{\gamma(\iota(j))}^{-1}T_{\omega(l)}^{-1}\psi_{\iota(l)} \\ &= \sum_k p_{\iota(j),\iota(k)}T_{\omega(j)^{-1}\gamma(\iota(j))^{-1}\omega(k)}\psi_k.\end{aligned}$$

For the last equality we substitute $l = \iota(k)$ and use the first identity in Lemma 7.1. Now using Proposition 7.4, the second identity in Lemma 7.1 and formula (13) one can see that the right-hand side equals $(P^*\psi)_j$. ■

8. Proof of the main theorem

To prove our main result, Theorem A, we use a new theorem on pointwise convergence for powers of a Markov operator, Theorem 8.6, which is stated in Section 8.1; its proof is postponed to Section 9. The result is an elaboration of that used in [14] under weaker assumptions. In order to apply Theorem 8.6 to the operators defined by (11) some work is needed to check that these assumptions hold.

In Section 8.2 we check Assumptions 8.2, 8.3 in the ergodic case, that is, when the sigma-algebra $\mathcal{I}_{G_0^2}$ is trivial. The remaining Assumption 8.4 is checked in Section 8.3, concluding the proof of Theorem A in the ergodic case. Finally, in Section 8.4 we deal with the general, non-ergodic, case.

8.1. General theorem on pointwise convergence

Let (Z,η) be a Lebesgue probability space, and let Q be a measure-preserving Markov operator on $L^1(Z,\eta)$. In order to state our convergence result, we need the following assumptions.

Assumption 8.1. *There exists a decomposition $Q = VW$, where V and W are measure-preserving Markov operators, so that $Q^* = WV$.*

Assumption 8.2. *For every $n \in \mathbb{N}$ the equation $Q^n\psi = \psi$ has only constant solutions in $L^2(Z,\eta)$.*

Assumption 8.3. *There exists $m \in \mathbb{N}$ such that the equation $(Q^*)^m Q^m\psi = \psi$ has only constant solutions in $L^2(Z,\eta)$.*

Assumption 8.4. *There exists a sequence of operators A_n, a constant $C > 0$, and $a, b \in \mathbb{N}$ so that for all $n \geq m_0 := \lceil a/2 \rceil$ the following inequality holds for any non-negative $\varphi \in L^1(Z, \eta)$:*

$$W Q^{2n-a} \varphi \leq C \sum_{j=-b}^{b} (Q^*)^n Q^{n+j} \varphi + A_n \varphi. \tag{14}$$

Here W is the operator from Assumption 8.1. The operators $A_n : L^1(Z,\eta) \to L^1(Z,\eta)$ map non-negative functions into non-negative ones, and for any $p \in [1, \infty]$ map $L^p(Z,\eta)$ to itself. Moreover, $\|A_n\|_{L^p} \leq \alpha_n$, with $\sum_{n=m_0}^{\infty} \alpha_n < \infty$.

Remark 8.5. Applying $V' = QV$ to both sides of (14), we arrive at the inequality

$$Q^{2n-a'} \varphi \leq C V' \sum_{j=-b}^{b} (Q^*)^n Q^{n+j} \varphi + A'_n \varphi \tag{14'}$$

with the same estimates on the norms of the operators A'_n. We will use both (14) and (14′) below.

Theorem 8.6. *Let $Q : L^1(Z,\eta) \to L^1(Z,\eta)$ be a measure-preserving Markov operator acting on a Lebesgue probability space (Z,η) and satisfying Assumptions 8.1–8.4. Then for every function $\varphi \in L \log L(Z,\eta)$ the sequence $Q^n \varphi$ converges almost surely and in L^1 to $\int_Z \varphi \, d\eta$ as $n \to \infty$.*

As remarked above, the proof of this theorem is deferred to Section 9.

We now proceed to check that the above assumptions hold in our case.

8.2. Checking Assumptions 8.2 and 8.3

Let P, U be the Markov operators defined in (11) and define

$$Q = P^2, \quad V = PU, \quad W = UP. \tag{15}$$

In this subsection we check Assumptions 8.2 and 8.3 of Theorem 8.6 for these Q, U, V in the case in which the sigma-algebra $\mathcal{I}_{G_0^2}$ is trivial. To do this we express the equations from these assumptions in terms of the components φ_j, $j \in \Xi$, of a function $\varphi \in L^2(Y, \nu)$.

Proposition 8.7. *Let P be the Markov operator defined by (11). Then the following hold.*

(1) *A function $\varphi \in L^2(Y,\nu)$ is a solution to the equation $P^k \varphi = \varphi$ if and only if for any admissible sequence $i_0 \to i_1 \to \cdots \to i_k$ of states we have*

$$\varphi_{i_0} = T^{-1}_{\gamma(i_0)} \ldots T^{-1}_{\gamma(i_{k-1})} \varphi_{i_k}. \tag{16}$$

(2) *For $k \ge N_0$, where N_0 is defined in Corollary 6.6, a function $\varphi \in L^2(Y,\nu)$ is a solution to $(P^*)^k P^k \varphi = \varphi$ if and only if, for any admissible sequences $i_0 \to i_1 \to \dots \to i_k$ and $j_0 \to j_1 \to \dots \to j_k$ with $i_0 = j_0$, we have*

$$T^{-1}_{\gamma(i_1)} \dots T^{-1}_{\gamma(i_{k-1})} \varphi_{i_k} = T^{-1}_{\gamma(j_1)} \dots T^{-1}_{\gamma(j_{k-1})} \varphi_{j_k}.$$

Proof. We consider each assertion separately.

(1) The equation $P^k \varphi = \varphi$ is equivalent to

$$\varphi_{i_0} = (P^k \varphi)_{i_0} = \sum_{i_1,\dots,i_k} p_{i_0 i_1} \dots p_{i_{k-1} i_k} T^{-1}_{\gamma(i_0)} \dots T^{-1}_{\gamma(i_{k-1})} \varphi_{i_k},$$

hence, since the T_g's are unitary,

$$\|\varphi_{i_0}\|_{L^2} \le \sum_{i_1,\dots,i_k} p_{i_0 i_1} \dots p_{i_{k-1} i_k} \|\varphi_{i_k}\|_{L^2}. \tag{17}$$

Multiplying these inequalities by p_{i_0} and summing them up for all $i_0 \in \Xi$, we obtain

$$\sum_{i_0} p_{i_0} \|\varphi_{i_0}\|_{L^2} \le \sum_{i_k} \Bigg[\sum_{i_0,\dots,i_{k-1}} p_{i_0} p_{i_0 i_1} \dots p_{i_{k-1} i_k} \Bigg] \|\varphi_{i_k}\|_{L^2} = \sum_{i_k} p_{i_k} \|\varphi_{i_k}\|_{L^2}.$$

Therefore, for each i_0 inequality (17) is indeed an equality, and the vector $(\|\varphi_i\|_{L^2})_{i \in \Xi}$ is a right eigenvector of the stochastic matrix π^k, where the matrix $\pi = (p_{ij})$ is defined by (8). Corollary 6.6 and the Perron–Frobenius theorem yield that $(1,\dots,1)$ is the only eigenvector of π with non-negative coordinates up to scaling, hence all φ_i have the same L^2-norm.

Finally, in the Hilbert space $L^2(X,\mu)$ the triangle inequality (17) attains equality only if all non-zero summands are proportional to each other with positive coefficients, whence $\varphi_{i_0} = c \cdot T^{-1}_{\gamma(i_0)} \dots T^{-1}_{\gamma(i_{n-1})} \varphi_{i_n}$. Calculating the L^2-norms of both sides, we get $c = 1$.

(2) Similarly, $(P^*)^k P^k \varphi = \varphi$ yields

$$\varphi_{j_k} = \sum_{\substack{j_0,\dots,j_{k-1} \\ i_1,\dots,i_k}} \Bigg[\frac{p_{j_0} p_{j_0 j_1} \dots p_{j_{k-1} j_k}}{p_{j_k}} p_{j_0 i_1} p_{i_1 i_2} \dots p_{i_{k-1} i_k} \\ \times T_{\gamma(j_{k-1})} \dots T_{\gamma(j_1)} T^{-1}_{\gamma(i_1)} \dots T^{-1}_{\gamma(i_{k-1})} \varphi_{i_k} \Bigg].$$

The remaining proof is the same as in the first statement: $(\|\varphi_i\|)_{i \in \Xi}$ is a right eigenvector of the stochastic matrix $(\pi^*)^k \pi^k$ with positive entries, where we have $(\pi^*)_{ij} = p_j p_{ji} / p_i$; hence the L^2-norms of all φ_i's are equal, and the same argument with the triangle inequality completes the proof. ∎

Lemma 8.8. *Let M and N_0 be defined as in Proposition 6.2 and Corollary 6.6. Then for any $l \geq l^* := \max(2M, N_0)$, the following holds: if a function $\varphi \in L^2(Y, \nu)$ satisfies equalities*

$$T^{-1}_{\gamma(i_1)} \dots T^{-1}_{\gamma(i_{l-1})} \varphi_{i_l} = T^{-1}_{\gamma(j_1)} \dots T^{-1}_{\gamma(j_{l-1})} \varphi_{j_l} \tag{18}$$

for all admissible sequences $i_0 \to i_1 \to \dots \to i_l$, $j_0 \to j_1 \to \dots \to j_l$ with $i_0 = j_0$, then $\varphi(x,k)$ does not depend on $k \in \Xi$: $\varphi(x,k) = \varphi^\circ(x)$, and $\varphi^\circ(x)$ is G_0^2-invariant.

Remark 8.9. If (18) holds for all pairs of sequences of a given length l, then it holds for any pair of sequences

$$i_0 \to i_1 \to \dots \to i_{l'}, \quad j_0 \to j_1 \to \dots \to j_{l'}$$

of length $l' \leq l$ with $i_0 = j_0$. Indeed, append an arbitrary prefix $i_{-(l-l')} \to \dots \to i_0$ to these sequences and apply (18) to the resulting sequences of length l. One can see that

$$T^{-1}_{\gamma(i_{-(l-l')+1})} \dots T^{-1}_{\gamma(i_0)}$$

cancels out and we arrive at (18) for the original sequences of length l'.

Let us first deduce Assumptions 8.2 and 8.3 from Lemma 8.8.

Corollary 8.10. *In the ergodic case, that is, assuming $\mathcal{I}_{G_0^2}$ is trivial, Assumptions 8.2 and 8.3 hold for the operator Q defined by (11) and (15).*

Proof. We divide the proof into two steps.

(1) Suppose that $Q^n \varphi = \varphi$. Choose s such that $l = 2ns \geq l^*$. Then

$$P^l \varphi = (Q^n)^s \varphi = \varphi.$$

Therefore, (18) holds, as both sides are equal to $T_{\gamma(i_0)} \varphi_{i_0}$ by (16). Lemma 8.8 then implies that all φ_j are equal to the same function φ°, where φ° is G_0^2-invariant and hence, by ergodicity, constant.

(2) Suppose that $(Q^*)^m Q^m \varphi = \varphi$, where m satisfies $2m \geq l^*$. Proposition 8.7 implies that (18) holds for φ with $l = 2m$, so φ is constant. ■

It remains to prove Lemma 8.8.

Proof of Lemma 8.8. For every $e \in G_0$, let $\underline{\mathcal{H}}^e$, $\underline{\mathcal{T}}^e$, $\underline{i}^e$, $\underline{j}^e$, $\mathbf{h}(e)$, and $\mathbf{t}(e)$ be defined as in Proposition 6.2. Recall that $\mathcal{T}_0^e = \mathcal{H}_0^e = \mathcal{R}$ and define $g_e, h_e \in G$ such that $\mathcal{T}^e_{M-1} = g_e \mathcal{R}$, $\mathcal{H}^e_{-M+1} = h_e \mathcal{R}$, whence

$$g_e = \gamma(i_0^e)^{-1} \dots \gamma(i_{M-2}^e)^{-1}, \quad h_e = \gamma(j_{-1}^e) \dots \gamma(j_{-M+1}^e).$$

Denote $\psi_e = T_{g_e} \varphi_{\mathbf{t}(e)}$.

Take any e_1, e_2 and choose $\widehat{e} \neq e_1^{-1}, e_2^{-1}$. Let $\underline{\mathcal{S}}^\alpha$ ($\alpha = 1, 2$) be the paths from the first part of Proposition 6.3 applied to e_α and $\widehat{e}$, and let $\underline{k}^\alpha = (k^\alpha_{-M} \to \cdots \to k^\alpha_{M-1})$ be the corresponding sequences of states. Then

$$\mathcal{S}^\alpha_{-M+1} = \mathcal{H}^{\widehat{e}}_{-M+1} = h_{\widehat{e}}\mathcal{R}, \quad \mathcal{S}^\alpha_{M-1} = \mathcal{T}^{e_\alpha}_{M-1} = g_{e_\alpha}\mathcal{R},$$

hence

$$g_{e_\alpha} = h_{\widehat{e}}\gamma(k^\alpha_{-M+1})^{-1} \ldots \gamma(k^\alpha_{M-2})^{-1}.$$

Therefore, (18) for the sequences $\underline{k}^\alpha$ yields

$$T_{h_{\widehat{e}}^{-1}} T_{g_{e_1}} \varphi_{\mathbf{t}(e_1)} = T_{h_{\widehat{e}}^{-1}} T_{g_{e_2}} \varphi_{\mathbf{t}(e_2)},$$

whence $\psi_{e_1} = \psi_{e_2}$. We thus obtain that all ψ_e are equal to the same function ψ°.

Now again take any e_1, e_2, choose $\widehat{e} \neq e_1^{-1}, e_2^{-1}$, and apply the same argument to the paths from the second part of Proposition 6.3 for $\widehat{e}$, e_α and $\widetilde{e}_\alpha$. We have that

$$T_{h_{\widehat{e}}^{-1} e_1} \psi_{\widetilde{e}_1} = T_{h_{\widehat{e}}^{-1} e_1 g_{\widetilde{e}_1}} \varphi_{\mathbf{t}(\widetilde{e}_1)} = T_{h_{\widehat{e}}^{-1} e_2 g_{\widetilde{e}_2}} \varphi_{\mathbf{t}(\widetilde{e}_2)} = T_{h_{\widehat{e}}^{-1} e_2} \psi_{\widetilde{e}_2},$$

Therefore,

$$T_{e_1^{-1} e_2} \psi^\circ = \psi^\circ,$$

so ψ° is G_0^2-invariant. Then the function $T_h \psi^\circ$ is independent of $h \in G_0$; denote it by $\psi^\bullet$. The function $\psi^\bullet$ is also G_0^2-invariant and $T_h \psi^\bullet = \psi^\circ$ for all $h \in G_0$.

Finally, fix $\widetilde{\imath} = \mathbf{t}(e)$ and take any $\widetilde{\jmath} \in \Xi$. With $N = N_0$ as in Corollary 6.6, we can consecutively choose $i_{N-1}, \ldots, i_1, i_0$ so that the sequence

$$i_0 \to \cdots \to i_{N-1} \to i_N = \widetilde{\imath}$$

is admissible, and then, since $(\Pi^{N_0})_{i_0 \widetilde{\jmath}} > 0$, we can choose $j_1, \ldots, j_{N-1}$ such that

$$i_0 = j_0 \to j_1 \to \cdots \to j_{N-1} \to j_N = \widetilde{\jmath}$$

is also admissible. Using (18) for these sequences and taking into account the definition of ψ_e, we see that

$$\varphi_{\widetilde{\jmath}} = T_{\gamma(j_{N-1}) \ldots \gamma(j_1)\gamma(i_1)^{-1} \ldots \gamma(i_{N-1})^{-1} \gamma(i^e_{M-2}) \ldots \gamma(i^e_0)} \psi^\circ.$$

The number $L = 2N + M - 3$ of the generators in the product is the same for all $\widetilde{\jmath}$, hence all the $\varphi_{\widetilde{\jmath}}$ are the same; they are equal to either ψ° or $\psi^\bullet$ depending on the parity of L. ■

8.3. Checking Assumption 8.4

Lemma 8.11. *Assumption 8.4 holds for the operators defined by (11) and (15). Precisely, inequality (14) holds for $a = 6$ and $b = 2$.*

The proof rests on a rather complicated geometric statement which is needed to compare terms on the two sides of estimate (14) in Assumption 8.4. The underlying meaning of this geometric statement, made precise in Lemma 8.12, is the following. Consider a thickened path $\underline{\mathcal{S}} = (\mathcal{S}_0 = \mathcal{A}, \dots, \mathcal{S}_{2n} = \mathcal{B})$. Then if this thickened path does not belong to a small set of "exceptional" paths, it can be embedded into a triangle $\mathcal{ABC}$ of thickened paths such that (1) the lengths of $\mathcal{AC}$ and $\mathcal{CB}$ are not more than $n + \mathrm{const}$, and (2) the triangle has, informally speaking, "zero angles" in all its vertices. In the simplest case of free group "zero angle" in vertex $\mathcal{A}$ means the coincidence of the levels of the paths $\mathcal{AB}$ and $\mathcal{AC}$ that are adjacent to $\mathcal{A}$. One can easily see that here we have no exceptional paths: choose $\mathcal{C}$ to be any neighbor of $\mathcal{S}_n$ other than $\mathcal{S}_{n-1}$ and $\mathcal{S}_{n+1}$, then $\mathcal{CA} = (\mathcal{C}, \mathcal{S}_n, \mathcal{S}_{n-1}, \dots, \mathcal{S}_0)$ and $\mathcal{CB} = (\mathcal{C}, \mathcal{S}_n, \mathcal{S}_{n+1}, \dots, \mathcal{S}_{2n})$ have length $n + 1$.

In the general case the states of our Markov chain are not uniquely recovered from the configurations of the domains, so we say that "zero angle" in $\mathcal{A}$ means the coincidence of the first elements in the sequences of states generating $\mathcal{AB}$ and $\mathcal{AC}$. Another amendment in the formal statement of the lemma is that we deal with any sequences generates by our Markov chain, not only with thickened paths.

To find such a triangle we cut the sequence $\mathcal{AB}$ near its midpoint and then modify the halves to construct $\mathcal{AC}$ and $\mathcal{CB}$. This is done by the lemmas from Section 5, which allow us to keep track of the generating sequences of states throughout the modification process.

Denote by $\mathbf{P}_r$ the set of all admissible sequences $\underline{i} = (i_0 \to \dots \to i_r)$ of states in the Markov chain and let $\lambda > 1$ be the Perron–Frobenius eigenvalue of its adjacency matrix Π.

Lemma 8.12. *For all sufficiently large N, there exists an exceptional subset*

$$E_{2N-1} \subset \mathbf{P}_{2N-1}$$

with $\#E_{2N-1} = O(\lambda^N)$, such that the following holds for every $\underline{i} \in \mathbf{P}_{2N-1} \setminus E_{2N-1}$: there exist $\alpha \in \{1, 2, 3, 4\}$, $\beta \in \{-1, 0, 1, 2\}$, and admissible sequences

$$\underline{j} = (j_l)_{l=0}^{N-\beta+\alpha-1}, \quad \underline{k} = (k_l)_{l=0}^{N+\beta+\alpha-1}$$

with the following properties:

(i) $j_0 = k_0$, $j_{N-\beta+\alpha-1} = \iota(i_0)$, $k_{N+\beta+\alpha-1} = i_{2N-1}$.

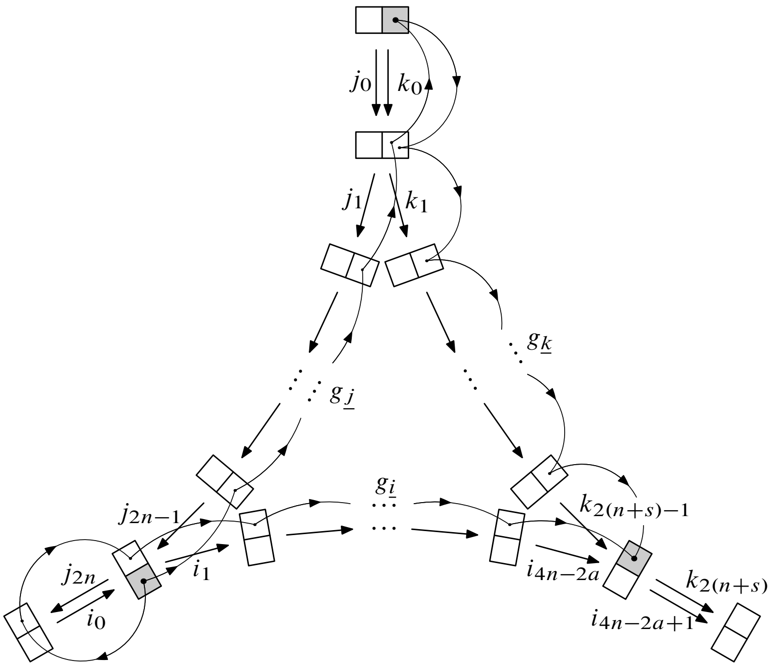


Figure 15. Illustrating Lemma 8.12 and formula (22).

(ii) *Let* $\underline{\mathcal{S}} = (\mathcal{S}_0, \dots, \mathcal{S}_{2N})$ *be any sequence of domains generated by* $\underline{i}$ *and let*

$$\underline{\mathcal{U}} = (\mathcal{U}_0, \dots, \mathcal{U}_{N-\beta+\alpha}), \quad \underline{\mathcal{V}} = (\mathcal{V}_0, \dots, \mathcal{V}_{N+\beta+\alpha})$$

be the unique sequences generated by $\underline{j}$ *and* $\underline{k}$*, respectively, with the property that* $\mathcal{U}_{N-\beta+\alpha} = \mathcal{S}_0$*,* $\mathcal{V}_{N+\beta+\alpha} = \mathcal{S}_{2N}$*. Then* $\mathcal{U}_0 = \mathcal{V}_0$*.*

(iii) $j_1 \neq k_1$.

Moreover, the mapping $\underline{i} \mapsto (\underline{j}, \underline{k})$ *with* $(\underline{j}, \underline{k})$ *satisfying* (i)–(iii) *is injective.*

This lemma is illustrated in Figure 15. Every state from the sequences $\underline{i}$, $\underline{j}$, $\underline{k}$ is represented by a straight arrow, while the past and the future domains of the state are shown as the pairs of squares near the start and the end of this arrow. Other details of this figure, including the numbering for elements of $\underline{j}$ and $\underline{k}$, which is different from that in the statement of the lemma, are discussed below when proving equality (22).

Remark 8.13. In the lemma, β can be set to zero except in the special case of Remark 4.5.

Statements (i) and (ii) remain true if the same admissible sequence $\underline{t} = (t_l)_{l=-s}^{0}$ with $t_0 = j_0 = k_0$ is prepended to both $\underline{j}$ and $\underline{k}$. Statement (iii) requires $\underline{j}$ and $\underline{k}$ to be chosen without this common initial part. The possibility of adding or removing such a common initial segment will be used later.

Proof of Lemma 8.11 assuming Lemma 8.12. The values of a and b in the statement of Lemma 8.11 are in fact $b = \max |\beta|$, $a = b + \max \alpha$, where the possible values of α and β are described in the course of the proof of Lemma 8.12; see Claims 8.14 and 8.17.

Let us examine the terms on each side of (14). First, for a non-negative function $\varphi \in L^1(Y, \nu)$, we have

$$\begin{aligned}(WQ^{2n-a}\varphi)_l &= (UP^{4n-2a+1}\varphi)_l \\ &= \sum_{i_1,\dots,i_{4n-2a+1}} p_{\iota(l),i_1} p_{i_1 i_2} \cdots p_{i_{4n-2a} i_{4n-2a+1}} \\ &\qquad \times T^{-1}_{\omega(l)} T^{-1}_{\gamma(\iota(l))} T^{-1}_{\gamma(i_1)} \cdots T^{-1}_{\gamma(i_{4n-2a})} \varphi_{i_{4n-2a+1}}. \end{aligned} \tag{19}$$

The coefficient in a term of the last sum is non-zero if and only if the sequence $\iota(l) \to i_1 \to \cdots \to i_{4n-2a+1}$ is admissible. Since (p_{ij}) is the matrix for the Parry measure, formula (9) yields

$$(WQ^{2n-a}\varphi)_l \le \widetilde{C}_1 \lambda^{-4n} \sum_{\substack{\underline{i} \in \mathbf{P}_{4n-2a+1}, \\ i_0 = \iota(l)}} T_{\omega(i_0)} T^{-1}_{\gamma(i_0)} T^{-1}_{\gamma(i_1)} \cdots T^{-1}_{\gamma(i_{4n-2a})} \varphi_{i_{4n-2a+1}}, \tag{20}$$

where we use $\omega(\iota(l)) = \omega(l)^{-1}$; see Lemma 7.1. Similarly,

$$\begin{aligned} &((Q^*)^n Q^{n+s}\varphi)_l \\ &\quad = \sum_{\substack{j_{2n-1},\dots,j_0, \\ k_1,\dots,k_{2n+2s}}} \frac{p_{j_0}}{p_l} p_{j_{2n-1} l} p_{j_{2n-2} j_{2n-1}} \cdots p_{j_0 j_1} p_{j_0 k_1} p_{k_1 k_2} \cdots p_{k_{2n+2s-1} k_{2n+2s}} \\ &\qquad\quad \times T_{\gamma(j_{2n-1})} \cdots T_{\gamma(j_0)} T^{-1}_{\gamma(j_0)} T^{-1}_{\gamma(k_1)} \cdots T^{-1}_{\gamma(k_{2n+2s-1})} \varphi_{k_{2n+2s}}, \end{aligned}$$

hence (9) yields the estimate

$$\begin{aligned} &((Q^*)^n Q^{n+s}\varphi)_l \ge \widetilde{C}_2 \lambda^{-4n} \\ &\quad \times \sum_{\substack{\underline{j} \in \mathbf{P}_{2n}, \underline{k} \in \mathbf{P}_{2n+2s}, \\ j_0 = k_0, j_{2n} = l}} T_{\gamma(j_{2n-1})} \cdots T_{\gamma(j_0)} T^{-1}_{\gamma(k_0)} T^{-1}_{\gamma(k_1)} \cdots T^{-1}_{\gamma(k_{2n+2s-1})} \varphi_{k_{2n+2s}}. \end{aligned} \tag{21}$$

Here $\widetilde{C}_2$ is chosen in such a way that this inequality holds for any s with $|s| \le b$. Apply Lemma 8.12 to a sequence $\underline{i}$ from (20) with $N = 2n - a + 1$. There are $O(\lambda^{2n})$ sequences in the exceptional set $E_{4n-2a+1}$, and the corresponding terms in (19) comprise $(A_n\varphi)_l$. Thus $\|A_n\| = O(\lambda^{-2n})$, hence the series $\sum_n \|A_n\|$ converges.

Suppose now that $\underline{i} \notin E_{4n-2a+1}$. Choose paths $\underline{\widetilde{j}} \in \mathbf{P}_{2n-a-\beta+\alpha}$ and $\underline{\widetilde{k}} \in \mathbf{P}_{2n-a+\beta+\alpha}$ as in Lemma 8.12. Denote $\eta = a - \alpha + \beta \geq 0$, consider any admissible sequence

$$\underline{t} = (t_{-\eta} \to \cdots \to t_0 = j_0 = k_0),$$

and adjoin $\underline{t}$ to both $\underline{\widetilde{j}}$ and $\underline{\widetilde{k}}$ to construct $\underline{j} \in \mathbf{P}_{2n}$ and $\underline{k} \in \mathbf{P}_{2n+2\beta}$.

Let us prove that the terms in (20) and (21) (with $s = \beta$) corresponding to this choice of $\underline{i}$, $\underline{j}$, and $\underline{k}$ are equal. Indeed, $l = \iota(i_0) = j_{2n}$ and $k_{2n+2s} = i_{4n-2a+1}$, so it remains to prove that

$$\underbrace{\omega(i_0)\gamma(i_0)^{-1}\gamma(i_1)^{-1}\ldots\gamma(i_{4n-2a}^{-1})}_{g_{\underline{i}}} = \underbrace{\gamma(j_{2n-1})\ldots\gamma(j_0)}_{g_{\underline{j}}}\underbrace{\gamma(k_0)^{-1}\ldots\gamma(k_{2n+2s-1})^{-1}}_{g_{\underline{k}}}, \tag{22}$$

where we define $g_{\underline{i}}$, $g_{\underline{j}}$, and $g_{\underline{k}}$ as shown.

Statements (i) and (ii) of Lemma 8.12 hold for $\underline{j}$ and $\underline{k}$, so let $\underline{\mathcal{S}}$, $\underline{\mathcal{U}}$, and $\underline{\mathcal{V}}$ be generated by $\underline{i}$, $\underline{j}$, $\underline{k}$ as in statement (ii). Let $h\mathcal{R}$ be the right domain in $\mathcal{S}_1$, then the definitions of $\gamma(\cdot)$ and $\omega(\cdot)$ in formula (7) imply that $h\omega(i_0)\mathcal{R}$ is the left domain in $\mathcal{S}_0$, $h\omega(i_0)\gamma(i_0)^{-1}\mathcal{R}$ is the left domain in $\mathcal{S}_1$, …, and $hg_{\underline{i}}\mathcal{R}$ is the left domain in $\mathcal{S}_{4n-2a+1}$.

On the other hand, as $\mathcal{S}_0 = \mathcal{U}_{2n+1}$ and $i_0 = \iota(j_{2n})$, we have that $\mathcal{S}_1 = \mathcal{U}_{2n}$ and $h\mathcal{R}$ is the left domain in $\mathcal{U}_{2n}$. The same argument as above gives us that $hg_{\underline{j}}\mathcal{R}$ is the left domain in $\mathcal{U}_0$, which coincides with the left domain in $\mathcal{V}_0$, so $hg_{\underline{j}}g_{\underline{k}}\mathcal{R}$ is the left domain in $\mathcal{V}_{2n+2s}$. But as

$$\mathcal{V}_{2n+2s+1} = \mathcal{S}_{4n-2a+2} \quad \text{and} \quad k_{2n+2s} = i_{4n-2a+1},$$

we obtain that the left domains in $\mathcal{V}_{2n+2s}$ and $\mathcal{S}_{4n-2a+1}$ coincide. Therefore,

$$hg_{\underline{i}}\mathcal{R} = hg_{\underline{j}}g_{\underline{k}}\mathcal{R},$$

so (22) holds. This is illustrated in Figure 15: the curved arrows link the domains $hg\mathcal{R}$, where g is an initial segment of either the left-hand or the right-hand side of (22); the shaded squares correspond to $g = \mathrm{id}$ (left), $g = g_{\underline{j}}$ (top), $g = g_{\underline{i}} = g_{\underline{j}}g_{\underline{k}}$ (right).

Therefore, we have proved that for every term in the right-hand side of (20) except those with $\underline{i} \in E_{4n-2a+1}$, there exists an equal term in the right-hand side of (21) for some s with $|s| \leq b$.

Finally, we check that different terms in (20) correspond to different terms in (21). Indeed, by the last statement of Lemma 8.12 different sequences $\underline{i} \in \mathbf{P}_{4n-2a+1} \setminus E_{4n-2a+1}$ correspond to different pairs $(\underline{\widetilde{j}}, \underline{\widetilde{k}})$. On the other hand, by item (iii) of this lemma the pair $(\underline{\widetilde{j}}, \underline{\widetilde{k}})$ is uniquely determined by the pair $(\underline{j}, \underline{k})$: to see this find maximal $s \geq 0$ such that $j_l = k_l$ for all $l = 0, \ldots, s$ and remove the common initial segment $(j_0, \ldots, j_{s-1})$ from $\underline{j}$ and $\underline{k}$. Therefore, different sequences $\underline{i}$ yield different pairs $(\underline{j}, \underline{k})$, hence

$$\sum_{\substack{\underline{i} \in \mathbf{P}_{4n-2a+1} \setminus E_{4n-2a+1}, \\ i_0 = \iota(l)}} T_{g_{\underline{i}}} \varphi_{i_{4n-2a+1}} \leq \sum_{s=-b}^{b} \sum_{\substack{\underline{j} \in \mathbf{P}_{2n}, \underline{k} \in \mathbf{P}_{2n+2s}, \\ j_0 = k_0, \, j_{2n} = l}} T_{g_{\underline{j}} g_{\underline{k}}} \varphi_{k_{2n+2s}}.$$

Combining this inequality with (20) and (21), we establish (14). ■

Proof of Lemma 8.12. The proof will be carried out in a number of steps. At various points we perform certain operations from Section 5 on the sequences in question and then check that the number of sequences for which this alters states so as to make the requirement (i) in the statement of the lemma impossible is $O(\lambda^N)$.

As in Lemma 5.1, let n_0 be the maximal value of $n(v)$ for all vertices of $\mathcal{R}$, $n_0 = 2$ if $\mathcal{R}$ has no vertices inside $\mathbb{D}$. From now on we assume that $N > n_0 + 5$; otherwise one can set $E_{2N-1} := \mathbf{P}_{2N-1}$. Take any $\underline{i} \in \mathbf{P}_{2N-1}$ and consider a sequence $\underline{\mathcal{S}} = (\mathcal{S}_0, \ldots, \mathcal{S}_{2N})$ that is generated by $\underline{i}$; let $s_l = \mathcal{S}_l \cap \mathcal{S}_{l+1}$.

Step 1. We begin by splitting $\underline{\mathcal{S}}$ at a suitable point $\mathcal{S}_{N-\beta}$ near $\mathcal{S}_N$, where β is chosen as in the next claim. Note that if $N(\mathcal{R}) \geq 4$ we can take $\beta = 0$ and the proof simplifies.

Claim 8.14. *There exists $\beta \in \{-1, 0, 1, 2\}$ and a domain $\mathcal{A} \in \mathcal{S}_{N-\beta}$ with a side $\tilde{s}$ not belonging to $s_{N-\beta-1} \cup s_{N-\beta}$. In the special case from Remark 4.5 we also require that $\tilde{s}$ has an end v that either belongs to $\partial\mathbb{D}$ or is incident to at most $n(v) - 1$ domains in $\underline{\mathcal{S}}$.*

Proof. Assume first that there exists a level $\mathcal{S}_{N-\beta}$, $\beta \in \{-1, 0, 1, 2\}$ which contains two fundamental domains. Then $\mathcal{S}_{N-\beta}$ has $2N(\mathcal{R})$ sides with at most six of them included in $s_{N-\beta-1} \cup s_{N-\beta}$. Any other side in $\partial\mathcal{S}_{N-\beta}$ can be chosen as $\tilde{s}$ with $\mathcal{A}$ being the domain in $\mathcal{S}_{N-\beta}$ adjacent to $\tilde{s}$.

Thus further consideration is needed only when $N(\mathcal{R}) = 3$ and both states $i_{N-\beta-1}$ and $i_{N-\beta}$ are of type E. The states of type E exists only if there are two adjacent vertices in $\partial\mathcal{R}$ that lie inside $\mathbb{D}$, hence we are in the special case from Remark 4.5. However, $(E \to E)$-transition needs a vertex u with $n(u) = 2$, and this is ruled out in this remark. We have thus proved that there exists a side

$$\tilde{s} \not\subset s_{N-\beta-1} \cup s_{N-\beta}.$$

Note that in the special case the side $\tilde{s}$ is not incident to the common vertex of two fundamental domains in $\mathcal{S}_{N-\beta}$, whence $\tilde{s}$ is non-compact.

Now assume that each $\mathcal{S}_{N-\beta}$, $\beta \in \{-1, 0, 1, 2\}$, contains only one domain. Then $s_{N-1} \cup s_N$ consists of two sides of $\mathcal{S}_N$. In the non-special case any other side in $\partial\mathcal{S}_N$ can be chosen as $\widetilde{s}$. In the special case this can fail if the side

$$s = \partial\mathcal{S}_N \setminus (s_{N-1} \cup s_N)$$

is compact. Then let u be the common vertex of s and s_{N-1}. If u is incident to $n(u) \geq 3$ levels of $\underline{\mathcal{S}}$, then u is incident to the only domain in $\mathcal{S}_{N-2}$. Therefore,

$$\widetilde{s} = \partial\mathcal{S}_{N-1} \setminus (s_{N-2} \cup s_{N-1})$$

is the only side of $\mathcal{S}_{N-1}$ non-adjacent to u, hence $\widetilde{s}$ is non-compact. ■

Step 2. Having split the sequence $\underline{\mathcal{S}}$ into two halves at $\mathcal{S}_{N-\beta}$, the next step is to narrow these halves, reducing $\mathcal{S}_{N-\beta}$ to the domain $\mathcal{A}$ chosen as above in Claim 8.14 (and fixed for the remainder of this proof). This we do by applying Lemma 5.2 to the sequences $(i_l)_{l=0}^{N-\beta-1}$ and $(\mathcal{S}_l)_{l=0}^{N-\beta}$ to obtain sequences

$$(j_l')_{l=0}^{N-\beta-1}, \quad (\mathcal{U}_l')_{l=0}^{N-\beta}$$

with $\mathcal{U}'_{N-\beta} = \mathcal{A}$. Similarly, from $(i_l)_{l=N-\beta}^{2N-1}$ and $(\mathcal{S}_l)_{l=N-\beta}^{2N}$, we obtain sequences

$$(k_l')_{l=N-\beta}^{2N-1}, \quad (\mathcal{V}_l')_{l=N-\beta}^{2N}$$

with $\mathcal{V}'_{N-\beta} = \mathcal{A}$.

Claim 8.15. *Let $E_{2N-1}^{(1)}$ be the set of sequences $\underline{i} \in \mathbf{P}_{2N-1}$ such that*

$$j_0' \neq i_0 \quad or \quad k_{2N-1}' \neq i_{2N-1}.$$

Then $\# E_{2N-1}^{(1)} = O(\lambda^N)$.

Proof. As one can see from the second statement of Lemma 5.2, $j_0' \neq i_0$ implies that

$$\mathcal{S}_{n_0-1} \neq \mathcal{U}'_{n_0-1}.$$

Hence, all states in the sequence $(i_l)_{l=n_0-1}^{N-\beta-1}$ are of types C and E_α, where $\alpha = L$ if $\mathcal{A} = \mathcal{L}_{N-\beta}$ and $\alpha = R$ if $\mathcal{A} = \mathcal{R}_{N-\beta}$. This means that the states $(i_l)_{l=n_0-1}^{N-\beta-1}$ are uniquely determined by $i_{N-\beta-1}$. Therefore, there are finitely many possibilities for $(i_l)_{l=0}^{N-\beta-1}$ and $O(\lambda^N)$ possibilities for $(i_l)_{l=N-\beta}^{2N-1}$. ■

Step 3. Assume now that $\underline{i} \notin E_{2N-1}^{(1)}$. The next step is to shift the numbering in the sequences constructed above and invert the first pair, after which we join a short head sequence as defined in Proposition 6.2 to the beginning of each of them.

Let M be the number from Proposition 6.2, then we define $\underline{j}''_+ = (j''_l)_{l=M}^{M+N-\beta-1}$, $\underline{\mathcal{U}}''_+ = (\mathcal{U}''_l)_{l=M}^{M+N-\beta}$ as

$$\mathcal{U}''_l = \mathcal{U}'_{M+N-\beta-l}, \quad j''_l = \iota(j'_{M+N-\beta-1-l})$$

and $\underline{k}''_+ = (k''_l)_{l=M}^{M+N+\beta-1}$, $\underline{\mathcal{V}}''_+ = (\mathcal{V}''_l)_{l=M}^{M+N+\beta}$ as

$$\mathcal{V}''_l = \mathcal{V}'_{N-\beta-M+l}, \quad k''_l = k'_{N-\beta-M+l}.$$

Further, let $\mathcal{A} = a\mathcal{R}$ and let $\widetilde{e}$ be the label on the side $\widetilde{s}$ inside $\mathcal{A}$. Define

$$\underline{j}''_- = \underline{k}''_- = (j^{\widetilde{e}}_{l-M})_{l=0}^{M-1}, \quad \underline{\mathcal{U}}''_- = \underline{\mathcal{V}}''_- = (a\mathcal{H}^{\widetilde{e}}_{l-M})_{l=0}^{M},$$

where $\underline{j}^{\widetilde{e}}$ and $\underline{\mathcal{H}}^{\widetilde{e}}$ are defined in Proposition 6.2.

Apply Lemma 5.4 to join $\underline{\mathcal{U}}''_-$ and $\underline{\mathcal{U}}''_+$. Except in the special case of Remark 4.5 this is possible because $j^{\widetilde{e}}_{-1}$ is type A_0 so that the path $\underline{\mathcal{U}}''_-$ adds only one domain incident to each end u of $\widetilde{s}$, so for the union of these paths the vertex u is either convex or minimally concave. In the special case of Remark 4.5 and compact side $\widetilde{s}$ this is amended as follows: $\underline{\mathcal{U}}''_-$ adds two domains to one of the ends of $\widetilde{s}$; by choosing $\mathcal{H}^{\widetilde{e}}$ to be either the path from Figure 13 (j) or its mirror image we make this end to be the vertex v from Claim 8.14.

Thus by joining $\underline{\mathcal{U}}''_-$ to $\underline{\mathcal{U}}''_+$, we obtain new sequences and states which we rename as

$$\underline{\mathcal{U}} = (\mathcal{U}_l)_{l=0}^{N+M-\beta}, \quad \underline{j} = (j_l)_{l=0}^{N+M-\beta-1}.$$

We have $j_0 = j^{\widetilde{e}}_{-M}$ by the construction of the path $\underline{\mathcal{H}}^{\widetilde{e}}$. Similarly, we join $\underline{\mathcal{V}}''_-$ and $\underline{\mathcal{V}}''_+$ to obtain

$$\underline{\mathcal{V}} = (\mathcal{V}_l)_{l=0}^{N+M+\beta}, \quad \underline{k} = (k_l)_{l=0}^{N+M+\beta-1},$$

and we have $k_0 = j^{\widetilde{e}}_{-M} = j_0$, as well as $\mathcal{U}_0 = a\mathcal{H}^{\widetilde{e}}_{-M} = \mathcal{V}_0$.

Claim 8.16. *Let $E^{(2)}_{2N-1}$ be the set of sequences $\underline{i} \in \mathbf{P}_{2N-1} \setminus E^{(1)}_{2N-1}$ such that*

$$j_{N+M-\beta-1} \neq j''_{N+M-\beta-1} \quad \textit{or} \quad k_{N+M+\beta-1} \neq k''_{N+M+\beta-1}.$$

Then $\#E^{(2)}_{2N-1} = O(\lambda^N)$.

Proof. We first consider changes to states made at the joining step to see to what extent the joining changes the states j, see Lemma 5.4. As above, $j_{N+M-\beta-1} \neq j''_{N+M-\beta-1}$ implies

$$\mathcal{U}_{M+N-\beta-n_0} \neq \mathcal{U}''_{M+N-\beta-n_0}.$$

Assume that $\widetilde{s}$ lies on the left boundary of $\mathcal{S}_{N-\beta}$. Then each level $(\mathcal{U}''_l)_{l=M}^{M+N-\beta-n_0}$ contains only one domain and these domains are the consecutive domains adjacent to a geodesic segment on the right boundary of $\underline{\mathcal{U}}''_+$, that is, the left boundary of $\underline{\mathcal{U}}'$. Moreover, the states $(j''_l)_{l=M}^{M+N-\beta-n_0-1}$ are uniquely defined by j''_M, so there are finitely many $\underline{j}''_+$'s (or, equivalently, $\underline{j}'$'s) such that

$$j_{N+M-\beta-1} \neq j''_{N+M-\beta-1}.$$

Now we consider changes to states made at the narrowing step, for which we use Lemma 5.2. Let us show that each of these $\underline{j}'$'s can be obtained from finitely many $(i_l)_{l=0}^{N-\beta-1}$. Indeed, assume that $\mathcal{U}'_{N-\beta-3} \neq \mathcal{S}_{N-\beta-3}$, i.e. that the narrowing step changes at least the four last domains in $(\mathcal{S}_l)_{l=0}^{N-\beta}$. Then for $l = 1, 2, 3$, we have that

$$\mathcal{U}'_{N-\beta-l} = \mathcal{L}_{N-\beta-l}$$

and all $\partial_R \mathcal{U}'_{N-\beta-l}$ belong to the same geodesic segment; see Figure 10 (a).

We have shown that the same holds for

$$\partial_L \mathcal{U}'_{N-\beta-l} = \partial_R \mathcal{U}''_{M+l}, \quad l = 1, 2, 3$$

since the joining step adds domains to all levels up to $\mathcal{U}''_{M+N-\beta-n_0+1}$. Therefore, each of $\partial_{L,R} \mathcal{U}'_{N-\beta-l}$ is either a side or a vertex since the $l = 2$ region is joined to the $l = 1$ and $l = 3$ regions across one side only. This means that $N(\mathcal{R}) \leq 4$ and $\mathcal{R}$ is compact. Assumption 1.1 now yields that $N(\mathcal{R}) = 4$ and each of $\partial_{L,R} \mathcal{U}'_{N-\beta-l}$ is a segment. Thus $\partial \underline{\mathcal{U}}'$ has a straight angle at every vertex u of $\mathcal{U}'_{N-\beta-2}$. On the other hand, there are only two domains in $\underline{\mathcal{U}}'$ that are adjacent to u, hence $n(u) = 2$. This contradicts Assumption 1.1.

Therefore, $j_{N+M-\beta-1} \neq j''_{N+M-\beta-1}$ only for finitely many sequences $(i_l)_{l=0}^{N-\beta}$, and hence for $O(\lambda^N)$ sequences $\underline{i} = (i_l)_{l=0}^{2N-1}$. ■

Step 4. Assume now that $\underline{i} \notin E_{2N-1} := E^{(1)}_{2N-1} \cup E^{(2)}_{2N-1}$. Then we have

$$\begin{aligned} j_{N+M-\beta-1} &= j''_{N+M-\beta-1} = \iota(j'_0) = \iota(i_0), \\ \mathcal{U}_{N+M-\beta} &= \mathcal{U}''_{N+M-\beta} = \mathcal{U}'_0 = \mathcal{S}_0. \end{aligned}$$

Similarly, we have $k_{N+M+\beta-1} = i_{2N-1}$ and $\mathcal{V}_{N+M+\beta} = \mathcal{S}_{2N}$. Also we have seen above that $j_0 = k_0$ and $\mathcal{U}_0 = \mathcal{V}_0$. Therefore, statements (i) and (ii) of the lemma hold for the constructed sequences.

Step 5. The next claim allows us to prove both (iii) and the final statement of the lemma, that is, that the map $\underline{i} \to (\underline{j}, \underline{k})$ is injective. The claim itself will proved in Step 6 below.

Claim 8.17. *Assume that we have sequences $\mathcal{U}, \mathcal{V}$ as in* (i) *and* (ii) *with $\mathcal{U}_0 = \mathcal{V}_0$ and let M be as in Proposition 6.2 and Step 3 above. Then:*

(1) *Let s be the maximal number such that $j_l = k_l$ for $l = 0, \dots, s$. Then $s < M$.*

(2) *Let s' be the maximal number l such that $\mathcal{U}_l$ and $\mathcal{V}_l$ have a common domain. Then $s' = M$ and $\mathcal{U}_M \cap \mathcal{V}_M = \mathcal{A}$ with $\mathcal{A}$ as in Claim 8.14 above.*

(3) *Lemma 5.2 applied to the sequence $(\mathcal{U}_l)_{l=M}^{M+N-\beta}$ and the domain $\mathcal{A}$ yields the sequence $\underline{\mathcal{U}}''_+$. Similarly, $(\mathcal{V}_l)_{l=M}^{M+N+\beta}$ yields $\underline{\mathcal{V}}''_+$.*

(4) *Lemma 5.4 applied to the sequences $\underline{\mathcal{U}}'$, $\underline{\mathcal{V}}'$ produces the original sequence $\underline{\mathcal{S}}$.*

As we have noted in Remark 8.13, one needs to remove from $\underline{j}$ and $\underline{k}$ their common initial segment to satisfy statement (iii). Item (1) of the claim means that, after removing their common initial segment, the sequences $\underline{j}$ and $\underline{k}$ belong to $\mathbf{P}_{N\pm\beta+\alpha-1}$ with $\alpha = M - s \in \{1, 2, 3, 4\}$, proving (iii).

Now assuming Claim 8.17, let us check that the map $\underline{i} \mapsto (\underline{j}, \underline{k})$ is injective. Suppose given $\underline{j}$ and $\underline{k}$ satisfying (i) and consider the sequences $\underline{\mathcal{U}}$ and $\underline{\mathcal{V}}$ constructed as in (ii), so that $\mathcal{U}_0 = \mathcal{V}_0$. The domain $\mathcal{A}$ is identified uniquely from (2) of the claim, then by (3) one can restore $\underline{\mathcal{U}}''_+$ and $\underline{\mathcal{V}}''_+$ (and hence $\underline{\mathcal{U}}'$ and $\underline{\mathcal{V}}'$), and by (4) the original sequence $\underline{\mathcal{S}}$. Thus we have a unique sequence of domains $\underline{\mathcal{S}}$ generated by $\underline{i}$, together with its initial and final states

$$i_0 = \iota(j_{N-\beta+\alpha-1}), \quad i_{2N-1} = k_{N+\beta+\alpha-1}.$$

Hence, we can uniquely restore the whole sequence $\underline{i}$ as in item (2) of the first part in the proof of Theorem 4.10 as required.

Step 6. Finally, we establish the last claim.

Proof of Claim 8.17. It is convenient to deal with sequences $\underline{i}$ which start and end in the states Ξ_S, Ξ_F, respectively. Thus we begin by showing that the statements of Claim 8.17 for suitably extended sequences imply the same statements for the original ones.

Consider any admissible sequences $(i_l)_{l=-\delta}^{0}$ and $(i_l)_{l=2N-1}^{2N-1+\varepsilon}$ such that $i_{-\delta} \in \Xi_S$, $i_{2N-1+\varepsilon} \in \Xi_F$ and denote

$$\hat{\underline{i}} = (i_l)_{l=-\delta}^{2N-1+\varepsilon}.$$

This extends the sequence $\underline{\mathcal{S}}$ to a sequence $\hat{\underline{\mathcal{S}}} = (\mathcal{S}_l)_{l=-\delta}^{2N+\varepsilon}$ generated by $\hat{\underline{i}}$. Provided that $\underline{i} \notin E_{2N-1}$, we can apply the above procedure of narrowing and joining to these extended sequences. For all sequences involved in this procedure we use the notation as above with an added hat, for example,

$$\hat{\underline{\mathcal{U}}}''_+ = (\hat{\mathcal{U}}''_l)_{l=M}^{M+N-\beta+\delta}.$$

Note that the narrowing and the joining in the original procedure does not change the terminal elements in $\underline{\mathcal{S}}$ and $\underline{i}$. Therefore, the same operations for the extended sequences do not modify any of the added segments. In other words, the sequences $\underline{\widehat{\jmath}}$, $\underline{\widehat{k}}'$, $\underline{\widehat{\mathcal{U}}}''_+$, etc. are the extensions of the corresponding sequences without hats by the segments $(i_l)_{l=-\delta}^{-1}$, $(i_l)_{l=2N}^{2N-1+\varepsilon}$, $(\mathcal{S}_l)_{l=-\delta}^{-1}$, and $(\mathcal{S}_l)_{l=2N+1}^{2N+\varepsilon}$, or their inversions.

Thus the statements of Claim 8.17 for the extended sequences imply the same statements for the original ones. From now on we deal with the extended sequences only.

Above we have constructed a Y-shaped combination of the three thickened paths

$$\underline{\widehat{\mathcal{U}}}''_+, \quad \underline{\widehat{\mathcal{V}}}''_+, \quad a\underline{\mathcal{H}}^{\widetilde{e}}, \tag{23}$$

all meeting in the domain $\mathcal{A}$. Now we construct a related Y-shaped triple of rays. To do this, consider generic points

$$O \in \operatorname{int} \mathcal{A}, \quad X_{\mathcal{U}} \in \partial_O \mathcal{U}''_{M+N-\beta+\delta}, \quad X_{\mathcal{V}} \in \partial_O \mathcal{V}''_{M+N+\beta+\varepsilon}, \quad X_{\mathcal{H}} \in \partial_O (a\mathcal{H}^{\widetilde{e}}_{-M})$$

such that the lines $\ell(OX_J)$, $J = \mathcal{U}, \mathcal{V}, \mathcal{H}$, do not contain any vertices of $\mathbf{T}_{\mathcal{R}}$, where we recall that ∂_O denotes those part of the boundary of the terminal domain in a path that is not shared with the adjacent domain of this path. The segments OX_J lie inside the corresponding convex sets:

$$OX_{\mathcal{U}} \subset \bigcup \underline{\widehat{\mathcal{U}}}''_+, \quad OX_{\mathcal{V}} \subset \bigcup \underline{\widehat{\mathcal{V}}}''_+, \quad OX_{\mathcal{H}} \subset \bigcup a\underline{\mathcal{H}}^{\widetilde{e}}.$$

These sets are the thickened paths between their ends hence each of these segments crosses all consecutive levels in its respective thickened path; see Section 3.2. In particular, these segments leave $\mathcal{A}$ via different sides: $OX_{\mathcal{U}}$ crosses a side from $s_{\mathcal{U}} := s_{N-\beta-1} \cap \partial\mathcal{A}$, $OX_{\mathcal{V}}$ crosses a side from $s_{\mathcal{V}} := s_{N-\beta} \cap \partial\mathcal{A}$, and $OX_{\mathcal{H}}$ crosses $\widetilde{s}$.

Define α_J to be the ray on $\ell(OX_J)$ that starts at O and contains X_J and let $\alpha_J^+ \subset \alpha_J$ be the ray starting at X_J. The rays α_J cut $\mathbb{D}$ into three sectors; denote the sector bounded by α_J and $\alpha_{J'}$ by $\Sigma_{JJ'}$.

From these definitions one can see that none of the curves $\widetilde{s}, s_{\mathcal{U}}, s_{\mathcal{V}}$ intersect with the "opposite" sector. Hence, every path in (23) intersects the "opposite" sector only in an appropriate part of the domain $\mathcal{A}$. Indeed, $\bigcup \underline{\widehat{\mathcal{U}}}''_+$ is a convex set. If $x \in \bigcup \underline{\widehat{\mathcal{U}}}''_+ \setminus \mathcal{A}$, there is a point $y \in Ox$ that belongs to $s_{\mathcal{U}}$. On the other hand, if $x \in \Sigma_{\mathcal{V}\mathcal{H}}$, then $y \in Ox \subset \Sigma_{\mathcal{V}\mathcal{H}}$ and we arrive at a contradiction.

The ends of $\widetilde{s}$ belong to $\Sigma_{\mathcal{U}\mathcal{H}}$ and $\Sigma_{\mathcal{V}\mathcal{H}}$, we denote them by $v_{\mathcal{U}}$, and $v_{\mathcal{V}}$, respectively. Then the domains added to $\underline{\widehat{\mathcal{U}}}''$ by the application of Lemma 5.4 belong to $\Sigma_{\mathcal{U}\mathcal{H}}$. Indeed, the set $\bigcup \underline{\widehat{\mathcal{U}}} \setminus \bigcup \underline{\widehat{\mathcal{U}}}''$ is a connected set that contains $v_{\mathcal{U}} \in \Sigma_{\mathcal{U}\mathcal{H}}$ on its boundary. This set cannot intersect the curve $X_{\mathcal{U}}OX_{\mathcal{H}}$ which lies inside $\bigcup \underline{\widehat{\mathcal{U}}}''$. And if it intersects, say, $\alpha_{\mathcal{U}}^+$, then the intersection $\alpha_{\mathcal{U}} \cap \bigcup \underline{\widehat{\mathcal{U}}}$ is a segment that goes beyond $X_{\mathcal{U}}$. This means that $\bigcup \underline{\widehat{\mathcal{U}}}$ contains the domain bordering $\mathcal{U}''_{M+N-\beta+\delta}$ at $X_{\mathcal{U}}$. But we have required that no domains adjacent to $\mathcal{U}''_{M+N-\beta+\delta}$ are added to $\underline{\widehat{\mathcal{U}}}''$.

We now pass to the proof of the statement of the claim.

(1) This statement follows directly from the second one: if $j_l = k_l$ for $l = 0, \ldots, M$, then $\mathcal{U}_0 = \mathcal{V}_0$ yields $\mathcal{U}_l = \mathcal{V}_l$ for $l = 1, \ldots, M+1$, hence $\mathcal{U}_{M+1} \cap \mathcal{V}_{M+1}$ is nonempty.

(2) Let us show that

$$\bigcup \underline{\widehat{\mathcal{U}}} \cap \bigcup \underline{\widehat{\mathcal{V}}} = \bigcup a \underline{\mathcal{H}}^{\tilde{e}}.$$

Indeed, the domains in $\bigcup \underline{\widehat{\mathcal{U}}}$ fall into three classes:

(a) those from $a\underline{\mathcal{H}}^{\tilde{e}}$,

(b) those from $\bigcup \underline{\widehat{\mathcal{U}}}''_+ \setminus \mathcal{A}$,

(c) those added by the joining.

The first two classes are disjoint since $\underline{\widehat{\mathcal{U}}}$ is a thickened path between its ends, so its different levels do not intersect. The domains in $\bigcup \underline{\widehat{\mathcal{V}}}$ are similarly classified into the classes (a), (b′), and (c′). The classes (b) and (b′) belong to the two different halves of the thickened path $\underline{\widehat{\mathcal{S}}}$ and hence do not intersect. The classes (b) ∪ (c) and (c′) do not intersect since the former contains no domains intersecting $\Sigma_{\mathcal{V}\mathcal{H}}$, while the latter lies in this sector. Hence, the intersection $\bigcup \underline{\widehat{\mathcal{U}}} \cap \bigcup \underline{\widehat{\mathcal{V}}}$ consists of the domains of the class (a) only.

(3) This follows directly from Corollary 5.5.

(4) A joining as in (3) is a minimal convex union of fundamental domains that contains both $\underline{\widehat{\mathcal{U}}}''_+$ and $\underline{\widehat{\mathcal{V}}}''_+$, hence it lies inside $\bigcup \underline{\widehat{\mathcal{S}}}$. Since $\underline{i} \notin E_{2N-1}$, the union

$$\bigcup \underline{\widehat{\mathcal{U}}}''_+ \cap \bigcup \underline{\widehat{\mathcal{V}}}''_+ = \bigcup \underline{\widehat{\mathcal{U}}}' \cap \bigcup \underline{\widehat{\mathcal{V}}}'$$

contains all domains in $\mathcal{S}_l$ with $l \le n_0 - 1$ or $l \ge 2N - n_0 + 1$. Therefore, the joining adds no domains to these levels and hence yields an admissible sequence $\underline{\hat{\imath}}$ with $\mathbf{i}_l = i_l$ for $l \le 0$ and $l \ge 2N-1$. Then each of $\underline{\hat{\imath}}$ and $\underline{\hat{\imath}}$ belongs to $\mathbf{P}^{S \to F}_{2N-1+\delta+\varepsilon}$ and generates the thickened path between $\mathcal{S}_{-\delta}$ and $\mathcal{S}_{2N+\varepsilon}$, hence $\underline{\hat{\imath}} = \underline{\hat{\imath}}$ by Theorem 4.10. ■

This completes the proof of Lemma 8.12. ■

8.4. Conclusion of the proof of Theorem A

Take an ergodic decomposition of the measure μ with respect to the action of the subgroup of G generated by $G_0^2 = \{g_1 g_2 : g_1, g_2 \in G_0\}$ and consider an ergodic G_0^2-invariant measure $\tilde{\mu}$.

Note that in general the operator P does not preserve the measure $\tilde{\mu} \times p$, but the operators Q, V, W defined by (15) do, as they contain only terms of the form

$f \circ T_{g_1} \circ T_{g_2}$ for $g_1, g_2 \in G_0$. Formula (12) then yields

$$\widetilde{\mathbf{S}}_{2n}(f) = \lambda^{2n-1} \sum_{j \in \Xi_S} h_j (Q^{n-1} V \varphi^{(f)})_j.$$

Note also that $\#S(2n)$ equals the number of paths from Ξ_S to Ξ_F of length $2n$, thus

$$\#S(2n) = \sum_{i \in \Xi_S,\, j \in \Xi_F} (\Pi^{2n-1})_{ij} = C\lambda^{2n-1}(1 + o(1)),$$

whence

$$\mathbf{S}_{2n}(f) = \widetilde{C} \sum_{j \in \Xi_S} h_j (Q^{n-1} V \varphi^{(f)})_j \cdot (1 + o(1)).$$

Now we apply Theorem 8.6 to the operators (15) acting on the space $L^1(Y, \widetilde{\nu})$, where $\widetilde{\nu} = \widetilde{\mu} \times p$. Recall that we have checked Assumptions 8.2–8.4 for these operators in Corollary 8.10 and Lemma 8.11. Hence, we obtain that the following holds for $\widetilde{\mu}$-almost every x:

- $\mathbf{S}_{2n}(f)(x)$ converges to some limit, which we denote as $\widetilde{f}(x)$.
- $\widetilde{f}(x) = \widetilde{f}(T_{g_1 g_2} x)$ for any $g_1, g_2 \in G_0$.

The second item results from the fact that $\widetilde{f}$ is constant $\widetilde{\mu}$-almost everywhere.

Therefore, the set of $x \in X$ such that these two conditions hold, is of full measure with respect to every convex combination of the ergodic measures, in particular, with respect to the initial measure μ. Thus $\lim_{n \to \infty} \mathbf{S}_{2n}(f)(x)$ exists μ-almost surely and is G_0^2-invariant. On the other hand, for every $A \in \mathcal{I}_{G_0^2}$ one has

$$\int_A f \, d\mu = \int_A \mathbf{S}_{2n}(f) \, d\mu \to \int_A \widetilde{f} \, d\mu,$$

whence $\widetilde{f} = \mathsf{E}(f | \mathcal{I}_{G_0^2})$. The proof of Theorem A is now complete. ∎

9. Proof of Theorem 8.6

In this section we prove Theorem 8.6, which gives conditions for the pointwise convergence of powers of a Markov operator. This result is a generalization of Theorem 1 in [14], and the proof here follows the same general scheme. After defining the space of trajectories corresponding to Q, we prove first that Q is mixing, next that the tail sigma-algebra of the space of trajectories is trivial, and finally use this to prove convergence for functions in $L \log L$ both in L^1 and pointwise.

9.1. The space of trajectories

Recall that $Q\colon L^1(Z,\eta) \to L^1(Z,\eta)$ is a measure-preserving Markov operator.

The *space of trajectories* corresponding to Q is the space $(\mathbf{Z}, \mathbb{P}_Q)$, where $\mathbf{Z} = Z^{\mathbb{Z}}$ with the usual Borel sigma-algebra $\mathcal{B}_{\mathbf{Z}}$, and the measure $\mathbb{P}_Q$ is defined below. This is essentially an application of the Ionescu Tulcea extension theorem, where the stochastic kernels depend only on the previous element of the trajectory and are the same:

$$\mathbb{P}_Q(z, A) = \mathbb{P}_{Q,z}(A) := Q[\mathbf{1}_A](z). \tag{24}$$

For the details of this construction we refer the reader to [35, Chapter 14]. However, the direct application of this approach faces the following difficulty: the right-hand side of (24) is defined for a fixed A up to a modification on a set of z of zero measure, so we cannot assert that $\mathbb{P}_{Q,z}$ is a sigma-additive measure.

We circumvent this problem by another approach to defining $\mathbb{P}_Q$ below. First, to motivate our definition, let us pretend for a moment that $\mathbb{P}_{Q,z}$ is indeed a sigma-additive measure on Z for any z. Then for an integral with respect to this measure we have

$$\int_{w\in Z} f(w)\, d\mathbb{P}_{Q,z}(w) = Q[f](z) \tag{25}$$

(for $f = \mathbf{1}_A$ this is (24), then use linearity and the monotone convergence theorem).

Now if we say that the conditional distribution of z_n with respect to $(z_m, \dots, z_{n-1})$ should be equal to $\mathbb{P}_{Q,z_{n-1}}$, we get the following formula for the probability of a cylinder set:

$$\begin{aligned}&\mathbb{P}_Q\{z_m \in A_m, \dots, z_n \in A_n\}\\ &= \int_{z_m\in A_m}\left[\int_{z_{m+1}\in A_{m+1}}\left[\cdots\left[\int_{z_n\in A_n} d\mathbb{P}_{Q,z_{n-1}}(z_n)\right]\cdots\right]d\mathbb{P}_{Q,z_m}(z_{m+1})\right]d\eta(z_m)\end{aligned}$$

(compare with [35, Theorem 14.22]). This formula can be rewritten as follows:

$$\begin{aligned}&\mathbb{P}_Q\{z_m \in A_m, \dots, z_n \in A_n\}\\ &\quad = \mathbb{P}_m^n(A_m\times\cdots\times A_n) := \mathsf{E}\big(\mathbf{1}_{A_m}\cdot Q(\mathbf{1}_{A_{m+1}}\cdot Q(\dots Q(\mathbf{1}_{A_n})\dots))\big).\end{aligned} \tag{26}$$

Indeed, the innermost integral is equal to $\mathbb{P}_{Q,z_{n-1}}(A_n) = Q(\mathbf{1}_{A_n})(z_{n-1})$, then we apply (25) for all integrals going from inside out.

Now we may *define* the measure $\mathbb{P}_Q$ as the measure with finite-dimensional distributions $\mathbb{P}_m^n$ given by (26). Let us check that these $\mathbb{P}_m^n$ satisfy assumptions of the Kolmogorov extension theorem.

Lemma 9.1. *The following assertions hold.*

(1) *For any $\varphi \in L^1(Z,\eta)$, we have $\mathsf{E}(Q\varphi) = \mathsf{E}\varphi$.*

(2) $\mathbb{P}_m^n$ *is a finitely-additive measure on the semi-ring of cylinders.*

(3) *If* $B_k \subset A_k$*, then*

$$\mathbb{P}_m^n(A_m \times \cdots \times A_n) - \mathbb{P}_m^n(B_m \times \cdots \times B_n) \le \sum_{k=m}^{n} \eta(A_k) - \eta(B_k).$$

(4) *The measure* $\mathbb{P}_m^n$ *is* σ*-additive.*

(5) *The distributions* $\mathbb{P}_m^n$ *are consistent:*

$$\begin{aligned}\mathbb{P}_m^n(A_m \times \cdots \times A_n) &= \mathbb{P}_{m-1}^n(Z \times A_m \times \cdots \times A_n)\\ &= \mathbb{P}_m^{n+1}(A_m \times \cdots \times A_n \times Z).\end{aligned}$$

Proof. We consider each point in turn.

(1) Since both sides are L^1-continuous, it is sufficient to consider $\varphi \in L^\infty(Z, \eta)$: $-C \le \varphi \le C$. Then $C - \varphi \ge 0$, hence

$$\begin{aligned}C - \mathsf{E}(Q\varphi) = \mathsf{E}(Q(C-\varphi)) &= \|Q(C-\varphi)\|_{L^1}\\ &\le \|C-\varphi\|_{L^1} = \mathsf{E}(C-\varphi) = C - \mathsf{E}\varphi.\end{aligned}$$

Therefore, $\mathsf{E}(Q\varphi) \ge \mathsf{E}\varphi$, and the same argument for $-\varphi$ yields $\mathsf{E}(Q\varphi) \le \mathsf{E}\varphi$.

(2) As usual, this is reduced to the case

$$\mathbb{P}_m^n(C_1) + \mathbb{P}_m^n(C_2) = \mathbb{P}_m^n(C_1 \sqcup C_2),$$

where $C_{1,2}$ have the same projections on all coordinates except one, and this case is clear.

(3) This follows from the inclusion

$$A_m \times \cdots \times A_n \setminus B_m \times \cdots \times B_n \subset \bigcup_{k=m}^{n} Z \times \cdots \times Z \times (A_k \setminus B_k) \times Z \times \cdots \times Z$$

and the first statement of the lemma.

(4) Since (Z, η) is a Lebesgue space, we may assume that it is a union of a segment and an at most countable set of atoms. Then the usual proof works: let $\widehat{C} = \bigsqcup_{i=1}^{\infty} C_i$, then

$$\sum_{i=1}^{\infty} \mathbb{P}_m^n(C_i) \le \mathbb{P}_m^n(\widehat{C})$$

follows from the finite additivity, and to obtain the opposite inequality, find open cylinders $D_i \supset C_i$ with $\mathbb{P}_m^n(D_i) \le \mathbb{P}_m^n(C_i) + \varepsilon/2^i$ and a compact cylinder $\widehat{D} \subset \widehat{C}$ with $\mathbb{P}_m^n(\widehat{D}) \ge \mathbb{P}_m^n(\widehat{C}) - \varepsilon$. These cylinders are constructed coordinate-wise using the estimate from the previous item. Then $\widehat{D}$ is covered by D_i's and hence by a finite number of them. Finite additivity then yields

$$\mathbb{P}_m^n(\widehat{D}) \leq \sum_{i=1}^{N} \mathbb{P}_m^n(D_i) \leq \sum_{i=1}^{\infty} \mathbb{P}_m^n(D_i),$$

hence

$$\mathbb{P}_m^n(\widehat{C}) \leq \sum_{i=1}^{\infty} \mathbb{P}_m^n(C_i) + 2\varepsilon.$$

It remains to take a limit as $\varepsilon \to +0$. This proves the σ-additivity on the semi-ring of cylinders, and the Carathédory extension theorem then extends $\mathbb{P}_m^n$ to a σ-additive measure on the Borel σ-algebra on Z^{n-m+1}.

(5) This is straightforward using the first statement of this lemma. ■

Therefore, $\mathbb{P}_Q$ defined by (26) exists by the Kolmogorov extension theorem.

It is also clear from the definition that *the left shift map* $\sigma\colon \mathbf{Z} \to \mathbf{Z}$, $(\sigma(\mathbf{z}))_n = z_{n+1}$ preserves the measure $\mathbb{P}_Q$.

We can clearly define a measure $\mathbb{P}_{Q^*}$ in a similar way. The following calculation relates $\mathbb{P}_Q$ and $\mathbb{P}_{Q^*}$. We have

$$\begin{aligned}\mathbb{P}_Q\{z_m \in A_m, \ldots, z_n \in A_n\} &= \langle \mathbf{1}_{A_m}, Q(\mathbf{1}_{A_{m+1}} \cdot Q(\ldots Q(\mathbf{1}_{A_n})\ldots))\rangle \\ &= \langle Q^*(\mathbf{1}_{A_m}), \mathbf{1}_{A_{m+1}} \cdot Q(\ldots Q(\mathbf{1}_{A_n})\ldots)\rangle \\ &= \langle \mathbf{1}_{A_{m+1}} \cdot (Q^*(\mathbf{1}_{A_m})), Q(\ldots Q(\mathbf{1}_{A_n})\ldots)\rangle \\ &= \cdots = \mathbb{P}_{Q^*}\{z_{-n} \in A_n, \ldots, z_{-m} \in A_m\}.\end{aligned}$$

In other words, $\mathbb{P}_{Q^*}$ is a pullback of $\mathbb{P}_Q$ under the time-reversal map

$$(z_n)_{n=-\infty}^{\infty} \mapsto (z_{-n})_{n=-\infty}^{\infty}.$$

We now derive an important result concerning conditional expectations. Let $\mathcal{F}_k^l$, $k, l \in \mathbb{Z} \cup \{+\infty, -\infty\}$ be the minimal complete sigma-algebras such that all functions $\pi_j\colon \mathbf{z} = (z_n) \mapsto z_j$ are measurable for $k \leq j \leq l$. For brevity we write $\mathcal{F}_n^n = \mathcal{F}_n$. Let us also recall that the *tail sigma-algebra* is defined as

$$\mathcal{F}_{\text{tail}} = \bigcap_{n=0}^{\infty} \mathcal{F}_n^{\infty}.$$

For any function $\varphi \in L^1(Z, \eta)$ we define the functions $\varphi^r \in L^1(\mathbf{Z}, \mathbb{P}_Q)$, $r \in \mathbb{Z}$ by the formula $\varphi^r(\mathbf{z}) = \varphi(z_r)$. The next lemma shows how to find the conditional expectation of a function on $\mathbf{Z}$ that depends on only one coordinate with respect to the sigma-algebra generated by some other coordinates.

Lemma 9.2. *For any $\varphi \in L^1(Z,\eta)$ and $n > 0$, we have*

$$\begin{aligned} \mathsf{E}(\varphi^r|\mathcal{F}_{-\infty}^{-n+r})(\mathbf{z}) &= \mathsf{E}(\varphi^r|\mathcal{F}_{-n+r})(\mathbf{z}) = (Q^n\varphi)(z_{-n+r}), \\ \mathsf{E}(\varphi^r|\mathcal{F}_{n+r}^{+\infty})(\mathbf{z}) &= \mathsf{E}(\varphi^r|\mathcal{F}_{n+r})(\mathbf{z}) = ((Q^*)^n\varphi)(z_{n+r}). \end{aligned} \tag{27}$$

Note that this lemma proves the starting point of our heuristic approach above, the formula (24): the first equality in (27) for $n = 1$ and $\varphi = \mathbf{1}_A$ yields

$$\mathbb{P}_Q(z_r \in A \mid z_{r-1}, z_{r-2}, \dots) = Q[\mathbf{1}_A](z_{r-1}).$$

Proof of Lemma 9.2. Let us check that

$$\mathsf{E}(\varphi^r|\mathcal{F}_{-\infty}^{-n+r})(\mathbf{z}) = (Q^n\varphi)(z_{-n+r}).$$

The sets of the form $\{z_m \in A_m, \dots, z_{-n+r} \in A_{-n+r}\}$ for all $m \le -n+r$ and all Borel A_j's generate the sigma-algebra $\mathcal{F}_{-\infty}^{-n+r}$, hence it is sufficient to check that for all such sets B, we have

$$\mathsf{E}(\mathbf{1}_B(\mathbf{z})\cdot\varphi(z_r)) = \mathsf{E}(\mathbf{1}_B(\mathbf{z})(Q^n\varphi)(z_{-n+r})). \tag{28}$$

To do this, observe that

$$\mathsf{E}_{\mathbb{P}_Q}(\mathbf{1}_{A_m}(z_m)\cdots\mathbf{1}_{A_{k-1}}(z_{k-1})\cdot\psi(z_k)) = \mathsf{E}_\eta(\mathbf{1}_{A_m}\cdot Q(\dots Q(\mathbf{1}_{A_{k-1}}\cdot Q(\psi))\dots)).$$

Indeed, for $\psi = \mathbf{1}_{A_k}$ this is (26); the general case follows by the linearity and L^1-continuity of both sides. Therefore, both sides of (28) are equal to

$$\mathsf{E}_\eta\big(\mathbf{1}_{A_m}\cdot Q(\dots Q(\mathbf{1}_{A_{-n-1+r}}\cdot Q(\mathbf{1}_{A_{-n+r}}\cdot Q^n(\varphi)))\dots)\big).$$

Also, since $\mathcal{F}_{-n+r} \subset \mathcal{F}_{-\infty}^{-n+r}$ and we have seen that $\mathsf{E}(\varphi^r|\mathcal{F}_{-\infty}^{-n+r})$ is $\mathcal{F}_{-n+r}$-measurable, we obtain $\mathsf{E}(\varphi^r|\mathcal{F}_{-\infty}^{-n+r}) = \mathsf{E}(\varphi^r|\mathcal{F}_{-n+r})$. This proves the first formula in (27), and the second one follows by time reversal as above. ∎

We have seen in the previous lemma that the conditional expectation of a function depending only on r-th coordinate, $r > k$, with respect to the "past" σ-algebra $\mathcal{F}_{-\infty}^{k}$ depends only on z_k, the last coordinate in this interval. This statement can be extended to any function that is measurable with respect to the "future" σ-algebra $\mathcal{F}_{k+1}^{\infty}$.

Lemma 9.3. *Assume that a function $\Phi \in L^1(\mathbf{Z},\mathbb{P}_Q)$ is $\mathcal{F}_{k+1}^{\infty}$-measurable. We then have that $\mathsf{E}(\Phi|\mathcal{F}_{-\infty}^{k})$ depends on z_k only.*

Proof. If $\Phi = \mathbf{1}_C$, where $C = \{z_{k+1} \in C_{k+1}, \dots, z_{k+s} \in C_{k+s}\}$, we use the same argument as for (28) and obtain

$$\mathsf{E}(\Phi|\mathcal{F}_{-\infty}^{k})(\mathbf{z}) = Q(\mathbf{1}_{C_{k+1}}\cdot Q(\mathbf{1}_{C_{k+2}}\cdot Q(\dots Q(\mathbf{1}_{C_{k+s}})\dots)))(z_k).$$

The general case follows by linearity and the fact that $\Phi_n \to \Phi$ in L^1 yields

$$\mathsf{E}(\Phi_n|\mathcal{G}) \to \mathsf{E}(\Phi|\mathcal{G}) \quad \text{in } L^1.$$

∎

9.2. Mixing of the operator Q

We start by proving mixing for $\widetilde{Q} = Q^m$ where m is defined in Assumption 8.3.

Lemma 9.4. *Let $\widetilde{Q}$ be a measure-preserving Markov operator on $L^1(Z, \eta)$ such that the equation $\widetilde{Q}^* \widetilde{Q}\varphi = \varphi$ has only constant solutions in $L^2(Z, \eta)$. Then for any $\varphi, \psi \in L^2(Z, \eta)$, we have*

$$\langle \widetilde{Q}^n \varphi, \psi \rangle = \int_Z \widetilde{Q}^n \varphi \cdot \bar{\psi}\, d\eta \to \int_Z \varphi\, d\eta \int_Z \bar{\psi}\, d\eta \quad \text{as } n \to \infty. \tag{29}$$

Proof. The statement follows from the mixing of the shift map σ in the trajectory space $(\mathbf{Z}, \mathbb{P}_{\widetilde{Q}})$. To obtain the latter we shall prove that σ has the K-property: that is, there exists a sub-sigma-algebra $\mathcal{K}$ of the Borel sigma-algebra $\mathcal{B}_{\mathbf{Z}}$ such that

$$\mathcal{K} \subset \sigma\mathcal{K}, \quad \bigvee_{n=0}^{\infty} \sigma^n \mathcal{K} = \mathcal{B}_{\mathbf{Z}}, \quad \bigcap_{n=0}^{\infty} \sigma^{-n} \mathcal{K} = \{\varnothing, \mathbf{Z}\}.$$

By the Rokhlin–Sinai theorem (see [44] and [26, Chapter 18]) the K-property is equivalent to the triviality of the Pinsker sigma-algebra $\Pi(\sigma)$ (the smallest sigma-algebra containing all measurable partitions of zero entropy). Consider $\mathcal{F}_- = \mathcal{F}_{-\infty}^0$. Then

$$\sigma\mathcal{F}_- \subset \mathcal{F}_-, \quad \bigvee_{k\in\mathbb{Z}} \sigma^k \mathcal{F}_- = \mathcal{B}_{\mathbf{Z}}.$$

Thus $\Pi(\sigma^{-1}) \subset \mathcal{F}_-$ (see, e.g., [26, Lemma 18.7.3]). Similarly, for $\mathcal{F}_+ = \mathcal{F}_0^\infty$ one has $\Pi(\sigma) \subset \mathcal{F}_+$. Therefore,

$$\Pi(\sigma) = \Pi(\sigma^{-1}) \subset \mathcal{F}_- \cap \mathcal{F}_+ = \mathcal{F}_0.$$

We have proved that any $\Pi(\sigma)$-measurable function $\varphi \in L^2(\mathbf{Z}, \mathbb{P}_{\widetilde{Q}})$ depends only on the zeroth coordinate: $\varphi(\mathbf{z}) = \varphi_0(z_0)$; with the notation of Section 9.1, $\varphi = (\varphi_0)^0$. More generally, $\varphi(\mathbf{z}) = \varphi_k(z_k)$.

Now we can calculate $\mathsf{E}(\varphi|\mathcal{F}_{-1})$ in two ways. On the one hand, φ is $\mathcal{F}_{-1}$-measurable, so it equals $\varphi = \varphi_{-1}(z_{-1})$, on the other hand, it equals

$$\mathsf{E}((\varphi_0)^0|\mathcal{F}_{-1}) = (\widetilde{Q}\varphi_0)(z_{-1})$$

by (27). Hence, $\varphi_{-1} = \widetilde{Q}\varphi_0$. Similarly, from $\mathsf{E}(\varphi|\mathcal{F}_0) = \varphi = \mathsf{E}((\varphi_{-1})^{-1}|\mathcal{F}_0)$, we obtain $\varphi_0 = \widetilde{Q}^*\varphi_{-1}$. Therefore,

$$\varphi_0 = \widetilde{Q}^* \widetilde{Q} \varphi_0,$$

hence by assumption of the lemma $\varphi_0 = \text{const}$, thus $\Pi(\sigma)$ is trivial. ∎

Corollary 9.5. *The operator Q is also mixing, that is,* (29) *holds for Q instead of $\widetilde{Q}$.*

Proof. The sequence $(\langle Q^n\varphi, \psi\rangle)_{n\geq 0}$ is the union of subsequences $(\langle Q^{nm+r}\varphi, \psi\rangle)_{n\geq 0}$, each of which converges to the desired limit by Lemma 9.4 applied to the pair of functions $(Q^r\varphi, \psi)$. ∎

9.3. Triviality of the tail sigma-algebra

The next step is to prove that the tail sigma-algebra for Q is trivial. First, we prove that the tail sigma-algebra cannot be *totally non-trivial*, that is, it cannot contain infinitely many different sets (up to sets of measure zero). The proof follows that of Lemma 6 in [14], which is a version of the 0–2 law in the form of Kaimanovich [33].

Lemma 9.6. *For a measure-preserving Markov operator R on $L^1(Z,\eta)$ the following holds. If the tail sigma-algebra of R is totally non-trivial then for any $b \in \mathbb{N}$ and any $\varepsilon > 0$ there exist non-negative functions $\varphi, \psi \in L^\infty(Z,\eta)$ with averages equal to 1 such that*

$$\limsup_{n\to\infty}\langle (R^*)^{n+b}\varphi, (R^*)^n\psi\rangle_{L^2(Z,\eta)} + \cdots + \langle (R^*)^{n-b}\varphi, (R^*)^n\psi\rangle_{L^2(Z,\eta)} < \varepsilon. \quad (30)$$

Proof. Let $(\mathbf{Z}, \mathbb{P}_R)$ be the corresponding trajectory space. If $\mathcal{F}_{\text{tail}}$ contains infinitely many subsets, it contains a subset of arbitrarily small measure. Indeed, split $\mathbf{Z} = A_1^{(2)} \sqcup A_2^{(2)}$, where each set $A_i^{(2)} \in \mathcal{F}_{\text{tail}}$ has non-zero measure. Then at least one of these parts can be split into two sets of non-zero measure (otherwise $\mathcal{F}_{\text{tail}}$ contains only finitely many sets, the union of some of the $A_i^{(2)}$). Repeating this procedure, we get $\mathbf{Z} = A_1^{(n)} \sqcup \cdots \sqcup A_n^{(n)}$. Then the measure of at least one of $A_j^{(n)}$ is not more than $1/n$.

Take any set $A\in\mathcal{F}_{\text{tail}}$ with $\mathbb{P}_R(A)<1/(2b+1)$. Then the set $B=\mathbf{Z}\setminus\bigcup_{s=-b}^{b}\sigma^s(A)$ has positive measure. Denote

$$\Phi(\mathbf{z}) = \mathbf{1}_A(\mathbf{z})/\mathbb{P}_R(A), \quad \Psi(\mathbf{z}) = \mathbf{1}_B(\mathbf{z})/\mathbb{P}_R(B).$$

Observe that Φ and Ψ are non-negative $\mathcal{F}_{\text{tail}}$-measurable functions, bounded by some constant M, and with expectations equal to 1. Moreover, $(\Phi\circ\sigma^{-j})\cdot\Psi = 0$ for $j = -b, \dots, b$.

Set $\varphi_k = \mathsf{E}(\Phi|\mathcal{F}_{-\infty}^k)$, $\psi_k = \mathsf{E}(\Psi|\mathcal{F}_{-\infty}^k)$. By Lemma 9.3, $\varphi_k(\mathbf{z})$ depends only on z_k, so abusing notation we use the same symbol φ_k for the corresponding function in $L^1(Z,\eta)$. For example, we will write $\varphi_k\circ\sigma^j(\mathbf{z}) = \varphi_k(z_{k+j})$.

Clearly, φ_k and ψ_k are non-negative and bounded by M. Therefore, the Martingale convergence theorem gives that $\varphi_k \to \Phi$, $\psi_k \to \Psi$ in $L^1(\mathbf{Z}, \mathbb{P}_R)$. Moreover,

$$\varphi_k(z_{k-j}) = \varphi_k\circ\sigma^{-j}(\mathbf{z}) \quad \text{and} \quad \varphi_k\circ\sigma^{-j} \to \Phi\circ\sigma^{-j}.$$

Hence,

$$\mathsf{E}(\varphi_k(z_{k-j})|\mathcal{F}_{\text{tail}}) \to \mathsf{E}(\Phi \circ \sigma^{-j}|\mathcal{F}_{\text{tail}}) = \Phi \circ \sigma^{-j}(\mathbf{z}), \quad \mathsf{E}(\psi_k(z_k)|\mathcal{F}_{\text{tail}}) \to \Psi(\mathbf{z})$$

in $L^1(\mathbf{Z}, \mathbb{P}_R)$. Since all these functions are bounded by the same constant M, for large k, we have that

$$\int_{\mathbf{Z}} \mathsf{E}(\varphi_k(z_{k-j})|\mathcal{F}_{\text{tail}})\mathsf{E}(\psi_k(z_k)|\mathcal{F}_{\text{tail}})\, d\mathbb{P}_R < \frac{\varepsilon}{2b+1}.$$

Applying the second formula in (27) to $\varphi_k^{k-j}(\mathbf{z}) = \varphi_k(z_{k-j})$ and $n + j$ in place of n, we obtain

$$\mathsf{E}(\varphi_k(z_{k-j})|\mathcal{F}_{n+k}^{\infty}) = [(R^*)^{n+j}\varphi_k](z_{n+k}).$$

Hence, for any $j = -b, \dots, b$, we have

$$\begin{aligned}
&\int_Z [(R^*)^{n+j}\varphi_k](z_{n+k}) \cdot [(R^*)^n \psi_k](z_{n+k})\, d\eta \\
&\quad = \int_{\mathbf{Z}} \mathsf{E}(\varphi_k(z_{k-j})|\mathcal{F}_{n+k}^{\infty}) \cdot \mathsf{E}(\psi_k(z_k)|\mathcal{F}_{n+k}^{\infty})\, d\mathbb{P}_R \\
&\quad \to \int_{\mathbf{Z}} \mathsf{E}(\varphi_k(z_{k-j})|\mathcal{F}_{\text{tail}}) \cdot \mathsf{E}(\psi_k(z_k)|\mathcal{F}_{\text{tail}})\, d\mathbb{P}_R < \frac{\varepsilon}{2b+1} \quad \text{as } n \to \infty.
\end{aligned}$$

Therefore, the functions φ_k and ψ_k for large k satisfy (30). ■

Lemma 9.7. *Under the assumptions of Theorem 8.6 the tail sigma-algebra for Q^* cannot be totally non-trivial.*

Proof. Assuming the contrary, the inequality (30) in Lemma 9.6 for $R = Q^*$ yields that for some non-negative functions φ, ψ with their averages equal to 1 and for all sufficiently large n we have

$$\langle (Q^{n+b} + \cdots + Q^{n-b})\varphi, Q^n \psi \rangle_{L^2(Z,\eta)} < \varepsilon. \tag{31}$$

On the other hand, by Assumption 8.4, the left-hand side of (31) is not less than

$$\frac{1}{C}\langle W Q^{2n-a}\varphi - A_n\varphi, \psi\rangle = \frac{1}{C}\langle Q^{2n-a}\varphi, W^*\psi\rangle - \frac{1}{C}\langle A_n\varphi, \psi\rangle \to \frac{1}{C} + 0.$$

Here we use Corollary 9.5; note that the average values of both φ and $W^*\psi$ are equal to 1. Therefore, for large n the left-hand side of (31) is larger than $1/C - \varepsilon$, so taking $\varepsilon < 1/2C$ we arrive at a contradiction. ■

Lemma 9.8. *Under the assumptions of Theorem 8.6 the tail sigma-algebra for Q cannot be totally non-trivial.*

Proof. Consider the trajectory space $(\mathbf{Z}, \mathbb{P})$ for the infinite sequence $\dots, V, W, V, W, \dots$ of Markov operators, that is,

$$\mathbb{P}(z_{2n+1} \in A \mid z_{2n}) = V[\mathbf{1}_A](z_{2n}), \quad \mathbb{P}(z_{2n+2} \in A \mid z_{2n+1}) = W[\mathbf{1}_A](z_{2n+1}).$$

In other words, we use the construction from Section 9.1, but with (26) replaced by

$$\mathbb{P}\{z_m \in A_m, \dots, z_n \in A_n\} = \mathsf{E}\big(\mathbf{1}_{A_m} \cdot R_m(\mathbf{1}_{A_{m+1}} \cdot R_{m+1}(\dots R_{n-1}(\mathbf{1}_{A_n})\dots))\big), \quad (32)$$

where $R_{2k} = V$, $R_{2k+1} = W$ for all $k \in \mathbb{Z}$. In fact, Lemma 9.1 holds for the finite-dimensional distributions (32) with any sequence of Markov operators (R_k). In our case, we have that

$$\begin{aligned}&\mathbb{P}\{z_{2k} \in A_k, z_{2(k+1)} \in A_{k+1}, \dots, z_{2l} \in A_l\}\\ &\quad = \mathbb{P}_Q\{z_k \in A_k, z_{k+1} \in A_{k+1}, \dots, z_l \in A_l\},\end{aligned}$$

hence the projection $\pi_0 \colon \mathbf{z} = (z_n) \mapsto (z_{2n})$ maps the trajectory space $(\mathbf{Z}, \mathbb{P})$ to the trajectory space $(\mathbf{Z}, \mathbb{P}_Q)$ for the operator $Q = VW$. Similarly, $\pi_1 \colon \mathbf{z} = (z_n) \mapsto (z_{2n+1})$ maps it to the trajectory space for $Q^* = WV$. Therefore, the total non-triviality of the tail sigma-algebras in the trajectory spaces for Q and Q^* is equivalent respectively to that of the sigma-algebras

$$\mathcal{F}_{\text{tail},0} = \bigcap_n \bigvee_{2k \geq n} \mathcal{F}_{2k} \quad \text{and} \quad \mathcal{F}_{\text{tail},1} = \bigcap_n \bigvee_{2k+1 \geq n} \mathcal{F}_{2k+1}$$

in the trajectory space $(\mathbf{Z}, \mathbb{P})$. Since we already know that $\mathcal{F}_{\text{tail},1}$ cannot be totally non-trivial, it is sufficient to prove that

$$\mathcal{F}_{\text{tail},j} = \mathcal{F}_{\text{tail},\mathbf{Z}} := \bigcap_n \bigvee_{k \geq n} \mathcal{F}_k.$$

Clearly, $\mathcal{F}_{\text{tail},j} \subset \mathcal{F}_{\text{tail},\mathbf{Z}}$. Let us prove the converse inclusion. Consider any $A \in \mathcal{F}_{\text{tail},\mathbf{Z}}$ and check that, say, $A \in \mathcal{F}_{\text{tail},0}$. Indeed, $A \in \bigvee_{m \geq 2n} \mathcal{F}_m$ for every n, and we can eliminate any finite number of $\mathcal{F}_k$ with odd k from this formula:

$$A \in \mathcal{F}_{2n} \vee \mathcal{F}_{2n+2} \vee \dots \vee \mathcal{F}_{2(n+s-1)} \vee \bigvee_{m \geq 2(n+s)} \mathcal{F}_m. \quad (33)$$

Consider the conditional probability $\mathbb{P}(\cdot \mid z_{2n}, z_{2n+2}, \dots)$ with respect to the sigma-algebra $\bigvee_{k \geq n} \mathcal{F}_{2k}$. As (33) shows, with respect to this conditional probability A depends only on the "odd tail" $\bigvee_{k \geq n+s} \mathcal{F}_{2k+1}$. But since the odd coordinates $z_{2n+1}, \dots, z_{2(n+s)+1}, \dots$ are independent for fixed even coordinates $z_{2n}, \dots, z_{2(n+s)}, \dots$, by Kolmogorov's 0–1 law, we obtain that A is trivial with respect to this conditional probability, so A is measurable with respect to $\bigvee_{k \geq n} \mathcal{F}_{2k}$, and hence $A \in \mathcal{F}_{\text{tail},0}$. ■

Lemma 9.9. *Under the assumptions of Theorem 8.6 the tail sigma-algebra for Q is trivial.*

Proof. It remains to eliminate the case in which $\mathcal{F}_{\text{tail}}$ contains only finitely many different sets. Assume that $\mathbf{Z} = A_1 \sqcup \cdots \sqcup A_r$, $r > 1$, where each $A_j \in \mathcal{F}_{\text{tail}}$ has no non-trivial subsets belonging to $\mathcal{F}_{\text{tail}}$. The shift map σ interchanges these subsets, whence for $A = A_1$, there exists n such that $\sigma^n A = A$. As in Lemma 9.6, we define $\Phi = \mathbf{1}_A / \mathbb{P}_Q(A)$ and $\varphi_k(z_k) = \varphi_k(\mathbf{z}) = \mathsf{E}(\Phi | \mathcal{F}^k_{-\infty})$. Then

$$\mathsf{E}(\Phi|\mathcal{F}^k_{-\infty}) \circ \sigma^n = \mathsf{E}(\Phi \circ \sigma^n | \mathcal{F}^{k+n}_{-\infty}) = \mathsf{E}(\Phi|\mathcal{F}^{k+n}_{-\infty}) = \varphi_{k+n},$$

hence

$$\begin{aligned}\varphi_{k+n}(z_k) &= \varphi_{k+n} \circ \sigma^{-n}(\mathbf{z}) = \mathsf{E}(\Phi|\mathcal{F}^k_{-\infty}) \\ &= \mathsf{E}(\mathsf{E}(\Phi|\mathcal{F}^{k+n}_{-\infty})|\mathcal{F}^k_{-\infty}) = \mathsf{E}(\varphi_{k+n}(z_{k+n})|\mathcal{F}^k_{-\infty}) = [Q^n \varphi_{k+n}](z_k).\end{aligned}$$

Thus we arrive at the equation $\varphi_{k+n}(z_k) = [Q^n \varphi_{k+n}](z_k)$ and Assumption 8.2 implies that φ_{k+n} is constant. Taking averages, we get $\mathsf{E}(\varphi_{k+n}) = \mathsf{E}(\Phi) = 1$, thus $\varphi_l \equiv 1$ for all l. But this contradicts the convergence $\varphi_l \to \Phi \not\equiv 1$, which was obtained in proof of Lemma 9.6. ∎

9.4. Convergence

Proposition 9.10 (see [33] and [14, Propositions 4, 5]). *For a measure-preserving Markov operator R on (Z, η) with trivial tail sigma-algebra, we have*

$$R^n \varphi \to \int_Z \varphi \, d\eta,$$

where the convergence takes place in L^1 for $\varphi \in L^1(Z,\eta)$ and in L^2 for $\varphi \in L^2(Z,\eta)$.

It remains to prove almost everywhere pointwise convergence for functions in L^p, $p > 1$ and in $L \log L$. Recall that the norm in $L \log L(Z, \eta)$ can be defined by the *Orlicz–Luxemburg norm*

$$\|\varphi\|_{L \log L} = \inf\left\{c : \int_Z \frac{|\varphi|}{c} \cdot \log\left(\frac{|\varphi|}{c} + e\right) d\eta \le 1\right\},$$

see, for example, [51]. In particular, since

$$\|\varphi\|_{L^1} = \inf\left\{c : \int_Z (|\varphi|/c) \, d\eta \le 1\right\},$$

we have $\|\varphi\|_{L^1} \le A \|\varphi\|_{L \log L}$ for some constant A. More generally, we have the following maximal inequalities:

Lemma 9.11 ([14, Lemma 8]). *For a measure-preserving Markov operator R on (Z,η) for any $p>1$, there exists a constant $A_p>0$ such that, for any non-negative function $\varphi\in L^p(Z,\eta)$, we have*

$$\Big\|\sup_{n\geq 0}(R^*)^n R^n\varphi\Big\|_{L^p}\leq A_p\|\varphi\|_{L^p}.$$

Similarly, there exists a constant $A_{\log}>0$ such that, for any non-negative function $\varphi\in L\log L(Z,\eta)$, we have

$$\Big\|\sup_{n\geq 0}(R^*)^n R^n\varphi\Big\|_{L^1}\leq A_{\log}\|\varphi\|_{L\log L}.$$

Corollary 9.12. *For any $s\in\mathbb{Z}$, the following inequalities hold:*

$$\Big\|\sup(R^*)^n R^{n+s}\varphi\Big\|_{L^p}\leq A_p\|\varphi\|_{L^p},\quad \Big\|\sup(R^*)^n R^{n+s}\varphi\Big\|_{L^1}\leq A_{\log}\|\varphi\|_{L\log L},\tag{34}$$

the suprema here being taken over all n such that both n and $n+s$ are non-negative.

Proof. We treat the $L\log L$ norm, the L^p norms are similar. For $s>0$, apply the lemma to $R^s\varphi$ and use the inequality $\|R^s\varphi\|_{L\log L}\leq\|\varphi\|_{L\log L}$. For $s<0$, we use the inequality

$$(R^*)^{|s|}\Big[\sup_{n\geq 0}(R^*)^n R^n\varphi\Big]\geq\sup_{n\geq 0}(R^*)^{n+|s|}R^n\varphi$$

from which we obtain the same inequality for the L^1-norms of both sides. Note also that the L^1-norm of the left-hand side does not exceed $\|\sup_{n\geq 0}(R^*)^n R^n\varphi\|_{L^1}$. Hence,

$$A_{\log}\|\varphi\|_{L\log L}\geq\Big\|\sup_{n\geq 0}(R^*)^n R^n\varphi\Big\|\geq\Big\|\sup_{n\geq 0}(R^*)^{n+|s|}R^n\varphi\Big\|,$$

and it remains to replace n with $n+s$. ■

Proof of Theorem 8.6. It remains to prove pointwise convergence. Combining (34) for $R=Q$ and Assumption 8.4 in the form (14′), for any non-negative function $\varphi\in L\log L(Z,\eta)$, we obtain

$$\begin{aligned}\Big\|\sup_{n\geq n_0}Q^{2n-a'}\varphi\Big\|_{L^1}&\leq C\Big\|V'\Big(\sup_{n\geq n_0}\sum_{j=-b}^{b}(Q^*)^n Q^{n+j}\varphi\Big)\Big\|_{L^1}+\Big\|\sup_{n\geq n_0}A'_n\varphi\Big\|_{L^1}\\&\leq(2b+1)A_{\log}C\|\varphi\|_{L\log L}+\sum_{n\geq n_0}\|A'_n\varphi\|_{L^1}\\&\leq B_{\log}\|\varphi\|_{L\log L}.\end{aligned}\tag{35}$$

Decomposing a function φ into its positive and negative parts we obtain (35) for all real-valued $\varphi\in L\log L(Z,\eta)$ with a larger $B_{\log}$. The same estimates hold for the L^p-norm with $p>1$.

Now consider a real-valued function $\varphi \in L^2(Z, \eta)$ with zero average. Applying (35) to $(Q^{2k}\varphi)$, we have

$$\Big\| \sup_{m \geq n_0 + k} Q^{2m-a'}\varphi \Big\|_{L^2} = \Big\| \sup_{n \geq n_0} Q^{2n+2k-a'}\varphi \Big\|_{L^2} \leq B_2 \|Q^{2k}\varphi\|_{L^2}.$$

Since the right-hand side tends to zero by Proposition 9.10, the sequence $Q^{2m-a'}\varphi$ tends to zero almost everywhere and in L^2 as $m \to \infty$.

We now extend pointwise convergence to all $\varphi \in L \log L$. Namely, for a real-valued function $\varphi \in L \log L(Z, \eta)$ with zero average, consider $\varphi' \in L^2(Z, \eta)$ with zero average such that

$$\|\varphi - \varphi'\|_{L \log L} \leq \varepsilon / B_{\log}.$$

Then almost surely, we have

$$\limsup_{n\to\infty} |Q^{2n-a'}\varphi(z)| \leq \limsup_{n\to\infty} |Q^{2n-a'}\varphi'(z)| + \limsup_{n\to\infty} |Q^{2n-a'}(\varphi - \varphi')(z)|.$$

By convergence for functions in L^2, the first term in the right-hand side equals zero, while, by the maximal inequality, the second satisfies

$$\Big\| \limsup_{n\to\infty} |Q^{2n-a'}(\varphi - \varphi')(z)| \Big\|_{L^1} \leq B_{\log} \|\varphi - \varphi'\|_{L \log L} \leq \varepsilon.$$

Therefore, we have $\limsup |Q^{2n-a'}\varphi(z)| \leq \delta$ outside a set of measure less than ε/δ for any $\delta > 0$. Taking $\varepsilon = 1/l^2$ and then $\delta = 1/l$ with $l \to \infty$, we obtain that this upper limit equals zero almost everywhere. The convergence in L^1 follows from the same decomposition:

$$\|Q^{2n-a'}\varphi(z)\|_{L^1} \leq \|Q^{2n-a'}\varphi'\|_{L^1} + \|Q^{2n-a'}(\varphi - \varphi')(z)\|_{L^1},$$

where the first term tends to zero even with the L^2-norm instead of L^1, and the second term is less than $\|\varphi - \varphi'\|_{L^1} \leq \varepsilon / B_{\log}$.

Finally, combining the convergence $Q^{2m-a'}\varphi \to 0$ already obtained with the same convergence for $Q\varphi$ in place of φ, we conclude that $Q^n\varphi \to 0$ almost everywhere as claimed. ∎

Funding. The research of A. Bufetov on this project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 647133 (ICHAOS). A. Klimenko's research was partially funded by Russian Fund of Basic Research (grants No. 18-31-20031 and 18-51-15010).

References

[1] V. I. Arnol'd and A. L. Krylov, Uniform dixion of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain. *Dokl. Akad. Nauk SSSR* **148** (1963), 9–12 Zbl 0237.34008 MR 0150374

[2] A. F. Beardon, An introduction to hyperbolic geometry. In *Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989)*, pp. 1–33, Oxford Sci. Publ., Oxford University Press, New York, 1991 Zbl 0755.30001 MR 1130171

[3] J. S. Birman and C. Series, Dehn's algorithm revisited, with applications to simple curves on surfaces. In *Combinatorial group theory and topology (Alta, Utah, 1984)*, pp. 451–478, Ann. of Math. Stud. 111, Princeton University Press, Princeton, NJ, 1987 Zbl 0624.20033 MR 895628

[4] L. Bowen, Invariant measures on the space of horofunctions of a word hyperbolic group. *Ergodic Theory Dynam. Systems* **30** (2010), no. 1, 97–129 Zbl 1205.37007 MR 2586347

[5] L. Bowen, A. Bufetov, and O. Romaskevich, Mean convergence of Markovian spherical averages for measure-preserving actions of the free group. *Geom. Dedicata* **181** (2016), 293–306 Zbl 1355.37011 MR 3475750

[6] L. Bowen and A. Nevo, Geometric covering arguments and ergodic theorems for free groups. *Enseign. Math. (2)* **59** (2013), no. 1-2, 133–164 Zbl 1373.37012 MR 3113602

[7] L. Bowen and A. Nevo, A horospherical ratio ergodic theorem for actions of free groups. *Groups Geom. Dyn.* **8** (2014), no. 2, 331–353 Zbl 1347.37008 MR 3231218

[8] L. Bowen and A. Nevo, Amenable equivalence relations and the construction of ergodic averages for group actions. *J. Anal. Math.* **126** (2015), 359–388 Zbl 1358.37015 MR 3358037

[9] L. Bowen and A. Nevo, Von Neumann and Birkhoff ergodic theorems for negatively curved groups. *Ann. Sci. Éc. Norm. Supér. (4)* **48** (2015), no. 5, 1113–1147 Zbl 1359.37007 MR 3429477

[10] R. Bowen and C. Series, Markov maps associated with Fuchsian groups. *Inst. Hautes Études Sci. Publ. Math.* **50** (1979), 153–170 Zbl 0439.30033 MR 556585

[11] A. I. Bufetov, Ergodic theorems for actions of several mappings. *Russian Math. Surveys* **54** (1999), no. 4, 835–836 Zbl 0981.28015 MR 1741287

[12] A. I. Bufetov, Operator ergodic theorems for actions of free semigroups and groups. *Funct. Anal. Appl.* **34** (2000), no. 4, 239–251 Zbl 0983.22006 MR 1818281

[13] A. I. Bufetov, Markov averaging and ergodic theorems for several operators. In *Topology, ergodic theory, real algebraic geometry*, pp. 39–50, Amer. Math. Soc. Transl. Ser. 2 202, American Mathematical Society, Providence, RI, 2001 Zbl 1002.47004 MR 1819180

[14] A. I. Bufetov, Convergence of spherical averages for actions of free groups. *Ann. of Math. (2)* **155** (2002), no. 3, 929–944 Zbl 1028.37001 MR 1923970

[15] A. I. Bufetov, M. Khristoforov, and A. Klimenko, Cesàro convergence of spherical averages for measure-preserving actions of Markov semigroups and groups. *Int. Math. Res. Not. IMRN* (2012), no. 21, 4797–4829 Zbl 1267.22003 MR 2993436

[16] A. I. Bufetov and A. Klimenko, Maximal inequality and ergodic theorems for Markov groups. *Proc. Steklov Inst. Math.* **277** (2012), no. 1, 27–42 Zbl 1302.47068 MR 3052262

[17] A. I. Bufetov and A. Klimenko, On Markov operators and ergodic theorems for group actions. *European J. Combin.* **33** (2012), no. 7, 1427–1443 Zbl 1281.37004 MR 2923460

[18] A. I. Bufetov, A. Klimenko, and C. Series, Convergence of spherical averages for actions of Fuchsian groups. 2018, arXiv:1805.11743v1

[19] A. I. Bufetov, A. Klimenko, and C. Series, A symmetric Markov coding and the ergodic theorem for actions of Fuchsian groups. *Math. Res. Rep.* **1** (2020), 5–14 Zbl 07544434 MR 4387447

[20] A. I. Bufetov and C. Series, A pointwise ergodic theorem for Fuchsian groups. *Math. Proc. Cambridge Philos. Soc.* **151** (2011), no. 1, 145–159 Zbl 1219.22008 MR 2801319

[21] D. Calegari and K. Fujiwara, Combable functions, quasimorphisms, and the central limit theorem. *Ergodic Theory Dynam. Systems* **30** (2010), no. 5, 1343–1369 Zbl 1217.37025 MR 2718897

[22] J. W. Cannon, The combinatorial structure of cocompact discrete hyperbolic groups. *Geom. Dedicata* **16** (1984), no. 2, 123–148 Zbl 0606.57003 MR 758901

[23] W. J. Floyd and S. P. Plotnick, Growth functions on Fuchsian groups and the Euler characteristic. *Invent. Math.* **88** (1987), no. 1, 1–29 Zbl 0608.20036 MR 877003

[24] K. Fujiwara and A. Nevo, Maximal and pointwise ergodic theorems for word-hyperbolic groups. *Ergodic Theory Dynam. Systems* **18** (1998), no. 4, 843–858 Zbl 0919.22002 MR 1645314

[25] É. Ghys and P. de la Harpe, Espaces métriques hyperboliques. In *Sur les groupes hyperboliques d'après Mikhael Gromov (Bern, 1988)*, pp. 27–45, Progr. Math. 83, Birkhäuser, Boston, MA, 1990 Zbl 0731.20025 MR 1086650

[26] E. Glasner, *Ergodic theory via joinings*. Math. Surveys Monogr. 101, American Mathematical Society, Providence, RI, 2003 Zbl 1038.37002 MR 1958753

[27] A. Gorodnik and A. Nevo, *The ergodic theory of lattice subgroups*. Ann. of Math. Stud. 172, Princeton University Press, Princeton, NJ, 2010 Zbl 1186.37004 MR 2573139

[28] R. I. Grigorchuk, Pointwise ergodic theorems for actions of free groups. In *Proc. Tambov Workshop in the Theory of Functions*, 1986

[29] R. I. Grigorchuk, Ergodic theorems for actions of free groups and free semigroups. *Math. Notes* **65** (1999), no. 5–6, 654–657 Zbl 0957.22006 MR 1716245

[30] R. I. Grigorchuk, An ergodic theorem for actions of a free semigroup. *Proc. Steklov Inst. Math.* **231** (2000), no. 4(231), 113–127 Zbl 1172.37303 MR 1841754

[31] M. Gromov, Hyperbolic groups. In *Essays in group theory*, pp. 75–263, Math. Sci. Res. Inst. Publ. 8, Springer, New York, 1987 Zbl 0634.20015 MR 919829

[32] Y. Guivarc'h, Généralisation d'un théorème de von Neumann. *C. R. Acad. Sci. Paris Sér. A-B* **268** (1969), 1020–1023 Zbl 0176.11703 MR 251191

[33] V. A. Kaimanovich, Measure-theoretic boundaries of Markov chains, 0-2 laws and entropy. In *Harmonic analysis and discrete potential theory (Frascati, 1991)*, pp. 145–180, Plenum, New York, 1992 MR 1222456

[34] J. G. Kemeny and J. L. Snell, *Finite Markov chains*. Undergrad. Texts Math., Springer, New York-Heidelberg, 1976 Zbl 0328.60035 MR 0410929

[35] A. Klenke, *Probability theory*. Universitext, Springer, London, 2008 Zbl 1141.60001 MR 2372119

[36] G. A. Margulis, A. Nevo, and E. M. Stein, Analogs of Wiener's ergodic theorems for semisimple Lie groups. II. *Duke Math. J.* **103** (2000), no. 2, 233–259 Zbl 0978.22006 MR 1760627

[37] A. Nevo, Harmonic analysis and pointwise ergodic theorems for noncommuting transformations. *J. Amer. Math. Soc.* **7** (1994), no. 4, 875–902 Zbl 0812.22005 MR 1266737

[38] A. Nevo, Pointwise ergodic theorems for actions of groups. In *Handbook of dynamical systems. Vol. 1B*, pp. 871–982, Elsevier, Amsterdam, 2006 Zbl 1130.37310 MR 2186253

[39] A. Nevo and E. M. Stein, A generalization of Birkhoff's pointwise ergodic theorem. *Acta Math.* **173** (1994), no. 1, 135–154 Zbl 0837.22003 MR 1294672

[40] A. Nevo and E. M. Stein, Analogs of Wiener's ergodic theorems for semisimple groups. I. *Ann. of Math. (2)* **145** (1997), no. 3, 565–595 Zbl 0884.43004 MR 1454704

[41] D. Ornstein, On the pointwise behavior of iterates of a self-adjoint operator. *J. Math. Mech.* **18** (1968/1969), 473–477 Zbl 0182.47103 MR 0236354

[42] V. I. Oseledec, Markov chains, skew products and ergodic theorems for "general" dynamic systems. *Th. Prob. App.* **10** (1965), 551–557 MR 0189123

[43] M. Pollicott and R. Sharp, Ergodic theorems for actions of hyperbolic groups. *Proc. Amer. Math. Soc.* **141** (2013), no. 5, 1749–1757 Zbl 1266.28008 MR 3020860

[44] V. A. Rokhlin and Ya. G. Sinai, Construction and properties of invariant measurable partitions. *Soviet Math. Dokl.* **2** (1961), 1611–1614. Zbl 0161.34301

[45] G.-C. Rota, An "Alternierende Verfahren" for general positive operators. *Bull. Amer. Math. Soc.* **68** (1962), 95–102 Zbl 0116.10403 MR 133847

[46] C. Series, The infinite word problem and limit sets in Fuchsian groups. *Ergodic Theory Dynam. Systems* **1** (1981), no. 3, 337–360 Zbl 0483.30029 MR 662473

[47] C. Series, The growth function of a Fuchsian group. In *Papers presented to E. C. Zeeman*, pp. 281–291, University of Warwick, 1988

[48] C. Series, Geometrical methods of symbolic coding. In *Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989)*, pp. 125–151, Oxford Sci. Publ., Oxford University Press, New York, 1991 Zbl 0752.30026 MR 1130175

[49] T. Tao, Failure of the L^1 pointwise and maximal ergodic theorems for the free group. *Forum Math. Sigma* **3** (2015), article no. e27 Zbl 1359.37010 MR 3482275

[50] M. Wroten, The eventual Gaussian distribution of self-intersection numbers on closed surfaces. 2014, arXiv:1405.7951

[51] A. Zygmund, *Trigonometric series: Vols. I, II*. Cambridge University Press, London-New York, 1968 MR 0236587

Received 11 March 2021.

Alexander I. Bufetov
Aix–Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 39 rue F. Joliot Curie, Marseille, France; and Steklov Mathematical Institute of Russian Academy of Sciences, Gubkina str. 8, 119991, Moscow, Russia; bufetov@mi-ras.ru

Alexey Klimenko
Steklov Mathematical Institute of Russian Academy of Sciences, Gubkina str. 8, 119991, Moscow; and National Research University Higher School of Economics, Usacheva str. 6, 119048, Moscow, Russia; klimenko@mi-ras.ru

Caroline Series
Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK;
c.m.series@warwick.ac.uk