THE HÖLDER PROPERTY FOR THE SPECTRUM OF TRANSLATION FLOWS IN GENUS TWO

ALEXANDER I. BUFETOV AND BORIS SOLOMYAK

Abstract. The main result of this paper, Theorem 1.1, establishes the H"older Property for the spectrum of generic translation flows corresponding to Abelian differentials with one zero of order two on surfaces of genus two.

1. Introduction and the Formulation of the main result.

Let $M$ be a compact orientable surface. To a holomorphic one-form $\omega$ on $M$ one can assign the corresponding vertical flow $h_t^+$ on $M$, i.e., the flow at unit speed along the leaves of the foliation $\mathcal{H}(\omega) = 0$. The vertical flow preserves the measure $m = i(\omega \wedge \overline{\omega})/2$, the area form induced by $\omega$. By a theorem of Katok [15], the flow $h_t^+$ is never mixing. The moduli space of abelian differentials carries a natural volume measure, called the Masur-Veech measure [19], [22]. For almost every Abelian differential with respect to the Masur-Veech measure, Masur [19] and Veech [22] independently and simultaneously proved that the flow $h_t^+$ is uniquely ergodic. Under additional assumptions on the combinatorics of the abelian differentials, weak mixing for almost all translation flows has been established by Veech in [23] and in full generality by Avila and Forni [2]. The spectrum of translation flows is therefore almost surely continuous and always has a singular component. No quantitative results have, however, previously been obtained about the spectral measure. Sinai [personal communication] posed the following

Problem. Find the local asymptotic for the spectral measure of translation flows.

The aim of this paper is to give first quantitative estimates on the spectrum of translation flows. Let $\mathcal{H}(2)$ be the moduli space of abelian differentials on a surface of genus 2 with one zero of order two. The natural smooth Masur-Veech measure on the stratum $\mathcal{H}(2)$ is denoted by $\mu_2$. Our main result is that for almost all abelian differentials in $\mathcal{H}(2)$, the spectral measures of Lipschitz functions with respect to the corresponding translation flows have the Hölder property. The spectral measure corresponding to a function $f$ is denoted by $\sigma_f$.

Theorem 1.1. There exists $\gamma > 0$ such that for $\mu_2$-almost every abelian differential $(M, \omega) \in \mathcal{H}(2)$ the following holds. For any Lipshitz function $f$ on $M$ and any $B > 0$ there exist constants $C = C(f, B), r_0 = r_0(f, B)$ such that for all $x \in [B^{-1}, B]$ and $r \in (0, r_0)$ we have

$$\sigma_f((x - r, x + r)) \leq C \cdot r^\gamma.$$
The proof uses the symbolic formalism of [8], namely, the representation of translation flows by flows along orbits of the asymptotic equivalence relation of a Markov compactum. The Hölder property is then reformulated as a statement on Diophantine approximation involving the incidence matrices of the graphs forming the Markov compactum that codes the translation flow. These incidence matrices are, in turn, realizations of a renormalization cocycle, isomorphic, under our symbolic coding, to the Kontsevich-Zorich cocycle over the Teichmüller flow on the moduli space of abelian differentials. By the Oseledec multiplicative ergodic theorem, the cocycle admits a decomposition into Oseledec subspaces corresponding to the distinct Lyapunov exponents. Our argument in this paper requires that the Kontsevich-Zorich cocycle admit two positive Lyapunov exponents and not have zero Lyapunov exponents; this is the reason why our present argument only works for the stratum $H(2)$.

We stress that our proof only works for the Masur-Veech smooth measure. This can informally be explained as follows. A translation flow can be represented as a suspension flow over an interval exchange transformation with a piecewise constant roof function. Take a measure on a stratum, invariant under the action of the Teichmüller flow. If one fixes an interval exchange transformation, then the invariant measure yields a conditional measure on the polyhedron of admissible tuples of heights of the rectangles. Our argument in its present form requires that this measure be absolutely continuous with respect to the natural Lebesgue measure; this property only holds for the Masur-Veech smooth measure.

We note that the Hölder exponent $\gamma > 0$ obtained in Theorem 1.1 can be given explicitly, but it is very small and certainly not sharp, so we do not pursue this. It is an interesting question to determine the sharp exponent. We can say more about the Hölder exponent at $x = 0$, see Remark 3.4.

The paper is organized as follows. In Section 2, we introduce the necessary setup of Markov compacta and Bratteli-Vershik (BV) transformations, as well the alternative, closely related framework, based on sequences of substitutions, or $S$-adic systems (see e.g. [8], [21] for further background). We will be working in this “symbolic” framework for most of the paper, only returning to translation flows in the last Section 9. Estimates of twisted Birkhoff sums and integrals in terms of matrix product are considered in Section 3, which builds on our paper [10]. The difference is that [10] was concerned with a single substitution, or equivalently, with a stationary Bratteli-Vershik diagram. In Section 4 we state the main theorem for random BV-systems, satisfying a certain list of conditions, among which the key condition is a uniform large deviations estimate. That theorem is proved in Sections 5-8. It should be emphasized that there are substantial technical difficulties in the transition from the stationary framework of [10] to the non-stationary setting of this paper. In the case of a single substitution matrix we could rely on estimates of the Vandermonde matrix, its determinant and its inverse. In this paper, we need similar estimates
for the cocycle matrices. The Oseledets Theorem controls norms of these matrices and angles between different subspaces only up to subexponential errors. It is precisely in order to control these errors that we need the assumption that only two Lyapunov exponents be positive. In Section 8 we use a variant of the “Erdős-Kahane argument” that had originated in the theory of Bernoulli convolution measures, see [11, 17]. Our argument in Section 8 builds on that of [10, Section 4], but again, the situation is much more subtle in the non-stationary case. Finally, in Section 9 we conclude the proof by explaining how the symbolic coding of translation flows gives rise to suspension flows over BV-maps and by checking all the conditions required. The key probabilistic condition is derived from a (slight generalization of) large deviation estimate for the Teichmüller flow in [7].

Acknowledgements. We are deeply grateful to G. Forni for helpful comments on the preliminary version of the paper.

A. Bufetov has been supported by the A*MIDEX project (no. ANR-11-IDEX-0001-02) funded by the “Investissements d’Avenir” French Government program, managed by the French National Research Agency (ANR). A. Bufetov has also been supported in part by the Grant MD-2859.2014.1 of the President of the Russian Federation, by the Programme “Dynamical systems and mathematical control theory” of the Presidium of the Russian Academy of Sciences, by the ANR under the project “VALET” of the Programme JCJC SIMI 1, and by the RFBR grants 11-01-00654, 12-01-31284, 12-01-33020, 13-01-12449.

B. Solomyak has been supported by NSF grants DMS-0968879 and DMS-1361424.

2. Preliminaries: Markov compacta, Bratteli-Vershik transformations, and substitutions

Let $\mathcal{G}$ be the set of all oriented graphs on $m$ vertices such that there is an edge starting at every vertex and an edge ending at every vertex (we allow loops and multiple edges). For an edge $e$ we denote by $I(e)$ and $F(e)$ the initial and final vertices of $e$ appropriately. Assume we are given a sequence $\{\Gamma_n\}_{n \in \mathbb{Z}}$ of graphs belonging to $\mathcal{G}$. To this sequence we associate the Markov compactum of paths in our sequence of graphs:

$$X = \{ \vec{r} = (e_n)_{n \in \mathbb{Z}} : e_n \in \mathcal{E} (\Gamma_n), \; F(e_{n+1}) = I(e_n) \}.$$ 

We will also need the one-sided Markov compactum $X_+$ (respectively $X_-$), defined in the same way with elements $(e_n)_{n \geq 1}$ (respectively $(e_n)_{n \leq 0}$). A one-sided sequence of graphs in $\mathcal{G}$ can also be considered as a Bratteli diagram of (finite) rank $m$. We then view its vertices as being arranged in levels $n \geq 0$, so that the graph $\Gamma_n$ connects the vertices of level $n$ to vertices of level $n - 1$. (Note that in some papers the direction of the edges is reversed.) Let $A_n(X) = A(\Gamma_n)$ be the incidence matrix of the graph $\Gamma_n$ given by the formula

$$A_{ij}(\Gamma) = \# \{ e \in \mathcal{E}(\Gamma) : I(e) = i, \; F(e) = j \}.$$
On the Markov compactum $X$ we define the “future tail” and “past tail” equivalence relations, in which two infinite paths are equivalent iff they agree from some point on (into the future or into the past).

There is a standard construction of telescoping (= aggregation): for any sequence $1 = n_0 < n_1 < n_2 < \cdots$ we “concatenate” the graphs $\Gamma_{n_j}, \ldots, \Gamma_{n_{j+1}-1}$ to obtain $\Gamma_j \in \mathcal{G}$.

**Standing Assumption.** The sequence $\Gamma_n$ (after appropriate telescoping) contains infinitely many occurrences of a single graph $\Gamma$ with a strictly positive incidence matrix, both in the past and in the future.

In this case, as is well-known since the work of Furstenberg (see e.g. (16.13) in [15]), the Markov compactum $X^+$ is uniquely ergodic, which means that there is a unique invariant probability measure for the “future tail” equivalence relation. We need to develop this point in more detail. In fact, in this case we have (see [8, 1.9.5]) that there exist strictly positive vectors $\tilde{z}(\ell), \tilde{u}(\ell)$, for $\ell \in \mathbb{Z}$, such that

\begin{align}
\tilde{z}(\ell) &= A^\ell_\ell \tilde{z}(\ell+1), \quad \ell \in \mathbb{Z}; \\
\bigcap_{n \in \mathbb{N}} A^\ell_{\ell+1} \cdots A^\ell_{\ell+n} \mathbb{R}_+^m &= \mathbb{R}_+^\ell \tilde{z}(\ell), \quad \ell \in \mathbb{Z}; \\
\tilde{u}(\ell) &= A^\ell_\ell \tilde{u}(\ell-1), \quad \ell \in \mathbb{Z}; \\
\bigcap_{n \in \mathbb{N}} A^\ell_{\ell-1} \cdots A^\ell_{\ell-n} \mathbb{R}_+^m &= \mathbb{R}_+^\ell \tilde{u}(\ell), \quad \ell \in \mathbb{Z}.
\end{align}

The vectors are normalized by $|\tilde{z}(0)|_1 = 1, \quad \langle \tilde{z}(0), \tilde{u}(0) \rangle = 1$.

As already mentioned, the Markov compactum $X^+$ is then uniquely ergodic, with the unique tail-invariant probability measure $\nu_+$ given by

\begin{equation}
\nu_+(X^+_j) = \tilde{z}_j(0), \quad \nu_+([\epsilon_1 \cdots \epsilon_\ell]) = \tilde{z}_j(\epsilon_\ell),
\end{equation}

where $X^+_j$ is the set of one-sided paths $\epsilon_1 \epsilon_2 \cdots \in X_+$ such that $F(\epsilon_1) = j$ and $[\epsilon_1 \cdots \epsilon_\ell]$ is the cylinder set corresponding to the finite path. The advantage of working with 2-sided Bratteli diagrams, which is one of the key ideas of [8], is that one can similarly define the “dual” measure $\nu_-$ on the set of “negative paths” $X_-$, invariant under the “past tail” equivalence relation. Then $\nu = \nu_+ \times \nu_-$ is a probability measure on $X$.

Now suppose that a linear ordering (Vershik’s ordering) is defined on the set $\{\epsilon \in \mathcal{E}(\Gamma_\ell) : I(\epsilon) = i\}$ for all $i \leq m$ and $\ell \in \mathbb{Z}$. This induces a partial lexicographic ordering $\sigma$ on the Markov compactum $X$; more precisely, two paths are comparable if they agree for $n \geq t$ for some $t \in \mathbb{Z}$. (Also two paths in $X_-$ are comparable if the end at the same vertex.) Restricting this ordering to the 1-sided compactum $X_+$, we obtain the adic, or Bratteli-Vershik (BV) transformation $\mathfrak{T}$, defined as the immediate successor of a path $\mathfrak{r}$ in the ordering $\sigma$. (See also [16].) Let $\text{Max}(\sigma)$ be the
set of paths in $X_+$ such that its every edge is maximal. It is easy to see that $\text{card}(\text{Max}(\mathfrak{o})) \leq m$. Similarly define the set of minimal paths $\text{Min}(\mathfrak{o})$, and let $\tilde{X}_+$ be the set of paths, which are not tail equivalent to any of the paths in $\text{Min}(\mathfrak{o}) \cup \text{Max}(\mathfrak{o})$. Then the $\mathbb{Z}$-action $\{\mathfrak{T}^n\}_{n \in \mathbb{Z}}$ is well-defined on $\tilde{X}_+$. Since we are excluding a countable set, the action is defined almost everywhere with respect to any non-atomic measure; certainly, $\nu_+$-a.e. in the uniquely ergodic case. We similarly define $\tilde{X}$ as the set of bi-infinite paths in $X$ which are not forward tail-equivalent to any of the maximal or minimal paths. Note that invariant measures for the future tail equivalence relation are precisely the invariant measures for the BV map, so we get unique ergodicity of the system $(X_+, \mathfrak{T})$ under our standing assumption.

We shall also consider suspension flows over BV-systems, with a piecewise-constant roof function depending only on the vertex at the level 0. More precisely, let $X_+$ be a one-sided Markov compactum with a Vershik ordering and BV-transformation $\mathfrak{T}$. For a strictly positive vector $\vec{s} = (s_1, \ldots, s_m)$ define the roof function $\phi_{\vec{s}}$ to be equal to $s_j$ on the cylinder set $X_j^+$. The resulting space will be denoted $X_{\vec{s}}$:

$$X_{\vec{s}} := \bigsqcup_{j=1}^{m} X_j^+ \times [0, s_j] / \sim, \quad \text{with } (e, \phi_{\vec{s}}(e)) \sim (\mathfrak{T}e, 0),$$

on which we consider the usual suspension flow $\{h_t\}_{t \in \mathbb{R}}$. It preserves the measure induced by the $\mathfrak{T}$-invariant measure $\nu_+$ on $X_+$ and the Lebesgue measure on $\mathbb{R}$. We will need the following

**Lemma 2.1.** [8, Section 2.5] Given a Vershik ordering on a uniquely ergodic Markov compactum $X$ with the unique invariant measure $\nu$, there is a symbolic flow $(X, g_t^+)_{t \in \mathbb{R}}$, defined on $\tilde{X}$, so $\nu$-a.e. on $X$, which is measurably conjugate to the suspension flow $(X_{\vec{s}}, h_t)_{t \in \mathbb{R}}$ over $(X_+, \mathfrak{T})$, with $s_j = u_j^{(1)}$, $j \leq m$. Moreover, the conjugating map $\mathfrak{F} : X \to X_{\vec{s}}$ is given by

$$\mathfrak{F}(\mathfrak{r}) = (P_+\mathfrak{r}, t), \quad \text{where } t = \nu_-(\{a \in X_- : I(a_0) = F(e_1), a \preceq P_- e\}),$$

where $P_+, P_-$ are the truncations from $X$ to $X_+$, $X_-$ respectively and $\preceq$ is the Vershik order. The map $\mathfrak{F}$ is well-defined on $\tilde{X}$ and its inverse is well-defined over $\tilde{X}_+$.

2.1. **Weakly Lipschitz functions.** Following [9] [8], we consider the space of weakly Lipschitz functions on a uniquely ergodic Markov compactum $X$ with the probability measure $\nu_+$ invariant for the “forward tail” equivalence relation and a Vershik ordering $\mathfrak{o}$. Recall that $\tilde{X}$ denotes the set of paths in $x$ that are not (forward) tail equivalent to any of the maximal or minimal paths in the Vershik ordering. We say that $f$ is weakly Lipschitz and write $f \in \text{Lip}_+^w(X)$ if $f$ is defined and bounded on $\tilde{X}$, and there exists $C > 0$ such that for all $\mathfrak{r}, \mathfrak{r}' \in \tilde{X}$, satisfying $e_k = e_k'$ for $-\infty < k \leq n$, with $n \in \mathbb{N}$, we have

$$|f(\mathfrak{r}) - f(\mathfrak{r}')| \leq C \nu_+(|e_1 \ldots e_n|). \quad (2.4)$$
The norm in $\text{Lip}_w^+(X)$ is defined by

\begin{equation}
\|f\|_L := \|f\|_{\text{sup}} + \widetilde{C},
\end{equation}

where $\widetilde{C}$ is the infimum of the constants in (2.4).

We analogously define the space $\text{Lip}_w^+(X^\mathfrak{s})$ of weakly Lipschitz functions on the space $X^\mathfrak{s}$ of the suspension flow over $(X_+, \mathfrak{T})$ with the roof function determined by the vector $\mathfrak{s} \in \mathbb{R}_m^+$. Namely, $f \in \text{Lip}_w^+(X^\mathfrak{s})$ if it is defined and bounded on $\tilde{X}_+$ and there exists $C > 0$ such that for all $\tilde{e}, \tilde{e}' \in \tilde{X}_+$, satisfying $e_k = e'_k$ for $k \leq n$, with $n \in \mathbb{N}$, and all $t \in [0, s_{F(e_1)}]$, we have

\begin{equation}
|f(\tilde{e}, t) - f(\tilde{e}', t)| \leq C\nu_+([e_1 \ldots e_n]).
\end{equation}

The norm $\|f\|_L$ is defined as in (2.5).

Note that weakly Lipschitz functions are not Lipschitz in the “transverse” direction, corresponding to the “past” in the 2-sided Markov compactum and to the vertical direction in the space of the suspension flow. Note also that if $f \in \text{Lip}_w^+(X)$, then $f \circ \mathfrak{s}^{-1} \in \text{Lip}_w^+(X^\mathfrak{s})$, with the same norm, where $\mathfrak{s}$ is defined in Lemma 2.1.

2.2. Substitutions. Along with the Markov compactum and BV-transformation, it is convenient to use the language of substitutions. Consider the alphabet $\mathcal{A} = \{1, \ldots, m\}$, which is identified with the vertex set of all the graphs $\Gamma_n$. A substitution is a map $\zeta : \mathcal{A} \to \mathcal{A}^+$, extended to $\mathcal{A}^+$ and $\mathcal{A}^\mathfrak{s}$ by concatenation. Given a Vershik ordering $\sigma$ on a 1-sided Bratteli diagram $\{\Gamma_j\}_{j \geq 1}$, the substitution $\zeta_j$ takes every $b \in \mathcal{A}$ into the word in $\mathcal{A}$ corresponding to all the vertices to which there is a $\Gamma_j$-edge starting at $b$, in the order determined by $\sigma$. Formally, the length of the word $\zeta_j(b)$ is

$$|\zeta_j(b)| = \sum_{a=1}^m A_{b,a}(\Gamma_j),$$

and the substitution itself is given by

\begin{equation}
\zeta_j(b) = u_1^{b,j} \ldots u_{|\zeta_j(b)|}^{b,j}, \quad b \in \mathcal{A}, \quad j \geq 0,
\end{equation}

where $(b, u_i^{b,j}) \in \mathcal{E}(\Gamma_j)$, listed in the linear order prescribed by $\sigma$. Substitutions, extended to $\mathcal{A}^+$, can be composed in the usual way as transformations $\mathcal{A}^+ \to \mathcal{A}^+$. We will use the notation

$$\zeta^{[n_1, n_2]} := \zeta_{n_1} \circ \ldots \circ \zeta_{n_2}, \quad n_2 \geq n_1,$$

and

$$\zeta^{[n]} := \zeta^{[1, n]}, \quad n \geq 1.$$

Given a substitution $\zeta$, its substitution matrix is defined by

$$S_\zeta(a, b) := \# \text{ symbols } a \text{ in } \zeta(b).$$
Observe that $S \circ \zeta_1 \circ \zeta_2 = S \circ \zeta_1 S \circ \zeta_2$. We will denote $S_n = S \circ \zeta_n$. Notice that the transpose $S^t_n$ is exactly the incidence matrix $A_n = A(\Gamma_n)$ by the definition of $\zeta_n$. We will use the notation $S^{[n_1, n_2]} = S^{\bigcirc \zeta_{n_1, n_2}}$ and $S^{[n]} = S^{\bigcirc \zeta_{[n]}}$.

Next, we associate to any $\mathbf{e} \in \mathbb{X}^+$, its “horizontal sequence” in the alphabet $\mathcal{A}$, defined by

$$h(\mathbf{e}) := x = (x_n)_{n \in \mathbb{Z}}, \quad \text{where } x_n = x_n(\mathbf{e}) = b \text{ whenever } F(\mathcal{I}^n(\mathbf{e})) = b, \quad n \in \mathbb{Z},$$

that is, we keep track of the vertex at level zero, applying the BV-transformation. The horizontal sequence is just the symbolic dynamics of $T$ with respect to the 0-level cylinder partition. We get a full 2-sided infinite sequence $h(\mathbf{e}) = x = (x_n)_{n \in \mathbb{Z}}$ whenever the (2-sided) orbit of $\mathbf{e}$ under $\mathcal{I}$ does not hit a maximal or a minimal path. (By our assumptions, no orbit of $\mathcal{I}$ can hit both.) Thus $h : \mathbb{X}_+ \to \mathcal{A}^\mathbb{Z}$ is well-defined. We obtain, by definition, the following commutative diagram:

$$\begin{array}{ccc}
\mathbb{X}_+ & \overset{T}{\longrightarrow} & \mathbb{X}_+ \\
\downarrow{h} & & \downarrow{h} \\
\mathcal{A}^\mathbb{Z} & \overset{T}{\longrightarrow} & \mathcal{A}^\mathbb{Z}
\end{array}$$

where $T$ is the left shift on $\mathcal{A}^\mathbb{Z}$. Of course, the map $h$ is far from being surjective. In order to understand its image, it is useful to have a more explicit algorithm for $h(\mathbf{e})$. Suppose that the path $\mathbf{e} \in \mathbb{X}_+$ goes through the vertices $b_0, b_1, \ldots$, that is, $b_n = F(\mathbf{e}_{n+1})$. Recalling the definition of the substitutions $\zeta_n$ we can write

$$\zeta_n(b_n) = u_{n-1} b_{n-1} v_{n-1}, \quad n \geq 1,$$

where $u_{n-1}$ and $v_{n-1}$ are words, possibly empty. Note that there may be more than one occurrence of $b_{n-1}$ in $\zeta_n(b_n)$, but we choose the representation (2.8) according to the edge $\mathbf{e}_n$. Consider the following sequence of words $U_n$, $n \geq 0$, defined inductively. We start with

$$U_0 = u_0 b_0 v_0,$$

where the dot . separates negative and positive coordinates. Then $U_{n+1}$ is obtained from $U_n$ inductively, by appending $\zeta_{n+1}(u_{n+1})$ from the left and $\zeta_{n+1}(v_{n+1})$ from the right. If we disregard the location of the dot, we simply have

$$U_n = \zeta_1 \circ \cdots \circ \zeta_{n+1}(b_{n+1}) = \zeta^{[n+1]}(b_{n+1}), \quad n \geq 0.$$

When we take the location of the dot into account, typically, the words $U_n$ will “grow” to infinity, both left and right, to a limiting 2-sided sequence, which is exactly $h(\mathbf{e})$:

$$h(\mathbf{e}) = \ldots \zeta_2(u_2) \zeta_1(u_1) u_0 b_0 v_0 \zeta_1(v_1) \zeta_2(v_2) \ldots$$
The other alternative is that it grows to infinity only on one side, which happens if and only if $e$ is tail equivalent to either minimal or a maximal path. Denote

$$\mathcal{Y} := \mathcal{h}(\overline{X}_+)$$

Now the following is clear.

**Lemma 2.2.** The space $\mathcal{Y} \subset \mathbb{A}^\mathbb{Z}$ is exactly the set of 2-sided sequences $x$ with the property that any subword of $x$ appears as a subword of $\zeta^{[n]}(b)$ for some $b \in \mathbb{A}$ and $n \geq 1$.

**Remark.** Dynamical systems $(\mathcal{Y}, T)$ have been studied under the name of $S$-adic systems. They were originally introduced by S. Ferenczi [12], with the additional assumption that there are finitely many different substitutions in the sequence $\{\zeta_j\}$; however, more recently this restriction has been removed, see e.g. [6].

Let $h_i^{(n)}$ be the number of finite paths $e_1, \ldots, e_n$ such that $I(e_n) = i$. (This is the height of the Rokhlin tower for the BV-map.) We get a sequence of real vectors $\vec{h}^{(n)}$ which satisfy the equations:

$$(2.10) \quad \vec{h}^{(n+1)} = A_{n+1} \vec{h}^{(n)} = S_{n+1} \vec{h}^{(n)}, \quad h_i^{(n)} = |\zeta^{[n]}(i)|, \quad n \geq 1.$$ 

Let $s$ be the left shift transformation on the 1-sided Markov compactum: $s(e_1, e_2, e_3, \ldots) = (e_2, e_3, \ldots)$. We thus obtain a sequence of Markov compacta $X_+^{(\ell)}$ for $\ell \geq 0$, with $X_+^{(0)} := X_+$, so that $s : X_+^{(\ell)} \to X_+^{(\ell+1)}$ for all $\ell$. The Vershik ordering and BV-transformation are naturally induced on the whole family. We can then consider the horizontal sequence map

$$\mathcal{h} : \overline{X}_+^{(\ell)} \to \mathbb{A}^\mathbb{Z}.$$ 

Its image, denoted by $\mathcal{Y}^{(\ell)}$, is described similarly to $\mathcal{Y} = \mathcal{Y}^{(0)}$, as the set of all sequences in $\mathbb{A}^\mathbb{Z}$ whose every subword occurs as a block in $\zeta^{[\ell+1,n]}(b)$ for some $n \geq \ell + 1$ and $b \in \mathbb{A}$.

A substitution $\zeta$ acts on $\mathbb{A}^\mathbb{Z}$ as follows:

$$\zeta(\ldots a_{-1}a_0a_1 \ldots) = \ldots \zeta(a_{-1}\zeta(a_0)\zeta(a_1) \ldots)$$

Definitions imply that we have a sequence of surjective maps $\zeta_{\ell} : \mathcal{Y}^{(\ell)} \to \mathcal{Y}^{(\ell-1)}$, $\ell \geq 1$. How are the horizontal sequences $x = \mathcal{h}(\epsilon) \in \mathcal{Y}^{(0)}$ and $x' = \mathcal{h}(s\epsilon) \in \mathcal{Y}^{(1)}$ related? It follows from definitions (and the explicit formulas for $\mathcal{h}$) that

$$x = T^{k-1} \zeta_1(x'),$$

where $k$ is the number (rank) of the edge $e_0$ in the Vershik ordering. Of course, similar formulas relate $\mathcal{h}(\epsilon)$ and $\mathcal{h}(s\epsilon)$ for $\epsilon \in X_+^{(\ell)}$. (Recall that $T$ denotes the left shift on $\mathbb{A}^\mathbb{Z}$.)
3. Estimating the growth of exponential sums and matrix products

Our goal is to estimate the growth of “twisted Birkhoff sums” \( \sum_{j=0}^{N-1} e^{-2\pi i\omega j} f(\vec{T}^j \vec{w}) \) for the BV-transformation \( \vec{T} \) on a 1-sided Markov compactum and the corresponding “twisted Birkhoff integrals” for suspension flows over \( \vec{T} \). We start with the simplest possible functions \( f(\vec{T}) \) which depend only on the 0-th vertex of the path (i.e. on \( \vec{w}(\vec{T})_0 \)). Every such function, of course, is a linear combination of characteristic functions of cylinder sets determined by the 0-th vertex. More precisely, for \( a \in \mathcal{A} \) denote by \([a]\) the set of all \( \vec{T} \in X_+ \) for which \( F(\vec{w}_1) = a \), and consider the corresponding characteristic function \( 1_{[a]} \). Then

\[ \Pi_{[a]}(\vec{T}^j \vec{w}) = \delta_{x_j,a}, \]

the “Kronecker delta,” where \( (x_j)_{j \in \mathbb{Z}} \) is the horizontal sequence \( \vec{w}(\vec{T}) \). Thus we are led to sums of the form \( \sum_{j=0}^{N-1} \delta_{x_j,a} e^{-2\pi i\omega j} \).

3.1. Setting up matrix products. For a word \( v = v_0 v_1 \ldots v_{N-1} \in \mathcal{A}^+ \) let

\[ \Phi_a(v, \omega) = \sum_{j=0}^{N-1} \delta_{v_j,a} e^{-2\pi i\omega j}, \]

where \( \delta_{v_j,a} \) is Kronecker \( \delta \) (equal to 1 if \( v_j = a \) and 0 otherwise). We need to estimate from above the growth of \( |\Phi_a(\zeta^{[n]}(b), \omega)| \), as \( n \to \infty \).

Observe that for any two words \( u, v \) and the concatenated word \( uv \) we have

\[ \Phi_a(uv, \omega) = \Phi_a(u, \omega) + e^{-2\pi i\omega|u|} \Phi_a(v, \omega). \]

Recalling (2.7), we can write

\[ \zeta^{[n]}(b) = \zeta^{[n-1]}(\zeta(b)) = \zeta^{[n-1]}(u_{1}^{b,n}) \ldots \zeta^{[n-1]}(u_{|\zeta(b)|}^{b,n}), \quad n \geq 1, \]

where we use the convention \( \zeta^{[0]} := Id \). Hence (3.2) implies for all \( a, b \in \mathcal{A} \):

\[ \Phi_a(\zeta^{[n]}(b), \omega) = \sum_{j=1}^{\zeta(b)} \exp \left[ -2\pi i\omega \left( |\zeta^{[n-1]}(u_{1}^{b,n}) \ldots u_{j-1}^{b,n})| \right) \right] \Phi_a(\zeta^{[n-1]}(u_{j}^{b,n}), \omega), \quad n \geq 1, \]

(for \( j = 1 \), the exponential reduces to \( \exp(0) = 1 \) by definition). For \( n \geq 0 \) let

\[ \Psi^{(a)}_n(\omega) := \begin{pmatrix} \Phi_a(\zeta^{[n]}(1), \omega) \\ \vdots \\ \Phi_a(\zeta^{[n]}(m), \omega) \end{pmatrix} \quad \text{and} \quad \Pi_n(\omega) = [\Psi^{(1)}_n(\omega), \ldots, \Psi^{(m)}_n(\omega)], \]

where \( \Pi_n(\omega) \) is the \( m \times m \) matrix-function specified by its column vectors. It follows that

\[ \Pi_n(\omega) = M_n(\omega) \Pi_{n-1}(\omega), \quad n \geq 1, \]
where $M_n(\omega)$ is an $m \times m$ matrix-function, whose matrix elements are trigonometric polynomials given by

$$
(M_n(\omega))(b, c) = \sum_{j \leq |\zeta_n(b)|} \exp\left[-2\pi i \omega \left(|\zeta^{[n-1]}(u_1^{b,n} \ldots u_{j-1}^{b,n})|\right)\right], \ n \geq 1.
$$

Note that $M_n(0) = S_n$, the transpose of the $n$-th substitution matrix, for all $n \geq 1$. Since $\vec{\Psi}_0^{(a)}(\omega) = \vec{e}_a$ (the basis vector corresponding to $a \in A$), it follows from (3.4) that

$$
\vec{\Psi}_n^{(a)}(\omega) = M_n(\omega)M_{n-1}(\omega) \cdots M_1(\omega)\vec{e}_a
$$

and

$$
\Pi_n(\omega) = M_n(\omega)M_{n-1}(\omega) \cdots M_1(\omega).
$$

### 3.2. Estimating matrix products.

**Definition 3.1.** A word $v$ is called a **good return word** for the substitution $\zeta$ if $v$ starts with some symbol $e$ (which can be any element of $A$) and $vc$ occurs in the word $\zeta(b)$ for every $b \in A$. Denote by $GR(\zeta)$ the set of good return words for $\zeta$.

Denote by $\vec{1}$ the vector of all 1’s. Below we denote by $\|x\|$ the distance from $x \in \mathbb{R}$ to the nearest integer. This is a standard notation in Diophantine approximation theory. We will sometimes use $\| \cdot \|$ to denote a vector or matrix norm (without a subscript), in which case it will always be the Euclidean norm and the corresponding operator norm; this should be clear from the context and not lead to a confusion.

**Proposition 3.2.** Let $X_+$ be a one-sided Markov compactum with a Vershik ordering, and let $\zeta_j$ be the corresponding sequence of substitutions, given by (2.7). Suppose that there exists a substitution $\zeta$, such that $Q = S_\zeta$ is strictly positive and there is an infinite increasing sequence $k_n$ satisfying

$$
\zeta_{k_n} = \zeta_{k_n+1} = \zeta.
$$

Then there exists $c_1 \in (0, 1)$, depending only on the substitution $\zeta$, such that for all $a, b \in A$ and $N \in \mathbb{N}$, $\omega \in [0, 1)$,

$$
|\Phi_a(\zeta^{[N]}(b), \omega)| \leq \|S^{[N]}\|_1 \prod_{k_n \leq N-1} \left(1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta^{[k_n]}(v)|\|_2\right).
$$

In fact, we can take

$$
c_1 = (2m)^{-1} \min_{i,j} Q_{ij} \cdot (\max_{i,j} Q_{ij})^{-2}.
$$

**Proof.** This is similar to the proof of (3.8) in [10] Proposition 3.2, but there are a number of new technical details. In view of (3.3), it suffices to estimate $\vec{\Psi}_N^{(a)}(\omega) = M_N(\omega)\vec{\Psi}_N^{(a)}(\omega)$. We will use the following notation:
for vectors $\vec{x}, \vec{y} \in \mathbb{R}^m$, the inequality $\vec{x} \leq \vec{y}$ means componentwise inequality, and similarly for real-valued matrices;

- the operation of taking absolute values of all entries for a vector $\vec{x}$ and a matrix $A$ will be denoted $|\vec{x}|$ and $|A|$, respectively.

It is clear that for any, generally speaking, rectangular matrices $A, B$ such that the product $AB$ is well-defined, we have

$$ (AB)^{|.|} \leq |A|^{|.|} B^{|.|}. $$

We fix $\omega$ and omit it from notation, so that $M_k \equiv M_k(\omega)$. Observe that (3.10) and (3.6) imply

$$ (\Psi_k^{(a)}(\omega))_{|.|} = (M_k \cdots M_1 e_{a})_{|.|} \leq M_k^{|.|} \cdots M_1^{|.|} e_a \leq M_k^{|.|} \cdots M_1^{|.|} \bar{1}. $$

For an arbitrary $\vec{x} = [x_1, \ldots, x_m]^t > \vec{0}$ and $k \geq 0$, we have $S_k^t \vec{x} > \vec{0}$ (since the substitution of any symbol is non-empty), and, trivially,

$$ M_k^{|.|} \vec{x} \leq S_k^t \vec{x}. $$

Now suppose $\zeta_{k_n+1} = \zeta$. By the definition of a good return word, for any $v \in GR(\zeta)$ and $b \in A$, we can write

$$ \zeta(b) = p^{(b)} vcq^{(b)}, $$

where $p^{(b)}$ and $q^{(b)}$ are words, possibly empty, and $v$ starts with $c$. We are going to estimate the absolute value of $M_{k_n+1}(b, c)$ given by (3.5). Note that $M_{k_n+1}(b, c)$ is a trigonometric polynomial with $S''_{\zeta}(b, c)$ exponential terms and all coefficients equal to one. By (3.5) and (3.13), the expression for $M_{k_n+1}(b, c)$ includes the terms $e^{-2\pi i \omega [\zeta^{[k_n]}(p^{(b)})]} + e^{-2\pi i \omega [\zeta^{[k_n]}(p^{(b)} v)]}$, hence

$$ |M_{k_n+1}(b, c)| \leq S''_{\zeta}(b, c) - 2 + |1 + e^{-2\pi i \omega [\zeta^{[k_n]}(v)]}|. $$

From the inequality

$$ |1 + e^{2\pi i \tau}| \leq 2 - \frac{1}{2} \|\tau\|^2, \quad \tau \in \mathbb{R}, $$

we have

$$ |M_{k_n+1}(b, c)| \leq S''_{\zeta}(b, c) - \frac{1}{2} \|\omega [\zeta^{[k_n]}(v)]\|^2. $$
For an arbitrary $\vec{x} = [x_1, \ldots, x_m]^t > 0$ and $b \in \mathcal{A}$, using (3.15) we can estimate
\[
(M_{k_n+1}^{(i)})_b = \sum_{j=1}^m |M_{k_n+1}(b, j)| x_j \leq \sum_{j=1}^m S^\xi_t(b, j) x_j - \frac{1}{2} \|\omega|\zeta^{[k_n]}(v)\|_2^2 x_c \leq \left(1 - c_2 \psi(\vec{x}) \|\omega|\zeta^{[k_n]}(v)\|_2^2 \right) \cdot \sum_{j=1}^m S^\xi_t(b, j) x_j \leq \left(1 - c_2 \psi(\vec{x}) \|\omega|\zeta^{[k_n]}(v)\|_2^2 \right) \cdot (S^\xi_{k_n+1} \vec{x})_b, \quad (3.16)
\]
where
\[
c_2 = \frac{1}{2m \max_{i,j} S^\xi_t(i, j)} = \frac{1}{2m \max_{i,j} Q_{ij}} \quad \text{and} \quad \psi(\vec{x}) = \frac{\min_j x_j}{\max_j x_j}.
\]
Thus,
\[
M_{k_n+1}^{(i)} \vec{x} \leq \left(1 - c_2 \psi(\vec{x}) \|\omega|\zeta^{[k_n]}(v)\|_2^2 \right) \cdot S^\xi_{k_n+1} \vec{x},
\]
and $v \in GR(\zeta)$ is arbitrary. We apply the last inequality with
\[
\vec{x}_n := (S^{[k_n]})^t \bar{1} = Q^t (S^{[k_n-1]})^t \bar{1},
\]
taking into account that $\zeta_{k_n} = \zeta$. Since the matrix $Q$ is strictly positive, it is easy to see that
\[
\psi(Q \vec{x}) \leq \frac{\min_{i,j} Q_{ij}}{\max_{i,j} Q_{ij}}.
\]
Therefore,
\[
(3.17) \quad M_{k_n+1}^{(i)} (S^{[k_n]})^t \bar{1} \leq \left(1 - c_1 \max_{v \in GR(\xi)} \|\omega|\zeta^{[k_n]}(v)\|_2^2 \right) \cdot (S^{[k_n+1]})^t \bar{1},
\]
where $c_1$ is given by (3.9), where we used that $(S^{[k_n+1]})^t \bar{1} = S^t_{k_n+1} (S^{[k_n]})^t \bar{1}$. Now we apply (3.11) using (3.17) for $k = k_n + 1$ and (3.12) otherwise, to obtain
\[
(\bar{\Psi}_N^{(q)}(\omega))^{(i)} \leq \prod_{k_n \leq N-1} \left(1 - c_1 \max_{v \in GR(\zeta)} \|\omega|\zeta^{[k_n]}(v)\|_2^2 \right) \cdot (S^{[N]})^t \bar{1}.
\]
Finally, note that the maximal component of $(S^{[N]})^t \bar{1}$ is the maximal column sum of $S^{[N]}$, which is $\|S^{[N]}\|_1$, and the proof is complete. We emphasize that both $\zeta_{k_n} = \zeta$ and $\zeta_{k_n+1} = \zeta$ were used in an essential way. \qed
3.3. The case of suspension flows. Here we extend the estimates of the previous subsection to the case of suspension flows over BV-transformations, with a piecewise-constant roof function.

Since our goal is to obtain estimates of spectral measures, we recall how they are defined for flows. Given a measure-preserving flow \((Y, h_t, \mu)_{t \in \mathbb{R}}\) and a test function \(f \in L^2(Y, \mu)\), there is a finite positive Borel measure \(\sigma_f\) on \(\mathbb{R}\) such that

\[
\int_{-\infty}^{\infty} e^{2\pi i \omega \tau} d\sigma_f(\omega) = \langle f \circ h_\tau, f \rangle \quad \text{for } \tau \in \mathbb{R}.
\]

In order to obtain local bounds on the spectral measure, we can use growth estimates of the “twisted Birkhoff integral”

\[
S_R^{(y)}(f, \omega) := \int_0^R e^{-2\pi i \omega \tau} f \circ h_\tau(y) d\tau.
\]

The following lemma is standard; a proof may be found in [10, Lemma 4.3].

**Lemma 3.3.** Suppose that for some fixed \(\omega \in \mathbb{R}, R_0 > 0, \alpha \in (0,1)\) we have

\[
\sup_{y \in Y} |S_R^{(y)}(f, \omega)| \leq C_1 R^\alpha \text{ for all } R \geq R_0.
\]

Then

\[
\sigma_f([\omega - r, \omega + r]) \leq \pi^2 C_1^2 r^{2(1-\alpha)} \text{ for all } r \leq (2R_0)^{-1}.
\]

**Remark 3.4.** 1. Estimates of twisted Birkhoff sums have been used for a number of different dynamical systems recently; in particular, see the work of Forni and Ulcigrai [14] on the Lebesgue spectrum for smooth time changes of the horocycle flow.

2. For \(\omega = 0\) the formula (3.18) reduces to the usual Birkhoff integral, for which sharp estimates and asymptotics are known in a number of cases. It should be possible to obtain precise asymptotics of the spectral measure at zero, governed by the second Lyapunov exponent, analogously to [10, Theorem 6.2].

Now let \(X_+\) be a one-sided Markov compactum with a Vershik ordering and BV-transformation \(\mathfrak{T}\). For a strictly positive vector \(\vec{s} = (s_1, \ldots, s_m)\) we define the roof function \(\phi_{\vec{s}}\) to be equal to \(s_a\) on the cylinder set \(X_{a+}\), as in Section 2, and obtain the suspension flow \((\mathfrak{F}_{\vec{s}}, h_t)\).

Recall that for \(\vec{t} \in X_+\) (minus a countable exceptional set) we defined its horizontal sequence \(h(\vec{t}) = (x_n)_{n \in \mathbb{Z}}\), in such a way that the BV-transformation intertwines the left shift. Similarly, we can associate to \((\vec{t}, t) \in \mathfrak{F}_{\vec{s}}\) a tiling of the line \(\mathbb{R}\): a symbol \(a\) corresponds to a closed line segment of length \(s_a\) (labeled by \(a\)), and these line segments are “strung together” according to the symbolic sequence \(h(\vec{t})\). The tile corresponding to \(x_0\) should contain the origin at the distance \(t\) from the left endpoint. This defines a map \(\tilde{h}\) from \(\mathfrak{F}_{\vec{s}}\) to a “tiling space,” which intertwines the flow \(h_\tau\) and the left shift by \(\tau\).
We start with test functions depending only on the cylinder set $X_a$ and the height $t$. More precisely, given some functions $\psi_a \in C([0, s_a]), a \in \mathcal{A}$, let

$$f = \sum_{a \in \mathcal{A}} c_a f_a, \text{ with } f_a(\vec{r}, t) = 1_{X_a} \psi_a(t), \text{ where } X_a = X_a \times [0, s_a].$$

For a word $v$ in the alphabet $\mathcal{A}$ denote by $\vec{\ell}(v) \in \mathbb{Z}^m$ its “population vector” whose $j$-th entry is the number of $j$’s in $v$, for $j \leq m$. We will need the “tiling length” of $v$ defined by

$$|v|_\vec{s} := \langle \vec{\ell}(v), \vec{s} \rangle.$$

For $v = v_0 \ldots v_{N-1} \in \mathcal{A}^+$ let

$$\Phi^{\vec{s}}_a(v, \omega) = \sum_{j=0}^{N-1} \delta_{v_j, a} \exp(-2\pi i \omega|v_0 \ldots v_j|_\vec{s}).$$

Then a straightforward calculation shows

$$S_R^{\vec{f}(0)}(f_a, \omega) = \hat{\psi}_a(\omega) \cdot \Phi^{\vec{s}}_a(x[0, N - 1], \omega) \text{ for } R = |x[0, N - 1]|_{\vec{s}},$$

where $(x_n)_{n \in \mathbb{Z}} = h(\vec{r})$. In order to estimate the growth of the expression in (3.21), we need an analog of (3.8).

**Proposition 3.5.** Suppose, as in Proposition 3.2, that there exists a substitution $\zeta$ with a strictly positive substitution matrix, such that $\zeta_{k_n} = \zeta_{k_n+1} = \zeta$ for a sequence $k_n \uparrow \infty$. Then for all $a, b \in \mathcal{A}, N \in \mathbb{N}, \vec{s} > \vec{0}$, and $\omega \in \mathbb{R},$

$$|\Phi^{\vec{s}}_a(\zeta^{[N]}(b), \omega)| \leq \|S^{[N]}\|_1 \prod_{k_n \leq N-1} \left(1 - c_1 \cdot \max_{v \in G(a, \zeta)} \|\omega|\zeta^{[k_n]}(v)|_{\vec{s}}|^2\right),$$

where $c_1$ is given by (3.9).

**Proof.** The proof is identical to that of Proposition 3.2. We just have to replace the usual length of words by their tiling length, defined in (3.19). More precisely, we use matrices $\tilde{M}_a$ obtained from (3.5) by replacing $| \cdot |$ with $| \cdot |_{\vec{s}}$ under the exp sign. \qed

More generally, suppose that $f$ on $\mathcal{X}$ is a “cylindrical function of level $\ell$”, that is, its value depends only on the first $\ell$ edges of the path $\vec{r}$ and on the height $t$. It is then convenient to represent $h_{\tau}$ as a suspension flow with a different height function, based on the decomposition (disjoint in measure)

$$\mathcal{X}^{\vec{s}} = \bigcup_{a \in \mathcal{A}} \mathcal{X}_a^{(\ell)}, \text{ where } \mathcal{X}_a^{(\ell)} = \{(\vec{r}, t) \in \mathcal{X}^{\vec{s}} : \vec{r} \in X_+, x_\ell = F(\xi_{\ell+1}) = a\}.$$

The BV-transformation $\Upsilon$ “changes” a vertex $a$ at the $\ell$-th level after $h_0^{(\ell)} = |\zeta^{[\ell]}(a)|$ iterates. Thus, after we enter the cylinder $\mathcal{X}_a^{(\ell)}$, the flow $h_{\tau}$ stays in it for the time equal to

$$s_0^{(\ell)} := |\zeta^{[\ell]}(a)|_{\vec{s}} = |S^{[\ell]}|^s_a.$$

(3.23)
More precisely, if $F(c_{t+1}) = a$ and $(\mathcal{F}, t) \in X$, then

$$
(3.24) \quad \exists t' \in [0, s] \text{ such that } (\mathcal{F}, t) = h_{t'}(c_{t+1} \ldots c_{t+1}c_{t+2} \ldots, 0),
$$

where $c_{t+1} \ldots c_{t}$ is the minimal path in the Bratteli diagram from the vertex $a$ on level $\ell$ to the level $0$. Observe that the horizontal sequence of a path $c_{t+1} \ldots c_{t+1}c_{t+2} \ldots$, with $F(c_{t+1}) = a$, begins with $\zeta^{[\ell]}(a)$ and can be written as $\zeta^{[\ell]}(x^{(\ell)})$ for some $x^{(\ell)} \in A$. (In fact, $x^{(\ell)} = h(\sigma^{\ell}) \in Y^{(\ell)}$, see Section 2). To summarize this discussion, for any real-valued continuous cylindrical function $f$ of level $\ell$ on $X$ there exist $c_{a} \in \mathbb{R}$ and $\psi^{(\ell)}_{a} \in C([0, s^{(\ell)}])$, $a \in A$, such that

$$
(3.25) \quad f = \sum_{a \in A} c_{a}f^{(\ell)}_{a}, \quad \text{where } f^{(\ell)}_{a}(\mathcal{F}, t) = \mathbb{1}_{x^{(\ell)}_{a}}(t'), \quad \text{with } t' \text{ from } (3.24).
$$

Now we can also write a generalization of (3.21). Denote $\bar{s}^{(\ell)} = (s^{(\ell)}_{a})_{a \in A}$ and assume $\mathcal{F}' = c_{1} \ldots c_{\mathcal{F}+1}c_{\mathcal{F}+2} \ldots$, with $F(c_{\mathcal{F}+1}) = a$. Then

$$
(3.26) \quad S^{(\mathcal{F}, 0)}(f^{(\ell)}_{a}, \omega) = \psi^{(\ell)}(\omega) \cdot \Phi^{(\ell)}_{a}(x^{(\ell)}_{a}[0, N-1], \omega) \quad \text{for } R = |x^{(\ell)}_{a}[0, N-1]|_{\bar{s}^{(\ell)}},
$$

where $(x_{n})_{n \geq 0} = \zeta^{[\ell]}(x_{n}^{(\ell)})_{n \geq 0}$ and $(x_{n})_{n \in \mathbb{Z}} = h(\mathcal{F})$.

**Corollary 3.6.** Under the assumptions of Proposition 3.5, for any $\ell \geq 0$, $a, b \in A$, $n \geq \ell + 1$, $\bar{s} > \bar{s}'$, and $\omega \in \mathbb{R}$, we have

$$
(3.27) \quad |\Phi^{(\ell)}_{a}(\zeta^{[\ell+1,n]}_{b}, \omega)| \leq \|S^{(\ell+1,n)}_{1} \cdot \prod_{\ell+1 \leq k_{p} \leq n-1} \left(1 - c_{1} \cdot \max_{v \in GR(\zeta)} ||\omega|\zeta^{[k_{p}]}(v)||_{\bar{s}^{(\ell)}}^{2}\right),
$$

where $c_{1}$ is given by (3.5).

**Proof.** It is immediate from Proposition 3.5 by shifting the indices that

$$
|\Phi^{(\ell)}_{a}(\zeta^{[\ell+1,n]}_{b}, \omega)| \leq \|S^{(\ell+1,n)}_{1} \cdot \prod_{\ell+1 \leq k_{p} \leq n-1} \left(1 - c_{1} \cdot \max_{v \in GR(\zeta)} ||\omega|\zeta^{[\ell+1,k_{p}]}(v)||_{\bar{s}^{(\ell)}}^{2}\right).
$$

Now it remains to note that

$$
|\zeta^{[\ell+1,k_{p}]}(v)|_{\bar{s}^{(\ell)}} = \langle \bar{\ell}(\zeta^{[\ell+1,k_{p}]}(v)), \bar{s}^{(\ell)} \rangle = \langle S^{[\ell+1,k_{p}]}(\bar{\ell}(v)), (S^{\ell})^{t} \bar{s} \rangle = \langle S^{[k_{p}]}(\bar{\ell}(v)), \bar{s} \rangle = \langle \bar{\ell}(\zeta^{[k_{p}]}(v)), \bar{s} \rangle = |\zeta^{[k_{p}]}(v)|_{\bar{s}}.
$$

$\square$
Proposition 3.7. Under the assumptions of Proposition 3.3, for any \( \ell \geq 0 \), \( a \in A \), \( N \in \mathbb{N} \), \( s > 0 \), \( x^{(\ell)} \in \mathcal{Y}^{(\ell)} \), and \( \omega \in \mathbb{R} \), we have,

\[
(3.28) \quad |\Phi_{n}^{(\ell)}(x^{(\ell)}[0,N-1],\omega)| \leq 2 \sum_{j=\ell}^{n-1} \|S_{j+1}\|_{1} \cdot \|S^{[\ell+1,j]}\|_{1} \cdot \prod_{\ell+1 \leq k_{p} \leq j-1} (1 - c_{1} \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta^{(k_{p})}(v)|_{x}^{2}),
\]

where \( c_{1} \) is given by \((3.9)\)

\[
(3.29) \quad \min_{b \in A} |\zeta^{[\ell+1,n]}(b)| \leq N \leq \max_{b \in A} |\zeta^{[\ell+1,n+1]}(b)|.
\]

Here we let \( S^{[\ell+1,\ell]} = I \).

Proof. We can assume that \( \ell = 0 \) without loss of generality, so that \( x^{(0)} = x = \theta(\epsilon) \in \mathcal{Y} \). The following well-known lemma is immediate from the description of \( \mathcal{Y}^{(0)} = \mathcal{Y} \) (see the construction of the horizontal sequence in Section 2).

Lemma 3.8. Let \( x \in \mathcal{Y} \) and \( N \geq 1 \). Then

\[
(3.30) \quad x[0,N-1] = u_{0}\zeta^{[1]}(u_{1}) \cdots \zeta^{[n]}(u_{n})\zeta^{[n-1]}(v_{n}) \cdots \zeta^{[1]}(v_{1})v_{0},
\]

where \( n \geq 0 \) and \( u_{j}, v_{j}, j = 0, \ldots, n \), are respectively proper prefixes and suffixes of the words \( \zeta_{j+1}(b), b \in A \). The words \( u_{j}, v_{j} \) may be empty, except that at least one of \( u_{n}, v_{n} \) is nonempty.

Note that \( n \) satisfies \((3.29)\) (with \( \ell = 0 \)). It remains to apply Corollary 3.6 (with \( \ell = 0 \)) to each term. The factor \( 2 \|S_{j+1}\|_{1} \) in \((3.28)\) appears, because \( |u_{j}|, |v_{j}| \leq \max_{b} |\zeta_{j+1}(b)| = \|S_{j+1}\|_{1} \). \( \square \)

4. Random BV-transformations: Statement of the theorem and plan of the proof

Here we consider dynamical systems generated by a random sequence of Markov compacta. In order to state our results, we need some preparation; specifically, the Oseledets Theorem.

Recall that \( \mathfrak{G} \) denotes the set of all oriented graphs on \( m \) vertices such that there is an edge starting at every vertex and an edge ending at every vertex (we allow loops and multiple edges). We also assume that each graph is equipped with a Vershik ordering. Let \( \Omega \) be the space of sequences of graphs:

\[
\Omega = \{a = \ldots a_{-n} \ldots a_{0}a_{1} \ldots a_{n} \ldots, a_{i} \in \mathfrak{G}, i \in \mathbb{Z}\}.
\]

For \( a \in \Omega \) we denote by \( X(a) \) and \( X_{+}(a) \) the Markov compacta corresponding to \( a \) according to the rule \( \Gamma_{n} = a_{n}, n \in \mathbb{Z} \). We denote by \( \sigma \) the left shift on \( \Omega \). Let \( \mathbb{P} \) be an ergodic \( \sigma \)-invariant probability measure on \( \Omega \) satisfying the following

Conditions:

(C1) There exists \( \Gamma_{1} \in \mathfrak{G} \) such that all the entries of the matrix \( A(\Gamma_{1}) \) are positive and

\[
(4.1) \quad \mathbb{P} \{a : a_{1} = \Gamma_{1}\} > 0.
\]

(C2) The matrices \( A(a_{n}) \) are almost surely invertible with respect to \( \mathbb{P} \).
The functions \( a \mapsto \log(1 + \|A^\pm_1(a_1)\|) \) are integrable.

(Here and below \( \|A\| \) denotes the Euclidean operator norm of the matrix.)

Observe that (C3), together with the Birkhoff ergodic theorem, immediately gives

\[
\lim_{n \to \infty} n^{-1} \log(1 + \|A(a_n)\|) = 0 \quad \text{for } \mathbb{P}\text{-a.e. } a \in \Omega.
\]

We obtain a measurable cocycle \( \mathbb{A} : \Omega \to GL(m,\mathbb{R}) \), defined by \( \mathbb{A}(a) = A(a_1) \), called the normalization cocycle. Denote

\[
\mathbb{A}(n, a) = \begin{cases} 
\mathbb{A}(\sigma^{-1} a) \cdot \ldots \cdot \mathbb{A}(a), & n > 0; \\
\text{Id}, & n = 0; \\
\mathbb{A}^{-1}(\sigma^{-n} a) \cdot \ldots \cdot \mathbb{A}^{-1}(\sigma^{-1} a), & n < 0,
\end{cases}
\]

so that

\[
\mathbb{A}(n, a) = A(a_{n-1}) \cdots A(a_1), \quad n \geq 1.
\]

As in Section 2, we consider the sequence of substitutions

\[
\zeta_k = \zeta_k(a) = \zeta(a_k), \quad a \in \Omega, \quad k \in \mathbb{Z},
\]

and their substitution matrices \( S_k(a) = S(\zeta_k(a)). \) (Recall that all graphs \( a_k \) are equipped with a Vershik ordering.) Thus

\[
\mathbb{A}(n, a) = S^t_{\zeta_k}(a) \cdots S^t_1(a) = (S^{[n]}^t)(a), \quad n \geq 1.
\]

By the Oseledets Theorem \[20\] (for a detailed survey, see Barreira-Pesin \[5\]), there exist Lyapunov exponents \( \theta_1 > \theta_2 > \ldots > \theta_r \) and, for \( \mathbb{P}\text{-a.e. } a \in \Omega \), a direct-sum decomposition

\[
\mathbb{R}^m = E_{a_1} \oplus \cdots \oplus E_{a_r}
\]

that depends measurably on \( a \in \Omega \) and satisfies the following:

(i) for \( \mathbb{P}\text{-a.e. } a \in \Omega \), any \( n \in \mathbb{Z} \), and any \( i = 1, \ldots, r \) we have

\[
\mathbb{A}(n, a) E_{a_i} = E_{a_{i+n}}.
\]

(ii) for any \( v \in E_{a_i}, v \neq 0 \), we have

\[
\lim_{|n| \to \infty} \frac{\log \|\mathbb{A}(n, a)v\|}{n} = \theta_i.
\]

(iii) \( \lim_{|n| \to \infty} \frac{1}{n} \log \angle \left( \bigoplus_{i \in I} E_{a_{i+n}}^i, \bigoplus_{j \in J} E_{a_{j+n}}^j \right) = 0 \) whenever \( I \cap J = \emptyset \).

Denote by \( \sigma_f \) the spectral measure for the system \((X, h_t)\) with the test function \( f \) (assuming the system is uniquely ergodic). Now we can state our theorem.

**Theorem 4.1.** Let \((\Omega, \mathbb{P}, \sigma)\) be an invertible ergodic measure-preserving system, satisfying conditions (C1)-(C3) above. Consider the cocycle \( \mathbb{A}(n, a) \) as above. Assume that
Remark. It is clear that condition \((C1)\) follows from assumption (b), but we chose to state \((C1)\) explicitly, since this is the condition which appears in the literature and implies unique ergodicity. The unique ergodicity of the system \((X_+,\mathfrak{F})\) for \(P_+-a.e.\) \(a^+\) under the given assumptions is well-known and goes back to the work of Furstenberg \cite{furstenberg} (see the beginning of Section 2).

The scheme of the proof is as follows: we are going to use Lemma \ref{lemma:3.3} and Proposition \ref{prop:3.7}. Then, roughly speaking, we need to show that for \(P\)-a.e. sequence of substitutions, for Lebesgue a.e. \(\bar{s}\), the distance from \(\omega|\zeta^{(k_0)}(v)|\bar{s}\) to the nearest integer (for some choice of a good return word
which may depend on $n$ and with $\omega$ bounded away from zero and infinity) is bounded below for a positive frequency of $n$’s. It is convenient to split the proof into two parts, separating the two “almost every”. The first part is more probabilistic, showing that certain assumptions on the sequence of substitutions $\zeta_j$ hold $\mathbb{P}$-almost surely. In the second part we fix a generic sequence $\zeta_n$ and obtain estimates for a.e. $\vec{s}$. This is done using the so-called “Erdős-Kahane argument.”

5. Exponential tails

By the definition of the measure $\mathbb{P}_+$ as projection of $\mathbb{P}$, taking $\mathbb{P}_+$-a.e. sequence $a^+ \in \Omega_+$ is equivalent to taking $\mathbb{P}$-a.e. $a \in \Omega$ and considering its future $a^+ = (a_n)_{n \geq 1}$.

For $a \in \Omega$ we consider the sequence of substitutions $\zeta_n(a)$, $n \geq 1$. Let $1 \leq k_1 < k_2 < \cdots < k_n < \cdots$ be the set of all indices, such that

$$\zeta_{k_n} = \zeta_{k_n + 1} = \zeta;$$

in particular, $\sigma^{k_n}a$ is the $n$-th return to $\Omega_q$ under $\sigma$ for $a \in \Omega_q$. Denote also

$$w_n := k_{n+1} - k_n = \ell_q(\sigma^{k_{n+1}}a), \ n \geq 1,$$

and

$$(5.1) \quad W_n = W_n(a) := \log \|\Lambda(\ell_q(\sigma^{k_{n+1}}a), \sigma^{k_n}a))\| = \log \|S_{k_{n+1}}(a) \cdots S_{k_n}(a)\|.$$  

Since all the matrices in the product have non-negative integer entries and a.s. invertible, and the first and last one are equal to $Q$ with strictly positive entries, we have $W_n > 0$, $n \geq 1$, for $\mathbb{P}$-a.e. $a$. This will always be assumed below, without loss of generality.

**Proposition 5.1.** Under the assumptions of Theorem 4.1, there exist positive constants $L_1, L_2$ such that for $\mathbb{P}$-a.e. $a$, the following holds:

(A1) For all $N$ sufficiently large we have

$$\# \{n : k_n \leq N\} \geq \lfloor N/L_1 \rfloor,$$

or equivalently,

$$(5.2) \quad k_n \leq L_1 n, \ n \geq n_0(a);$$

(A2) for any $\delta > 0$, for all $n$ sufficiently large (depending on $\delta$),

$$(5.3) \quad w_n \leq \delta n;$$

and

(A3) for any $\delta > 0$, for all $N$ sufficiently large (depending on $\delta$),

$$(5.4) \quad \max \left\{ \sum_{n \in \Psi} W_n : \Psi \subset \{n : k_{n+1} \leq N\}, \ |\Psi| \leq \delta N \right\} \leq L_2 \cdot \log(1/\delta) \cdot \delta N.$$  

We will prove the proposition at the end of the section, but first point out the following.
Remark 5.2. It follows from (A3) and (A1) that for any $\tilde{\delta} > 0$, for all $n$ sufficiently large,
\begin{equation}
W_n \leq \tilde{\delta} n.
\end{equation}

Indeed, in (5.4) we just need to take $\delta > 0$ such that $L_2 \cdot \log(1/\delta) \cdot \delta < \tilde{\delta}/(2L_1)$, and then $\Psi = \{n\}$, which clearly satisfies the condition $1 = |\Psi| \leq \delta k_{n+1} = N$ for $n$ sufficiently large. Then (5.4) and (5.2) yield
\[ W_n \leq \tilde{\delta} k_{n+1}/(2L_1) \leq \tilde{\delta} L_1(n + 1)/(2L_1) \leq \tilde{\delta} n, \]
for $n$ sufficiently large, as desired. $\square$

Since we are interested in the returns of the orbit $\sigma^n(a)$ to the set $\Omega_q$, it is convenient to introduce the (random) symbolic sequence $\{a_n\}_{n \in \mathbb{Z}}$, such that
\[ a_n = \begin{cases} 1, & \text{if } \sigma^n(a) \in \Omega_q; \\ 0, & \text{otherwise}. \end{cases} \]

Observe that
\[ \ell_q(a) = \min\{n \geq 1 : a_{-n} = 1\}. \]

It follows from ergodicity of $\sigma$ and $\mathbb{P}(\Omega_q) > 0$ that $\mathbb{P}$-a.s. the sequence $a_n$ contains infinitely many 1’s (both forward and backward).

**Lemma 5.3.** We have for all $n \geq 2$:
\[ \mathbb{P}\{\ell_q(a) \geq n\} \leq \frac{Ce^{-\varepsilon(n-1)}}{1 - e^{-\varepsilon}}, \]
where $C$ and $\varepsilon$ are from (4.5).

**Proof.** Observe that
\[ \mathbb{P}\{\ell_q(a) \geq n\} = \sum_{j=0}^{\infty} \mathbb{P}(a_{-n+1}, \ldots, a_{j-1} = 0, a_j = 1) \]
\[ = \sum_{j=0}^{\infty} \mathbb{P}(a_{-n+1}, \ldots, a_{j-1} = 0 | a_j = 1) \mathbb{P}(a_j = 1). \]

Note that $\mathbb{P}(a_{-n+1}, \ldots, a_{j-1} = 0 | a_j = 1) \leq Ce^{-\varepsilon(n+j-1)}$ by (4.3) and $\sigma$-invariance of $\mathbb{P}$, hence
\[ \mathbb{P}\{\ell_q(a) \geq n\} \leq C \sum_{j=0}^{\infty} e^{-\varepsilon(n+j-1)} = \frac{Ce^{-\varepsilon(n-1)}}{1 - e^{-\varepsilon}}, \]
as desired. $\square$

Let
\begin{equation}
\bar{k}_1^N(a) = \max\{n < N : a_n = 1\},
\end{equation}

(5.6)
that is,
\[ \tilde{k}^N_1(a) = N - \ell_q(\sigma^N a). \]

We continue inductively, setting
\[ \tilde{k}^N_j = \tilde{k}^N_j(a) = \max\{n < \tilde{k}^N_{j-1} : a_n = 1\} \quad \text{for } j \geq 2; \]
in other words, \( \tilde{k}^N_j \) is the \( j \)-th index going \textbf{backward} from \( N \) for which \( a_{\tilde{k}^N_j} = 1 \). Thus \( \tilde{k}^N_j \) is a decreasing sequence in \( j \), which will eventually become negative. Since \( a \) is \( 2 \)-sided and \( q \) occurs infinitely often, this is well-defined a.s. for all \( N \) and \( j \in \mathbb{N} \). Now let
\[ \tilde{w}^N_j = \tilde{k}^N_j - \tilde{k}^N_{j+1} > 0, \quad j \geq 1, \]
equivalently,
\[ (5.7) \quad \tilde{w}^N_j = \ell_q(\sigma^N \sigma^N \sigma^{N-j} a). \]

Further, let
\[ (5.8) \quad \tilde{W}^N_j = \log \| A_{\ell_q(\sigma^N \sigma^N \sigma^{N-j} a)} \| = \log \| S^t_{\tilde{k}^N_j} \cdots S^t_{\tilde{k}^N_{j+1}} \|. \]

\textbf{Lemma 5.4.} We have for all \( N \) and \( j \geq 1 \):
\[ (5.9) \quad \mathbb{E} \left[ \exp(\varepsilon \tilde{w}^N_j) \right] < C \]
and
\[ (5.10) \quad \mathbb{E} \left[ \exp(\varepsilon \tilde{W}^N_j) \right] < C, \]
where \( \varepsilon > 0 \) and \( C > 1 \) are the constants from (4.3).

\textbf{Proof.} We start with the proof of (5.9). Clearly, \( \tilde{k}^N_j \leq N - j \), so
\[ \mathbb{E} \left[ \exp(\varepsilon \tilde{w}^N_j) \right] = \sum_{i=j}^{\infty} \mathbb{E} \left[ \exp(\varepsilon \tilde{w}^N_j) \mid \tilde{k}^N_j = N - i \right] \cdot \mathbb{P} \left[ \tilde{k}^N_j = N - i \right] \]
\[ = \sum_{i=j}^{\infty} \mathbb{E} \left[ \exp(\varepsilon \ell_q(\sigma^{N-i} a)) \mid \tilde{k}^N_j = N - i \right] \cdot \mathbb{P} \left[ \tilde{k}^N_j = N - i \right], \]
in view of (5.7). Condition (4.5) implies
\[ \mathbb{E} \left[ \exp(\varepsilon \ell_q(a)) \right] \mid a_0 = 1, a_1 \ldots a_s = b_1 \ldots b_s < C \quad \text{for any } b_1 \ldots b_s \in \{0,1\}^s, \]
hence
\[ (5.12) \quad \mathbb{E} \left[ \exp(\varepsilon \ell_q(\sigma^{N-i} a)) \right] \mid a_{N-i} = 1, a_{N-i+1} \ldots a_N = b_1 \ldots b_i < C \quad \text{for any } b_1 \ldots b_i \in \{0,1\}^i, \]
by the shift-invariance of the measure \( \mathbb{P} \). Since the event \( [\tilde{k}^N_j = N - i] \) is a union of cylinder sets
\[ [a_{N-i} = 1, a_{N-i+1}, \ldots, a_N = b_1 \ldots b_i], \]
over all sequences $b_1 \ldots b_i \in \{0, 1\}^i$ containing exactly $(j - 1)$ ones, \eqref{5.12} implies

$$E \left[ \exp[\varepsilon \ell_q(\sigma^{N-i}a)] \mid \bar{k}_j^N = N - i \right] < C,$$

and then \eqref{5.11} yields \eqref{5.9}.

The proof of \eqref{5.10} is almost the same. We have

$$E \left[ \exp(\varepsilon \tilde{W}_j^N) \right] = \sum_{i=j}^{\infty} E \left[ \exp(\varepsilon \tilde{W}_j^N) \mid \bar{k}_j^N = N - i \right] \cdot P \left[ \bar{k}_j^N = N - i \right].$$

Condition \eqref{4.6} implies

$$E \left[ \|A(\ell_q(a), \sigma_{N-i}(\sigma N-a))\|_\varepsilon \mid \bar{k}_j^N = N - i \right] \cdot P \left[ \bar{k}_j^N = N - i \right].$$

Therefore,

$$E \left[ e^{\varepsilon \tilde{W}_j^N} \mid \bar{w}_N^N, \ldots, \bar{w}_{j_1}^N \right] < C.$$
provided $K \geq 2\varepsilon^{-1} \log C$. Let $S_n = \sum_{i=1}^n X_i$. Now,

$$E[\varepsilon S_n] = \sum_b E[\varepsilon S_{n-1} = b] \cdot P[\varepsilon S_{n-1} = b]$$

$$= \sum_b b \cdot E[\varepsilon X_n | \varepsilon S_{n-1} = b] \cdot P[\varepsilon S_{n-1} = b]$$

$$\leq e^{-\varepsilon K/2} \sum_b b \cdot P[\varepsilon S_{n-1} = b] = e^{-\varepsilon K/2} E[\varepsilon S_{n-1}],$$

taking (5.15) into account. Iterating the last inequality yields

$$E[\varepsilon S_n] \leq e^{-\varepsilon Kn/2},$$

and since $P[S_n \geq 0] \leq E[\varepsilon S_n]$, the estimate (5.13) is proved.

The proof of (5.14) is the same, except that $\tilde{W}{N}_j, \ldots, \tilde{W}{N}_{j-1}$ are determined not just by the symbolic sequence, but by $\zeta(a_{N,j}), \ldots, \zeta(a_{N})$, hence by the future, as far as $\tilde{W}_j$ is concerned, so we have

$$E[\varepsilon \tilde{W}_j | \tilde{W}_j, \ldots, \tilde{W}_j] < C.$$

Proof of Proposition 5.4. First we show that (A1) holds $P$-almost surely. Pick an integer $K \geq 2\varepsilon^{-1} \log C$ and let $L_1 = K + 1$. To ensure (A1) we just need that

$$\tilde{k}^N_{\lfloor N/L_1 \rfloor} = N - \left( \ell_q(\sigma^N a) + \sum_{j=1}^{[N/L_1]} \bar{w}_j^N \right) > 0$$

for $N$ sufficiently large. We have

$$\left\{ \tilde{k}^N_{\lfloor N/L_1 \rfloor} \leq 0 \right\} \subset \left\{ \ell_q(\sigma^N a) \geq \lfloor N/L_1 \rfloor \right\} \cup \left\{ \sum_{j=1}^{[N/L_1]} \bar{w}_j^N \geq K \lfloor N/L_1 \rfloor \right\},$$

hence by Lemma 5.3 and (5.13),

$$P\left\{ \tilde{k}^N_{\lfloor N/L_1 \rfloor} \leq 0 \right\} \leq C e^{-\varepsilon([N/L_1]-1)} + e^{-\varepsilon K \lfloor N/L_1 \rfloor}/2.$$

Borel-Cantelli implies that $P$-almost surely, $\tilde{k}^N_{\lfloor N/L_1 \rfloor} > 0$ for all sufficiently large $N$, so (A1) holds.

Now let $\delta > 0$ and we will verify (A2). In view of (A1), we can assume that $k_n \leq nL_1$ for $n$ sufficiently large. Condition (5.9), Markov inequality, and Borel-Cantelli yield that $P$-almost surely, for all $N$ sufficiently large, we have

$$\bar{w}_1^N \leq \frac{\delta}{3L_1} N.$$
It remains to note that $w_n = \tilde{w}_1^N$ for $N = k_{n+1} + 1$, hence

$$w_n = \tilde{w}_1^N \leq \frac{\delta}{3L_1} (k_{n+1} + 1) \leq \frac{\delta}{3L_1} ((n + 1)L_1 + 1) \leq \delta n,$$

for all $n$ sufficiently large, as desired.

Now let us verify (A3). Consider the event

$$W(N, \delta, K) = \left\{ \max_{\Psi \subset \{1, \ldots, N\} \atop |\Psi| \leq \delta N} \sum_{n \in \Psi} \tilde{W}_n^N \geq K(\delta N) \right\}$$

Observe that

$$\left\{ \sum_{n \in \Psi} W_n : \Psi \subset \{n : k_{n+1} \leq N\} \right\} \subset \left\{ \sum_{n \in \Psi} \tilde{W}_n^N : \Psi \subset \{1, \ldots, N\} \right\}.$$

Indeed, the right-hand side counts the same sums backwards, over a larger collection of subsets. We are using here that $k_N \leq N$, so $\tilde{k}_N^N \leq 0$, and thus $\tilde{W}_n^N$, $n = 1, \ldots, N$, necessarily exhaust all of $W_n$, with $k_{n+1} \leq N$, in view of (5.1) and (5.8). Then we have for $K \geq 2 \log C/\varepsilon$,

$$\mathbb{P}(W(N, \delta, K)) \leq \sum_{\Psi \subset \{1, \ldots, N\} \atop |\Psi| \leq \delta N} \mathbb{P}\left[ \sum_{n \in \Psi} \tilde{W}_n^N \geq K(\delta N) \right] \leq \sum_{i \leq \delta N} \binom{N}{i} e^{-\varepsilon K(\delta N)/2},$$

in view of Lemma 5.5. By Stirling, there exists $C' > 1$ such that

$$(5.16) \sum_{i \leq \delta N} \binom{N}{i} \leq \exp [C'\delta \log(1/\delta)N] \text{ for } \delta < e^{-1} \text{ and all } N > 1.$$

Therefore,

$$\sum_{i \leq \delta N} \binom{N}{i} \leq \exp[-\varepsilon K(\delta N)/4] \text{ for } K = \frac{4C'}{\varepsilon} \log(1/\delta),$$

whence, by Borel-Cantelli, the event $W(N, \delta, L_2 \log(1/\delta))$ does not occur for all $N$ sufficiently large, with

$$L_2 = \varepsilon^{-1} \max(4C', 2 \log C),$$

which means that condition (A3) holds. \qed
6. Linear algebra and the choice of good return words

In this section, as well as the next one, we continue to work with a \( P \)-generic 2-sided sequence \( a \), whose positive half is exactly a \( P_+ \)-generic one-sided sequence.

Under the assumptions of Theorem 4.1 for \( P \)-a.e. \( a \), the sequence of substitutions \( \zeta_n(a) \), \( n \in \mathbb{Z} \), satisfies several conditions. First of all, we can assume that the point \( a \) is generic for the Oseledets Theorem; that is, assertions (i)-(iii) from Section 4 hold. In view of the assumption (a) of Theorem 4.1, we can fix unit basis vectors \( \vec{e}_j^{(n)} \), \( j = 1, 2 \), for the one-dimensional subspaces \( E_{\sigma_n}^{j} \), \( j = 1, 2 \), \( n \geq 0 \), such that

\[
(S^n[n])^{t} \vec{e}_j^{(0)} = A(n,j)\vec{e}_j^{(n)} \quad \text{for some} \quad A(n,j) > 0.
\]

By (ii) in Oseledets Theorem, we have

\[
\frac{1}{n} \log A(n,j) \to \theta_j, \quad j = 1, 2.
\]

We start with an observation about the Lyapunov-Oseledets basis \( \{\vec{e}_j^{(n)}\}_{j=1}^{2} \) of the unstable subspace. Transpose substitution matrices \( S^t \) are non-negative. Thus,

\[
\| (S^n[n])^{t} \vec{e} \| \leq \| (S^n[n])^{t} \vec{e}^{(n)} \|,
\]

hence \( \vec{e}_1^{(n)} \in \mathbb{R}_+^m \) (the positive cone) for all \( n \geq 0 \), and so

\[
(6.2) \quad \vec{e}_1^{(n)} \in S^t_n \mathbb{R}_+^m.
\]

On the other hand, under our assumptions the image of the positive cone \( (S^n[n])^{t} \mathbb{R}_+^m \) shrinks to a single direction exponentially fast. (The fact that the cone shrinks to a single direction is equivalent to unique ergodicity, see (5.2) and Veech [22, 23, 24, 25].) It follows that the basis vector \( \vec{e}_2^{(n)} \) does not lie in \( \mathbb{R}_+^m \) for all \( n \) (otherwise we would get a contradiction with (iii) in Oseledets Theorem). Combined with (6.2), this implies that if \( S_n = Q \) (a strictly positive matrix), then the angle between \( \vec{e}_1^{(n)} \) and \( \vec{e}_2^{(n)} \) is bounded away from zero by a constant depending only on \( Q \).

We will need an elementary fact from linear algebra.

**Lemma 6.1.** Let \( B = \{\vec{x}_j\}_{j \leq m} \) be a basis of \( \mathbb{R}^m \), and let \( \{\xi_1, \ldots, \xi_r\} \subset \mathbb{R}^m \) be a linearly independent set, with \( r \leq m \). Then there exists a subset \( \{x_i\}_{i \in I} \subset B \) of cardinality \( r \) such that

\[
|D_I| := \left| \det \left( \langle \vec{x}_i, \vec{\xi}_j \rangle \right)_{i \in I, j \leq r} \right| \geq C_B \| \vec{\xi}_1 \wedge \cdots \wedge \vec{\xi}_r \|,
\]

where \( C_B \) depends only on the basis \( B \).

**Proof.** Let \( T \) be the linear isomorphism which takes the standard basis \( \{\vec{e}_j\}_{j \leq m} \) of \( \mathbb{R}^m \) into \( B \). Then

\[
D_I = \det \left( \langle \vec{x}_i, \vec{\xi}_j \rangle \right)_{i \in I, j \leq r} = \det \left( \langle T\vec{e}_i, \vec{\xi}_j \rangle \right)_{i \in I, j \leq r} = \det \left( \langle \vec{e}_i, T^* \vec{\xi}_j \rangle \right)_{i \in I, j \leq r}
\]
The latter determinant is the order-\(r\) minor of the matrix whose columns are \(\mathcal{T}^*\tilde{\xi}_j\), \(j = 1, \ldots, r\), corresponding to the rows indexed by \(I\). Thus,

\[
\sum_{\#I = r} |D_I|^2 = \|\mathcal{T}^*\tilde{\xi}_1 \wedge \cdots \wedge \mathcal{T}^*\tilde{\xi}_r\|^2 \geq \|(\Lambda^r \mathcal{T}^*)^{-1}\|^{-1} \|\tilde{\xi}_1 \wedge \cdots \wedge \tilde{\xi}_r\|^2.
\]

We are using here that \(\mathcal{T}^*\) is invertible, hence its exterior power is invertible. Thus,

\[
\max\{ |D_I| : \#I = r \} \geq \|(\Lambda^r \mathcal{T}^*)^{-1}\|^{-1/2} \left(\frac{m}{r}\right)^{-1/2} \|\tilde{\xi}_1 \wedge \cdots \wedge \tilde{\xi}_r\|,
\]

and the proof is complete. \qed

We now return to our theorem, in which \(r = 2\). Let \(\{u_j\}_{j \leq m}\) be the good return words from the Assumption (c) of Theorem 4.1. We will choose a sequence of words \(v_n \in \{u_j\}_{j \leq m}\), depending on our generic \(a \in \Omega\). We let \(n = k_n(a) \uparrow \infty\) be such that \(S(\zeta_{k_n}) = S(\zeta_{1+k_n}) = \mathcal{Q}\). For \(i \leq m\) consider

\[
\Theta_n := \begin{pmatrix}
A(k_n, 1)\langle \vec{\ell}(v_n), e^{(k_n)}_1 \rangle & A(k_n, 2)\langle \vec{\ell}(v_n), e^{(k_n)}_2 \rangle \\
A(k_n+1, 1)\langle \vec{\ell}(v_{n+1}), e^{(k_n+1)}_1 \rangle & A(k_n+1, 2)\langle \vec{\ell}(v_{n+1}), e^{(k_n+1)}_2 \rangle
\end{pmatrix}.
\]

Below \(C_{\zeta}\) denotes a constant depending only on the substitution \(\zeta\).

**Lemma 6.2.** For \(\mathbb{P}\)-a.e. \(a \in \Omega\) we can choose the words \(v_n \in \{u_j\}_{j \leq m}\), so that for all \(n \geq 1\),

\[
\|\Theta_n^{-1}\|_{\infty} \leq C_{\zeta} \cdot \frac{\max\{A(k_{n+j}, i) : j = 0, 1; i = 1, 2\}}{A(k_n, 1)A(k_n+1, 2)}
\]

and

\[
\|\Theta_{n+1}\Theta_n^{-1}\|_{\infty} \leq C_{\zeta} \cdot \frac{\max_{j=0,1,2} A(k_{n+j}, 1)}{A(k_n, 1)} \cdot \frac{\max_{j=0,1,2} A(k_{n+j}, 2)}{A(k_n, 2)}.
\]

**Proof.** We are going to choose \(v_n\) inductively. Pick \(v_1\) arbitrarily, and suppose \(v_1, \ldots, v_n\) have been chosen. For \(i \leq m\) consider

\[
\Delta_i := \begin{vmatrix}
A(k_n, 1)\langle \vec{\ell}(v_n), e^{(k_n)}_1 \rangle & A(k_n, 2)\langle \vec{\ell}(v_n), e^{(k_n)}_2 \rangle \\
A(k_n+1, 1)\langle \vec{\ell}(u_i), e^{(k_n+1)}_1 \rangle & A(k_n+1, 2)\langle \vec{\ell}(u_i), e^{(k_n+1)}_2 \rangle
\end{vmatrix}.
\]

Observe that

\[
\det \begin{pmatrix}
\langle \vec{\ell}(u_i), e^{(k_n+1)}_1 \rangle & \Delta_i \\
\langle \vec{\ell}(u_j), e^{(k_n+1)}_1 \rangle & \Delta_j
\end{pmatrix} = A(k_n, 1)A(k_n+1, 2)\langle \vec{\ell}(v_n), e^{(k_n)}_1 \rangle D_{ij},
\]

where

\[
D_{ij} := \begin{vmatrix}
\langle \vec{\ell}(u_i), e^{(k_n+1)}_1 \rangle & \langle \vec{\ell}(u_i), e^{(k_n+1)}_2 \rangle \\
\langle \vec{\ell}(u_j), e^{(k_n+1)}_1 \rangle & \langle \vec{\ell}(u_j), e^{(k_n+1)}_2 \rangle
\end{vmatrix}.
\]
Note that \( \xi_1 := e_{1}^{(k_n+1)} \in Q' \mathbb{R}^m \) and \( \xi_2 := e_{2}^{(k_n+1)} \not\in \mathbb{R}^m \) by the assumption \( S(\zeta_{k_n+1}) = Q \) and the comments above. Thus, the angle between \( \xi_1 \) and \( \xi_2 \) is bounded away from zero, uniformly in \( n \). Hence we can apply Lemma 6.1 to these vectors and find \( i \neq j \) such that
\[
|D_{ij}| \geq c_3 > 0,
\]

independent of \( n \). Note that for all \( i \leq m \),
\[
0 < c_4 \leq |\langle \ell(u_i), e_{i}^{(k_n)} \rangle| \leq C_5 := \max_{i \leq m} \|\ell(u_i)\|_2 < \infty
\]

for some positive constant \( c_4 = c_4(Q) \) independent of \( n \), since \( \ell(u_i), i \leq m \), are positive integer vectors. It follows from (6.6) and (6.7) that
\[
\max_i |\Delta_i| \geq \frac{c_4 c_5}{2} A(k_n, 1) A(k_{n+1}, 2).
\]

We choose \( v_{n+1} \in \{u_i\}_{i \leq m} \) to maximize \( |\Delta_i| \). Denote \( \Delta^{(n)} = \det(\Theta_n) \). As a result, we will have for all \( n \geq 1 \):
\[
|\Delta^{(n)}| \geq \frac{c_4 c_5}{2} A(k_n, 1) A(k_{n+1}, 2),
\]

which implies (6.4). Also, a direct calculation, combined with (6.7) and (6.8), yields (6.5). \( \square \)

**Corollary 6.3.** For \( \mathbb{P} \)-a.e. \( a \in \Omega \) we can choose the words \( v_n \in \{u_j\}_{j \leq m} \), so that for any \( \delta_1 > 0 \) there exists \( n_0 \in \mathbb{N} \) such that for all \( n \geq n_0 \),
\[
\|\Theta_n^{-1}\|_\infty \leq C_\zeta \exp[-(\theta_2 - \delta_1)k_n]
\]

and
\[
\|\Theta_{n+1}\Theta_n^{-1}\|_\infty \leq C_\zeta \exp[2(W_n + W_{n+1})],
\]

where \( C_\zeta \) is the constant from Lemma 6.2 and \( W_n \) are defined in (5.1).

**Proof.** This is a combination of the last lemma, Oseledets Theorem, and Proposition 5.1. First we prove (6.9). By Oseledets Theorem, for \( \mathbb{P} \)-a.e. \( a \in \Omega \), for all \( n \) sufficiently large,
\[
\exp[(\theta_i - \delta_1/4)k_n] \leq A(k_n, i) \leq \exp[(\theta_i + \delta_1/4)k_n], \quad i = 1, 2.
\]

We will use (6.4), where clearly the maximum in the numerator is (eventually) attained for \( i = 1 \), to obtain for \( n \) sufficiently large, keeping in mind that \( w_n = k_{n+1} - k_n \):
\[
\|\Theta_n^{-1}\|_\infty \leq C_\zeta \exp[(\theta_1 + \delta_1/4)k_{n+1} - (\theta_1 - \delta_1/4)k_n - (\theta_2 - \delta_1/4)k_{n+1}]
= C_\zeta \exp[(\theta_1 - \theta_2 + \delta_1/2)w_n - (\theta_2 - 3\delta_1/4)k_n].
\]

Using (A2) from Proposition 5.1 we can ensure that for \( n \) sufficiently large,
\[
(\theta_1 - \theta_2 + \delta_1/2)w_n \leq (\delta_1/4)k_n,
\]
and (6.9) follows. Next, let us verify (6.10). Equation (6.1) implies that

\[ A(k_{n+1}, j) e_j^{(k_{n+1})} = S_{k_{n+1}}(a) \cdots S_{k_1+1}(a) A(k_n, j) e_j^{(k_n)}, \quad j = 1, 2, \]

hence \( A(k_{n+1}, j) / A(k_n, j) \leq e^{W_n} \) by (5.1). Now (6.10) follows from (6.5). \( \square \)

7. Continuation of the proof of Theorem 4.1

Fix a \( \mathbb{P} \)-generic point \( a \) and the corresponding sequence of substitutions \( \zeta_n \), as in the last section. We assume that the conclusions of Proposition 5.1 hold. Without loss of generality, assume that all the entries of \( Q \) are at least 2, where \( Q = S_\zeta \). Further, denote by \( K \) the set \( \{ k_n \}_{n=1}^\infty \), that is, \( i \in K \) iff \( \zeta_i = \zeta_{i+1} = \zeta \).

**Proposition 7.1.** For \( \mathbb{P} \)-a.e. \( a \in \Omega \), for any \( \eta \in (0, 1) \), there exists \( \ell_\eta = \ell_\eta(a) \in \mathbb{N} \), such that for all \( \ell \geq \ell_\eta \) and any bounded cylindrical function \( f^{(\ell)} \) of level \( \ell \), for any \( (\vec{r}, t) \in X^\ell \), with \( \vec{r} \in X_+(a) \), and \( \omega \in \mathbb{R} \), assuming \( \ell \in K \),

\[
|S_{\vec{R}}^{(\ell)}(f^{(\ell)}, \omega)| \leq \overline{C}_2 \|f^{(\ell)}\|_\infty \left( R^{1/2} + R^{1+\eta} \prod_{\ell+1 \leq k_p < \log R \over 1} (1 - c_1 \cdot \max_{v \in GR(\zeta)} ||\zeta^{(k_p)}(v)||_2^2) \right),
\]

for all \( R \geq e^{8\theta_1 \ell} \), where \( \overline{C}_2 \) depends only on \( Q \).

**Proof.** Without loss of generality we can assume that \( f^{(\ell)} = f^{(\ell)}_a \) for some \( a \in A \), as in (3.25) and find \( \vec{r}' \) and \( t' \) as in (3.24). Since \( (\vec{r}, t) = h_{t'}(\vec{r}', 0) \), we have

\[
|S_{\vec{R}}^{(\ell)}(f^{(\ell)}_a, \omega)| = \left| \int_0^R e^{-2\pi i \omega \tau} f^{(\ell)}_a(\vec{r}_\tau, 0) \, d\tau \right| = \left| \int_{t'}^{t'+t} e^{-2\pi i \omega \tau} f^{(\ell)}_a(\vec{r}, 0) \, d\tau \right|.
\]

Recall that \( \vec{s}^{(\ell)} = (S^{(\ell)})^{(\ell)} \vec{s} \), and we let \( s^{(\ell)}_{\max} \) and \( s^{(\ell)}_{\min} \) be the maximal and minimal components of the vector \( \vec{s}^{(\ell)} \), respectively. Note that \( |t'| \leq s^{(\ell)}_{\max} \), so we obtain

\[
|S_{\vec{R}}^{(\ell)}(f^{(\ell)}_a, \omega)| = \left| \int_0^R e^{-2\pi i \omega \tau} f^{(\ell)}_a(\vec{r}_\tau, 0) \, d\tau \right| \leq 2 \|f^{(\ell)}\|_\infty s^{(\ell)}_{\max}.
\]

Next, consider \( x^{(\ell)} \in Y^{(\ell)} \) as in (3.26) and take the maximal \( N \) such that \( R' := |x^{(\ell)}[0, N-1]|_{\vec{s}^{(\ell)}} \leq R \). Then \( |R - R'| \leq s^{(\ell)}_{\max} \), hence

\[
|S_{\vec{R}}^{(\ell)}(f^{(\ell)}_a, \omega) - S_{\vec{R}}^{(\ell,0)}(f^{(\ell)}_a, \omega)| \leq 2 \|f^{(\ell)}\|_\infty s^{(\ell)}_{\max},
\]

and for \( S_{\vec{R}}^{(\ell)}(f^{(\ell)}_a, \omega) \) the formula in (3.26) applies (with \( R \) replaced by \( R' \)). Thus, the combined error in the above estimates (7.2), (7.3) is bounded by \( 3 \|f^{(\ell)}\|_\infty s^{(\ell)}_{\max} \). By Oseledets, we can make sure that \( \ell_\eta \) is such that

\[
\ell^{-1} \log \|\vec{s}^{(\ell)}\|_1 - \theta_1 \leq \theta_1 \ell / 10, \quad \text{for all } \ell \geq \ell_\eta.
\]
Then
\[ s_{\max}^{(\ell)} \leq \|S^{(\ell)}\|_1 \leq e^{\theta_1 \ell(1+\eta/10)} \leq e^{\theta_1 \ell-1.1} < R^{1/2}, \]
for \( \ell \geq \ell_\eta \) and \( R \geq e^{\theta_1 \ell} \). Taking \( \tilde{C}_2 \geq 3 \), we have accounted for the first (trivial) term in the right-hand side of (7.1). Thus it suffices to consider the case of (3.26). Since \( R' \leq R \) and \( |R - R'| \leq s_{\max}^{(\ell)} < R^{1/2} \), and \( (R - R')^2 \geq R^{3/4} \) for \( R \geq 9 \), the proposition will follow from the following lemma. \( \square \)

**Lemma 7.2.** For \( \mathbb{P} \)-a.e. \( a \in \Omega \), for any \( \eta \in (0,1) \), there exists \( \ell_\eta = \ell_\eta(a) \in \mathbb{N} \), such that for all \( \ell \geq \ell_\eta \) and any bounded cylindrical function \( f^{(\ell)} \) of level \( \ell \), for any \( \bar{\tau} \in X_+ (a) \) such that \( b(\bar{\tau}) = \zeta^{[\ell]}(x^{[\ell]}) \), with \( x^{(\ell)} \in Y^{[\ell]} \) and \( \omega \in \mathbb{R} \), assuming \( \ell \in \mathbb{K} \),
\[
|S_R^{(\bar{\tau},0)}(f^{(\ell)}, \omega)| \leq \tilde{C}_2 \|f^{(\ell)}\|_\infty R^{1+\eta} \prod_{\ell+1 \leq k_p < \frac{\log R}{\theta_1}} (1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta^{[k_p]}(v)|x\|^2),
\]
where \( \tilde{C}_2 \) depends only on \( Q \), whenever
\[
R = |x^{(\ell)}[0,N - 1]|_{\pi^R} \geq e^{\theta_1 \ell}.\]

**Proof.** Again we can assume without loss of generality that \( f^{(\ell)} = f_a^{(\ell)} \), where \( f_a^{(\ell)}(\bar{\tau}, t) = \mathbb{I}_{x^{(\ell)}_a}(\psi_a(t')) \), with \( t' \in [0, s_a^{(\ell)}] \). Recall the formula (3.26) which applies here:
\[
\tilde{S}_R^{(\bar{\tau},0)}(f_a^{(\ell)}, \omega) = \tilde{\psi}_a(\omega) \cdot \Phi_a^{(\ell)}(x^{(\ell)}[0,N - 1], \omega).
\]
First observe that
\[
|\tilde{\psi}_a(\omega)| \leq \|\psi_a\|_1 \leq \|\psi_a\|_{\infty} s_a^{(\ell)} \leq \|f_a^{(\ell)}\|_\infty s_{\max}^{(\ell)}.
\]
Next we apply (3.28):
\[
|\Phi_a^{(\bar{\tau})}(x^{(\ell)}[0,N - 1], \omega)| \leq 2 \sum_{j=\ell}^{n} \|S_{j+1}^{(\ell+1,j)}\|_1 \cdot \prod_{\ell+1 \leq k_p \leq j-1} (1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta^{[k_p]}(v)|x\|^2),
\]
where
\[
\min_{b \in A} |\zeta^{[\ell+1,n]}(b)| \leq N \leq 2 \max_{b \in A} |\zeta^{[\ell+1,n+1]}(b)|.
\]
By (4.2), we can assume that
\[
\|S_{j+1}\|_1 \leq e^{\theta_1 \eta/10} \quad \text{for all } j \geq \ell_\eta.
\]

**Claim.** Suppose that \( \zeta \) and \( \zeta' \) are two substitutions on \( A \), such that \( S_\zeta = Q \) has all entries \( \geq 2 \). Then for all \( a, b \in A \) we have \( |\zeta' \circ \zeta(a)| \geq 2 |\zeta'(b)| \).

**Proof** of the Claim is immediate, since for all \( a, b \in A \) the word \( \zeta(a) \) contains at least two letters \( b \) by assumption. \( \square \)
It follows from the Claim that

\begin{equation}
\|S^{[\ell+1, j]}\|_1 \leq 2\|S^{[\ell+1, \tilde{\ell}]}\|_1, \quad \text{where } \tilde{\ell} = \min\{i \in \mathcal{K} : i > j\}.
\end{equation}

Let \(\tilde{n} = k_q\). Recall that \(\ell \in \mathcal{K}\), so \(\ell = k_i\) for some \(i\). We have from (7.7), (7.9), and (7.10):

\begin{equation}
|\Phi_\alpha^{(\ell)}(x^{(\ell)}[0, N - 1], \omega)| \leq 4 \sum_{p=i}^{q-1} \|S^{[\ell+1, k_{p+1}]}\|_1 \cdot \sum_{j=k_{p+1}}^{k_{p+1}} e^{\eta/10} \times \\
\times \prod_{j=i}^{p} \left(1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta_j(v)|_x\|_s^2\right) \\
(7.11) \leq 4 \sum_{p=i}^{q-1} \|S^{[\ell+1, k_{p+1}]}\|_1 \cdot k_{p+1} e^{k_{p+1} \eta/10} \times \\
\times \prod_{j=i}^{p} \left(1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta_j(v)|_x\|_s^2\right),
\end{equation}

estimating trivially the “inner sum” from above. Now observe that \(\|S^{[\ell+1, k_{p+2}]}\|_1/\|S^{[\ell+1, k_{p+1}]}\|_1 \geq 2\) by (7.10), obviously \(k_{p+2} e^{k_{p+2} \eta/10}/k_{p+1} e^{k_{p+1} \eta/10} > 1\), and

\[1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta_j(v)|_x\|_s^2 \geq 3/4,\]

since \(c_1 \leq 1/4\) by (3.9). It follows that the ratio of consecutive terms in (7.11) is at least 3/2, so that the last term dominates, and we can estimate the sum from (7.11) by a geometric series with ratio 2/3 times the last term to obtain

\begin{equation}
|\Phi_\alpha^{(\ell)}(x^{(\ell)}[0, N - 1], \omega)| < 12 \|S^{[\ell+1, \tilde{n}]}\|_1 \cdot \tilde{n} e^{\eta \tilde{n}/10} \cdot \prod_{\ell+1}^{n-1} \\
\end{equation}

where

\[\prod_{\ell+1}^{n-1} := \prod_{\ell+1 \leq k_p \leq n-1} \left(1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta_j(v)|_x\|_s^2\right) = \prod_{j=i}^{q-1} \left(1 - c_1 \cdot \max_{v \in GR(\zeta)} \|\omega|\zeta_j(v)|_x\|_s^2\right).\]

Recall that \(\tilde{k}_q^n = \max\{j < n : j \in \mathcal{K}\} = k_{q-1}\). In view of the properties (A1) and (A2) from Proposition 5.1, we can assume that

\begin{equation}
\tilde{n} - \tilde{k}_q^n \leq n(\eta/10) \quad \text{for } n > \ell \geq \ell_\eta.
\end{equation}

Also, by Remark 5.2 and the definition of \(W_{q-1} = \log \|S^{[\tilde{k}_q^n + 1, \tilde{n}]}\|_1\), we can assume without loss of generality that

\begin{equation}
\|S^{[\tilde{k}_q^n + 1, \tilde{n}]}\|_1 \leq e^{n \theta_1 (\eta/10)} \quad \text{for } n > \ell \geq \ell_\eta.
\end{equation}
Thus we can estimate
\[(7.15) \quad \|S^{[\ell+1,\bar{n}]}\|_1 \leq \|S^{[\ell+1,\bar{k}_1^\eta]}\|_1 \leq \|S^{[\ell+1,\bar{k}_1^\eta]}\|_1 \leq e^{n\theta_1(\eta/10)}.
\]
Also,
\[
\hat{n}e^{\hat{n}\theta_1(\eta/10)} \leq n(1 + \eta/10)e^{n(1+\eta/10)\theta_1(\eta/10)} \leq e^{n\theta_1(\eta/5)},
\]
for \(n > \ell > \ell_\eta\) (everything depends on \(\theta_1\) and other Lyapunov exponents of course). Before we proceed further, let us collect what we have so far:
\[
(7.16) \quad |S_R^{[\ell,0]}(f^{(\ell)}_a, \omega)| \leq 12\|f^{(\ell)}_a\|_\infty \cdot s^{(\ell)}_{\max} \cdot \|S^{[\ell+1,\bar{k}_1^\eta]}\|_1 \cdot e^{n\theta_1(3\eta/10)} \cdot \prod(\ell + 1, n - 1).
\]

Next we need to relate \(N, n, \) and \(R\). First observe that
\[
(7.17) \quad R = \max|\{x^{(\ell)}[0, N - 1]\}|_{s^{(\ell)}} \in [N s^{(\ell)}_{\min}, N s^{(\ell)}_{\max}].
\]
Note that
\[
(7.18) \quad s^{(\ell)}_{\max} \leq C_Q s^{(\ell)}_{\min}, \quad \text{with} \quad C_Q = \max_{i,j} Q_{ij} / \min_{i,j} Q_{ij},
\]
since \(s^{(\ell)}_{\max}\) and \(s^{(\ell)}_{\max}\) are the maximal and minimal, components of \(s^{(\ell)}_t = (S^{[\ell]}_t)^t s = Q^t (S^{[\ell-1]}_t)^t s\) respectively (recall also that \(S_\ell = Q\) by assumption). We have by \((7.8)\), for any \(b \in A\),
\[
(7.19) \quad N \geq \max_{a \in A} |\zeta^{[\ell+1,\eta]}(b)| \geq \max_{a \in A} |\zeta^{[\ell+1,\bar{k}_1^\eta]}(a)| = \|S^{[\ell+1,\bar{k}_1^\eta]}\|_1,
\]
using that \(\zeta^{[\ell+1,\eta]}(b) = \zeta^{[\ell+1,\bar{k}_1^\eta]} \circ \zeta^{[\bar{k}_1^\eta+1]} \circ \zeta^{[\bar{k}_1^\eta+2,\eta]}(b)\) and \(\zeta^{[\bar{k}_1^\eta+1]} = \zeta\), for which the substitution matrix \(Q\) is strictly positive. Therefore, by \((7.17)\), \((7.18)\), and \((7.19)\),
\[
(7.20) \quad \|S^{[\bar{k}_1^\eta]}\|_1 \leq \|S^{[\ell]}\|_1 \cdot \|S^{[\ell+1,\bar{k}_1^\eta]}\|_1 \leq C_Q R,
\]
By \((7.4)\) and \((7.13)\),
\[
\|S^{[\bar{k}_1^\eta]}\|_1 \geq e^{\bar{k}_1^\eta \theta_1(1-\eta/10)} \geq e^{n\theta_1(1-\eta/10)^2} \geq e^{n\theta_1(1-\eta/5)}.
\]
Combining this with \((7.20)\) yields
\[
e^{n\theta_1(3\eta/10)} \leq (C_Q R)^{(3\eta/10)(1-\eta/5)^{-1}} \leq (C_Q R)^\eta \leq C_Q R^{1+\eta}, \quad \text{for} \quad \eta \in (0, 1)
\]
(with a large “margin”), so that \((7.16)\) becomes, taking the second inequality of \((7.20)\) into account:
\[
(7.21) \quad |S_R^{[\ell,0]}(f^{(\ell)}_a, \omega)| \leq 12\|f^{(\ell)}_a\|_\infty C_Q^2 R^{1+\eta} \prod_{\ell+1}^{n-1}
\]
In order to conclude the proof of \((7.1)\), it remains to show that
\[
(7.22) \quad n \geq \frac{\log R}{3\theta_1}.
\]
By the upper bound in \((7.8)\) and \((7.4)\),
\[
(7.23) \quad N \leq 2\|S^{[\ell+1,n+1]}\|_1 \leq 2\|S^{[n+1]}\|_1 \leq 2e^{(n+1)\theta_1(1+\eta/10)} \leq e^{2n\theta_1} \quad \text{for} \quad n \geq \ell \geq \ell_\eta.
\]
The inequality \( \|S_{n+1}\|_1 \leq \|S_{n+1}\|_1 \) holds, since
\[
\|S_{n+1}\|_1 = \max_{b \in A} |\zeta_{n+1}(b)| = \max_{b \in A} |\zeta_{\ell+1}(b)|,
\]
and \( \zeta_{\ell} \) does not decrease the length of words. By (7.4),
\[
\max_{\ell} \leq e^{\theta_1(1+\eta/10)} \leq e^{2\theta_2} \quad \text{for } \ell \geq \ell_\eta.
\]
Since \( R \geq e^{6\theta_1} \) by the assumption of our lemma, we have
\[
N \geq R/s_{\max} \geq e^{4\theta_1}.
\]
On the other hand, \( N \geq e^{2n\theta_1} \) by (7.23), hence \( n \geq 2\ell \) and
\[
R \leq Ns_{\max} \leq e^{2n\theta_1+2\theta_1} \leq e^{3n\theta_1},
\]
confirming (7.22). The lemma, and hence the proposition, is proved completely. \( \square \)

Now we proceed with the proof of Theorem 4.1. Fix the sequence of good return words \( v_n \) from Lemma 6.2. Under the assumptions of Theorem 4.1,
\[
(7.24) \quad \vec{s} = \sum_{j=1}^{2} a_j \vec{e}_j^{(0)} + P_s \vec{s},
\]
where \( P_s \) is the projection to the stable subspace \( E_3^a \oplus \cdots \oplus E_r^a \) in (4.4). A statement “for a.e. \( \vec{s} \)” will follow from a statement “for a.e. \( (a_1, a_2) \)”, and since we can rescale and normalize, we can assume that \( a_1 = 1 \), so that
\[
a_2 = P_{a}^{2}(\vec{s})/P_{a}^{1}(\vec{s}).
\]
For \( n \in \mathbb{N} \), let
\[
(7.25) \quad \omega|\zeta_{\ell}(v_n)|_{\vec{s}} = K_{k_n} + \varepsilon_{k_n}, \quad \text{where } K_{k_n} \in \mathbb{N}, \ |\varepsilon_{k_n}| \leq 1/2,
\]
so that
\[
\| \omega|\zeta_{\ell}(v_n)|_{\vec{s}} \| = |\varepsilon_{k_n}|.
\]
We should keep in mind that \( K_{k_n} \) and \( \varepsilon_{k_n} \) depend on \( \omega \in [B^{-1}, B] \) and on \( \vec{s} \), although this is suppressed in notation to avoid clutter. Given positive \( \gamma \) and \( \delta \), define
\[
E_N(\gamma, \delta, B) := \left\{ a_2 : |a_2| \in [\beta^{-1}, \beta], \ \exists \omega \in [B^{-1}, B] : \text{card}\{ n \leq N : |\varepsilon_{k_n}| \geq \gamma \} < \delta N \right\},
\]
and
\[
\mathcal{E}(\gamma, \delta, B) := \limsup E_N(\gamma, \delta, B) = \bigcap_{N_0=1}^{\infty} \bigcup_{N=N_0}^{\infty} E_N(\gamma, \delta, B).
\]

Proposition 7.3. There exists \( \gamma > 0 \) such that for \( \mathbb{P}_+ - \text{a.e. } a_+ \in \Omega_+ \) we have
\[
\forall \epsilon > 0, \ \exists \delta_0 > 0, \ \forall \beta > 0, \ \forall B > 0 : \quad \delta < \delta_0 \quad \Rightarrow \quad \dim_H(\mathcal{E}(\gamma, \delta, B)) < \epsilon.
\]
In the remaining part of this section, we derive Theorem 4.1 from Proposition 7.3. Then in the next section we use the “Erdős-Kahane argument” to prove the proposition.

Proof of Theorem 4.1 assuming Proposition 7.3. In view of Lemma 3.3 it suffices to show

\[ |S^{(T)}_R(f, \omega)| \leq O(1) \cdot R^{1-\eta}, \quad R \geq R_0, \]

for some \( \eta \in (0, 1) \), uniformly in \((\mathbf{r}, t) \in \mathbb{X}^\mathbb{Z}\). We will specify \( \eta \) at the end of the proof, see (7.30).

Since \( f \) is weakly Lipschitz on \( \mathbb{X}^\mathbb{Z} \) (see Section 2.1), for almost every \( a_+ \), we can approximate \( f \), for any \( \ell \in \mathbb{N} \), by a function \( f^{(\ell)} \), which is cylindrical of level \( \ell \), and has sup-norm bounded by \( \|f\|_\infty \), so that

\[ \|f - f^{(\ell)}\|_\infty \leq \|f\|_L \cdot \nu_+([e_1 \ldots e_n]). \]

We can do this simply taking \( f^{(\ell)}(\mathbf{r}, t) := f(\mathbf{r}^{(\ell)}, t) \), where \( \mathbf{r}^{(\ell)} \) agrees with \( \mathbf{r} \) down to level \( \ell \) after which it is extended to infinity in any fixed way. By (2.3), (2.3), and (2.2), we have

\[ \lim_{n \to \infty} \frac{\log \nu_+([e_1 \ldots e_n])}{n} = -\theta_1, \]

\( \mathbb{P} \)-almost surely, hence for \( \ell \) sufficiently large we have

\[ \|f - f^{(\ell)}\|_\infty \leq \|f\|_L \cdot e^{-\frac{1}{2} \theta_1 \ell}. \]

Recall that \( S^{(T)}_R(f, \omega) = \int_0^R e^{-2\pi \omega t} f \circ h_T(\mathbf{r}, t) \, d\tau \). Let \( \kappa = \frac{2 \log R}{\theta_1} \), so that \( R = e^{\frac{1}{2} \theta_1 \kappa} \). By the property (A2), we can make sure that for \( R \geq R_0 \) there exists \( \ell \in \mathcal{K} \) (that is, \( \ell = k_n \) for some \( n \)) satisfying

\[ \eta \kappa \leq \ell \leq 2 \eta \kappa. \]

Then (7.27) yields

\[ |S^{(T)}_R(f, \omega) - S^{(T)}_R(f^{(\ell)}, \omega)| \leq R \cdot \|f\|_L \cdot e^{-\frac{1}{2} \theta_1 \ell} = O(1) \cdot e^{\frac{1}{2} \theta_1 (\kappa - \ell)} \leq O(1) \cdot e^{\frac{1}{2} \theta_1 \kappa (1 - \eta)} = O(1) \cdot R^{1-\eta}. \]

Thus, it is enough to obtain (7.26) for \( R = e^{\frac{1}{2} \theta_1 \kappa} \), with \( f \) replaced by \( f^{(\ell)} \). Therefore, we can apply Proposition 7.1. Recall the inequality (7.1), using the sequence of good return words \( \{v_n\} \):

\[ |S^{(T)}_R(f^{(\ell)}, \omega)| \leq \widetilde{C}_2 \|f^{(\ell)}\|_\infty \left( R^{1/2} + R^{1+\eta} \prod_{\ell + 1 \leq k_n \leq \log R \theta_1} (1 - c_1 \cdot \|\omega \| \varsigma^{[k_n]}(v_n)| \|_2)^2) \right), \]

for \( \ell \geq \ell_\eta \) and all \( R \geq e^{8\theta_1 \ell} \). We can ensure that \( \ell \geq \ell_\eta \) by taking \( R_0 \) sufficiently large, and \( R = e^{\frac{1}{2} \theta_1 \kappa} \geq e^{8\theta_1 \ell} \) will follow if \( \eta \leq 1/32 \), by (7.28). Since our goal is (7.26), we can discard the \( R^{1/2} \) term immediately. Now choose \( \gamma > 0 \) and \( \delta > 0 \) from Proposition 7.3 such that \( \dim_H(\mathcal{E}(\gamma, \delta, \beta, B)) < 1 \) for all \( \beta, B > 0 \). It is enough to verify

\[ \prod_{\ell + 1 \leq k_n \leq \log R \theta_1} (1 - c_1 \cdot \|\omega \| \varsigma^{[k_n]}(v_n)| \|_2)^2) \leq O(1) \cdot R^{-2\eta}, \quad \text{with} \quad R = e^{\frac{1}{2} \theta_1 \kappa} \geq R_0(\beta, B), \]

for all \( \ell \geq \ell_\eta \) and all \( R \geq e^{8\theta_1 \ell} \).
for all vectors $\vec{s}$, for which $a_2 \in [\beta^{-1}, \beta] \setminus \mathcal{E}(\gamma, \delta, B)$, with $|\omega| \in [B^{-1}, B]$, thus obtaining an even stronger than ‘almost every $\vec{s}$’ statement.

By definition, $a_2 \notin \mathcal{E}(\gamma, \delta, B)$ means $a_2 \notin E_N(\gamma, \delta, B)$ for all $N \geq N_0$, for some $N_0 \in \mathbb{N}$, depending on $\beta$ and $B$. Let

$$N = \left\lfloor \frac{\kappa}{4L_1} \right\rfloor \leq \frac{\log R}{L_1(4\theta_1)}.$$ 

By taking $R$ large enough, we can ensure that $N \geq N_0$. Recall that $k_N \leq L_1N$ for $N \geq n_0(a)$ by the condition (A1), so the product in (7.29) is less than or equal to

$$\prod_{n=\ell+1}^{N} (1 - c_1|\varepsilon_{k_n}|^2),$$

where we also use (7.25). By definition, $a_2 \notin E_N(\gamma, \delta, B)$ means that there are at least $\delta N$ numbers $n \in \{1, \ldots, N\}$ with $|\varepsilon_{k_n}| \geq \gamma$, hence the left-hand side of (7.29) is $\leq (1 - c_1\gamma^2)^{\delta N - \ell}$. Recalling that $N \geq \kappa/(4L_1) - 1$ and $\ell \leq 2\eta\kappa$, we see that it suffices to obtain

$$(1 - c_1\gamma^2)^{\kappa\frac{\delta}{16L_1} - 2\eta} \leq e^{\theta_1\kappa(-\eta)},$$

and this is ensured if we take

$$(7.30) \quad \eta = \min \left\{ \frac{\delta}{16L_1}, \frac{-\log(1 - c_1\gamma^2)}{8L_1\theta_1} \right\}.$$ 

The proof is complete. \hfill \Box

8. Erdős-Kahane method: Conclusion of the proof of Theorem 4.1

After some preparation, we prove here Proposition 7.3 which concludes the proof of Theorem 4.1. In view of (6.1) and (7.24), we have for $n \geq 1$, denoting $\xi_{k_n} = \langle \bar{\ell}(v_n), (S^{[k_n]}|P_s\vec{s})$:

$$|\zeta^{[k_n]}(v_n)|_{\vec{s}} = \langle \bar{\ell}(\zeta^{[k_n]}(v_n)), \vec{s} \rangle = \langle S^{[k_n]}\bar{\ell}(v_n), \vec{s} \rangle = \langle \bar{\ell}(v_n), (S^{[k_n]}|c_0^{(0)})$$

$$= \sum_{j=1}^{2} a_j A(k_n, j) \langle \bar{\ell}(v_n), e_j^{(k_n)} \rangle + \xi_{k_n}. \quad (8.1)$$

By the Assumption (a) of Theorem 4.1 $\limsup_{n \to \infty} n^{-1} \log \|S^{[n]}|P_s\| \leq \theta_3 < 0$, hence

$$\limsup_{n \to \infty} n^{-1} \log |\xi_{k_n}| \leq \theta_3 < 0. \quad (8.2)$$

Recall that we assumed $a_1 = 1$; recall also (7.25), and denote

$$\vec{a} = \left( \begin{array}{c} a_1 \\ a_2 \end{array} \right), \quad \vec{K}_n = \left( \begin{array}{c} K_{k_n} \\ K_{k_{n+1}} \end{array} \right), \quad \vec{\xi}_n = \left( \begin{array}{c} \xi_{k_n} \\ \xi_{k_{n+1}} \end{array} \right), \quad \vec{\varepsilon}_n = \left( \begin{array}{c} \varepsilon_{k_n} \\ \varepsilon_{k_{n+1}} \end{array} \right).$$

We need the matrices $\Theta_n$ defined in (6.3):

$$\Theta_n = \left( \begin{array}{cccc} A(k_n, 1) \langle \bar{\ell}(v_n), e_1^{(k_n)} \rangle & A(k_n, 2) \langle \bar{\ell}(v_n), e_2^{(k_n)} \rangle \\ A(k_{n+1}, 1) \langle \bar{\ell}(v_{n+1}), e_1^{(k_{n+1})} \rangle & A(k_{n+1}, 2) \langle \bar{\ell}(v_{n+1}), e_2^{(k_{n+1})} \rangle \end{array} \right).$$
The equations (7.25) for $n, n+1$, in view of (8.1), combine into
\[
\omega \Theta_n \vec{a} = \vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n,
\]
hence
\[
(8.3) \quad \vec{a} = \omega^{-1} \Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n).
\]
It follows that
\[
(8.4) \quad a_2 = \omega^{-1} [\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_2 \quad \text{and} \quad \omega^{-1} [\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_1 = 1
\]
where $[\cdot]_j$ denotes $j$-th component of a vector. Therefore,
\[
a_2 = \frac{[\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_2}{[\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_1} \approx \frac{[\Theta_n^{-1} \vec{K}_n]_2}{[\Theta_n^{-1} \vec{K}_n]_1}
\]
for large $n$. To estimate the error, we subtract the expressions and obtain
\[
|a_2 - [\Theta_n^{-1} \vec{K}_n]_2| \leq \frac{|[\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_2|}{|[\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_1|} + \frac{|[\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_1|}{|[\Theta_n^{-1} \vec{K}_n]_1|} \frac{|[\Theta_n^{-1} \vec{K}_n]_2|}{|[\Theta_n^{-1} \vec{K}_n]_1|}.
\]
We have
\[
|[\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_1| \in [B^{-1}, B] \quad \text{and} \quad|[\Theta_n^{-1} (\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n)]_2| \in [B^{-1} \beta^{-1}, B \beta]
\]
by (8.4), since we are dealing with the set $E(\gamma, \delta, \beta, B)$, for which $|\omega| \in [B^{-1}, B]$ and $a_2 \in [\beta^{-1}, \beta]$. Note that $||\vec{\varepsilon}_n||_\infty \leq \frac{1}{2}$ for all $n$, and $||\vec{\xi}_n||_\infty \leq (2B)^{-1}$ for $n \geq n_0(a, B)$ by (8.2), so that
\[
||\vec{\varepsilon}_n - \omega \vec{\xi}_n||_\infty \leq 1, \quad n \geq n_0(a, B).
\]
Choose $\delta_1 > 0$ such that $\theta_2 - \delta_1 > 0$ and $\theta_3 + \delta_1 < 0$ (there may be more conditions later). Then, combining the inequalities above with (6.9), we obtain for $n \geq n_1(a, B, \beta)$:
\[
|a_2 - [\Theta_n^{-1} \vec{K}_n]_2| \leq \frac{|[\Theta_n^{-1}]_\infty|}{B^{-1}} + \frac{|[\Theta_n^{-1}]_\infty|}{B^{-1}} \frac{B \beta}{B^{-1} - ||\Theta_n^{-1}||_\infty}
\]
(8.5)
\[
\leq O(1) \exp[-(\theta_2 - \delta_1)k_n].
\]
On the other hand, comparing (8.3) for $n$ and $n+1$ yields
\[
\vec{K}_{n+1} + \vec{\varepsilon}_{n+1} - \omega \vec{\xi}_{n+1} = \Theta_{n+1} \Theta_n^{-1} [\vec{K}_n + \vec{\varepsilon}_n - \omega \vec{\xi}_n],
\]
hence
\[
||\vec{K}_{n+1} - \Theta_{n+1} \Theta_n^{-1} \vec{K}_n||_\infty \leq ||\vec{\varepsilon}_{n+1}||_\infty + B||\vec{\xi}_{n+1}||_\infty + ||\Theta_{n+1} \Theta_n^{-1}||_\infty (||\vec{\varepsilon}_n||_\infty + B||\vec{\xi}_n||_\infty).
\]
Taking the second component of the vector in the left-hand side and using \((8.10)\), we obtain for \(n\) sufficiently large:

\[
\left| K_{n+2} - [\Theta_{n+1} \Theta_n^{-1} \bar{K}_n] \right| \leq (1 + C_\zeta \exp[2(W_n + W_{n+1})] \times \max\{|\varepsilon_{kn}|, |\varepsilon_{kn+1}|, |\varepsilon_{kn+2}|\} + B \max\{|\xi_{kn}|, |\xi_{kn+1}|, |\xi_{kn+2}|\}.
\]

(8.6)

Let

\[
M_n := 1 + C_\zeta \exp[2(W_n + W_{n+1})] \quad \text{and} \quad \rho_n = \frac{1}{4M_n}.
\]

(8.7)

Note that \((5.5)\) implies:

\[
\text{for any } \tilde{\delta} > 0, \quad M_n \leq \exp\left(\tilde{\delta} \cdot n\right) \quad \text{for } n \text{ sufficiently large.}
\]

(8.8)

**Lemma 8.1.** For all \(n, n_2 = n_2(a, B, \beta)\) we have the following, independent of \(\omega \in [B^{-1}, B]\) and \(\bar{s}\), satisfying \(a_2 = P^2_a(\bar{s})/P^1_a(\bar{s}) \in [\beta^{-1}, \beta]\):

(i) Given \(K_{n}, K_{n+1}\), there are at most \(2M_n + 1\) possibilities for the integer \(K_{n+2};\)

(ii) if \(\max\{|\varepsilon_{kn}|, |\varepsilon_{kn+1}|, |\varepsilon_{kn+2}|\} < \rho_n\), then \(K_{n+2}\) is uniquely determined by \(K_{n}, K_{n+1}\).

**Proof.** For part (i), we just use that \(\|\varepsilon_n\| \leq \frac{1}{2}\) for all \(n\) and \(\|\tilde{\xi}_n\| \leq (2B)^{-1}\) for \(n\) sufficiently large, and that the number of integer points in an interval of length \(2M_n\) is at most \(2M_n + 1\).

In part (ii) we claim that \(K_{n+2}\) belongs to a neighborhood of radius less than \(\frac{\delta}{2}\), centered at \([\Theta_{n+1} \Theta_n^{-1} \bar{K}_n]_2\), under the given assumptions, for \(n\) sufficiently large. We have \(M_n \rho_n = 1/4\), so it remains to make sure that

\[
\max\{|\xi_{kn}|, |\xi_{kn+1}|, |\xi_{kn+2}|\} \leq \rho_n
\]

for \(n\) sufficiently large. But \(\rho_n \geq (1/4) e^{-\bar{\delta}n}\) by \((8.8)\), whereas \(\|\tilde{\xi}_n\| \leq e^{(\theta_3 + \delta_1)k_n} \leq e^{(\theta_3 + \delta_1)L_1 n}\) for \(n\) sufficiently large by \((8.2)\) and \((5.7)\). (Recall that \(\theta_3 + \delta_1 < 0\).) Taking \(\tilde{\delta} < |\theta_3 + \delta_1|/L_1\) ensures that \(\|\tilde{\xi}_n\|\) is dominated by \(\rho_n\), as desired. \(\Box\)

**Proof of Proposition 7.3.** Let \(\tilde{E}_N(\delta, \beta, B)\) be defined by

\[
\tilde{E}_N(\delta, \beta, B) := \left\{a_2 : |a_2| \in [\beta^{-1}, \beta], \exists \omega \in [B^{-1}, B] : \text{card}\{n \leq N : \max\{|\varepsilon_{kn}|, |\varepsilon_{kn+1}|, |\varepsilon_{kn+2}|\} \geq \rho_n\} < \delta N\right\}.
\]

First we claim that \(\mathbb{P}_+\)-almost surely,

\[
\tilde{E}_N(\delta, \beta, B) \supset E_N(\gamma, \delta/6, \beta, B)
\]

for \(N\) sufficiently large, where

\[
\gamma = (1/4)(1 + C_\zeta e^{2K})^{-1}, \quad \text{with} \quad K = 25L_2 \log(1/\delta).
\]

(8.10)
Thus, by (8.11), the number of balls of radius $\rho_n$ for all $k_n \in \Gamma_N$. Observe that there are fewer than $\delta N/6$ integers $n \leq N$ for which $W_n + W_{n+1} > K$, for $N$ sufficiently large. Indeed, otherwise we can find $\Psi \subset \{1, \ldots, N\}$, with $|\Psi| \geq \delta N/12$, such that $W_n > K/2$ for $n \in \Psi$, hence

$$\sum \{W_n : n \in \Psi\} \geq K\delta N/24,$$

which contradicts (A3) for $K > 24L_2 \log(1/\delta)$. In view of (8.7) and (8.10), it follows that

$$\text{card}\{n \in \Gamma_N : \rho_n \geq \gamma\} \geq \delta N/6.$$

Thus $a_2 \notin E_N(\gamma, \delta/6, \beta, B)$ which confirms (8.3).

It follows that it is enough to estimate the dimension of

$$\tilde{E} := \tilde{E}(\delta, \beta, B) := \bigcap_{N_0=1}^{\infty} \bigcup_{N=N_0}^{\infty} \tilde{E}_N(\delta, \beta, B) \supset \mathcal{E}(\gamma, \delta/6, \beta, B).$$

Suppose $a_2 \in \tilde{E}_N := \tilde{E}_N(\delta, \beta, B)$; choose $\omega$ from the definition of $\tilde{E}_N$, and find the corresponding sequence $K_{k_n}, \varepsilon_{k_n}$. We have from (8.5) that $a_2$ is covered by an interval of radius

$$(8.11) \quad O(1) \exp[-(\theta_2 - \delta_1)k_N] \leq O(1) \exp[-(\theta_2 - \delta_1)N],$$

centered at $[\Theta_n^{-1}\tilde{K}_n]/2/[\Theta_n^{-1}\tilde{K}_n]_1$. Thus, it suffices to estimate the number of sequences $K_{k_n}, n \leq N$, which may arise. Let $\Psi_N$ be the set of $n \in \{1, \ldots, N\}$ for which we have

$$\max\{|\varepsilon_{k_n}|, |\varepsilon_{k_{n+1}}|, |\varepsilon_{k_{n+2}}|\} \geq \rho_n.$$ 

By the definition of $\tilde{E}_N$ we have $|\Psi_N| \leq \delta N$. There are $\sum_{i \leq \delta N} \binom{N}{i}$ such sets. For a fixed $\Psi_N$ the number of possible sequences $\{K_{k_n}\}$ is at most

$$B_N := \prod_{n \in \Psi_N} (2M_n + 1),$$

times the number of “beginnings” $K_{k_1}, \ldots, K_{k_{k_2}}$, by Lemma 8.1. The number of possible “beginnings” is clearly bounded, independent of $N$ (in view of the a priori bounds on $\omega$ and $s$). By the definition of $M_n$ and (A3), we have, for $N$ sufficiently large:

$$B_N \lesssim \exp \left(C'' \sum_{n \in \Psi_N} (W_n + W_{n+1}) \right) \leq \exp \left[\tilde{L} \log(1/\delta)(\delta N)\right].$$

Thus, by (8.11), the number of balls of radius $O(1) \exp[-(\theta_2 - \delta_1)N]$ needed to cover $\tilde{E}$ is at most

$$(8.12) \quad O_{\beta,B}(1) \cdot \sum_{i \leq \delta N} \binom{N}{i} \exp \left[\tilde{L} \log(1/\delta)(\delta N)\right] \leq O_{\beta,B}(1) \cdot \exp \left[\tilde{L} + C' \log(1/\delta)(\delta N)\right],$$

where
using (5.16) in the last inequality. Since \( \delta \log(1/\delta) \to 0 \) as \( \delta \to 0 \), we can choose \( \delta_0 > 0 \) so small that \( \delta < \delta_0 \) implies
\[
\left[ (\tilde{L} + C') \log(1/\delta)(\delta N) \right] < \epsilon(\theta_2 - \delta_1)N,
\]
whence \( \tilde{E} \) has Hausdorff dimension less than \( \epsilon \), as desired. The proof of Proposition 7.3 and hence of Theorem 4.1 is complete.

\[\square\]

9. Derivation of Theorem 1.1 from Theorem 4.1

Consider our surface \( M \) of genus 2. By the results of [8, Section 4] there is a correspondence between almost every translation flow and an element \( a \in \Omega \) (space of 2-sided Markov compacta), such that the (uniquely ergodic) flow is measure-theoretically conjugate to the uniquely ergodic flow \( (X(a), g) \) and hence to the suspension flow \( (\mathcal{X}^\mathbf{s}, h_t) \) over the Vershik map \( (X_+(a), \mathcal{T}) \) with the roof function corresponding to an appropriate vector \( \mathbf{s} \), see Lemma 2.1. By construction (see [8]), this correspondence intertwines the Teichmüller flow on the space of Abelian differentials and a measure-preserving system \( (\Omega, \mathbb{P}, \sigma) \), as considered at the beginning of our Section 4. The key point here is that the Masur-Veech measure on the space of abelian differentials is taken, under this correspondence, to a measure mutually absolute continuous with the product of the measure \( \mathbb{P}_+ \) on \( \Omega_+ \) and the Lebesgue measure on the 3-dimensional set of possible vectors \( \mathbf{s} \) defining the suspension.

A bit more precisely: as is well-known, the translation flow on the surface can be realized as a suspension flow over an interval exchange transformation (IET), see [33] for details. Veech [22] constructed, for any connected component of a stratum \( \mathcal{H} \), a measurable finite-to-one map from the space \( \mathcal{V}(\mathcal{R}) \) of zippered rectangles corresponding to the Rauzy class \( \mathcal{R} \), to \( \mathcal{H} \), which intertwines the Teichmüller flow on \( \mathcal{H} \) and a flow \( P_t \) that Veech defined on \( \mathcal{V}(\mathcal{R}) \). Section 4.3 of [8] constructs the symbolic coding of the flow on \( \mathcal{V}(\mathcal{R}) \) as the flow on a space of Markov compacta \( \Omega \). Observe that in our case the stratum \( \mathcal{H}(2) \) is connected and corresponds to the Rauzy class of the IET with permutation \((4, 3, 2, 1)\). Furthermore, the Kontsevich-Zorich cocycle over the Teichmüller flow is mapped, under this correspondence into the renormalization cocycle for the Markov compacta. In general, this mapping is not bijective (it corresponds to passing from absolute to relative real cohomologies in the manifold \( M \)), but in our case the kernel is trivial, since the manifold has only one singularity and therefore there are no saddle connections. For background and complete details the reader is referred to [8] and [33].

Now let \( f \) be a Lipschitz function on \( M \) with an abelian differential \( \omega \). Under the symbolic coding from [8], it is mapped into a weakly Lipschitz function on \( \mathcal{X}(a) \) and then to a weakly Lipschitz function \( f^\mathbf{s} \) on \( \mathcal{X}^\mathbf{s} \) with appropriate \( \mathbf{s} \), to which Theorem 4.1 applies. Note that its norm \( \|f^\mathbf{s}\|_L \) is dominated by \( \|f\|_L \) for all \( \mathbf{s} \). In order to reduce Theorem 1.1 to Theorem 4.1 we must now check that the assumptions of Theorem 4.1 hold for the left shift on the space of
Markov compacta endowed with the push-forward of the Masur-Veech smooth measure under the isomorphism of [8].

The condition (a) on the Lyapunov spectrum from Theorem [4.1] follows from results of Forni [13] in our case (later Avila and Viana [4] proved this for an arbitrary genus \( \geq 2 \)). The condition (b) about the strictly positive matrix \( Q \) follows from the properties of substitutions associated to Rauzy operations, see Veech [22], [23] (see also [8, Section 4.3.4] for an exposition). Using the “aggregation” (“telescoping”) operation on the Markov compactum, we can assume, without loss of generality that all the entries of matrix \( Q \) are \( \geq 2 \), as was needed in Section 7.

To be more precise, we recall the construction of the Markov compactum and Bratteli-Vershik realization of the translation flow from [8]. The symbolic representation of the translation flow on the 2-sided Markov compactum is obtained as the natural extension of the 1-sided symbolic representation for the IET which we now describe. An interval exchange is denoted by \((\lambda, \pi)\), where \( \pi \) is the permutation of \( m \) subintervals and \( \lambda \) is the vector of their lengths. The well-known Rauzy induction (operations “a” and “b”) proceeds by inducing on a smaller interval, so that the first return map is again an exchange of \( m \) intervals. The Rauzy graph is a directed labeled graph, whose vertices are permutations of IET’s and the edges lead to permutations obtained by applying one of the operations. Moreover, the edges are labeled by the type of the operation (“a” or “b”). As is well-known, for almost every IET, there is a corresponding infinite path in the Rauzy graph, and the length of the interval on which we induce tends to zero. For any finite “block” of this path, we have a pair of intervals \( J \subset I \) and IET’s on them, denoted \( T_I \) and \( T_J \), such that both are exchanges of \( m \) intervals and \( T_J \) is the first return map of \( T_I \) to \( J \). Let \( I_1, \ldots, I_m \) be the subintervals of the exchange \( T_I \) and \( J_1, \ldots, J_m \) the subintervals of the exchange \( T_J \). Let \( r_i \) be the return time for the interval \( J_i \) into \( J \) under \( T_I \), that is, \( r_i = \min\{k > 0 : T^k J_i \subset J\} \). Represent \( I \) as a Rokhlin tower over the subset \( J \) and its induced map \( T_J \), and write

\[
I = \bigsqcup_{i=1,\ldots,m,k=0,\ldots,r_i-1} T^k J_i.
\]

By construction, each of the “floors” of our tower, that is, each of the subintervals \( T^k J_i \), is a subset of some, of course, unique, subinterval of the initial exchange, and we define an integer \( n(i,k) \) by the formula

\[
T^k J_i \subset I_{n(i,k)}.
\]

To the pair \( I, J \) we now assign a substitution \( \zeta_{I,J} \) on the alphabet \( \{1, \ldots, m\} \) by the formula

\[
(9.1) \quad \zeta_{I,J} : i \rightarrow n(i,0)n(i,1)\ldots n(i, r_i - 1).
\]

This is the sequence of substitutions arising from the Bratteli-Vershik realization of an IET.

Now, in order to obtain the condition (b) from Theorem [4.1] about a strictly positive matrix, it is necessary to perform an “aggregation” (“telescoping”) operation. Since we consider a random Bratteli-Vershik system and apply the Oseledets Theorem, we should be somewhat careful with it,
in order to preserve stationarity and ergodic-theoretic properties. To this end, fix an admissible word corresponding to a closed loop in the Rauzy graph, and consider the first return map to the cylinder set defined by this loop. Let $\zeta$ be the associated substitution. In the induced system, all substitutions will be of the form $\zeta' \circ \zeta$, with varying substitutions $\zeta'$, corresponding to the “waiting time” until the next occurrence of $\zeta$. Note that $\zeta'$ is an identity map, whenever the loop is repeated. Since we are inducing on a subset of positive measure, it is clear that the Lyapunov spectrum for our cocycle over the first return map still satisfies condition (a). Since any repetition of the loop is admissible, the substitution $\zeta$ itself, and any fixed power of it, will appear infinitely often in almost every sequence of the induced system, hence condition (b) from Theorem 4.1 is satisfied as well.

Condition (c) is verified in the next lemma. Substitutions obtained from finite paths in the Rauzy graph will be called admissible.

**Lemma 9.1.** There exists a substitution $\zeta$, with $Q = S_\zeta$, admissible in the Rauzy class, and **good return words** $u_1, \ldots, u_m \in GR(\zeta)$, such that $\{\vec{\ell}(u_j) : j \leq m\}$ is a basis of $\mathbb{R}^m$.

We start with a simple preliminary claim.

**Lemma 9.2.** There exists a letter $c$ and an admissible substitution $\eta$ such that $\eta(j)$ starts with $c$, for all letters $j \leq m$.

**Proof.** Indeed, start with an arbitrary loop in the Rauzy graph such that the corresponding renormalization matrix has all entries positive. Consider the interval exchange transformation with periodic Rauzy-Veech expansion obtained by going along the loop repeatedly (it is known from [22] that such an IET exists). As the number of passages through the loop grows, the length of the interval forming phase space of the new interval exchange (the result of the induction process) goes to zero. In particular, after sufficiently many moves, this interval will be completely contained in the first subinterval of the initial interval exchange — but this means, in view of (9.1) that $n(i, 0) = 1$ for all $i$, and hence the resulting substitution $\eta$ satisfies all the requirements of the lemma, with $c = 1$.

**Proof of Lemma 9.1.** By construction, the matrix $S_\eta$ of the substitution $\eta$ given by Lemma 9.2 has all entries strictly positive, hence $\eta(i)$ contains all letters $j$, for any $i \leq m$. We can always replace the substitution $\eta$ by its positive power $\eta^k$, since $\eta$ corresponds to a loop in the Rauzy graph. Note that, for every $i \leq m$, the word $\eta^2(i)$ is a concatenation of all words $\eta(j)$, $j \leq m$, in some order, maybe with repetitions, all of which begin with $c$. Therefore, for every $i \leq m$, the word $\eta^3(i)$ contains every $\zeta(j)$, $j \leq m$, followed by another $\zeta(j')$, also starting with $c$. It follows that $u_j := \eta(j)$ is a good return word for $\zeta := \eta^3$, for every $j \leq m$. The population vector $\vec{\ell}(\eta(j))$ is the $j$-th column vector of $S_\eta$. As is well-known, the matrices corresponding to Rauzy operations are invertible, which implies that the columns of $S_\eta$ span $\mathbb{R}^m$. The proof is complete.
The only remaining, key conditions to check are (4.5) and (4.6). They will be derived from a variant of the exponential estimate for return times of the Teichmüller flow into compact sets. For large compact sets of special form, this estimate is due to Athreya [1], whereas in the general form it was established in [7] and independently in [3]. We will use the notation of [7]. We stress that the symbolic coding of the Rauzy-Veech induction map on the space of interval exchange transformations used in [7] corresponds to the symbolic coding of the Teichmüller flow as a suspension flow over the shift on the space of Markov compacta: indeed, the Rauzy-Veech expansion precisely identifies an interval exchange transformation with a Bratteli-Vershik automorphism (cf. [8]).

The symbol $\Delta(\mathcal{R})$ stands for the space of interval exchange transformations whose permutation lies in a given Rauzy class (fixed and omitted from notation); the symbolic space $\Omega^Z_{A,B}$ is the one-sided topological Markov chain over a countable alphabet that realizes the symbolic coding of the Rauzy-Veech-Zorich induction map; the space $\Omega^\mathcal{Z}_{A,B}$ is its natural extension, the corresponding two-sided topological Markov chain; in the notation of the previous sections, the space $\Omega^\mathcal{Z}_{A,B}$ corresponds to the space $\Omega$. The space $\Omega^\mathcal{Z}_{A,B}$ can also be viewed as the phase space of the natural extension of the Rauzy-Veech-Zorich induction, that is, the space of sequences of interval exchange transformations ordered by nonpositive integers:

$$(\lambda(0), \pi(0)), (\lambda(-1), \pi(-1)), \ldots, (\lambda(-k), \pi(-k)), \ldots$$

where $(\lambda(n), \pi(n))$ is the image of $(\lambda(n-1), \pi(n-1))$ under the Rauzy-Veech-Zorich induction map, in particular

$$\lambda(n) = \frac{A_n \lambda(n+1)}{A_n \lambda(n+1)},$$

where $A_n$ is the corresponding renormalization matrix and the symbol $|\lambda|$ stands for the sum of coordinates of a vector $\lambda$. In other words, the induction map that takes $(\lambda(n), \pi(n))$ to $(\lambda(n+1), \pi(n+1))$ consists in applying the matrix $A_n^{-1}$ and normalizing to unit length. (Note that $A_n$ are not the matrices $A_n$ from Section 4; rather, the latter are obtained from the former by taking products resulting from the inducing/aggregation, and a transpose.) Following [7], we also introduce the “non-normalized” lengths $\Lambda(n)$ inductively by the rule $\lambda(0) = \Lambda(0)$, $\Lambda(n) = A_{n+1} \Lambda(n+1)$ for $n < 0$. Informally, $\log |\Lambda(-n)|$ is the “Teichmüller time” corresponding to the discrete normalization “Rauzy-Veech-Zorich” time $n$.

We keep the symbol $\sigma$ for the shift on the space $\Omega_{A,B}$. Let $q$ be a word admissible in the Rauzy class, such that the resulting Rauzy-Veech renormalization matrix has all entries positive; it corresponds to the positive matrix (4.1) in Condition (C1) of Section 4. The symbol $\mathbb{P}$ denotes the Masur-Veech measure.

For $\omega \in \Omega^\mathcal{Z}_{A,B}$, let $\ell_q(\omega)$ be the (negative) time until the first appearance of the word $q$ in the sequence $\omega$. Let $\Omega_{A,B}(q)$ be the set of $\omega^+$ which start with $q$. 

Proposition 9.3. There exists $\epsilon > 0$ such that for any $\omega^+ \in \Omega_{A,B}(q)$ we have

$$\int_{\Omega_{A,B}^\omega} \exp(\epsilon \ell_q(\omega)) dP(\omega^+) < +\infty.$$ 

This will imply the corresponding condition (4.5), by passing to the induced system.

Further, denote by $L_q(\omega)$ the “Teichmüller time” until the first (negative) appearance of the word $q$, for $\omega$ with $\omega^+ \in \Omega_{A,B}(q)$. Note that the ratio of $\exp(L_q(\omega))$ and the norm of the corresponding renormalization matrix is bounded from above and from below by a constant depending only on $q$, since this is the time between two successive returns to $\Omega_q$ and so this renormalization matrix has the form $A(q)CA(q)$, where $C$ is a unimodular matrix with nonnegative entries and $A(q)$ has strictly positive entries. This product matrix, by definition, coincides with the (transpose of the) one from (4.6).

Proposition 9.4. There exists $\epsilon > 0$ such that for any $\omega^+ \in \Omega_{A,B}(q)$ we have

$$\int_{\Omega_{A,B}^\omega} \exp(\epsilon L_q(\omega)) dP(\omega^+) < +\infty.$$ 

Again, the corresponding condition (4.6) will follow as well, by passing to the induced system. Proposition 9.3 is a direct consequence of Proposition 9.4, since the the “Rauzy-Veech-Zorich time” $\ell_q(\omega)$ is dominated by the logarithm of the norm of the renormalization matrix, see Corollary 9 in [7], and as already mentioned, the norm of the renormalization matrix between two occurrences of the word $q$ is comparable to the exponential of the Teichmüller time $\exp(L_q(\omega))$.

Proof of Proposition 9.4. We argue in much the same way as in Section 11 of [7]. Our main tool will be Lemma 16 from [7] whose formulation we now recall:

**Lemma 9.5.** For any word $q \in W_{A,B}$ such that all entries of the matrix $A(q)$ are positive, there exist constants $K_0(q), p(q)$, depending only on $q$ and such that the following is true. For any $K \geq K_0$ and any $(\lambda, \pi) \in \Delta(\mathcal{R})$,

$$P\left(\exists n : (\lambda(-n), \pi(-n)) \in \Delta_q, |\Lambda(-n)| < K \right) | (\lambda, \pi) = (\lambda(0), \pi(0)) \geq p(q).$$

For a sequence $\omega^+ \in \Omega_{A,B}$ let $P(\omega|\omega^+)$ be the conditional distribution of $P$ on the set of all pasts conditioned on the future $\omega^+$. The next lemma is a reformulation of Lemma 7 from [7].

**Lemma 9.6.** There exists $K_0 \in \mathbb{N}$ and $C = C(\mathcal{R}) \geq 1$ such that the following holds for any $K > K_0$: for all $\omega^+ \in \Omega_{A,B}$ we have

$$P\left(|\Lambda(-1)| > K | \omega^+ \right) \leq \frac{C}{K}.$$
Define a random time $k_1(\omega)$ to be the first moment $n \geq 1$ such that $|\Lambda(-n)(\omega)| > K$. Note that the map

$$\tilde{\sigma}(\omega) = \sigma^{-k_1(\omega)}(\omega)$$

is invertible. Introduce a function $\eta : \Omega_{A,B}^Z \to \mathbb{N}$ by the formula

$$\eta(\omega) = \left\lfloor \frac{\log |\Lambda(-k_1(\omega))|}{\log K} \right\rfloor.$$

In other words, $\eta(\omega) = n$ if $K^n \leq |\Lambda(-k_1(\omega))| < K^{n+1}$. Combining Lemmas 9.5 and 9.6 we obtain

**Proposition 9.7.** For any $\omega^+ \in \Omega_{A,B}$ and any $r \in \mathbb{N}$ we have

$$\mathbb{P}\left( \{ \omega : \eta(\omega) = r, \, \omega_{-k_1(\omega)+1}, \ldots, \omega_0 \text{ does not contain } q | \omega^+ \} \right) \leq \frac{C}{K^{r-1}} \cdot (1 - p(q)).$$

Now, take a large $N$ and let $n_N(\omega) = \min\{n : \eta(\omega) + \cdots + \eta(\tilde{\sigma}^n(\omega)) \geq N\}$. Let also

$$k_n(\omega) = k_1(\omega) + \cdots + k_1(\tilde{\sigma}^n(\omega)).$$

Then we have, by the definition of $\eta$:

$$K^N \leq |\Lambda(-k_{n_N}(\omega))| < K^{n_N(\omega)+\eta(\omega)+\cdots+\eta(\tilde{\sigma}^n(\omega))}. \tag{9.2}$$

Clearly $n_N(\omega) \leq N$, and observe that for all $\omega^+ \in \Omega_{A,B}$,

$$\mathbb{P}(\mathcal{B}(N) | \omega^+) \leq \frac{C}{K^N},$$

where $\mathcal{B}(N) = \{ \omega : \eta(\tilde{\sigma}^{n_N}(\omega)) > N \}$, by Lemma 9.6. Consider the set

$$\tilde{\Omega}(N) = \{ \omega : \omega_{-k_{n_N}(\omega)+1}, \ldots, \omega_0 \text{ does not contain the word } q \} \bigcap \{ \omega : \eta(\tilde{\sigma}^{n_N}(\omega)) \leq N \}.$$

Note that

$$\{ \omega : L_q(\omega) > 3N \} \subset \tilde{\Omega}(N) \cup \mathcal{B}(N).$$

It suffices, therefore, to prove that there exists $\rho < 1$ such that

$$\mathbb{P}(\tilde{\Omega}(N) | \omega^+) \leq \rho^N.$$

We have from Proposition 9.7 for any $\ell \in \mathbb{N}$ and $n_1, \ldots, n_\ell \in \mathbb{N}$:

$$\mathbb{P}\left( \eta(\omega) = n_1, \eta(\tilde{\sigma}(\omega)) = n_2, \ldots, \eta(\tilde{\sigma}^\ell(\omega)) = n_\ell, \, \omega_{-k_\ell(\omega)+1}, \ldots, \omega_0 \text{ does not contain } q | \omega^+ \right) \leq (1 - p(q))^\ell \left( \frac{C}{K} \right)^{(n_1+\cdots+n_\ell)-\ell}. \tag{9.3}$$
Choose $K \in \mathbb{N}$ such that $1 - p(\mathbf{q}) + \frac{C}{K} < 1$. We have

$$
P(\hat{\Omega}(N) \mid \omega^+) \leq \sum_{\tilde{N}=N}^{2N} P(\hat{\Omega}(N) \mid \omega^+ \& \eta(\omega) + \cdots + \eta(\tilde{a}^{\tilde{N}}(\omega)) = \tilde{N}) \leq \sum_{\tilde{N}=N}^{2N} \left(1 - p(\mathbf{q}) + \frac{C}{K}\right) \tilde{N} \leq \rho^N,$$

using (9.3) and the binomial formula in the last line, and Proposition 9.4 is proved. Now Proposition 9.3 follows as well, and Theorem 1.1 is proved completely.

**References**

[1] Athreya, Jayadev S. Quantitative recurrence and large deviations for Teichmüller geodesic flow. *Geom. Dedicata* 119 (2006), 121–140.

[2] Avila, Artur; Forni, Giovanni. Weak mixing for interval exchange transformations and translation flows. *Annals of Mathematics* 165 (2007), 637–664.

[3] Avila, Artur; Gouëzel, Sébastien; Yoccoz, Jean-Christophe. Exponential mixing for the Teichmuller flow. *Publ. Math. Inst. Hautes Etudes Sci.* No. 104 (2006), 143–211.

[4] Avila, Artur; Viana, Marcelo. Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture. *Acta Math.* 198 (2007), no. 1, 1–56.

[5] Barreira, Luis; Pesin, Yakov. Nonuniform hyperbolicity. Dynamics of systems with nonzero Lyapunov exponents. *Encyclopedia of Mathematics and its Applications*, 115. Cambridge University Press, Cambridge, 2007. xiv+513 pp.

[6] Berthé, Valérie; Steiner, Wolfgang; Thuswaldner, Jörg M. Geometry, dynamics, and arithmetic of $S$-adic shifts. *Preprint* arXiv:1410.0331, 34 pages.

[7] Bufetov, Alexander I. Decay of correlations for the Rauzy–Veech–Zorich induction map on the space of interval exchange transformations and the Central Limit Theorem for the Teichmüller Flow on the moduli space of Abelian differentials. *Journal of the American Mathematical Society* 19 (2006), no. 3, 579–623.

[8] Bufetov, Alexander I. Limit theorems for suspension flows over Vershik’s automorphisms. *Russian Mathematical Surveys* 68 (2013), no. 5, 789–860.

[9] Bufetov, Alexander I. Limit theorems for translation flows. *Annals of Mathematics* 179 (2014), no. 2, 431–499.

[10] Bufetov, Alexander I.; Solomyak, Boris. On the modulus of continuity for spectral measures in substitution dynamics. *Advances in Mathematics* 260 (2014), 84–129.

[11] Erdős, Paul. On the smoothness properties of Bernoulli convolutions. *Amer. J. Math.* 62 (1940), 180–186.

[12] Ferenczi, Sébastien. Rank and symbolic complexity. *Ergodic Theory Dynam. Systems* 16 (1996), no. 4, 663–682.

[13] Forni, Giovanni. Deviation of ergodic averages for area-preserving flows on surfaces of higher genus. *Annals of Mathematics* (2) 155 (2002).

[14] Forni, Giovanni; Ulcigrai, Corinna. Time-changes of horocycle flows. *J. Mod. Dyn.* 6 (2012), no. 2, 251-273.

[15] Furstenberg, Harry. Stationary processes and prediction theory. *Annals of Mathematics Studies*, No. 44. Princeton University Press, Princeton, N.J. 1960.

[16] Ito, Shunji. A construction of transversal flows for maximal Markov automorphisms. *Tokyo Journal of Mathematics* 1 (1978), no. 2, 305–324.

[17] Kahane, Jean-Pierre. Sur la distribution de certaines séries aléatoires. *Colloque Th. Nombres [1969, Bordeaux], Bull. Soc. math. France*, Mémoire 25 (1971), 119–122.
Alexander I. Bufetov, Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373
Steklov Institute of Mathematics, Moscow
Institute for Information Transmission Problems, Moscow
National Research University Higher School of Economics, Moscow
Rice University, Houston TX
E-mail address: bufetov@mi.ras.ru

Boris Solomyak, Department of Mathematics, Bar-Ilan University, Ramat-Gan, Israel
University of Washington, Seattle, Washington, USA

E-mail address: solomyak@uw.edu