Mathematics
Research
Reports

Alexander I. Bufetov, Alexey Klimenko, Caroline Series

A symmetric Markov coding and the ergodic theorem for actions of Fuchsian Groups

Volume 1 (2020), p. 5-14.

<http://mrr.centre-mersenne.org/item/MRR_2020__1__5_0>

Mathematics Research Reports
Volume 1 (2020), 5–14

A symmetric Markov coding and the ergodic theorem for actions of Fuchsian Groups

Alexander I. Bufetov, Alexey Klimenko* & Caroline Series

(Recommended by Henk Bruin)

Abstract. The main result of this note is the pointwise convergence of spherical averages for measure-preserving actions of Fuchsian groups. The proof relies on a new self-inverse Markovian symbolic coding for Fuchsian groups and the method of Markov operators.

1. 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

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

The object of this note is to announce a new pointwise convergence result for the averages $\mathbf{S}_n(f)$ in the case in which G is a Fuchsian group and G_0 is a certain class of geometric generators as described below.

Let $\mathscr{R}$ be a fundamental domain for a Fuchsian group G, possibly with vertices or edges on the boundary of hyperbolic disc $\mathbb{D}$, and let $\mathbf{T}_{\mathscr{R}} = \{g\mathscr{R} : g \in G\}$ be the corresponding tessellation of $\mathbb{D}$. We say that $\mathscr{R}$ has *even corners* if the geodesic extension of every side of $\mathscr{R}$ is entirely contained in $\partial \mathbf{T}_{\mathscr{R}}$, that is, in the union of boundaries of all domains $g\mathscr{R} \in \mathbf{T}_{\mathscr{R}}$. Let $v \in \mathbb{D}$ be a vertex of $\mathbf{T}_{\mathscr{R}}$. If $\mathscr{R}$ has even corners, then the boundary of $\mathbf{T}_{\mathscr{R}}$ in a small neighbourhood of v consists of $n = n(v)$ geodesic segments intersecting at v and dividing our neighbourhood into $2n(v)$ sectors. Let $N(\mathscr{R})$ denote the number of sides of $\mathscr{R}$ inside $\mathbb{D}$. Here a side is understood to have a nonzero length; we allow that G_0 has order two elliptic elements in which case the fixed point is not counted as a vertex of $\mathscr{R}$. We need the following assumption on $\mathscr{R}$.

Assumption 1. *$\mathscr{R}$ has even corners and either $N(\mathscr{R}) \geq 5$ or*

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

Received November 18, 2018; accepted December 29, 2018.
2020 *Mathematics Subject Classification.* 20H10, 22D40, 37A30.
Keywords. Ergodic theorem, Fuchsian group, Markov coding, Markov operator, spherical averages.
* Corresponding author.

Suppose further that $\mathscr{R}$ is endowed with a set of side-pairing transformations $G_0 \subset G$ which identify sides of $\mathscr{R}$, mapping $\mathscr{R}$ to the domains of $\mathbf{T}_{\mathscr{R}}$ having a common side with $\mathscr{R}$. As is well known, G_0 is a symmetric set of generators for G. Our main result is the following:

Theorem 1.1. *Let G be a nonelementary Fuchsian group G and suppose it has a fundamental domain $\mathscr{R}$ satisfying Assumption 1, together with a symmetric set of side pairing transformations G_0 as above. Let G act on a Lebesgue probability space (X,μ) by measure-preserving transformations. Denote by $\mathscr{I}_{G_0^2}$ the σ-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*

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

where $L\log L(X,\mu) = \{f \in L^1 : \int |f| \log^+ |f| d\mu < \infty\}$ and $\mathbb{E}(f|\mathscr{I}_{G_0^2})$ is the expectation of f relative to the σ-algebra $\mathscr{I}_{G_0^2}$.

The condition that $\mathscr{R}$ have 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 $\mathscr{R}$. Thus Theorem 1.1 extends immediately to any presentation of a Fuchsian group for which one can find a deformed group G' which has a fundamental domain $\mathscr{R}'$ with the same pattern of sides and side-pairings and even corners, see [11, 4] and [18] for a detailed discussion.

The need to restrict to spheres $S(2n)$ 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. This appears to indicate that the condition on all relators having even length, as is implied by the even corner condition, may be essential.

A complete proof of Theorem 1.1 appears in [10] and we hope to publish a somewhat simplified version in due course.

A convenient method suggested by R. I. Grigorchuk [13], J.-P. Thouvenot (oral communication), and in [5] for proving ergodic theorems for actions of free semigroups and groups is to associate to the group action a Markov operator on a suitable function space. In [7] pointwise convergence is proved for Markovian spherical averages under the additional assumption that the Markov chain be reversible. The key step in [7] is the triviality of the tail σ-algebra for the corresponding Markov operator; this is proved using Rota's "Alternierende Verfahren" [16], that is to say, martingale convergence. The reduction of powers of the Markov operator to Rota's "Alternierende Verfahren" in [7] essentially relies on the reversibility of the Markov chain.

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 [12]. Another result in this direction was obtained in [2]; it states the mean convergence for analogues of spherical averages for an arbitrary Markov chain satisfying very mild conditions. It is not known whether a similar result holds for pointwise convergence in general.

Cesàro convergence of spherical averages for all measure-preserving actions of Markov semigroups, and, in particular, Gromov hyperbolic groups, was established in [8]; earlier partial results were obtained in [6]. For Fuchsian groups the Cesàro convergence is proven earlier in [11] using the Bowen–Series Markovian coding [4], see also [1, 17, 18]. In the special case of hyperbolic groups a shorter proof of the Cesàro convergence was later given by Pollicott and Sharp [15]. Using the method of amenable equivalence relations, L. Bowen and Nevo [3] established ergodic theorems for "spherical shells" in Gromov hyperbolic groups. For further background see, e. g., the surveys [14, 9].

2. The method

To establish Theorem 1.1, we develop the approach from [7] for free groups based on the reduction of convergence for spherical averages to convergence of the powers of a corresponding Markov operator. Recall that each element in a free group F_k on a symmetric set of generators $F_k^0 = \{a_1, \bar{a}_1, \ldots, a_k, \bar{a}_k\}$ (where $\bar{a}_i = a_i^{-1}$) can be uniquely represented as a shortest word in the generators, and that shortest words are precisely the reduced sequences of generators, namely those in which no generator is immediately followed by its inverse. Thus elements in the group can be uniquely characterised as the admissible finite sequences in a Markov chain on the state space $\Xi = F_k^0$ in which all transitions are admissible except for those of the form $g \to \bar{g}, g \in F_k^0$.

Using the Parry measure on the associated shift space, the method of [7] is to construct an associated Markov operator P on a suitable space of functions on $\Xi \times X$, and reduce the convergence of the averages $\mathbf{S}_{2n}(f)$ to that of powers of P. To prove this latter convergence, a crucial point in the argument is the relation between P and its adjoint P^*. Namely, there is a unitary Markov involution $U = U^{-1} = U^*$ such that $P^* = UPU$. This relation between P and P^* stems from the symmetry or self-inverse property of the underlying Markov chain, that is, if a sequence $j_0 \to \cdots \to j_k$ corresponding to an element $g \in F_k$ is admissible, then so is its inverse $\iota(j_k) \to \cdots \to \iota(j_0)$, where ι is the involution which sends a generator to its inverse. This is, of course, nothing other than the observation that the inverse of a reduced word is itself reduced.

The problem with extending the method of [7] to more general Fuchsian groups is that the Bowen–Series coding of [4], which was used to establish Cesàro convergence of spherical averages for Fuchsian groups in [11], is not symmetric in the above sense. This difficulty stems from the fact that in general shortest representations of a group element are no longer unique. Indeed to make a unique choice of shortest word, one has to define a 'direction', clockwise or anticlockwise, of travel around each vertex v of the tessellation $\mathbf{T}_{\mathscr{R}}$. The inverse of such sequence necessarily travels in the opposite direction around v, in contradiction to the previous choice.

The central ingredient of the proof of Theorem 1.1 is a new method of coding for Fuchsian groups which avoids this difficulty and which has the required symmetry. This coding is constructed using a variant of a coding introduced by Matthew Wroten [19]. Wroten's idea is to code all possible representations of a group element as a shortest word simultaneously. Thus a finite admissible cylinder in our Markov shift represents all possible shortest paths between two regions of $\mathbf{T}_{\mathscr{R}}$, and the state space is endowed with an involution which is directly derived from inverting all of these paths together, thus avoiding the need for choices as above. Using this involution, one obtains an associated unitary involution U with the above property $P^* = UPU$.

Given this construction, the proof of convergence for spherical averages follows the outline of the proof from [7]. This is a general result for Markov operators with the above property, together with two further assumptions about the corresponding Markov operator. First, the equations $P^n\varphi = \varphi$ and $(P^*)^m P^m \varphi = \varphi$ should have only constant solutions in L^2. The second assumption is an inequality between the operators UP^{2n-a} and $(P^*)^{n+b} P^{n+b'}$. For free groups in [7] it takes the form $UP^{2n-1}\varphi \le c(P^*)^n P^n \varphi$ with some $c > 0$, while in our case it says that there exist constants $a, b \in \mathbb{N}$, $c > 0$, a sequence (α_n) of positive numbers with $\sum_n \alpha_n < \infty$, and a sequence of operators A_n that map the cone of nonnegative functions to itself and with $\|A_n\|_{L^p} \le \alpha_n$ for any $p \in [1,\infty]$ such that

$$UP^{2n-a}\varphi \le c\sum_{j=-b}^{b}(P^*)^n P^{n+2j}\varphi + A_n\varphi. \tag{2.1}$$

The first assumptions about constant solutions are not very hard to check given our coding; see Lemma 5.1 below for further discussion of the inequality (2.1).

3. The new coding

We now describe the coding in more detail. Consider the adjacency graph for the fundamental domains (that is, copies $g\mathscr{R}$ of the original fundamental domain $\mathscr{R}$) of $\mathbf{T}_{\mathscr{R}}$, which coincides with the Cayley graph for the group G with respect to the generating set G_0. (Two domains are taken to be adjacent if they meet along some side of nonzero length; domains which touch only at a vertex are not considered adjacent.) Note also that the shortest representations $g = g_1 \dots g_n$, $g_i \in G_0$ of an element $g \in G$ bijectively correspond to the shortest paths from $\mathscr{R}$ to $g\mathscr{R}$, namely, each product $g_1 \dots g_n$ corresponds to the path of domains $\mathscr{R}, g_1\mathscr{R}, g_1 g_2\mathscr{R}, \dots, g\mathscr{R}$. (Note that the multiplication by successive group elements here is on the right.)

We define the *thickened path* $\underline{\mathscr{L}}$ from $g\mathscr{R}$ to $g'\mathscr{R}$ to be the union of all those domains in $\mathbf{T}_{\mathscr{R}}$ which are traversed by some shortest path of domains going from $g\mathscr{R}$ to $g'\mathscr{R}$. Every domain $h\mathscr{R} \in \underline{\mathscr{L}}$ is endowed with an *index*, namely the distance in the Cayley graph from $g\mathscr{R}$ to $h\mathscr{R}$, equivalently the length of the word $g^{-1}h$. We will refer to the set of all domains in $\underline{\mathscr{L}}$ with index k as a *level* of the thickened path and denote it $[\underline{\mathscr{L}}]_k$. As we will discuss below, under Assumption 1 every level contains at most two domains.

Up to the group action, we may take the starting domain of any thickened path $\underline{\mathscr{L}} = ([\underline{\mathscr{L}}]_0, \dots, [\underline{\mathscr{L}}]_n)$ to be $\mathscr{R}$. The path is then uniquely defined by the sequence of arrangements $\hat{\jmath}_i$ of pairs $([\underline{\mathscr{L}}]_i, [\underline{\mathscr{L}}]_{i+1})$ of its adjacent levels. It is not hard to see that there are only finitely many different possible arrangements. One can think of this as a process of building up the thickened path step by step; at each step, there are only finitely ways to extend what is already built as a thickened path. Modulo a minor adjustment, we take the states of our Markov chain to be the set of possible arrangements of these pairs of levels, which we denote by $\widehat{\Xi}$. Only certain transitions between consecutive pairs of levels, that is between certain arrangements, are possible; this will define the transition matrix for our chain. Our coding will generate these sequences of arrangements in the following way.

Lemma 3.1. *There exists a Markov chain with a set of states Ξ, subsets $\Xi_S, \Xi_F \subset \Xi$ and a map $\pi\colon \Xi \to \widehat{\Xi}$ such that the following holds. Consider the set*

$$\mathscr{P}_n^{S,F} = \{j = (j_0 \to \dots \to j_{n-1}) \text{ is an admissible sequence, } j_0 \in \Xi_S,\, j_{n-1} \in \Xi_F\}.$$

Then the map $(j_0 \to j_1 \to \dots \to j_{n-1}) \mapsto (\pi(j_0), \pi(j_1), \dots, \pi(j_{n-1}))$ is a bijection from $\mathscr{P}_n^{S,F}$ to the set of sequences $(\hat{\jmath}_0, \dots, \hat{\jmath}_{n-1})$ that are sequences of arrangements for the adjacent levels for all shortest paths of length n.

The states Ξ differ from the possible arrangements $\widehat{\Xi}$ only in a minor way as explained in the next section. The sets Ξ_S, Ξ_F denote possible start and end states: not all arrangements can be the initial or final level in a thickened path from one region to another, as will also become clear in the next section.

Since each element of G is uniquely represented by the thickened path from $\mathscr{R}$ to $g\mathscr{R}$ we obtain:

Corollary 1. *The above map* $(j_0 \to j_1 \to \cdots \to j_{n-1}) \mapsto (\pi(j_0), \pi(j_1), \ldots, \pi(j_{n-1}))$ *is a bijection between admissible sequences of* Ξ *starting and ending in* Ξ_S, Ξ_F *respectively, and the elements of* G *of length* n.

Observe that the set $S(g,h)$ of all possible shortest paths

$$\beta = g\mathscr{R}, gg_1\mathscr{R}, gg_1g_2\mathscr{R}, \ldots, gg_1 \ldots g_n\mathscr{R}$$

from $g\mathscr{R}$ to $h\mathscr{R}, h = gg_1 \ldots g_n$ is clearly self-inverse, in the sense that if $\beta \in S(g,h)$, then the same path read backwards, namely $\beta^{-1} = gg_1 \ldots g_n\mathscr{R}, \ldots, gg_1\mathscr{R}, g\mathscr{R} \in S(h,g)$, is represented by the word $\bar{g}_n \ldots \bar{g}_1$. This gives rise to an involution $\iota : \Xi \to \Xi$ which inverts shortest paths, interchanges Ξ_S and Ξ_F, and from which we derive the definition of a unitary involution U with the crucial required property $P^* = UPU$.

4. The arrangements

We now give some further details of the possible arrangements of levels and the transitions allowed by our coding, also explaining the distinction between Ξ and $\widehat{\Xi}$.

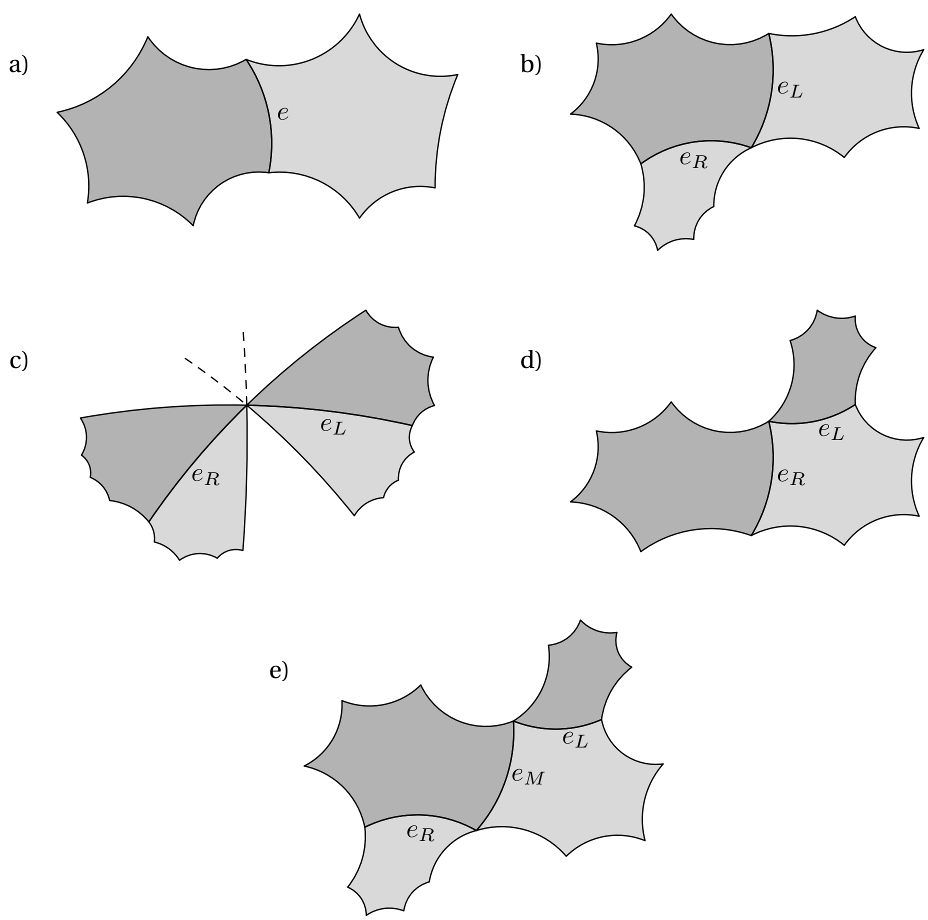

FIGURE 1. Configurations for pairs of adjacent levels:
a) $A(e)$, b) $B(e_L, e_R)$, c) $C_2(e_L, e_R)$, d) $D(e_L, e_R)$, e) $E_R(e_L, e_M, e_R)$.
The dark grey and the light grey domains represent the 'past' and the 'future' levels respectively.

The hyperbolicity of the tessellation $\mathbf{T}_{\mathscr{R}}$ implies that any thickened path has uniformly bounded width, independent of the distance between the two domains it joins. More precisely, there is a uniform bound on the number of domains in a given level. A crucial consequence of the property of even corners is that thickened paths are in fact much narrower than this: as can be shown by the methods of [1], every level contains either one or two domains, which in the latter case share a common vertex v. In this latter case the two sectors of $\mathbf{T}_{\mathscr{R}}$ around v occupied by these two domains separate the remaining sectors into two components, each of which contains an odd number of sectors. This leads to five possible types of arrangements for a pair of adjacent levels in thickened paths. This is shown in Figure 1, in which the domains shaded dark grey have level k and those shaded light grey are level $k+1$ for some k. We think of the thickened path as progressing from the dark grey to the light grey level, so refer to these as the 'past' and 'future' domains respectively. Thus for example in Figure 1(c), which is shown with $n(v)=4$, the two 'future' domains shaded light grey separate the set of all sectors around v into one collection of three domains in the past and one of one domain in the future.

To encode these configurations we endow each side s of $\mathbf{T}_{\mathscr{R}}$ with two labels. Namely, if $g\mathscr{R}$ and $h\mathscr{R}$ are adjacent regions with s their common side, then the group element $g^{-1}h$ which carries $g\mathscr{R}$ into $h\mathscr{R}$ is by definition an element of the generating set G_0, as is $h^{-1}g$ which carries $h\mathscr{R}$ into $g\mathscr{R}$. (The labelling is chosen to be G-invariant, so that the element which carries $g\mathscr{R}$ into $h\mathscr{R}$ must be the same as that which carries $\mathscr{R}$ into $g^{-1}h\mathscr{R}$.) We place the label $g^{-1}h$ on the $h\mathscr{R}$-side of s and the label $h^{-1}g$ on its $g\mathscr{R}$-side. This labelling is chosen so that the path $\beta = g\mathscr{R}, gg_1\mathscr{R}, gg_1g_2\mathscr{R},\dots, gg_1\dots g_n\mathscr{R}$ cuts in order the sides labelled inside the successive 'future' domains by $g_1, g_2,\dots, g_n$, in accordance with the conventions of [4] and [1].

The arrangements shown in Figures 1(a),(b),(d) which contain only one domain in the 'past' or 'future' level (types (a,b), and (a,d) respectively) are uniquely defined by the labels on the sides separating these two levels. We denote these configurations as $A(e)$, $B(e_L, e_R)$, $D(e_L, e_R)$, where $e, e_L, e_R \in G_0$ are the labels on the future level side as shown.

In the configuration shown in Figure 1 (c) all four of the 'past' and 'future' domains have a common vertex v. The other sectors at v are separated into two connected components each with an odd number of sectors. If the component adjacent to the 'past' domains contains $2k-1$ sectors, then we denote this configuration as type $C_k(e_L, e_R)$, $1 \le k < n(v)-1$, again with e_L, e_R as shown.

Finally, in the configuration shown in Figure 1 (e) the 'right past' domain has common sides with both 'future' domains, while the 'left past' domain has a common side only with the 'left future' domain. We denote this case as $E_R(e_L, e_M, e_R)$. The case symmetric with respect to the left and right is denoted by $E_L(e_L, e_M, e_R)$.

Note that the labels of the configurations should satisfy some constraints stemming from the combinatorics of the partition. For example, if we know the type of configuration (B, C_k, D, E_L, or E_R) and one of the labels (say, e_L), then we can uniquely determine all other labels. Nevertheless, we retain this excessive notation due to its symmetry. Let $\widehat{\Xi}$ be the set of all arrangements $A(e)$, $B(e_L, e_R)$, $C_k(e_L, e_R)$, $D(e_L, e_R)$, $E_{L,R}(e_L, e_M, e_R)$ satisfying these conditions on labels.

Observe that the configurations from $\widehat{\Xi}$ cannot be regarded as the states of a Markov chain which generate all possible thickened paths. Indeed, consider a vertex v with $n = n(v) \ge 3$. Let $s_0, s_1, \dots, s_{2n-1}$ be the sides incident to v, in the anticlockwise ordering (we continue this numbering modulo $2n$), and let $\mathscr{R}_j$ be the domain between s_j and s_{j+1}. Finally, let e_j be the label on $\mathscr{R}_j$-side of s_j. Then every pair $(A(e_j), A(e_{j+1}))$ encodes a thickened path $(\mathscr{R}_{j-1}, \mathscr{R}_j, \mathscr{R}_{j+1})$ containing three sectors around v, so the transition $A(e_j) \to A(e_{j+1})$ should be allowed. On the other hand, the sequence $(A(e_j), \dots, A(e_{j+n-1}))$

cannot appear in a thickened path, since such thickened path would have to contain a sequence of levels of the form $(h\mathscr{R}_{j-1}, h\mathscr{R}_j, \dots, h\mathscr{R}_{j+n-1})$. But one can reach $h\mathscr{R}_{j+n-1}$ from $h\mathscr{R}_{j-1}$ in the same number of steps following through the sequence $(h\mathscr{R}_{j-1}, h\mathscr{R}_{j-2}, \dots, h\mathscr{R}_{j-n-1})$, so these domains should belong to the thickened path as well, meaning that the levels cannot all be of type A.

To circumvent this difficulty we endow the arrangements of type A with some additional information as follows. For any label e consider a side s_e of $\mathscr{R}$ with outside label e. Let $v_L(e), v_R(e)$ be the left and right ends of this side when looking from the inside of $\mathscr{R}$; if the left (respectively, right) end of s_e belongs to $\partial\mathbb{D}$, then $v_L(e)$ (respectively, $v_R(e)$) is undefined.

Consider a thickened path $\underline{\mathscr{S}}$. If $([\underline{\mathscr{S}}]_k, [\underline{\mathscr{S}}]_{k+1})$ forms a configuration $A(e)$, with $s_k = [\underline{\mathscr{S}}]_k \cap [\underline{\mathscr{S}}]_{k+1}$, we define numbers $i_{\pm,L}, i_{\pm,R}$ which record the number of 'past' and 'future' domains in $\underline{\mathscr{S}}$ which meet at $v_L(s_k), v_R(s_k)$ respectively. Precisely, $i_{-,L}$ is the number of levels $[\underline{\mathscr{S}}]_m$ which contain $v_L(s_k)$ for $m \le k$, while $i_{+,L}$ is the number of levels $[\underline{\mathscr{S}}]_m$ which contain $v_L(s_k)$ for $m \ge k+1$. Define $i_{\pm,R}$ similarly. If the vertex $v_J(s_k), J \in \{L, R\}$ is not defined, we set $i_{\pm,J} = 1$.

Note that it is not possible to have $i_{+,L} > 1$ and $i_{+,R} > 1$ (or $i_{-,L} > 1$ and $i_{-,R} > 1$) simultaneously: these conditions mean that both sides of $[\underline{\mathscr{S}}]_{k+1}$ adjacent to s_k are incident to some domain from $[\underline{\mathscr{S}}]_{k+2}$, hence these two domains from $[\underline{\mathscr{S}}]_{k+2}$ do not share a common vertex. The same argument as above shows that $\underline{\mathscr{S}}$ cannot contain more than $n(v_J(s_k))$ domains adjacent to $v_J(s_k)$, $J = L, R$. Therefore, $i_{-,J} + i_{+,J} \le n(v_J(s_k))$. This yields the following possibilities for these four indices.

- $A_0(e)$: 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 A_L.
- $A_{LR}[i_-, i_+](e)$: here $i_{-,L} = i_-$, $i_{+,R} = i_+$, and $i_{+,L} = i_{-,R} = 1$. The conditions on the indices $i_\pm$ are $2 \le i_- \le n(v_L(e)) - 1$, $2 \le i_+ \le n(v_R(e)) - 1$.
- $A_{RL}[i_-, i_+](e)$: symmetric to A_{LR}.

We denote by Ξ the set of all states $A_{\dots}[\dots](e), e \in G_0$ enumerated above, and all configurations of types B, C, D, and E from $\widehat{\Xi}$. This is the set of states of our Markov chain. The projection $\pi\colon \Xi \to \widehat{\Xi}$ simply removes the added information for the states of type A. The sets Ξ_S and Ξ_F are now defined as follows:

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

It will be seen that these are indeed the only possibilities for the initial or final level of a thickened path joining two domains, as such paths necessarily contain only one domain. Thus they have to be treated as initial and final states in our representation of all elements of G.

The set of transitions in our Markov chain arises from two classes of constraints. First, if we have a transition $j \to j'$, then the future level of j should be the same as the past level of j' up to the group action. Moreover, if $(\mathscr{S}_-, \mathscr{S}_+)$ and $(\mathscr{S}_+, \mathscr{S}_{++})$ represent $\pi(j)$ and $\pi(j')$ respectively, then $\mathscr{S}_- \cap \mathscr{S}_+$ and $\mathscr{S}_+ \cap \mathscr{S}_{++}$ cannot have a common side. For example, let $j = B(e_L, e_R)$, let $(\mathscr{S}_-, \mathscr{S}_+)$ be its representation, and let v be the common vertex of two domains in $\mathscr{S}_+$. Then, provided $n(v) > 2$, the next state should be $C_1(\dots)$, and the next states are $C_2(\dots), \dots, C_{n(v)-2}(\dots)$.

Second, there are constraints ensuring that the boundary of a thickened path is convex. Such convexity is an elaboration of the example of a cycle round a vertex in the previous section. In fact it can be shown using the methods of [1] or [10] that the thickened path joining domains $g\mathscr{R}$ to $h\mathscr{R}$ is the smallest union of domains which is geodesically convex and which contains any given shortest path between the same domains. This is closely related to another important consequence of the property of even corners, namely that every hyperbolic geodesic joining a point in $g\mathscr{R}$ to one in $h\mathscr{R}$ cuts through precisely $|g^{-1}h|+1$ domains, see [1].

Constraints coming from this convexity condition are mostly conditions on the indices of A-states. For example, if $j = A_L[i_-, i_+](e)$ with $i_+ \geq 3$, then $j' = A_L[i_- + 1, i_+ - 1](e')$. Similarly, if $i_+ = 2$ there are three possible transitions: $j' = A_L[i_- + 1, 1](e')$, $j' = A_{LR}[i_- + 1, i'_+](e')$ or $j' = B(e', e'')$. One special case for this class of constraints is the $D \to B$ transition: if $\mathscr{S}_- \cap \mathscr{S}_+$ and $\mathscr{S}_+ \cap \mathscr{S}_{++}$ have a common vertex u, then one needs to require $n(u) > 2$.

The required involution ι of the state space Ξ comes from reversing the direction of paths and hence arrangements in the obvious way. Thus for example, $\iota(A_0(e)) = A_0(\bar{e})$; $\iota(A_L[i_-, i_+](e) = A_R[i_+, i_-](\bar{e})$ while $\iota(B(e_L, e_R)) = D(\bar{e}_R, \bar{e}_L)$ and $\iota(C_i(e_L, e_R) = C_{n(v)-i-1}(\bar{e}_R, \bar{e}_L)$. It is easy to check that an involution thus defined exchanges Ξ_S with Ξ_F.

It is clear that every thickened path is generated by this Markov chain in the way stated in Lemma 3.1. But the converse part of this lemma, that is, that every sequence of arrangements generated by this Markov chain corresponds to a thickened path, is among the main technical difficulties of the proof.

5. The final ingredients

Our first observation is that the operator P has to be defined relative to the action of a single element of G, while the states Ξ of our coding record arrangements in a thickened path. We get around this by projecting each state to the label of its left edge e_R, so that the thickened path is projected to the sequence of group elements which record the labels of the sides which meet its left boundary. Since it is only the product of these elements which is important in the sum $\mathbf{S}_n(f)$, this allows one to construct a well defined operator P whose powers record the required sums $\sum_{g\in S(n)} f \circ T_g$.

It remains to check the assumptions in the convergence theorem on Markov operators described above. The proofs that the associated operators P and P^*P have only constant solutions are fairly standard. The proof of the inequality (2.1) requires more work, in which the first step is to derive the following geometric interpretation.

Lemma 5.1. *There exist constants $\theta < 1$, $a, b \in \mathbb{N} \cup \{0\}$ such that the following holds. Let $\underline{i} = (i_0 \to \cdots \to i_{2n-1})$ be an admissible sequence of states and let $([\mathscr{R}]_0, \ldots, [\mathscr{R}]_{2n})$ be a sequence of levels representing $\underline{i}$. Then either $\underline{i}$ belongs to an exceptional set E_n, whose proportion in the set of all admissible sequences is at most θ^n, or the following holds.*

There exist a number $\beta \in \{-b, \ldots, b\}$ and admissible sequences $\underline{j} = (j_0 \to \cdots \to j_{n-1+a+\beta})$, $\underline{k} = (k_0 \to \cdots \to k_{n-1+a})$ such that $j_0 = i_0$, $j_{n-1+a+\beta} = k_{n-1+a}$, $k_0 = \iota(i_{2n-1})$, and $\underline{j}$ and $\underline{k}$ can be represented respectively by sequences $([\mathscr{Q}]_0, \ldots, [\mathscr{Q}]_{n+a+\beta})$ and $([\mathscr{S}]_0, \ldots, [\mathscr{S}]_{n+a})$ with $[\mathscr{Q}]_0 = [\mathscr{R}]_0$, $[\mathscr{S}]_0 = [\mathscr{R}]_{2n}$, and $[\mathscr{Q}]_{n+a+\beta} = [\mathscr{S}]_{n+a}$.

In other words, this lemma states that almost every segment of a thickened path can be included in a "triangle of thickened paths" with "zero angles" at all vertices. Two vertices are at the two ends $[\mathscr{R}]_0, [\mathscr{R}]_{2n}$ of the given sequence $\underline{i} = (i_0 \to \cdots \to i_{2n-1})$, with the third vertex $[\mathscr{Q}]_{n+a+\beta} = [\mathscr{S}]_{n+a}$ approximately half way between. The remainder terms $A_n\varphi$ in (2.1) are needed to account for the paths from the exceptional sets E_n. A short amount of experimentation will easily convince the reader of the meaning and veracity of this lemma in the special case of the free group F_2 with $a = 1, b = 0$ and $E_n = \varnothing$.

Acknowledgements

A.B.'s research is supported by the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme, grant 647133 (ICHAOS), by the Agence Nationale de Recherche, project ANR-18-CE40-0035, and by the Russian Foundation for Basic Research, grant 18-31-20031. A.K.'s research is partially supported by Laboratory of Dynamical Systems and Applications NRU HSE, of the Ministry of science and higher education of the RF grant ag. No. 075-15-2019-193, and by the Russian Foundation for Basic Research grants 18-31-20031 and 18-51-15010.

References

[1] J. S. Birman & C. Series, "Dehn's algorithm revisited, with applications to simple curves on surfaces", in *Combinatorial group theory and topology (Alta, Utah, 1984)*, Ann. of Math. Stud., vol. 111, Princeton Univ. Press, Princeton, NJ, 1987, p. 451-478.

[2] L. Bowen, A. Bufetov & O. Romaskevich, "Mean convergence of Markovian spherical averages for measure-preserving actions of the free group", *Geom. Dedicata* **181** (2016), p. 293-306.

[3] L. Bowen & A. Nevo, "Von Neumann and Birkhoff ergodic theorems for negatively curved groups", *Ann. Sci. Éc. Norm. Supér. (4)* **48** (2015), no. 5, p. 1113-1147.

[4] R. Bowen & C. Series, "Markov maps associated with Fuchsian groups", *Inst. Hautes Études Sci. Publ. Math.* (1979), no. 50, p. 153-170.

[5] A. I. Bufetov, "Operator ergodic theorems for actions of free semigroups and groups", *Funct. Anal. Appl* **34** (2000), no. 4, p. 239-251.

[6] A. I. Bufetov, "Markov averaging and ergodic theorems for several operators", in *Topology, ergodic theory, real algebraic geometry*, Amer. Math. Soc. Transl. Ser. 2, vol. 202, Amer. Math. Soc., Providence, RI, 2001, p. 39-50.

[7] ———, "Convergence of spherical averages for actions of free groups", *Ann. of Math. (2)* **155** (2002), no. 3, p. 929-944.

[8] A. I. Bufetov, M. Khristoforov & A. Klimenko, "Cesàro convergence of spherical averages for measure-preserving actions of Markov semigroups and groups", *Int. Math. Res. Not. IMRN* (2012), no. 21, p. 4797-4829.

[9] A. I. Bufetov & A. Klimenko, "On Markov operators and ergodic theorems for group actions", *European J. Combin.* **33** (2012), no. 7, p. 1427-1443.

[10] A. I. Bufetov, A. Klimenko & C. Series, "Convergence of spherical averages for actions of Fuchsian Groups", preprint, 2018, `https://arxiv.org/abs/1805.11743`.

[11] A. I. Bufetov & C. Series, "A pointwise ergodic theorem for Fuchsian groups", *Math. Proc. Cambridge Philos. Soc.* **151** (2011), no. 1, p. 145-159.

[12] K. Fujiwara & A. Nevo, "Maximal and pointwise ergodic theorems for word-hyperbolic groups", *Ergodic Theory Dynam. Systems* **18** (1998), no. 4, p. 843-858.

[13] R. I. Grigorchuk, "An ergodic theorem for actions of a free semigroup", *Proc. Steklov Inst. Math.* **231** (2000), no. 4, p. 113-127.

[14] A. Nevo, "Pointwise ergodic theorems for actions of groups", in *Handbook of dynamical systems. Vol. 1B*, Elsevier B. V., Amsterdam, 2006, p. 871-982.

[15] M. Pollicott & R. Sharp, "Ergodic theorems for actions of hyperbolic groups", *Proc. Amer. Math. Soc.* **141** (2013), no. 5, p. 1749-1757.

[16] G.-C. Rota, "An "Alternierende Verfahren" for general positive operators", *Bull. Amer. Math. Soc.* **68** (1962), p. 95-102.

[17] C. Series, "The infinite word problem and limit sets in Fuchsian groups", *Ergodic Theory Dynam. Systems* **1** (1981), no. 3, p. 337-360.

[18] ———, "Geometrical methods of symbolic coding", in *Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989)*, Oxford Sci. Publ., Oxford Univ. Press, New York, 1991, p. 125-151.

[19] M. Wroten, "The Eventual Gaussian Distribution for Self-Intersection Numbers on Closed Surfaces", preprint, 2014, `https://arxiv.org/abs/1405.7951`.

Alexander I. Bufetov: Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373,
39 rue F. Joliot Curie, 13453 Marseille Cedex 13, France; and
Steklov Mathematical Institute of Russian Academy of Sciences,
Gubkina str. 8, 119991, Moscow, Russia
E-mail: bufetov@mi-ras.ru

Alexey Klimenko: Steklov Mathematical Institute of Russian Academy of Sciences,
Gubkina str. 8, 119991, Moscow, Russia; and
National Research University Higher School of Economics, Usacheva str. 6, 119048, Moscow, Russia
E-mail: klimenko@mi-ras.ru

Caroline Series: Mathematics Institute, University of Warwick,
Coventry CV4 7AL, UK
E-mail: C.M.Series@warwick.ac.uk