JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 19, Number 3, Pages 579–623
S 0894-0347(06)00528-5
Article electronically published on February 22, 2006

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

ALEXANDER I. BUFETOV

Se non quel tanto che n'accende il sole.
Michelangelo Buonarroti

1. Introduction

The aim of this paper is to prove a stretched-exponential bound for the decay of correlations for the Rauzy-Veech-Zorich induction map on the space of interval exchange transformations (Theorem 4). A corollary is the Central Limit Theorem for the Teichmüller flow on the moduli space of abelian differentials with prescribed singularities (Theorem 10).

The proof of Theorem 4 proceeds by the method of Sinai [13] and Bunimovich–Sinai [14]: the induction map is approximated by a sequence of Markov chains satisfying the Doeblin condition. The main "loss of memory" estimate is Lemma 4.

1.1. Interval exchange transformations.

Let m be a positive integer. Let π be a permutation on m symbols. The permutation π will always be assumed *irreducible*, which means that $\pi\{1,\ldots,k\} = \{1,\ldots,k\}$ only if $k = m$.

Let λ be a vector in $\mathbb{R}^m_+$, $\lambda = (\lambda_1,\ldots,\lambda_m)$, $\lambda_i > 0$ for all i. Denote

$$|\lambda| = \sum_{i=1}^{m} \lambda_i.$$

Consider the half-open interval $[0, |\lambda|)$. Consider the points $\beta_1 = 0$, $\beta_i = \sum_{j<i} \lambda_j$, $\beta_1^\pi = 0$, $\beta_i^\pi = \sum_{j<i} \lambda_{\pi^{-1}j}$.

Denote $I_i = [\beta_i, \beta_{i+1})$, $I_i^\pi = [\beta_i^\pi, \beta_{i+1}^\pi)$. The length of I_i is λ_i, whereas the length of I_i^π is $\lambda_{\pi^{-1}i}$. Set

$$T_{(\lambda,\pi)}(x) = x + \beta^\pi_{\pi i} - \beta_i \text{ for } x \in I_i.$$

The map $T_{(\lambda,\pi)}$ is called an interval exchange transformation corresponding to (λ,π). The map $T_{(\lambda,\pi)}$ is an order-preserving isometry from I_i onto $I^\pi_{\pi(i)}$. We say that λ is *irrational* if there are no rational relations between $|\lambda|$, λ_1, λ_2, $\ldots$, λ_{m-1}.

Received by the editors October 6, 2004.
2000 *Mathematics Subject Classification.* Primary 37A25, 37F25, 37F30, 37E05, 60F05, 60J10.
Key words and phrases. Interval exchange transformations, Rauzy induction, speed of mixing, Teichmüller geodesic flow, central limit theorem.

Theorem 1 (Oseledets ([5]), Keane ([9])). *Let π be irreducible and λ irrational. Then for any $x \in [0, \sum_{i=1}^{m} \lambda_i)$, the set $\{T^n_{(\lambda,\pi)}x, n \geq 0\}$ is dense in $[0, \sum_{i=1}^{m} \lambda_i)$.*

1.2. **Rauzy operations** a **and** b**.** Let (λ, π) be an interval exchange. Assume that π is irreducible and λ is irrational. Following Rauzy [6], consider the induced map of (λ, π) on the interval $[0, |\lambda| - \min(\lambda_m, \lambda_{\pi^{-1}(m)}))$. The induced map is again an interval exchange of m intervals. For $i, j = 1, \dots, m$, denote by E_{ij} an $m \times m$ matrix of which the (i,j)-th element is equal to 1, all others to 0. Let E be the $m \times m$ identity matrix.

1.2.1. *Case a:* $\lambda_{\pi^{-1}m} > \lambda_m$. Define

$$A(a, \pi) = \sum_{i=1}^{\pi^{-1}(m)} E_{ii} + E_{m,\pi^{-1}m+1} + \sum_{i=\pi^{-1}m+1}^{m} E_{i,i+1};$$

$$a\pi(j) = \begin{cases} \pi j, & \text{if } j \leq \pi^{-1}m; \\ \pi m, & \text{if } j = \pi^{-1}m + 1; \\ \pi(j-1), & \text{other } j. \end{cases}$$

If $\lambda_{\pi^{-1}m} > \lambda_m$, then the induced interval exchange of $T_{(\lambda,\pi)}$ on the interval $[0, \sum_{i \neq m} \lambda_i)$ is $T_{(\lambda',\pi')}$, where $\lambda' = A(a,\pi)^{-1}\lambda$ and $\pi' = a\pi$.

1.2.2. *Case b:* $\lambda_m > \lambda_{\pi^{-1}m}$. Define

$$A(b, \pi) = E + E_{m,\pi^{-1}m};$$

$$b\pi(j) = \begin{cases} \pi j, & \text{if } \pi j \leq \pi m; \\ \pi j + 1, & \text{if } \pi m < \pi j < m; \\ \pi m + 1, & \text{if } \pi j = m. \end{cases}$$

If $\lambda_{\pi^{-1}m} < \lambda_m$, then the induced interval exchange of $T_{(\lambda,\pi)}$ on the interval $[0, \sum_{i \neq \pi^{-1}m} \lambda_i)$ is $T_{(\lambda',\pi')}$, where $\lambda' = A(b,\pi)^{-1}\lambda$ and $\pi' = b\pi$.

Note that operations a and b are invertible on the space of permutations; namely, we have:

$$a^{-1}\pi(j) = \begin{cases} \pi(j), & \text{if } j \leq \pi^{-1}(m); \\ \pi(j+1), & \text{if } \pi^{-1}(m) + 1 < j < m; \\ \pi(\pi^{-1}(m)+1), & \text{if } j = m; \end{cases}$$

$$b^{-1}\pi(j) = \begin{cases} \pi(j), & \text{if } \pi(j) \leq \pi(m); \\ m, & \text{if } j = \pi^{-1}(\pi(m)+1); \\ \pi(j) - 1, & \text{if } \pi(j) > \pi(m) + 1. \end{cases}$$

For $(\lambda, \pi) \in \Delta(\mathcal{R})$, denote

$$T_{a^{-1}}(\lambda, \pi) = (A(a^{-1}\pi, a)\lambda, a^{-1}\pi), \tag{1}$$

$$T_{b^{-1}}(\lambda, \pi) = (A(b^{-1}\pi, b)\lambda, b^{-1}\pi). \tag{2}$$

The interval exchange $T_{a^{-1}}(\lambda, \pi)$ is the preimage of (λ, π) under the operation a, and the interval exchange $T_{b^{-1}}(\lambda, \pi)$ is the preimage of (λ, π) under the operation b.

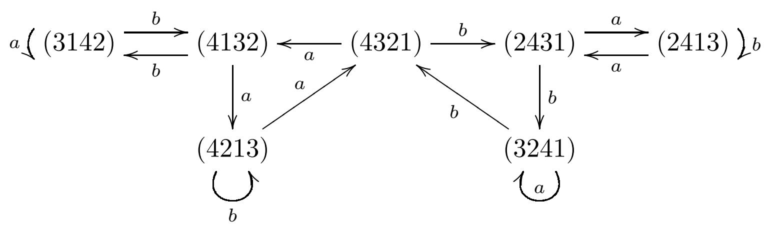
Normalize (dividing by $|\lambda| = \lambda_1 + \cdots + \lambda_m$) and set:

$$t_{a^{-1}}(\lambda, \pi) = \Big(\frac{A(a^{-1}\pi, a)\lambda}{|A(a^{-1}\pi, a)\lambda|}, a^{-1}\pi\Big), \tag{3}$$

$$t_{b^{-1}}(\lambda, \pi) = \Big(\frac{A(b^{-1}\pi, b)\lambda}{|A(b^{-1}\pi, b)\lambda|}, b^{-1}\pi\Big). \tag{4}$$

1.3. **Rauzy class and Rauzy graph.** If π is an irreducible permutation, then its *Rauzy class* is the set of all permutations that can be obtained from π by applying repeatedly the operations a and b; the Rauzy class of the permutation π is denoted $\mathcal{R}(\pi)$. The Rauzy class has a natural structure of an oriented labelled graph: namely, the permutations of the Rauzy class are the vertices of the graph, and if $\pi = a\pi'$, then we draw an edge from π to π' and label it by a, and if $\pi = b\pi'$, then we draw an edge from π to π' and label it by b. This labelled graph will be called the *Rauzy graph* of the permutation π.

For example, the Rauzy graph of the permutation (4321) is



For a permutation π, consider the set $\{a^n\pi, n \geq 0\}$. This set forms a cycle in the Rauzy graph which will be called the *a-cycle* of π. Similarly, the set $\{b^n\pi, n \geq 0\}$ will be called the *b-cycle* of π.

1.4. **The Rauzy-Veech-Zorich induction.** Denote

$$\Delta_{m-1} = \{\lambda \in \mathbb{R}_+^m : |\lambda| = 1\},$$
$$\Delta_\pi^+ = \{\lambda \in \Delta_{m-1}, \lambda_{\pi^{-1}m} > \lambda_m\}, \Delta_\pi^- = \{\lambda \in \Delta_{m-1}, \lambda_m > \lambda_{\pi^{-1}m}\},$$
$$\Delta(\mathcal{R}) = \Delta_{m-1} \times \mathcal{R}(\pi).$$

Define a map

$$\mathcal{T} : \Delta(\mathcal{R}) \to \Delta(\mathcal{R})$$

by

$$\mathcal{T}(\lambda, \pi) = \begin{cases} (\frac{A(\pi,a)^{-1}\lambda}{|A(\pi,a)^{-1}\lambda|}, a\pi), & \text{if } \lambda \in \Delta_\pi^-; \\ (\frac{A(\pi,b)^{-1}\lambda}{|A(\pi,b)^{-1}\lambda|}, b\pi), & \text{if } \lambda \in \Delta_\pi^+. \end{cases}$$

Each $(\lambda, \pi) \in \Delta(\mathcal{R})$ has exactly two preimages under the map $\mathcal{T}$, namely, $t_{a^{-1}}(\lambda, \pi)$ and $t_{b^{-1}}(\lambda, \pi)$ (see (3), (4)).

The set $\Delta(\mathcal{R})$ is a finite union of simplices. Let $\mathbf{m}$ be the Lebesgue measure on $\Delta(\mathcal{R})$ normalized in such a way that $\mathbf{m}(\Delta(\mathcal{R})) = 1$.

Theorem 2 (Veech [1])**.** *The map $\mathcal{T}$ has an infinite conservative ergodic invariant measure that is absolutely continuous with respect to Lebesgue measure on $\Delta(\mathcal{R})$.*

From this result Veech [1] derives that almost all (with respect to $\mathbf{m}$) interval exchange transformations are uniquely ergodic.

Denote

$$\Delta^+ = \bigcup_{\pi' \in \mathcal{R}(\pi)} \Delta^+_{\pi'},\ \Delta^- = \bigcup_{\pi' \in \mathcal{R}(\pi)} \Delta^-_{\pi'}.$$

Following Zorich [4], we define the function $n(\lambda, \pi)$ in the following way:

$$n(\lambda,\pi) = \begin{cases} \inf\{k > 0 : \mathcal{T}^k(\lambda,\pi) \in \Delta^-\}, & \text{if } \lambda \in \Delta^+_\pi; \\ \inf\{k > 0 : \mathcal{T}^k(\lambda,\pi) \in \Delta^+\}, & \text{if } \lambda \in \Delta^-_\pi. \end{cases}$$

Define

$$\mathcal{G}(\lambda,\pi) = \mathcal{T}^{n(\lambda,\pi)}(\lambda,\pi).$$

The map $\mathcal{G}$ will be referred to as *the Rauzy-Veech-Zorich induction map* [6, 1, 4].

For $(\lambda, \pi) \in \Delta(\mathcal{R})$, denote

$$t_{a^{-n}}(\lambda,\pi) = t^n_{a^{-1}}(\lambda,\pi),\quad t_{b^{-n}}(\lambda,\pi) = t^n_{b^{-1}}(\lambda,\pi), \tag{5}$$

$$T_{a^{-n}}(\lambda,\pi) = T^n_{a^{-1}}(\lambda,\pi),\quad T_{b^{-n}}(\lambda,\pi) = T^n_{b^{-1}}(\lambda,\pi). \tag{6}$$

Under the map $\mathcal{G}$, each interval exchange (λ, π) has countably many preimages:

$$\mathcal{G}^{-1}(\lambda,\pi) = \begin{cases} \{t_{a^{-n}}(\lambda,\pi), n \in \mathbb{N}\}, & \text{if } (\lambda,\pi) \in \Delta^+; \\ \{t_{b^{-n}}(\lambda,\pi), n \in \mathbb{N}\}, & \text{if } (\lambda,\pi) \in \Delta^-. \end{cases}$$

Theorem 3 (Zorich [4])**.** *The map $\mathcal{G}$ has an ergodic invariant probability measure that is absolutely continuous with respect to Lebesgue measure on $\Delta(\mathcal{R})$.*

Denote this invariant measure by ν; the probability with respect to ν will be denoted by $\mathbb{P}$. Let $\rho(\lambda, \pi)$ be the density of ν with respect to the Lebesgue measure $\mathbf{m}$. Zorich [4] showed that for any $\pi \in \mathcal{R}$ there exist two positive rational homogeneous of degree $-m$ functions ρ^+_π, ρ^-_π such that

$$\rho(\lambda,\pi) = \begin{cases} \rho^+_\pi(\lambda), & \text{if } \lambda \in \Delta^+_\pi; \\ \rho^-_\pi(\lambda), & \text{if } \lambda \in \Delta^-_\pi. \end{cases} \tag{7}$$

Remark 1. In particular, the invariant density is bounded from below: there exists a positive constant $C(\mathcal{R})$, depending on the Rauzy class only and such that $\rho(\lambda, \pi) \geq C(\mathcal{R})$ for any $(\lambda, \pi) \in \Delta(\mathcal{R})$.

The map $\mathcal{G}$ is not mixing: indeed, from the definition of $\mathcal{G}$, we have

$$\mathcal{G}(\Delta^+) = \Delta^-,\ \mathcal{G}(\Delta^-) = \Delta^+.$$

Let $\mathcal{B}$ be the Borel σ-algebra on $\Delta(\mathcal{R})$, and let $\mathcal{B}_n = \mathcal{G}^{-n}\mathcal{B}$. We have $\mathcal{B}_{n+2} \subset \mathcal{B}_n$. Recall [23] that *exactness* of the map $\mathcal{G}^2$ means, by definition, that the σ-algebra $\bigcap_{n=1}^\infty \mathcal{B}_{2n}$ is trivial [23] (in other words, that Kolmogorov's 0-1 law holds for the map $\mathcal{G}^2$).

Proposition 1. *The map $\mathcal{G}^2 : \Delta^+ \to \Delta^+$ is exact with respect to $\nu|_{\Delta^+}$.*

Proposition 1 is proven in Section 4; it implies strong mixing for the map $\mathcal{G}^2$.

1.5. **The main result.** *The Hilbert metric* on the simplex Δ_{m-1} is defined by the formula

$$d(\lambda, \lambda') = \log \frac{\max_i \frac{\lambda_i}{\lambda'_i}}{\min_i \frac{\lambda_i}{\lambda'_i}}. \tag{8}$$

(For the definition and the properties of the Hilbert metric, see pp. 16-20 of Viana's book [17] (in which this metric is ascribed to G.D. Birkhoff).)

Introduce a metric on $\Delta(\mathcal{R})$ by setting

$$d((\lambda,\pi),(\lambda',\pi')) = \begin{cases} 2 + d(\lambda,\lambda'), & \text{if } \pi \neq \pi'; \\ d(\lambda,\lambda'), & \text{if } \pi = \pi'. \end{cases}$$

For $\alpha > 0$, let H_α be the space of functions $\phi : \Delta(\mathcal{R}) \to \mathbb{R}$ such that if $d((\lambda,\pi),(\lambda',\pi')) \leq 1$, then $|\phi(\lambda,\pi) - \phi(\lambda',\pi')| \leq C d((\lambda,\pi),(\lambda',\pi'))^\alpha$ for some constant C, depending only on the function ϕ.

Define

$$C_{H_\alpha}(\phi) = \sup_{d((\lambda,\pi),(\lambda',\pi'))\leq 1} \frac{|\phi(\lambda,\pi) - \phi(\lambda',\pi')|}{d((\lambda,\pi),(\lambda',\pi'))^\alpha}.$$

The main result of this paper is

Theorem 4. *Let $\mathcal{G} : \Delta(\mathcal{R}) \to \Delta(\mathcal{R})$ be the Rauzy-Veech-Zorich induction map and let ν be the absolutely continuous invariant measure.*

Let $p > 2$. Then, for any $\alpha > 0$, there exist positive constants C, δ such that for any $\phi \in H_\alpha \cap L_p(\Delta^+(\mathcal{R}),\nu)$ and $\psi \in L_2(\Delta^+(\mathcal{R}),\nu)$ we have

$$|\int \phi \ \times \psi \circ \mathcal{G}^{2n} d\nu - \int \phi d\nu \int \psi d\nu| \leq C \exp(-\delta n^{1/6})(C_{H_\alpha}(\phi) + |\phi|_{L_p})(|\psi|_{L_2}).$$

Denote by $\mathcal{N}(0,\sigma)$ the Gaussian distribution with mean 0 and variance σ. By [7, 8, 17], we have

Corollary 1. *Let $\phi \in H_\alpha \cap L_p(\Delta(\mathcal{R})^+,\nu)$, $\int \phi d\nu = 0$. Assume that there does not exist $\psi \in L_2(\Delta(\mathcal{R})^+,\nu)$ such that $\phi = \psi \circ \mathcal{G}^2 - \psi$. Then there exists $\sigma > 0$, depending only on ϕ, and such that*

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \phi \circ \mathcal{G}^{2n} \xrightarrow{d} \mathcal{N}(0,\sigma) \text{ as } N \to \infty.$$

1.6. **Veech's space of zippered rectangles.** A zippered rectangle associated to the Rauzy class $\mathcal{R}$ is a quadruple (λ, h, a, π), where $\lambda \in \mathbb{R}^m_+, h \in \mathbb{R}^m_+, a \in \mathbb{R}^m, \pi \in \mathcal{R}$, and the vectors h and a satisfy the following equations and inequalities (one introduces auxiliary components $a_0 = h_0 = a_{m+1} = h_{m+1} = 0$, and sets $\pi(0) = 0$, $\pi^{-1}(m+1) = m+1$):

$$h_i - a_i = h_{\pi^{-1}(\pi(i)+1)} - a_{\pi^{-1}(\pi(i)+1)-1}, i = 0, \ldots, m, \tag{9}$$

$$h_i \geq 0, i = 1, \ldots, m, \ a_i \geq 0, i = 1, \ldots, m-1, \tag{10}$$

$$a_i \leq \min(h_i, h_{i+1}) \text{ for } i \neq m, i \neq \pi^{-1}m, \tag{11}$$

$$a_m \leq h_m, \ a_m \geq -h_{\pi^{-1}m}, \ a_{\pi^{-1}m} \leq h_{\pi^{-1}m+1}. \tag{12}$$

The area of a zippered rectangle is given by the expression

$$\lambda_1 h_1 + \cdots + \lambda_m h_m.$$

Following Veech, we denote by $\Omega(\mathcal{R})$ the space of all zippered rectangles, corresponding to a given Rauzy class $\mathcal{R}$ and satisfying the condition

$$\lambda_1 h_1 + \cdots + \lambda_m h_m = 1.$$

We shall denote by x an individual zippered rectangle.

Veech further defines a map $\mathcal{U}$ and a flow P^t on the space of zippered rectangles in the following way:

$$P^t(\lambda, h, a, \pi) = (e^t \lambda, e^{-t} h, e^{-t} a, \pi),$$

$$\mathcal{U}(\lambda, h, a, \pi) = \begin{cases} (A^{-1}(a,\pi)\lambda, A^t(a,\pi)h, a', a\pi), & \text{if } (\lambda,\pi) \in \Delta^-, \\ (A^{-1}(b,\pi)\lambda, A^t(b,\pi)h, a'', b\pi), & \text{if } (\lambda,\pi) \in \Delta^+, \end{cases}$$

where

$$a'_i = \begin{cases} a_i, & \text{if } j < \pi^{-1}m, \\ h_{\pi^{-1}m} + a_{m-1}, & \text{if } i = \pi^{-1}m, \\ a_{i-1}, & \text{other } i, \end{cases}$$

$$a''_i = \begin{cases} a_i, & \text{if } j < m, \\ -h_{\pi^{-1}m} + a_{\pi^{-1}m-1}, & \text{if } i = m. \end{cases}$$

Remark 2. As these formulas show, we have a representation

$$a' = U'a + v'(h), \ a'' = U''a + v''(h), \tag{13}$$

where U', U'' are orthogonal matrices, depending on the permutation π only, and $v'(h)$, $v''(h)$ are vectors, depending on π and h only.

The map $\mathcal{U}$ is invertible; $\mathcal{U}$ and P^t commute ([1]). Denote

$$\tau(\lambda, \pi) = (\log(|\lambda| - \min(\lambda_m, \lambda_{\pi^{-1}m})),$$

and for $x \in \Omega(\mathcal{R})$, $x = (\lambda, h, a, \pi)$, write

$$\tau(x) = \tau(\lambda, \pi).$$

Now define

$$\mathcal{Y}(\mathcal{R}) = \{x \in \Omega(\mathcal{R}) : |\lambda| = 1\}$$

and

$$\Omega_0(\mathcal{R}) = \bigcup_{x \in \mathcal{Y}(\mathcal{R}), 0 \le t \le \tau(x)} P^t x.$$

$\Omega_0(\mathcal{R})$ is a fundamental domain for $\mathcal{U}$ and, identifying the points x and $\mathcal{U}x$ in $\Omega_0(\mathcal{R})$, we obtain a natural flow, also denoted by P^t, on $\Omega_0(\mathcal{R})$.

The space $\Omega(\mathcal{R})$ has a natural Lebesgue measure class and so does the transversal $\mathcal{Y}(\mathcal{R})$. Veech [1] has proved the following theorem.

Theorem 5. *There exists a measure $\mu_{\mathcal{R}}$ on $\Omega(\mathcal{R})$, absolutely continuous with respect to Lebesgue measure, preserved by both the map $\mathcal{U}$ and the flow P^t and such that $\mu_{\mathcal{R}}(\Omega_0(\mathcal{R})) < \infty$.*

For $x \in \mathcal{Y}(\mathcal{R})$, define

$$\mathcal{S}(x) = \mathcal{U}P^{\tau(x)}(x).$$

The map $\mathcal{S}$ is a lift of $\mathcal{T}$ to the space of zippered rectangles: indeed, if

$$\mathcal{S}(\lambda, h, a, \pi) = (\lambda', h', a', \pi'),$$

then $(\lambda', \pi') = \mathcal{T}(\lambda', \pi')$.

Since $\mathcal{Y}(\mathcal{R})$ is a transversal to the flow, the measure $\mu_{\mathcal{R}}$ induces an absolutely continuous measure $\mu_{\mathcal{R}}^{(1)}$ on $\mathcal{Y}(\mathcal{R})$; since $\mu_{\mathcal{R}}$ is both $\mathcal{U}$-invariant and P^t-invariant, the measure $\mu_{\mathcal{R}}^{(1)}$ is $\mathcal{S}$-invariant. Since $\mu_{\mathcal{R}}(\Omega_0(\mathcal{R})) < \infty$, the measure $\mu_{\mathcal{R}}^{(1)}$ is conservative; it is, however, infinite (Veech [1]).

Zorich [4] constructed a different section for the flow P^t, for which the restricted measure has finite total mass.

Following Zorich [4], define

$$\Omega^+(\mathcal{R}) = \{x = (\lambda, h, a, \pi) : (\lambda, \pi) \in \Delta^+, a_m \geq 0\},$$
$$\Omega^-(\mathcal{R}) = \{x = (\lambda, h, a, \pi) : (\lambda, \pi) \in \Delta^-, a_m \leq 0\},$$
$$\mathcal{Y}^+(\mathcal{R}) = \mathcal{Y}(\mathcal{R}) \cap \Omega^+(\mathcal{R}),\ \mathcal{Y}^-(\mathcal{R}) = \mathcal{Y}(\mathcal{R}) \cap \Omega^-(\mathcal{R}),\ \mathcal{Y}^{\pm}(\mathcal{R}) = \mathcal{Y}^+(\mathcal{R}) \cup \mathcal{Y}^-(\mathcal{R}).$$

Take $x \in \mathcal{Y}^{\pm}(\mathcal{R})$, $x = (\lambda, h, a, \pi)$, and define

$$\mathcal{F}(x) = \mathcal{S}^{n(\lambda,\pi)}x.$$

The map $\mathcal{F}$ is a lift of the map $\mathcal{G}$ to the space of zippered rectangles: if

$$\mathcal{F}(\lambda, h, a, \pi) = (\lambda', h', a', \pi'),$$

then $(\lambda', \pi') = \mathcal{G}(\lambda', \pi')$.

We shall see, moreover, that the map $\mathcal{F}$ can be almost surely (with respect to Lebesgue measure) identified with the natural extension of the map $\mathcal{G}$ (Section 3).

If $x \in \mathcal{Y}^+$, then $\mathcal{F}(x) \in \mathcal{Y}^-$, and if $x \in \mathcal{Y}^-$, then $\mathcal{F}(x) \in \mathcal{Y}^+$. The map $\mathcal{F}$ is the induced map of $\mathcal{S}$ to the subset $\mathcal{Y}^{\pm}(\mathcal{R})$.

Since $\mathcal{Y}^{\pm}(\mathcal{R})$ is a transversal to the flow P^t, the measure $\mu_{\mathcal{R}}$ naturally induces an absolutely continuous measure $\overline{\nu}$ on $\mathcal{Y}^{\pm}(\mathcal{R})$; since $\mu_{\mathcal{R}}$ is both $\mathcal{U}$-invariant and P^t-invariant, the measure $\overline{\nu}$ is $\mathcal{F}$-invariant. Under the natural projection of $\mathcal{Y}^{\pm}(\mathcal{R})$ onto $\Delta(\mathcal{R})$, given by $(\lambda, h, a, \pi) \to (\lambda, \pi)$, the map $\mathcal{F}$ projects onto the map $\mathcal{G}$ and the measure $\overline{\nu}$ onto the measure ν ([4]).

Zorich [4] proved

Theorem 6. *The measure $\overline{\nu}$ is finite and ergodic for $\mathcal{F}$.*

Since the map $\mathcal{G}$ is exact (as is shown in Section 4), the map $\mathcal{F}$ satisfies the K-property of Kolmogorov, and, in particular, is strongly mixing. Decay of correlations is proven for the map $\mathcal{F}$ as well.

Introduce a metric on the space of zippered rectangles in the following way. Take two zippered rectangles $x = (\lambda, h, a, \pi)$ and $x' = (\lambda', h', a', \pi')$. Write

$$d((\lambda, h, a), (\lambda', h', a')) = \log \frac{\max_i \frac{\lambda_i}{\lambda'_i}, \frac{h_i}{h'_i}, \frac{|a_i|}{|a'_i|}, \frac{|h_i - a_i|}{|h'_i - a'_i|}}{\min_i \frac{\lambda_i}{\lambda'_i}, \frac{h_i}{h'_i}, \frac{|a_i|}{|a'_i|}, \frac{|h_i - a_i|}{|h'_i - a'_i|}}.$$

Define the metric on $\Omega(\mathcal{R})$ by

$$d(x, x') = \begin{cases} d((\lambda, h, a), (\lambda', h', a')) & \text{if } \pi = \pi' \text{ and } \frac{a_m}{a'_m} > 0; \\ 2 + d((\lambda, h, a), \lambda', (h', a')), & \text{otherwise.} \end{cases}$$

As above, for $\alpha > 0$, let H_α be the space of functions $\phi : \mathcal{Y}^{\pm}(\mathcal{R}) \to \mathbb{R}$ such that if $d(x, x') \leq 1$, then $|\phi(x) - \phi(x')| \leq C d(x, x')^\alpha$ for some constant C. Note that the distance $d(x, x')$ is not defined if $a_i = 0$ or $a'_i = 0$ for some $i = 1, \ldots, m$; nothing, therefore, is said about the values of a function from H_α at such points. This does not represent a problem, however, since we only need the space H_α for results concerning decay of correlations and the Central Limit Theorem, and are therefore dealing with functions defined almost everywhere.

Define

$$C_{H_\alpha}(\phi) = \sup_{d(x,x')\le 1} \frac{|\phi(x)-\phi(x')|}{d(x,x')^\alpha}.$$

Theorem 7. *Let $\mathcal{F} : \mathcal{Y}^{\pm}(\mathcal{R}) \to \mathcal{Y}^{\pm}(\mathcal{R})$ be the Rauzy-Veech-Zorich induction map on the space of zippered rectangles and let $\overline{\nu}_{\mathcal{R}}$ be the absolutely continuous invariant probability measure. Let $p > 2$. Then, for any $\alpha > 0$, there exist positive constants C, δ such that for any $\phi, \psi \in H_\alpha \cap L_p(\mathcal{Y}^{\pm}(\mathcal{R}), \overline{\nu}_{\mathcal{R}})$ we have*

$$\begin{aligned} &|\int \phi \ \times \psi \circ \mathcal{F}^{2n} d\overline{\nu}_{\mathcal{R}} - \int \phi d\overline{\nu}_{\mathcal{R}} \int \psi d\overline{\nu}_{\mathcal{R}}| \\ &\quad \le C \exp(-\delta n^{1/6})(C_{H_\alpha}(\phi) + |\phi|_{L_p})(C_{H_\alpha}(\psi) + |\psi|_{L_p}) \end{aligned}$$

Theorem 7 will be established simultaneosuly with Theorem 4. Indeed, the map $\mathcal{F}$ can be almost surely identified with the natural extension of the map $\mathcal{G}$, and the method of Markov approximations of Sinai [13] and Bunimovich–Sinai [14] allows us to obtain the decay of correlations for the invertible case simultaneously with that for the noninvertible one.

Since the flow P^t is a special flow over the map $\mathcal{F}$, by the theorem of Melbourne and Török [15], the decay of correlations for the map $\mathcal{F}$ allows us to obtain the Central Limit Theorem for the flow P^t.

Denote by X_t the derivative with respect to the flow P^t.

Theorem 8. *Let $p > 2$ and let $\phi \in H_\alpha(\Omega_0(\mathcal{R})) \cap L_p(\Omega_0(\mathcal{R}), \mu_{\mathcal{R}})$ satisfy $\int \phi d\nu = 0$. Assume that there does not exist $\psi \in L_2(\Omega_0(\mathcal{R}), \mu_{\mathcal{R}})$ such that $\phi = X_t\psi$. Then there exists $\sigma > 0$, depending only on ϕ, such that*

$$\frac{1}{\sqrt{T}} \int_0^T \phi \circ P^t \xrightarrow{d} \mathcal{N}(0,\sigma) \text{ as } T \to \infty.$$

This theorem will be proved in Section 16.

1.7. **Zippered rectangles and the moduli space of holomorphic differentials.** Let $g \ge 2$ be an integer. Take an arbitrary integer vector $\kappa = (k_1, \ldots, k_\sigma)$ such that $k_i > 0$, $k_1 + \cdots + k_\sigma = 2g - 2$.

Denote by $\mathcal{M}_\kappa$ the moduli space of Riemann surfaces of genus g endowed with a holomorphic differential of area 1 with singularities of orders $k_1, \ldots, k_\sigma$ (the *stratum* in the moduli space of holomorphic differentials). Denote by g_t the Teichmüller flow on $\mathcal{M}_\kappa$ (see [10], [21], [28], [29]). The flow g_t preserves a natural absolutely continuous probability measure on $\mathcal{M}_\kappa$ ([21], [1], [29]). We denote that measure by μ_κ.

A zippered rectangle naturally defines a Riemann surface endowed with a holomorphic differential. This correspondence preserves area. The orders of the singularities of ω are uniquely defined by the Rauzy class of the permutation π ([1]). For any $\mathcal{R}$ we thus have a map

$$\pi_{\mathcal{R}} : \Omega(\mathcal{R}) \to \mathcal{M}_\kappa,$$

where κ is uniquely defined by $\mathcal{R}$.

Veech [1] proved the following.

Theorem 9 (Veech). (1) *Up to a set of measure zero, the set $\pi_{\mathcal{R}}(\Omega_0(\mathcal{R}))$ is a connected component of $\mathcal{M}_\kappa$. Any connected component of any $\mathcal{M}_\kappa$ has the form $\pi_{\mathcal{R}}(\Omega_0(\mathcal{R}))$ for some $\mathcal{R}$.*

(2) *The map $\pi_{\mathcal{R}}$ is finite-to-one and almost everywhere locally bijective.*

(3) $\pi_{\mathcal{R}}(\mathcal{U}x) = \pi_{\mathcal{R}}(x)$.

(4) *The flow P^t on $\Omega_0(\mathcal{R})$ projects under $\pi_{\mathcal{R}}$ to the Teichmüller flow g_t on the corresponding connected component of $\mathcal{M}_\kappa$.*

(5) $(\pi_{\mathcal{R}})_*\mu_\kappa = \mu_{\mathcal{R}}$.

A detailed treatment of the relationship between Rauzy classes, zippered rectangles and connected components is given by M. Kontsevich and A. Zorich in [26].

Say that a function $\psi : \mathcal{M}_\kappa \to \mathbb{R}$ is *Hölder in the sense of Veech* if there exists a Hölder function $\phi : \Omega_0(\mathcal{R}) \to \mathbb{R}$ such that $\psi \circ \pi_{\mathcal{R}} = \phi$.

Remark 3. This definition has a natural interpretation in terms of the cohomological coordinates of Hubbard and Masur [28]. Indeed, under the map $\pi_{\mathcal{R}}$ the Veech coordinates on the space of zippered rectangles correspond, up to a linear change of variables, to the cohomological coordinates of Hubbard and Masur. Locally, one can associate a Hilbert metric to those coordinates. A function is Hölder in the sense of Veech if and only if it is Hölder with respect to that metric. Note that the thus defined local Hilbert distance between two elements in $\mathcal{M}_\kappa$ majorizes the Teichmüller distance between their underlying surfaces. Therefore, if a function $\phi : \mathcal{M}_\kappa \to \mathbb{R}$ is a lift of a smooth function from the underlying moduli space $\mathcal{M}_g$ of compact surfaces of genus g, then ϕ is Hölder in the sense of Veech.

Denote by $\mathcal{X}_t$ the derivative in the direction of the flow g_t.

Theorem 8 and Theorem 9 imply the following.

Theorem 10. *Let $\mathcal{H}$ be a connected component of $\mathcal{M}_\kappa$. Let $p > 2$, and let $\psi \in L_p(\mathcal{H}, \mu_\kappa)$ be Hölder in the sense of Veech and satisfy $\int \phi d\mu_\kappa = 0$. Assume that there does not exist $\psi \in L_2(\mathcal{H}, \mu_\kappa)$ such that $\phi = \mathcal{X}_t\psi$. Then there exists $\sigma > 0$ such that*

$$\frac{1}{\sqrt{T}} \int_0^T \phi \circ g_t dt \xrightarrow{d} \mathcal{N}(0, \sigma) \text{ as } T \to \infty.$$

1.8. **Outline of the Proof of Theorem 4.** First, one takes a subset of the space $\Delta(\mathcal{R})$ such that the induced map of $\mathcal{G}$ is uniformly expanding (namely, the set of all interval exchanges such that the renormalization matrix for them is a fixed matrix all of whose elements are positive; see Proposition 4. Note that the return map on such a subset is an essential element in Veech's proof of unique ergodicity [1]). Then one estimates the statistics of return times in this subset, in the spirit of Lai-Sang Young [11]. After that, the method of Markov approximations, due to Sinai [13], Bunimovich and Sinai [14], is used to complete the proof.

The paper is organized as follows. In Section 2, we state auxiliary propositions about unimodular matrices. In Section 3, following Veech [1] and Zorich [4], we construct symbolic dynamics for the Rauzy-Veech-Zorich induction map $\mathcal{G}$, compute its transition probabilities in the sense of Sinai [13], and identify the natural extension of $\mathcal{G}$ with $\mathcal{F}$. In Section 4, we establish the exactness of $\mathcal{G}^2$. In Section 6, we state the main Lemma 4, whose proof takes Sections 6 – 10. In the remainder of the paper we apply the Markov approximation method of Sinai [13], Bunimovich and Sinai [14], in order to obtain the decay of correlations for $\mathcal{G}$ and $\mathcal{F}$. In the final section, we apply the theorem of Melbourne and Török to obtain the Central Limit Theorem for the Teichmüller flow.

2. Matrices

Let A be an $m \times m$ matrix with positive entries. Denote

$$|A| = \sum_{i,j=1}^{m} A_{ij},$$
$$col(A) = \max_{i,j,k} \frac{A_{ij}}{A_{kj}},$$
$$row(A) = \max_{i,j,k} \frac{A_{ij}}{A_{ik}}.$$

Proposition 2. *Let Q be a matrix with positive entries, and let A be a matrix with nonnegative entries without zero columns or rows. Then all entries of the matrices AQ and QA are positive, and, moreover, we have*

$$row(AQ) \leq row(Q), col(QA) \leq col(Q).$$

Corollary 2. *Let Q be a matrix with positive entries and A a matrix with nonnegative entries without zero columns or rows. Then*

$$row(QAQ) \leq row(Q), col(QAQ) \leq col(Q).$$

These statements are proven by a straightforward computation.

Now let A be an $m \times m$ matrix with nonnegative entries and determinant 1. Consider the map $J_A : \Delta_{m-1} \to \Delta_{m-1}$ given by

$$J_A(\lambda) = \frac{A\lambda}{|A\lambda|}.$$

Then, with respect to any linear system of coordinates on the simplex Δ_{m-1}, we have (see, for example, [3]):

$$detDJ_A(\lambda) = \frac{1}{|A\lambda|^m}. \tag{14}$$

Suppose all entries of A are positive; then, for any $\lambda, \lambda' \in \Delta_{m-1}$, we have

$$row(A)^{-m} \leq \frac{detDJ_A(\lambda)}{detDJ_A(\lambda')} \leq row(A)^m, \tag{15}$$

whence we have the following.

Proposition 3. *Let $C_1 \subset \Delta_{m-1}$, $C_2 \subset \Delta_{m-1}$ be measurable sets, and let A be a matrix with positive entries and determinant 1. Then*

$$row(A)^{-m} \frac{\mathbf{m}(C_1)}{\mathbf{m}(C_2)} \leq \frac{\mathbf{m}(J_A(C_1))}{\mathbf{m}(J_A(C_2))} \leq row(A)^m \frac{\mathbf{m}(C_1)}{\mathbf{m}(C_2)}.$$

Finally, note the following well-known lemma (see, for example, [17], pp. 16-20).

Lemma 1. *Suppose all entries of the matrix A are positive. Then the map J_A is uniformly contracting with respect to the Hilbert metric on Δ_{m-1}.*

3. Symbolic dynamics for $\mathcal{G}$

First, following Veech [1] and Zorich [4], we describe a Markov partition and a symbolic dynamics for the map $\mathcal{G}^2$. Then we identify almost surely the induction map $\mathcal{F}$ on the space of zippered rectangles with the natural extension of $\mathcal{G}$, and, finally, we compute for $\mathcal{G}$ its transition probabilities in the sense of Sinai [25].

3.1. **A Markov partition.**

Let $\pi \in \mathcal{R}$, and let n be a positive integer. Set

$$\Lambda(a,n,\pi) = \{\lambda\colon \text{there exists } (\lambda',\pi') \text{ such that } \lambda' \in \Delta^+_{\pi'} \text{ and } (\lambda,\pi) = t_{a^{-n}}(\lambda',\pi')\},$$
$$\Delta(a,n,\pi) = \{(\lambda,\pi), \lambda \in \Lambda(a,n,\pi)\}.$$

In other words, $\Delta(a,n,\pi)$ is the set of interval exchange transformations such that the application of the Zorich induction results in the application of the a-operation n times.

The sets $\Delta(a,n,\pi)$ and $\Delta(a,n',\pi')$ are disjoint unless $n=n'$, $\pi=\pi'$, and

$$\Delta^-_\pi = \bigcup_{n=1}^{\infty} \Delta(a,n,\pi)$$

up to a set of measure zero (namely, a union of countably many hyperplanes on which Zorich induction is not defined).

If $\pi' = a^n\pi$, then we have

$$\mathcal{G}\Delta(a,n,\pi) = \Delta^+_{\pi'}.$$

Similarly, for $\pi \in \mathcal{R}$, and n a positive integer, set

$$\Lambda(b,n,\pi) = \{\lambda : \text{there exists } (\lambda',\pi') \text{ such that } \lambda' \in \Delta^-_{\pi'} \text{ and } (\lambda,\pi) = t_{b^{-n}}(\lambda',\pi')\},$$
$$\Delta(b,n,\pi) = \{(\lambda,\pi), \lambda \in \Lambda(b,n,\pi)\}.$$

In other words, $\Delta(b,n,\pi)$ is the set of interval exchange transformations such that the application of the Zorich induction results in the application of the b-operation n times.

The sets $\Delta(b,n,\pi)$ and $\Delta(b,n',\pi')$ are disjoint unless $n=n'$, $\pi=\pi'$, and

$$\Delta^+_\pi = \bigcup_{n=1}^{\infty} \Delta(b,n,\pi)$$

up to a set of measure zero (namely, a union of countably many hyperplanes on which the Zorich induction is not defined).

If $\pi' = b^n\pi$, then, clearly,

$$\mathcal{G}(\Delta(b,n,\pi)) = \Delta^-_{\pi'}.$$

Note that the sets $\Delta(a,n,\pi)$ and $\Delta(b,n',\pi')$ are always disjoint, since we have $\Delta(a,n,\pi) \subset \Delta^-_\pi$, $\Delta(b,n',\pi') \subset \Delta^+_{\pi'}$.

The sets $\Delta(a,n,\pi)$, $\Delta(b,n,\pi)$, for all $n>0$ and all $\pi \in \mathcal{R}$, form a Markov partition for $\mathcal{G}$.

3.2. **Words.** Consider the alphabet

$$\mathcal{A} = \{(c, n, \pi) | \ c = a \text{ or } b, n \in \mathbb{N}, \pi \in \mathcal{R}\}.$$

For $w_1 \in \mathcal{A}$, $w_1 = (c_1, n_1, \pi_1)$, we write $c_1 = c(w_1), \pi_1 = \pi(w_1), n_1 = n(w_1)$. For $w_1, w_2 \in \mathcal{A}$, $w_1 = (c_1, n_1, \pi_1)$, $w_2 = (c_2, n_2, \pi_2)$, define the function $B(w_1, w_2)$ in the following way: $B(w_1, w_2) = 1$ if $c_1^{n_1}\pi_1 = \pi_2$ and $c_1 \neq c_2$ and $B(w_1, w_2) = 0$ otherwise. Let

$$W_{\mathcal{A},B} = \{w = w_1 \dots w_n, w_i \in \mathcal{A}, B(w_i, w_{i+1}) = 1 \text{ for all } i = 1, \dots, n\}.$$

For $w_1 \in \mathcal{A}$, $w_1 = (c_1, n_1, \pi_1)$, set

$$A(w_1) = A(c_1, \pi_1)A(c_1, c_1\pi_1) \dots A(c_1, c_1^{n_1-1}\pi_1),$$

and for $w \in W_{\mathcal{A},B}$, $w = w_1 \dots w_n$, set

$$A(w) = A(w_1) \dots A(w_n).$$

The action of words from $W_{\mathcal{A},B}$ on permutations from $\mathcal{R}$ is introduced as follows.

If $w_1 \in \mathcal{A}$, $w_1 = (c_1, n_1, \pi_1)$, then we set $w_1\pi_1 = c_1^{n_1}\pi_1$. For permutations $\pi \neq \pi_1$, the symbol $w_1\pi$ is not defined. Furthermore, for $w \in W_{\mathcal{A},B}$, $w = w_1 \dots w_n$, we write

$$w\pi = w_n(w_{n-1}(\dots w_1\pi) \dots),$$

assuming the left-hand side of the expression is defined. Finally, if $\pi' = w\pi$, then we also write $\pi = w^{-1}\pi'$.

Say that $w_1 \in \mathcal{A}$ is compatible with $(\lambda, \pi) \in \Delta(\mathcal{R})$ if either

(1) $\lambda \in \Delta_\pi^+$, $c_1 = a$, and $a^{n_1}\pi_1 = \pi$, or
(2) $\lambda \in \Delta_\pi^-$, $c_1 = b$, and $b^{n_1}\pi_1 = \pi$.

Say that a word $w \in W_{\mathcal{A},B}$, $w = w_1 \dots w_n$ is compatible with (λ, π) if w_n is compatible with (λ, π). We shall also sometimes say that (λ, π) is compatible with w instead of saying that w is compatible with (λ, π). We can write

$$\mathcal{G}^{-n}(\lambda, \pi) = \{t_w(\lambda, \pi) : |w| = n \text{ and } w \text{ is compatible with } \ (\lambda, \pi)\}.$$

Recall that the set $\mathcal{G}^{-n}(\lambda, \pi)$ is infinite.

Now take $w \in W_{\mathcal{A},B}$, and for (λ, π) compatible with w, define

$$t_w(\lambda, \pi) = \left(\frac{A(w)\lambda}{|A(w)\lambda|}, w^{-1}\pi\right).$$

Consider also the map

$$T_w(\lambda, \pi) = (A(w)\lambda, w^{-1}\pi).$$

For $w_1 \in \mathcal{A}$, $w_1 = (c_1, n_1, \pi_1)$, we write $\Delta(w_1) = \Delta(c_1, n_1\pi_1)$.

For $w \in W_{\mathcal{A},B}$, $w = w_1 \dots w_n$, denote

$$\Delta(w) = \{t_w(\lambda, \pi) | \ (\lambda, \pi) \ \text{ is compatible with } w\}.$$

Then, by definition,

$$\Delta(w) = \{(\lambda, \pi) : (\lambda, \pi) \in \Delta(w_1), \mathcal{G}(\lambda, \pi) \in \Delta(w_2), \dots, \mathcal{G}^{n-1}(\lambda, \pi) \in \Delta(w_n)\}.$$

We shall also sometimes write Δ_w instead of $\Delta(w)$.

Suppose that a word $w \in W_{\mathcal{A},B}$ is compatible with both (λ, π) and (λ', π). Then

$$d(t_w(\lambda, \pi), t_w(\lambda', \pi)) \leq d((\lambda, \pi), (\lambda', \pi')).$$

If, moreover, all entries of the matrix $A(w)$ are positive, then, by Lemma 1, there exists $\alpha(w)$, $0 < \alpha(w) < 1$, such that

$$d(t_w(\lambda,\pi), t_w(\lambda',\pi)) \le \alpha(w) d((\lambda,\pi),(\lambda',\pi')).$$

We therefore have

Proposition 4. *Let $w \in W_{\mathcal{A},B}$ be such that all entries of the matrix $A(w)$ are positive. Then the return map of $\mathcal{G}$ on $\Delta(w)$ is uniformly expanding with respect to the Hilbert metric.*

3.3. **Sequences.** Now let

$$\Omega_{\mathcal{A},B} = \{\omega = \omega_1 \dots \omega_n \dots,\ \omega_n \in \mathcal{A}, B(\omega_n, \omega_{n+1}) = 1 \text{ for all } n \in \mathbb{N}\}$$

and

$$\Omega^{\mathbb{Z}}_{\mathcal{A},B} = \{\omega = \dots \omega_{-n} \dots \omega_1 \dots \omega_n \dots,\ \omega_n \in \mathcal{A}, B(\omega_n, \omega_{n+1}) = 1 \text{ for all } n \in \mathbb{Z}\}.$$

Denote by σ the right shift on both of these spaces.

There is a natural map $\Phi : \Delta \to \Omega_{\mathcal{A},B}$ given by the formula

$$\Phi(\lambda,\pi) = \omega_1 \dots \omega_n \dots$$

if

$$\mathcal{G}^n(\lambda,\pi) \in \Delta(\omega_n).$$

The measure ν projects under Φ to a σ-invariant measure on $\Omega_{\mathcal{A},B}$; probability with respect to that measure will be denoted by $\mathbb{P}$.

For $w \in \mathcal{W}_{\mathcal{A},B}$, $w = w_1 \dots w_n$, let

$$C(w) = \{\omega \in \Omega_{\mathcal{A},B} : \omega_1 = w_1, \dots, \omega_n = w_n\}.$$

We have then

$$\Delta(w) = \Phi^{-1}(C(w)).$$

W. Veech [1] has proved the following.

Proposition 5. *The map Φ is ν-almost surely bijective.*

We thus obtain a symbolic dynamics for the map $\mathcal{G}$.

3.4. **The natural extension.** Consider the natural extension for the map $\mathcal{G}$. The phase space is the space of sequences of interval exchanges; it will be convenient to number them by negative integers. We set:

$$\begin{aligned}\overline{\Delta}(\mathcal{R}) = \{\mathbf{x} = \dots (\lambda(-n), \pi(-n)), \dots, (\lambda(0), \pi(0))|\\ \mathcal{G}(\lambda(-n), \pi(-n)) = (\lambda(1-n), \pi(1-n)), n = 1, \dots\}.\end{aligned}$$

By definition of the natural extension, the map $\mathcal{G}$ and the invariant measure ν are extended to $\overline{\Delta}$ in the natural way, and the extended map is invertible. We shall still denote the probability with respect to the extended measure by $\mathbb{P}$.

We extend the map Φ to a map

$$\overline{\Phi} : \overline{\Delta} \to \Omega^{\mathbb{Z}}_{\mathcal{A},B},$$

$$\overline{\Phi}(\lambda) = \dots \omega_{-n} \dots \omega_0 \dots \omega_n \dots,$$

if $(\lambda(-n), \pi(-n)) \in \Delta(\omega_{-n})$, and $\mathcal{G}^n(\lambda(0), \pi(0)) \in \Delta(\omega_n)$. Since the map Φ : $(\Delta(\mathcal{R}), \nu) \to (\Omega_{\mathcal{A},B}, \mathbb{P})$ is almost surely bijective, by definition of the natural extension, the map $\overline{\Phi} : \overline{\Delta} \to \Omega^{\mathbb{Z}}_{\mathcal{A},B}$ is also almost surely bijective.

Now take a zippered rectangle $x \in \mathcal{Y}^{\pm}(\mathcal{R})$, $x = (\lambda, h, a, \pi)$. For $l \in \mathbb{Z}$, set $\mathcal{F}^l(x) = (\lambda(l), h(l), a(l), \pi(l))$ and define $\Phi' : \mathcal{Y}^{\pm} \to \overline{\Delta}$ by the formula

$$\Phi'(x) = (\ldots, (\lambda(-n), \pi(-n)), \ldots, (\lambda(0), \pi(0)).$$

As remarked in subsection 1.6, under the natural projection of $\mathcal{Y}^{\pm}(\mathcal{R})$ onto $\Delta(\mathcal{R})$, given by $(\lambda, h, a, \pi) \to (\lambda, \pi)$, the map $\mathcal{F}$ projects onto the map $\mathcal{G}$ and the measure $\overline{\nu}$ onto the measure ν. Therefore, by definition of the natural extension, the map Φ' is almost surely surjective. Our aim is now to show that it is also almost surely injective. It will be convenient to identify $\overline{\Delta}$ with $\Omega_{\mathcal{A},B}^{\mathbb{Z}}$ by the map $\overline{\Phi}$ and consider the map

$$\tilde{\Phi} : \mathcal{Y}^{\pm}(\mathcal{R}) \to \Omega_{\mathcal{A},B}^{\mathbb{Z}} \tag{16}$$

given by the formula $\tilde{\Phi} = \overline{\Phi} \circ \Phi'$, or, in other words, by the formula

$$\tilde{\Phi} : (\lambda, h, a, \pi) \ \to \ldots \omega_{-n} \ldots \omega_0 \ldots \omega_n \ldots,$$

where

$$(\lambda(n), \pi(n)) \in \Delta(\omega_n)$$

for all $n \in \mathbb{Z}$.

By definition, the map $\tilde{\Phi}$ is almost surely surjective and the measure $\tilde{\Phi}_* \overline{\nu}$ is exactly the probability measure $\mathbb{P}$ on the space $\Omega_{\mathcal{A},B}^{\mathbb{Z}}$.

To complete the identification of the dynamical systems $(\mathcal{Y}^{\pm}(\mathcal{R}), \overline{\nu}, \mathcal{F})$ to $(\Omega_{\mathcal{A},B}^{\mathbb{Z}}, \mathbb{P}, \sigma)$, it remains to show that almost surely there is at most one zippered rectangle corresponding to a given symbolic sequence.

Proposition 6. *Let $\mathbf{q} \in W_{\mathcal{A},B}$ be such that all entries of the matrix $A(\mathbf{q})$ are positive. Let $\omega \in \Omega_{\mathcal{A},B}^{\mathbb{Z}}$ be such that the word $\mathbf{q}$ occurs infinitely many times both in $\omega_0 \omega_1 \ldots \omega_n \ldots$ and in $\ldots \omega_{-n} \ldots \omega_0$. Then there exists at most one zippered rectangle corresponding to ω.*

Remark 4. Ergodicity of the map $\mathcal{G}$ with respect to the measure ν is equivalent to the ergodicity of the shift σ on the space $\Omega_{\mathcal{A},B}^{\mathbb{Z}}$ with respect to the measure $\mathbb{P}$. Therefore, $\mathbb{P}$-almost any sequence from $\Omega_{\mathcal{A},B}^{\mathbb{Z}}$ satisfies the assumptions of the proposition.

Proof of Proposition 6. Write

$$\omega = \ldots \omega_{-n} \ldots \omega_0 \ldots \omega_n \ldots,$$

and let (λ, h, a, π) be a zippered rectangle corresponding to ω; we want to show that (λ, h, a, π) is uniquely defined by ω.

First, since there are infinitely many appearances of the word $\mathbf{q}$ in the "future" $\omega_0 \ldots \omega_n \ldots$ of ω, from Proposition 4 we derive that (λ, π) is uniquely defined by $\omega_0 \ldots \omega_n \ldots$ (see also Veech [1]).

Denote $w(n) = \omega_{-n} \ldots \omega_0$, $(\lambda(-n), h(-n), a(-n), \pi(-n)) = \mathcal{F}^{-n}(\lambda, h, a, \pi)$.

For any n , the interval exchange $(\lambda(-n), \pi(-n))$ corresponds to the symbolic sequence $\omega_{-n} \ldots \omega_0 \ldots$, and, again, is uniquely defined by that sequence.

By definition of the map $\mathcal{F}$, we have

$$\lambda(-n) = \frac{A(w(n))\lambda}{|A(w(n))\lambda|}, \ h(-n) = (A(w(n))^t)^{-1} h \cdot |A(w(n))\lambda|.$$

Projectively, therefore, we have

$$\mathbb{R}_+ h \subset A(w(n))^t \mathbb{R}_+^m.$$

Since the subword $\mathbf{q}$ occurs infinitely many times, the intersection

$$\bigcap_{n=1}^{\infty} A(w(n))^t \mathbb{R}_+^m$$

consists of a single line and the vector h is therefore uniquely determined by the condition $\langle \lambda, h \rangle = 1$.

It remains to determine the vector a. By definition of the map $\mathcal{F}$, from the formula (13) we see that for any n there exists an orthogonal matrix $U(-n)$, uniquely determined by ω, and a vector $v(-n)$, uniquely determined by the vectors $h(-n), \dots, h(0)$ and ω, such that

$$\frac{U(-n)a(-n) + v(-n)}{|A(w(n))\lambda|} = a. \tag{17}$$

Now let n be a moment such that all $\lambda(-n)_i > \frac{1}{100||A(\mathbf{q})||m}$ (there are infinitely many such moments: indeed, any moment corresponding to an occurrence of the word $\mathbf{q}$ would be one of them). Then, by Veech's relations (9)–(12) defining a zippered rectangle, we have $|a(-n)_i| < 100||A(\mathbf{q})||m$ for all $i = 1, \dots, m$ and, since $|A(w(n))\lambda| \to \infty$ as $n \to \infty$, (17) implies that a is also uniquely determined by ω. The proof is complete. □

3.5. **Transition probabilities.** Take a sequence $c_1 \dots c_n \dots \in \Omega_{\mathcal{A},B}$. Following Sinai [25], consider the *transition probability*

$$\mathbb{P}(\omega_1 = c_1 | \omega_2 = c_2, \dots, \omega_n = c_n, \dots) = \lim_{n\to\infty} \frac{\mathbb{P}(c_1 c_2 \dots c_n)}{\mathbb{P}(c_2 \dots c_n)}.$$

In this subsection, we give a formula for this probability in terms of $(\lambda, \pi) = \Phi^{-1}(c_2 \dots c_n \dots)$.

Assume $w_1 \in \mathcal{A}$ is compatible with (λ, π). Denote

$$\mathbb{P}(w_1 | (\lambda, \pi)) = \mathbb{P}(((\lambda(-1), \pi(-1)) = t_{w_1}(\lambda(0), \pi(0)) | (\lambda(0), \pi(0)) = (\lambda, \pi)).$$

If $w_1 \in \mathcal{A}$ is compatible with (λ, π), from the definition of $\mathcal{G}$ and from (14) we have

$$\mathbb{P}(w_1 | (\lambda, \pi)) = \frac{\rho(t_{w_1}(\lambda, \pi))}{\rho(\lambda, \pi) |A(w_1)\lambda|^m}. \tag{18}$$

Since the invariant density is a homogeneous function of degree $-m$, we have

$$\rho(T_{w_1}(\lambda, \pi)) = \frac{\rho(t_{w_1}(\lambda, \pi))}{|A(w_1)\lambda|^m},$$

and we can rewrite (18) as follows:

$$\mathbb{P}(w_1 | (\lambda, \pi)) = \frac{\rho(T_{w_1}(\lambda, \pi))}{\rho(\lambda, \pi)}. \tag{19}$$

Let $w = w_1 \dots w_n$ be compatible with (λ, π). Denote

$$\mathbb{P}(w | (\lambda, \pi)) = \mathbb{P}((\lambda(-k), \pi(-k)) = t_{w_{n-k+1}}(\lambda(1-k), \pi(1-k)), k = 1, \dots, n | \\ (\lambda(0), \pi(0)) = (\lambda, \pi)).$$

From (18), by induction, we have

$$\mathbb{P}(w | (\lambda, \pi)) = \frac{\rho(t_w(\lambda, \pi))}{\rho(\lambda, \pi) |A(w)\lambda|^m}. \tag{20}$$

Since the invariant density is a homogeneous function of degree $-m$, we have

$$\rho(T_w(\lambda,\pi)) = \frac{\rho(t_w(\lambda,\pi))}{|A(w)\lambda|^m},$$

and we can rewrite (20) as follows:

$$\mathbb{P}(w|(\lambda,\pi)) = \frac{\rho(T_w(\lambda,\pi))}{\rho(\lambda,\pi)}. \tag{21}$$

Corollary 3. *There exists $C > 0$ such that the following is true. Suppose $w \in W_{\mathcal{A},B}$ is compatible with (λ,π). Then*

$$\mathbb{P}(w|(\lambda,\pi)) \geq \frac{C}{\rho(\lambda,\pi)|A(w)|^m}.$$

Proof. Recall that the invariant density is a positive homogeneous function of degree $-m$ and therefore is bounded from below: there exists $C > 0$ such that $\rho(\lambda,\pi) > C$ for all $(\lambda,\pi) \in \Delta(\mathcal{R})$. In particular, $\rho(t_w(\lambda,\pi)) > C$. Substituting into (20), we obtain the result.

For $\epsilon : 0 < \epsilon < 1$, let

$$\Delta_\epsilon = \{(\lambda,\pi) \in \Delta(\mathcal{R}), \min|\lambda_i| \geq \epsilon\}.$$

For any $\epsilon > 0$ there exists a constant $C(\epsilon)$ such that for any $(\lambda,\pi) \in \Delta_\epsilon$ we have $\rho(\lambda,\pi) < C(\epsilon)$. □

Corollary 4. *For any $\epsilon > 0$ there exists $C(\epsilon) > 0$ such that if $(\lambda,\pi) \in \Delta_\epsilon$, then, for every $w \in W_{\mathcal{A},B}$ compatible with (λ,π), we have*

$$\mathbb{P}(w|(\lambda,\pi)) \geq \frac{C(\epsilon)}{|A(w)|^m}.$$

4. Proof of the exactness

First, one notes that $\mathcal{G}^2$ acts without period on the Rauzy class $\mathcal{R}$, and then the proof follows the standard pattern [27, 17]: since almost every point of any measurable subset is a density point, bounded distortion estimates of Proposition 3 imply that if the measure of a tail event is positive, then it must be arbitrarily close to 1.

In more detail, observe that there exists an integer M such that for any $n > M$ and for any $\pi, \pi' \in \mathcal{R}$ there exist $k_1, \ldots, k_{2n}$ such that $a^{k_1}b^{k_2}\ldots a^{k_{2n-1}}b^{k_{2n}}\pi = \pi'$ (in other words, the action of $\mathcal{G}^2$ on the Rauzy class $\mathcal{R}$ is aperiodic). This follows from connectedness of the Rauzy graph and the fact that for any $\pi \in \mathcal{R}$ there exist n_1, n_2 such that $a^{n_1}\pi = b^{n_2}\pi = \pi$.

Now let α_0 be the partition of Δ^+ into Δ^+_π, $\pi \in \mathcal{R}$, and let α_n be the partition into the cylinders $\Delta(w)$, where $w \in W_{\mathcal{A},B}$, $|w| = 2n$. As usual, if $(\lambda,\pi) \in \Delta_{m-1}$, then $\alpha_n(\lambda,\pi)$ stands for the element of the partition α_n, containing (λ,π).

Lemma 2. *There exists $k > 0$ such that the following is true. Suppose $C \subset \Delta^+$, and there exists $\pi \in \mathcal{R}$ such that $\Delta^+_\pi \subset C$. Then $\mathcal{G}^{2k}C = \Delta^+(\mathcal{R})$.*

This lemma easily follows from aperiodicity; in turn, it immediately implies the following.

Lemma 3. *There exists $k > 0$ such that the following holds. For any $\varepsilon > 0$ there is $\delta > 0$ such that for any $C \subset \Delta^+(\mathcal{R})$ satisfying $\mathbf{m}(C \triangle \Delta^+_\pi) < \delta$ for some $\pi \in \mathcal{R}$, we have $\mathbf{m}(\mathcal{G}^{2k}C \triangle \Delta^+) < \varepsilon$.*

Lemma 3 is a straightforward corollary of Lemma 2 and its proof is omitted.

We are now ready to prove the exactness of the map $\mathcal{G}^2$. Suppose $C \subset \Delta^+$ is a $\mathcal{G}^2$-tail event, i.e., for any $n > 0$ there exists B_n such that $C = \mathcal{G}^{-2n} B_n$ and $0 < \nu(C) < 1$. Then $\nu(B_n) = \nu(C)$ and, by Lemma 3, we can assume that there exists $\varepsilon > 0$ such that for any $\pi \in \mathcal{R}$, we have

$$\mathbf{m}((\Delta^+ \setminus C) \cap \Delta^+_\pi) \geq \varepsilon. \tag{22}$$

Let $\mathbf{q} = q_1 \dots q_{2l}$ be a word such that the matrix $A(\mathbf{q})$ is positive.

For almost any $(\lambda, \pi) \in C$ we have

$$\lim_{n \to \infty} \frac{\mathbf{m}(\alpha_n(\lambda, \pi) \cap C)}{\mathbf{m}(\alpha_n(\lambda, \pi))} = 1. \tag{23}$$

Now let n be such that $\mathcal{G}^{2n-2l}(\lambda, \pi) \in \Delta(\mathbf{q})$. Denote $(\lambda', \pi') = \mathcal{G}^{2n}(\lambda, \pi)$. Let A be the corresponding renormalization matrix, that is, $\lambda = J_A \lambda'$. Then $A = A_1 A(\mathbf{q})$ for some (unimodular nonnegative integer) matrix A_1. We have $\alpha_n(\lambda, \pi) = J_A(\Delta^+_{\pi'})$. By Proposition 2, we have $row(A) \leq row(A(\mathbf{q}))$. By Proposition 3, from (22), we deduce that there exists ε', not depending on n, such that

$$\frac{\mathbf{m}(\alpha_n(\lambda, \pi) \cap (\Delta^+ \setminus C))}{\mathbf{m}(\alpha_n(\lambda, \pi))} \geq \varepsilon'.$$

But by ergodicity of the map $\mathcal{G}$, for almost any (λ, π) we can find infinitely many n such that $\mathcal{G}^{2n-2l}(\lambda, \pi) \in \Delta(\mathbf{q})$. We therefore arrive at a contradiction with (23), which gives the exactness of $\mathcal{G}^2$. Proposition 1 is proved.

5. The main lemma

We shall suppose from now on that the Rauzy class $\mathcal{R}$ is fixed and will often suppress it from notation. Take a point $\mathbf{x} \in \overline{\Delta}(\mathcal{R})$,

$$\mathbf{x} = \dots (\lambda(-n), \pi(-n)), \dots, (\lambda(0), \pi(0)),$$

where $\mathcal{G}(\lambda(-n), \pi(-n)) = (\lambda(1-n), \pi(1-n))$ for all n.

By ergodicity of the map $\mathcal{G}$, for $\mathbb{P}$-almost any point $\mathbf{x} \in \overline{\Delta}$, and a $\gamma : 0 < \gamma < \frac{1}{m}$ there exists n such that $(\lambda(-n), \pi(-n))$ satisfies the inequality $\lambda(-n)_i > \gamma$ for all i. Our aim in the following lemma is to estimate the "waiting time" n.

For $\epsilon : 0 < \epsilon < 1$, write, in the same way as above,

$$\Delta_\epsilon = \{(\lambda, \pi) \in \Delta(\mathcal{R}), \min |\lambda_i| \geq \epsilon\}.$$

We have then

Lemma 4. *There exist positive constants γ, K, p such that the following is true for any $\epsilon > 0$. Suppose $(\lambda, \pi) \in \Delta_\epsilon$. Then*

$$\mathbb{P}(\exists n \leq K |\log \epsilon|, (\lambda(-n), \pi(-n)) \in \Delta_\gamma | (\lambda(0), \pi(0)) = (\lambda, \pi)) \geq p.$$

Remark 5. The probability should of course be understood in the space $\overline{\Delta}$.

Corollary 5. *Let $\mathbf{q} \in W_{\mathcal{A},B}$, $\mathbf{q} = q_1 \dots q_l$ be such that all entries of the matrix $A(\mathbf{q})$ are positive. Then there exist positive constants $K(\mathbf{q}), p(\mathbf{q})$ such that the following is true for any $\epsilon > 0$. Suppose $(\lambda, \pi) \in \Delta_\epsilon$. Then*

$$\mathbb{P}(\exists n \leq K(\mathbf{q}) |\log \epsilon|, (\lambda(-n), \pi(-n)) \in \Delta(\mathbf{q}) | (\lambda(0), \pi(0)) = (\lambda, \pi)) \geq p(\mathbf{q}).$$

Proof of the Corollary. By Lemma 4, we may consider $(\lambda(-n), \pi(-n)) \in \Delta_\gamma$. Using Corollary 4, we can control the probability of any word w, compatible with $(\lambda(-n), \pi(-n))$. If $\mathbf{q}$ is compatible, we are done; if not, observe that any two vertices in the Rauzy graph can be joined by a path of conditional probability not less than a constant, depending only on γ, and the corollary is proved. □

The proof of Lemma 4 proceeds by getting rid of small intervals.

For $\gamma > 0$, $k \le m$, denote

$$\Delta_{\gamma,k} = \{(\lambda, \pi) : \exists i_1, \dots, i_k : \lambda_{i_1}, \dots, \lambda_{i_k} \ge \gamma\}$$

and

$$\Delta_{\gamma,k,\epsilon} = \{(\lambda, \pi) : \lambda_i \ge \epsilon \text{ for all } i = 1, \dots, m \text{ and } \exists i_1, \dots, i_k : \lambda_{i_1}, \dots, \lambda_{i_k} \ge \gamma\}.$$

Lemma 5. *There exist constants L, K, p, depending only on the Rauzy class, such that the following is true for any $\gamma > 0$, $\epsilon > 0$, $k < m$. Assume $(\lambda, \pi) \in \Delta_{\gamma,k,\epsilon}$.*

Then

$$\mathbb{P}(\exists n \le K|\log \epsilon| : (\lambda(-n), \pi(-n)) \in \Delta_{\gamma/L,k+1,\epsilon/L} | (\lambda(0), \pi(0)) = (\lambda, \pi)) \ge p.$$

Lemma 5 says that, departing from an interval exchange having k "large" subintervals of length not less than γ, with positive probability we can arrive at an interval exchange having $k+1$ subintervals of length not less than γ/L. Proceeding inductively, we make all subintervals "large". To prove Lemma 4, it thus suffices to establish Lemma 5.

Lemma 5 is proved in the next four sections.

6. An estimate on the number of Rauzy operations

Recall that, if $(\lambda, \pi) \in \Delta^+$, then the $\mathcal{G}$-preimages of (λ, π) are the exchanges $t_{a^{-n}}(\lambda, \pi)$, $n = 1, \dots$, whereas if $(\lambda, \pi) \in \Delta^-$, then the $\mathcal{G}$-preimages of (λ, π) are the exchanges $t_{b^{-n}}(\lambda, \pi)$, $n = 1, \dots$.

Denote

$$\mathbf{p}_n(\lambda, \pi) = \begin{cases} \mathbb{P}((\lambda(-1), \pi(-1)) = t_{a^{-n}}(\lambda, \pi) | (\lambda(0), \pi(0)) = (\lambda, \pi)), & \text{if } (\lambda, \pi) \in \Delta^+ ; \\ \mathbb{P}((\lambda(-1), \pi(-1)) = t_{b^{-n}}(\lambda, \pi) | (\lambda(0), \pi(0)) = (\lambda, \pi)), & \text{if } (\lambda, \pi) \in \Delta^-. \end{cases}$$

In other words, $\mathbf{p}_n(\lambda, \pi)$ is the probability that exactly n elementary Rauzy operations were applied in passing from $(\lambda(0), \pi(0))$ to $((\lambda(-1), \pi(-1))$.

Throughout this section, we shall be mainly interested in estimating, for a given $k > 0$, the quantity

$$\sum_{n>k} \mathbf{p}_n(\lambda, \pi),$$

that is, the probability that more than k Rauzy operations were applied.

For $\lambda \in \mathbb{R}^m_+$, set

$$T^{(\pi)}_{a^{-1}}(\lambda) = A(a^{-1}\pi, a)\lambda, \quad t^{(\pi)}_{a^{-1}}(\lambda) = \frac{A(a^{-1}\pi, a)\lambda}{|A(a^{-1}\pi, a)\lambda|},$$
$$T^{(\pi)}_{b^{-1}}(\lambda) = A(b^{-1}\pi, b)\lambda, \quad t^{(\pi)}_{b^{-1}}(\lambda) = \frac{A(b^{-1}\pi, b)\lambda}{|A(b^{-1}\pi, b)\lambda|},$$

and

$$T^{(\pi)}_{a^{-n}}(\lambda) = T^{(a^{1-n}\pi)}_{a^{-1}} \dots T^{(\pi)}_{a^{-1}}\lambda, \quad t^{(\pi)}_{a^{-n}}(\lambda) = t^{(a^{1-n}\pi)}_{a^{-1}} \dots t^{(\pi)}_{a^{-1}}\lambda,$$
$$T^{(\pi)}_{b^{-n}}(\lambda) = T^{(b^{1-n}\pi)}_{b^{-1}} \dots T^{(\pi)}_{b^{-1}}\lambda, \quad t^{(\pi)}_{b^{-n}}(\lambda) = t^{(b^{1-n}\pi)}_{b^{-1}} \dots t^{(\pi)}_{b^{-1}}\lambda,$$

so that we have

$$t_{a^{-n}}(\lambda,\pi) = (t^{(\pi)}_{a^{-n}}\lambda, a^{-n}\pi), \quad T_{a^{-n}}(\lambda,\pi) = (T^{(\pi)}_{a^{-n}}\lambda, a^{-n}\pi),$$
$$t_{b^{-n}}(\lambda,\pi) = (t^{(\pi)}_{b^{-n}}\lambda, b^{-n}\pi), \quad T_{b^{-n}}(\lambda,\pi) = (T^{(\pi)}_{b^{-n}}\lambda, b^{-n}\pi).$$

Lemma 6. *If $(\lambda,\pi) \in \Delta^+$, then, for any $N \geq 1$, we have*

$$\sum_{n=N+1}^{\infty} \mathbf{p}_n(\lambda,\pi) = \frac{\rho^+_{a^{-N}\pi}(T^{(\pi)}_{a^{-N}}(\lambda))}{\rho^+_\pi(\lambda)}.$$

If $(\lambda,\pi) \in \Delta^-$, then, for any $N \geq 1$, we have

$$\sum_{n=N+1}^{\infty} \mathbf{p}_n(\lambda,\pi) = \frac{\rho^-_{b^{-N}\pi}(T^{(\pi)}_{b^{-N}}(\lambda))}{\rho^-_\pi(\lambda)}.$$

Proof. We only consider the case $(\lambda,\pi) \in \Delta^+$. In this case, the formula (19) can be written as

$$\mathbf{p}_n(\lambda,\pi) = \frac{\rho^-_{a^{-n}\pi}(T^{(\pi)}_{a^{-n}}\lambda)}{\rho^+_\pi(\lambda)},$$

whence we can write

$$\rho^+_\pi(\lambda) = \sum_{n=1}^{\infty} \rho^-_{a^{-n}\pi}(T^{(\pi)}_{a^{-n}}\lambda). \tag{24}$$

The series in the left side converges for any $\lambda \in \mathbb{R}^m_+$. The formula thus is true for *any* permutation π and *any* $\lambda \in \mathbb{R}^m_+$: even if $\lambda \notin \Delta^+_\pi$, the equality (24), being an identity between rational functions, still holds.

Since, for any λ, we have

$$T^{(\pi)}_{a^{-n-N}}\lambda = T^{(a^{-N}\pi)}_{a^{-n}}(T^{(\pi)}_{a^{-N}}\lambda),$$

from (24) we obtain

$$\rho^+_{a^{-N}\pi}(T^{(\pi)}_{a^{-N}}\lambda) = \sum_{n=1}^{\infty} \rho^-_{a^{-n-N}\pi}\left(T^{(\pi)}_{a^{-n-N}}\lambda\right) = \rho^+_\pi(\lambda)(\sum_{n=N+1}^{\infty} \mathbf{p}_n(\lambda,\pi)),$$

and the lemma is proved. □

6.1. **Bounded growth.** Let $(\lambda,\pi) \in \Delta(\mathcal{R})$. Define

$$(\lambda^{(n)},\pi^{(n)}) = \begin{cases} t_{a^{-n}}(\lambda,\pi), & \text{if } (\lambda,\pi) \in \Delta^+ ; \\ t_{b^{-n}}(\lambda,\pi), & \text{if } (\lambda,\pi) \in \Delta^-; \end{cases}$$

$$(\Lambda^{(n)},\pi^{(n)}) = \begin{cases} T_{a^{-n}}(\lambda,\pi), & \text{if } (\lambda,\pi) \in \Delta^+ ; \\ T_{b^{-n}}(\lambda,\pi), & \text{if } (\lambda,\pi) \in \Delta^-. \end{cases}$$

We have

$$\mathcal{G}^{-1}(\lambda,\pi) = \{(\lambda^{(n)},\pi^{(n)}), n = 1,\dots\}$$

and

$$\mathbf{p}_n = \mathbb{P}((\lambda(-1),\pi(-1)) = (\lambda^{(n)},\pi^{(n)})|((\lambda(0),\pi(0)) = (\lambda,\pi)).$$

For any $n \in \mathbb{N}$, there exists $i(n) \in \{1, \dots, m\}$ such that

$$|\Lambda^{(n)}| - |\Lambda^{(n-1)}| = \lambda_{i(n)}.$$

If $(\lambda(-1), \pi(-1))$ is a $\mathcal{G}$-preimage of (λ, π) and $(\lambda(-1), \pi(-1)) = t_{c^{-n}}(\lambda, \pi)$, $c = a$ or b, then we define a vector $\Lambda(-1)$ by the relation $(\Lambda(-1), \pi(-1)) = T_{c^{-n}}(\lambda, \pi)$ (in other words, $(\Lambda(-1), \pi(-1))$ is the Zorich preimage without normalization).

Lemma 7. *There exists a constant $C(\mathcal{R})$, depending on the Rauzy class only, such that for any $K > 3$ and for any $(\lambda, \pi) \in \Delta(\mathcal{R})$ we have*

$$\mathbb{P}(|\Lambda(-1)| > K | (\lambda(0), \pi(0)) = (\lambda, \pi)) < \frac{C(\mathcal{R})}{K}.$$

Proof. For definiteness, assume $\lambda \in \Delta_\pi^-$ (the proof is completely identical in the other case). Then $\mathcal{G}$-preimages of (λ, π) are $(\lambda^{(n)}, \pi^{(n)}) = t_{b^{-n}}(\lambda, \pi)$, $n = 1, 2, \dots$.

By construction [4], the invariant density ρ_π^- has the form

$$\rho_\pi^-(\lambda) = \sum_{i=1}^{S} \frac{1}{l_{i1}(\lambda) l_{i2}(\lambda) \dots l_{im}(\lambda)},$$

where the functions l_{ij} are linear:

$$l_{ij}(\lambda) = a_{ij}^{(1)} \lambda_1 + \dots + a_{ij}^{(m)} \lambda_m,$$

and all $a_{ij}^{(r)}$ are nonnegative (in fact, $a_{ij}^{(r)} = 0$ or 1, but we do not need this fact here).

Let l be the length of the b-cycle of π, that is, the smallest such number that $b^l \pi = \pi$.

Since for any $k > 0$ we have $b^{-kl}\pi = \pi$, from Lemma 6 we obtain

$$\sum_{n=kl+1}^{\infty} \mathbf{p}_n(\lambda, \pi) = \frac{\rho_\pi^-(\Lambda^{(kl)})}{\rho_\pi^-(\lambda)}.$$

As noted above, for any $n > 0$ there exists $\lambda_{i(n)}$ such that

$$|\Lambda^{(n)}| - |\Lambda^{(n-1)}| = \lambda_{i(n)},$$

and, in fact,

$$\Lambda^{(n)} = (\lambda_1, \dots, \lambda_{m-1}, \lambda_m + \lambda_{i(1)} + \dots + \lambda_{i(n)}).$$

Since

$$\sum_{n=kl+1}^{\infty} \mathbf{p}_n(\lambda, \pi) \to 0 \text{ as } k \to \infty,$$

for any $i = 1, \dots, S$ there exists j such that $a_{ij}^{(m)} > 0$. Renumbering, if necessary, the linear forms l_{ij}, we may assume that $a_{i1}^{(m)} > 0$ for any i. Among the linear forms l_{i1}, $i = 1, \dots, S$, choose any one which maximizes the ratio $\dfrac{l_{i1}(\lambda)}{l_{i1}(\Lambda^{(kl)})}$. Let it be the form l_{p1}. We first apply inequality $l_{ij}(\Lambda) \geq l_{ij}(\lambda)$ (which is valid for any pair i, j) to all linear forms for which $i > 1$ to increase the numerator of the expression below, and then we use the maximality of the ratio $\dfrac{l_{p1}(\lambda)}{l_{p1}(\Lambda)}$ to obtain the following sequence of relations:

$$\frac{\rho_\pi^-(\Lambda^{(kl)})}{\rho_\pi^-(\lambda)} = \frac{\sum_{i=1}^S \frac{1}{l_{i1}(\Lambda^{(kl)})l_{i2}(\Lambda^{(kl)})\dots l_{im}(\Lambda^{(kl)})}}{\sum_{i=1}^S \frac{1}{l_{i1}(\lambda)l_{i2}(\lambda)\dots l_{im}(\lambda)}}$$

$$\le \frac{\sum_{i=1}^S \frac{1}{l_{i1}(\Lambda^{(kl)})l_{i2}(\lambda)\dots l_{im}(\lambda)}}{\sum_{i=1}^S \frac{1}{l_{i1}(\lambda)l_{i2}(\lambda)\dots l_{im}(\lambda)}} \le \frac{\frac{1}{l_{p1}(\Lambda^{(kl)})}}{\frac{1}{l_{p1}(\lambda)}} = \frac{l_{p1}(\lambda^{(kl)})}{l_{p1}(\Lambda)}.$$

Now denote $\epsilon = \min a_{i1}^{(m)}$ and $L = \max a_{i1}^{(r)}$. For any $\lambda \in \mathbb{R}_+^m$ we have then

$$\epsilon\lambda_m \le l_{i1}(\lambda) \le L|\lambda|,$$

whence

$$\frac{\rho_\pi^-(\Lambda^{(kl)})}{\rho_\pi^-(\lambda)} \le \frac{L}{\epsilon(\lambda_m + \lambda_{i(1)} + \dots + \lambda_{i(kl)})}. \tag{25}$$

Let N be the smallest number such that $|\Lambda(-N)| > K$ and let s be the largest integer such that $sl < N$. Then $|\Lambda(-sl)| > K-1$ (because all $\lambda_{i(sl+1)}, \dots, \lambda_{i(N)}$ are distinct) and $\lambda_m + \lambda_{i(1)} + \dots + \lambda_{i(sl)} > K-2$ (because $|\Lambda(-sl)| = 1 + \lambda_{i(1)} + \dots + \lambda_{i(sl)}$).

Therefore, by (25), we obtain

$$\frac{\rho_\pi^-(\Lambda^{(sl)})}{\rho_\pi^-(\lambda)} \le \frac{L}{\epsilon}\frac{1}{K-2},$$

whence

$$\sum_{n>N} \mathbf{p}_n(\lambda,\pi) \le \sum_{n>sl} \mathbf{p}_n(\lambda,\pi) < \frac{L}{\epsilon}\frac{1}{K-2},$$

and the lemma is proved. □

Lemma 8. *Suppose $(\lambda,\pi) \in \Delta^+$, and let l be the length of the a-cycle of π. Then, for any $k \ge 1$, we have*

$$\sum_{n=kl+1}^{\infty} \mathbf{p}_n(\lambda,\pi) \ge (\frac{\lambda_{\pi^{-1}m}}{\lambda_{\pi^{-1}m}+k})^m.$$

Suppose $(\lambda,\pi) \in \Delta^-$, and let l be the length of the b-cycle of π. Then, for any $k \ge 1$, we have

$$\sum_{n=kl+1}^{\infty} \mathbf{p}_n(\lambda,\pi) \ge (\frac{\lambda_m}{\lambda_m+k})^m.$$

Proof. Again, we only consider the case $(\lambda,\pi) \in \Delta^-$, as the proof of the other case is identical. We have

$$\sum_{n=kl+1}^{\infty} \mathbf{p}_n(\lambda,\pi) = \frac{\rho_\pi^-(\Lambda^{(kl)})}{\rho_\pi^-(\lambda)}.$$

Set $\Lambda^{(kl)} = (\Lambda_1^{(kl)}, \dots, \Lambda_m^{(kl)})$.

For $k=1$ we have $\Lambda_i^{(l)}=\lambda_i$ for $i<m$ and $\Lambda_m^{(l)}=\lambda_m+\lambda_{i(1)}+\dots+\lambda_{i(l)}$, and for arbitrary k by induction we obtain $\Lambda_i^{(kl)}=\lambda_i$ for $i<m$ and $\Lambda_m^{(kl)}=\lambda_m+k(\lambda_{i(1)}+\dots+\lambda_{i(l)})$.

Note that $\lambda_{i(1)}+\dots+\lambda_{i(l)}\le 1$ (since $i(1),\dots,i(l)$ are all distinct).

As in the proof of the previous lemma, write

$$\rho_\pi^-(\lambda)=\sum_{i=1}^{S}\frac{1}{l_{i1}(\lambda)l_{i2}(\lambda)\dots l_{im}(\lambda)},$$

whence

$$\frac{\rho_\pi^-(\Lambda^{(kl)})}{\rho_\pi^-(\lambda)}\ge\min_i\frac{l_{i1}(\lambda)l_{i2}(\lambda)\dots l_{im}(\lambda)}{l_{i1}(\Lambda^{(kl)})l_{i2}(\Lambda^{(kl)})\dots l_{im}(\Lambda^{(kl)})}. \tag{26}$$

For any linear form $l(\lambda)=a_1\lambda_1+\dots+a_m\lambda_m$, $a_i\ge 0$, we have

$$\frac{l(\lambda)}{l(\Lambda^{(kl)})}\ge\frac{\lambda_m}{\lambda_m+k(\lambda_{i(1)}+\dots+\lambda_{i(l))})}\ge\frac{\lambda_m}{\lambda_m+k},$$

and the lemma follows. □

7. An estimate on the probability of stopping

Lemma 9. *For any $\gamma>0$, there exists $c(\gamma)>0$ such that if $\lambda_{i(N)}>\gamma|\Lambda(-N)|$, then*

$$\frac{\mathbf{p}_N(\lambda,\pi)}{\sum_{n=N+1}^{\infty}\mathbf{p}_n(\lambda,\pi)}\ge c(\gamma).$$

Combining with Lemma 8, we immediately obtain the following corollary.

Corollary 6. *For any $\gamma>0$, there exists $c(\gamma)>0$ such that the following is true. Assume $(\lambda,\pi)\in\Delta^+$, $\lambda_{i(N)}>\gamma$, $\lambda_{\pi^{-1}m}>\gamma$. Then*

$$\mathbf{p}_N\ge\frac{c(\gamma)}{N^m}.$$

Similarly, if $(\lambda,\pi)\in\Delta^-$, $\lambda_{i(N)}>\gamma$, $\lambda_m>\gamma$, then

$$\mathbf{p}_N\ge\frac{c(\gamma)}{N^m}.$$

If $(\lambda,\pi)\in\Delta^+$, then, by the definition of $\mathbf{p}_n(\lambda,\pi)$ and by Lemma 6, we have

$$\mathbf{p}_N(\lambda,\pi)=\frac{\rho^-_{a^{-N}\pi}(T^{(\pi)}_{a^{-N}}(\lambda))}{\rho_\pi^+(\lambda)},$$

$$\sum_{n=N+1}^{\infty}\mathbf{p}_n(\lambda,\pi)=\frac{\rho^+_{a^{-N}\pi}(T^{(\pi)}_{a^{-N}}(\lambda))}{\rho_\pi^+(\lambda)},$$

and, therefore,

$$\frac{\mathbf{p}_N(\lambda,\pi)}{\sum_{n=N+1}^{\infty}\mathbf{p}_n(\lambda,\pi)}=\frac{\rho^-_{a^{-N}\pi}(T^{(\pi)}_{a^{-N}}(\lambda))}{\rho^+_{a^{-N}\pi}(T^{(\pi)}_{a^{-N}}(\lambda))}.$$

Lemma 9 follows now from the following.

Lemma 10. *For any $\gamma > 0$ there exists a constant $c(\gamma) > 0$ such that the following is true. Let $(\lambda, \pi) \in \Delta(\mathcal{R})$. If $\lambda_{\pi^{-1}m+1} > \gamma$, then*

$$\frac{\rho_\pi^-(\lambda)}{\rho_\pi^+(\lambda)} \geq c(\gamma).$$

If $\lambda_{\pi^{-1}(\pi(m)+1)} > \gamma$, then

$$\frac{\rho_\pi^+(\lambda)}{\rho_\pi^-(\lambda)} \geq c(\gamma).$$

The proof of Lemma 10 will take the remainder of this section.

First, we modify Veech's coordinates on the space of zippered rectangles. Take a zippered rectangle $(\lambda, h, a, \pi) \in \Omega(\mathcal{R})$ and introduce the vector $\delta = (\delta_1, \ldots, \delta_m) \in \mathbb{R}^m$ by the formula

$$\delta_i = a_{i-1} - a_i, \; i = 1, \ldots, m$$

(here we assume, as always, $a_0 = a_{m+1} = 0$).

Proposition 7. *The data (λ, π, δ) determine the zippered rectangle (λ, h, a, π) uniquely.*

Remark 6. The coordinates (λ, π, δ) on the space of zippered rectangles have a natural interpretation in terms of the cohomological coordinates of Hubbard and Masur [28]: namely, the λ_i are the real parts of the corresponding cycles, and the δ_i are (minus) the imaginary parts.

Proof of Proposition 7. For any $i = 1, \ldots, m$, we have

$$a_i = -\delta_1 - \cdots - \delta_i, \tag{27}$$

so the vector a is uniquely defined by δ. It remains to show that the vector h is uniquely defined by δ, and, to do this, we shall express the h through the a. First note that

$$h_{\pi^{-1}m} = a_{\pi^{-1}m} - a_m.$$

Now, if $i \neq \pi^{-1}m$, then $i = \pi^{-1}(k-1)$ for some $k \in \{1, \ldots, m\}$. The equation

$$h_i - a_i = h_{\pi^{-1}(\pi(i)+1)} - a_{\pi^{-1}(\pi(i)+1)-1} \tag{28}$$

then takes the form

$$h_{\pi^{-1}(k-1)} - a_{\pi^{-1}(k-1)} = h_{\pi^{-1}(k)} - a_{\pi^{-1}(k)-1},$$

or, equivalently,

$$h_{\pi^{-1}(k)} = a_{\pi^{-1}(k)-1} + h_{\pi^{-1}(k-1)} - a_{\pi^{-1}(k-1)}.$$

Since

$$h_{\pi^{-1}(1)} = a_{\pi^{-1}(1)-1},$$

by induction, we obtain

$$h_{\pi^{-1}(k)} = a_{\pi^{-1}(k)-1} + \sum_{l=1}^{k-1} (a_{\pi^{-1}(l)-1} - a_{\pi^{-1}(l)})$$

for any $k = 1, \ldots, m$, and Proposition 7 is proved. $\square$

The above computations give us the following expression for h in terms of δ:

$$h_{\pi^{-1}(k)} = -\sum_{i=1}^{\pi^{-1}k-1} \delta_i + \sum_{l=1}^{k-1} \delta_{\pi^{-1}(l)} \tag{29}$$

or, equivalently,

$$h_r = -\sum_{i=1}^{r-1} \delta_i + \sum_{l=1}^{\pi(r)-1} \delta_{\pi^{-1}(l)} \tag{30}$$

(here our convention is that $\sum_{i=1}^{0}(...) = 0$).

Rewriting in terms of δ the inequalities (9)–(12), defining the zippered rectangle, we obtain by a straightforward computation the following equivalent system:

$$\delta_1 + \cdots + \delta_i \le 0, \quad i = 1, \ldots, m-1,$$
$$\delta_{\pi^{-1}(1)} + \cdots + \delta_{\pi^{-1}(i)} \ge 0, \quad i = 1, \ldots, m-1.$$

The parameter $a_m = -(\delta_1 + \cdots + \delta_m)$ can be both positive and negative.

Introduce the following cones in $\mathbb{R}^m$:

$$K_\pi = \{\delta = (\delta_1, \ldots, \delta_m) : \delta_1 + \cdots + \delta_i \le 0, \delta_{\pi^{-1}1} + \cdots + \delta_{\pi^{-1}i} \ge 0, \\ i = 1, \ldots, m-1\},$$

$$K_\pi^+ = K_\pi \cap \{\delta : \sum_{i=1}^m \delta_i \le 0\}, \quad K_\pi^- = K_\pi \cap \{\delta : \sum_{i=1}^m \delta_i \ge 0\}.$$

We have established the following.

Proposition 8. *For $(\lambda, \pi) \in \Delta(\mathcal{R})$ and an arbitrary $\delta \in K_\pi$ there exists a unique zippered rectangle (λ, h, a, π) corresponding to the parameters (λ, π, δ).*

In what follows, we shall simply refer to the zippered rectangle (λ, π, δ).

Remark 7. It would be interesting to write down explicitly the generating vectors for the cones K_π, K_π^+, K_π^-; in particular, that would allow us to give an explicit expression for the invariant densities of Veech [1] and Zorich [4].

Denote by $Area(\lambda, \pi, \delta)$ the area of the zippered rectangle (λ, π, δ). We have:

$$\begin{aligned} Area(\lambda, \pi, \delta) &= \sum_{r=1}^m \lambda_r h_r = \sum_{r=1}^m \lambda_r\Big(-\sum_{i=1}^{r-1} \delta_i + \sum_{l=1}^{\pi(r)-1} \delta_{\pi^{-1}l}\Big) \\ &= \sum_{i=1}^m \delta_i\Big(-\sum_{r=i+1}^m \lambda_r + \sum_{r=\pi(i)+1}^m \lambda_{\pi^{-1}r}\Big) \end{aligned} \tag{31}$$

(again, our convention is that $\sum_{i=m+1}^m(...) = 0$ and $\sum_{i=1}^0(...) = 0$). A straightforward calculatation shows that in the coordinates (λ, π, δ) the Rauzy induction map is written as follows:

$$\mathcal{T}(\lambda, \pi, \delta) = \begin{cases} \Big(\frac{A(\pi,b)^{-1}\lambda}{|A(\pi,b)^{-1}\lambda|}, b\pi, A(\pi, b)^{-1}\delta \cdot |A(\pi, b)^{-1}\lambda|\Big), & \text{if } \lambda \in \Delta_\pi^+; \\ \Big(\frac{A(\pi,a)^{-1}\lambda}{|A(\pi,a)^{-1}\lambda|}, a\pi, A(\pi, a)^{-1}\delta \cdot |A(\pi, a)^{-1}\lambda|\Big), & \text{if } \lambda \in \Delta_\pi^-. \end{cases}$$

For $\lambda \in \mathbb{R}_+^m$, denote

$$K(\lambda,\pi) = K_\pi \cap \{\delta : Area(\lambda,\pi,\delta) \le 1\},$$
$$K^+(\lambda,\pi) = K_\pi^+ \cap \{\delta : Area(\lambda,\pi,\delta) \le 1\},$$
$$K^-(\lambda,\pi) = K_\pi^- \cap \{\delta : Area(\lambda,\pi,\delta) \le 1\}.$$

Denote by vol_m the Lebesgue measure in $\mathbb{R}^m$. Set

$$\mathbf{r}(\lambda,\pi) = vol_m(K(\lambda,\pi)), \mathbf{r}^+(\lambda,\pi) = vol_m(K^+(\lambda,\pi)), \mathbf{r}^-(\lambda,\pi) = vol_m(K^-(\lambda,\pi)).$$

By definition, the functions $\mathbf{r}, \mathbf{r}^+, \mathbf{r}^-$ are positive rational functions, homogeneous of degree $-m$.

Lemma 11. *For any (λ,π) the following relations are valid.*

(1) $\mathbf{r}^-(\lambda,\pi) = \mathbf{r}(T_{b^{-1}}(\lambda,\pi))$.
(2) $\mathbf{r}^+(\lambda,\pi) = \mathbf{r}(T_{a^{-1}}(\lambda,\pi))$.
(3) $\mathbf{r}(\lambda,\pi) = \mathbf{r}(T_{a^{-1}}(\lambda,\pi)) + \mathbf{r}(T_{b^{-1}}(\lambda,\pi))$.

Proof. If

$$\delta = (\delta_1,\dots,\delta_m) \in K^-(\lambda,\pi),$$

then

$$\tilde{\delta} = (\delta_1,\dots,\delta_{m-1},\delta_m + \delta_{\pi^{-1}m}) \in K(T_{b^{-1}}(\lambda,\pi)),$$

and vice versa. This gives a volume-preserving bijection between $K^-(\lambda,\pi)$ and $K(T_{b^{-1}}(\lambda,\pi))$, whence $\mathbf{r}^-(\lambda,\pi) = \mathbf{r}(T_{b^{-1}}(\lambda,\pi))$. The second assertion is proved in the same way, and the third follows from the first two.

Corollary 7.

$$\mathbf{r}^+(\lambda,\pi) = \sum_{n=1}^{\infty} \mathbf{r}^-(T_{a^{-n}}(\lambda,\pi)).$$
$$\mathbf{r}^-(\lambda,\pi) = \sum_{n=1}^{\infty} \mathbf{r}^+(T_{b^{-n}}(\lambda,\pi)).$$

We only prove the first assertion. We have

$$\begin{aligned}\mathbf{r}^+(\lambda,\pi) &= \mathbf{r}(T_{a^{-1}}(\lambda,\pi)) = \mathbf{r}^+(T_{a^{-1}}(\lambda,\pi)) + \mathbf{r}^-(T_{a^{-1}}(\lambda,\pi))\\ &= \mathbf{r}(T_{a^{-2}}(\lambda,\pi)) + \mathbf{r}^-(T_{a^{-1}}(\lambda,\pi)).\end{aligned}$$

Proceeding by induction,

$$\mathbf{r}^+(\lambda,\pi) = \sum_{n=1}^{N} \mathbf{r}^-(T_{a^{-n}}(\lambda,\pi)) + \mathbf{r}(T_{a^{-N-1}}(\lambda,\pi)).$$

Since

$$T_{a^{-N-1}}(\lambda,\pi) = (T^{(\pi)}_{a^{-N-1}}(\lambda), a^{-N-1}\pi),$$

and $|T^{(\pi)}_{a^{-N-1}}(\lambda)| \to \infty$ as $N \to \infty$, we obtain $\mathbf{r}(T_{a^{-N-1}}(\lambda,\pi)) \to 0$ as $N \to \infty$, and the corollary is proved.

Since the functions $\mathbf{r}, \mathbf{r}^+, \mathbf{r}^-$ are positive, rational and homogeneous of degree $-m$, Corollary 7 implies that, for some positive constant $C(\mathcal{R})$, depending only on the Rauzy class $\mathcal{R}$, we have

$$\rho^+(\lambda,\pi) = C(\mathcal{R})\mathbf{r}^+(\lambda,\pi), \quad \rho^-(\lambda,\pi) = C(\mathcal{R})\mathbf{r}^-(\lambda,\pi).$$

By construction, for any $\lambda \in \mathbb{R}_+^m$ we have

$$\mathbf{r}_\pi^+(\lambda_1, \ldots, \lambda_m) = \mathbf{r}_{\pi^{-1}}^-(\lambda_{\pi(1)}, \ldots, \lambda_{\pi(m)}).$$

In view of this observation, it suffices to prove only the first assertion of Lemma 10, as the second one follows automatically.

Take $\delta = (\delta_1, \ldots, \delta_m) \in \mathbb{R}^m$, and, for $\theta > 0$, define

$$J_\theta^{(m)}\delta = (\delta_1, \ldots, \delta_m + \theta),\ J_{-\theta}^{(\pi^{-1}m)}\delta = (\delta_1, \ldots, \delta_{\pi^{-1}m} - \theta, \ldots, \delta_m).$$

Proposition 9. *Let $\theta > 0$. If $\delta \in K_\pi$, then $J_\theta^{(m)}\delta \in K_\pi, J_{-\theta}^{(\pi^{-1}m)}\delta \in K_\pi$. If $\delta \in K_\pi^-$, then $J_\theta^{(m)}\delta \in K_\pi^-$. If $\delta \in K_\pi^+$, then $J_{-\theta}^{(\pi^{-1}m)}\delta \in K_\pi^+$.*

This follows directly from the definition of the cones K_π, K_π^-, K_π^+. From (31) we obtain

$$Area(\lambda, \pi, J_\theta^{(m)}\delta) = Area(\lambda, \pi, \delta) + \theta(\sum_{r=\pi(m)+1}^{m} \lambda_{\pi^{-1}r}),$$

$$Area(\lambda, \pi, J_{-\theta}^{(\pi^{-1}m)}\delta) = Area(\lambda, \pi, \delta) + \theta(\sum_{r=\pi^{-1}(m)+1}^{m} \lambda_r),$$

which implies

Proposition 10.

$$Area(\lambda, \pi, \delta) \le Area(\lambda, \pi, J_\theta^{(m)}\delta) \le Area(\lambda, \pi, \delta) + \theta|\lambda|,$$

$$Area(\lambda, \pi, \delta) \le Area(\lambda, \pi, J_{-\theta}^{(\pi^{-1}m)}\delta) \le Area(\lambda, \pi, \delta) + \theta|\lambda|.$$

For $s \in \mathbb{R}$ and a hyperplane of the form $\delta_1 + \cdots + \delta_m = s$, let vol_{m-1} stand for the induced $(m-1)$-dimensional volume form on the hyperplane.

Denote

$$K_{s,\pi} = K_\pi \cap \{\delta : \sum_{i=1}^{m} \delta_i = s\},$$

$$K_s(\lambda, \pi) = K(\lambda, \pi) \cap K_{s,\pi},$$

$$V_s(\lambda, \pi) = vol_{m-1}(K_s(\lambda, \pi)).$$

Denote by $\mathbf{a}_{\mathbf{max}}^-$ the maximal possible value of $\delta_1 + \cdots + \delta_m = -a_m$ in $K(\lambda, \pi)$, by $\mathbf{a}_{\mathbf{max}}^+$ the maximal possible value of $-(\delta_1 + \cdots + \delta_m) = a_m$ in $K(\lambda, \pi)$.

Proposition 11. *Assume $0 \le s \le \mathbf{a}_{\mathbf{max}}^-$. Then*

$$V_s(\lambda, \pi) \le (1+s)^{m-1} V_0(\lambda, \pi).$$

Proof. Indeed, if $(\lambda, \pi, \delta) \in K_s(\lambda, \pi)$, then Proposition 10 implies

$$(\lambda, \pi, \frac{J_{-s}^{(\pi^{-1}m)}\delta}{1+s}) \in K_0(\lambda, \pi),$$

and the assertion follows. □

Proposition 12. *Assume s, $0 \le s \le 1$ is such that $\frac{s}{1-s} \le \mathbf{a}_{\mathbf{max}}^-$. Then*

$$(\frac{1}{1-s})^{m-1} V_s(\lambda, \pi) \ge V_0(\lambda, \pi).$$

Denote $\theta = \frac{s}{1-s}$; then $s = \frac{\theta}{1+\theta}$. If $(\lambda, \pi, \delta) \in K_0(\lambda, \pi)$, then

$$(\lambda, \pi, \frac{J_\theta^{(m)}\delta}{1+\theta}) \in K_s(\lambda, \pi),$$

and, again, the assertion follows. Propositions 11, 12 imply

Lemma 12. *For any $C_1 > 0$ there exists $C_2 > 0$ such that the following is true. If $\mathbf{a}^-_{\mathbf{max}}(\lambda, \pi) < C_1$, then*

$$\mathbf{r}^-(\lambda, \pi) < C_2 V_0^{m-1}(\lambda, \pi).$$

If $\mathbf{a}^+_{\mathbf{max}}(\lambda, \pi) < C_1$, then

$$\mathbf{r}^+(\lambda, \pi) < C_2 V_0^{m-1}(\lambda, \pi).$$

Note that there exists $\epsilon > 0$, depending only on $\mathcal{R}$ and such that for any $(\lambda, \pi) \in \Delta(\mathcal{R})$, we have $\mathbf{a}^-_{\mathbf{max}} > \epsilon$, $\mathbf{a}^+_{\mathbf{max}} > \epsilon$. In conjunction with Propositions 11, 12, this observation implies

Lemma 13. *There exists a constant C_3 such that for any $(\lambda, \pi) \in \Delta(\mathcal{R})$, we have*

$$\mathbf{r}^-(\lambda, \pi) \geq C_3 V_0(\lambda, \pi), \;\; \mathbf{r}^+(\lambda, \pi) \geq C_3 V_0(\lambda, \pi).$$

Veech's inequalities for a zippered rectangle imply the inequality

$$-h_{\pi^{-1}m} \leq a_m \leq h_{\pi^{-1}(m)+1}$$

(one of these is part of the inequalities; the other easily follows from them and can also be seen from geometric considerations: the "zipper" a_m cannot go above the height of the $(\pi^{-1}(m)+1)$-th rectangle). We thus have

$$\mathbf{a}^+_{\mathbf{max}}(\lambda, \pi) \leq \frac{1}{\lambda_{\pi^{-1}(m)+1}}.$$

Estimating $\mathbf{r}^+$ from above by Lemma 12 and $\mathbf{r}^-$ from below by Lemma 13 now yields the following.

Corollary 8. *For any $C_4 > 0$ there exists $C_5 > 0$ such that if $\lambda_{\pi^{-1}(m)+1} > C_4$, then*

$$\frac{\mathbf{r}^+(\lambda, \pi)}{\mathbf{r}^-(\lambda, \pi)} < C_5,$$

Recalling that $\mathbf{r}^-(\lambda, \pi)$ is, up to a constant, equal to $\rho^-(\lambda, \pi)$ and that the invariant densities are bounded away from zero, we see that Lemma 10 is now proven completely.

8. Kerckhoff names

In the following two sections, we shall use Kerckhoff's convention of numbering the subintervals of an interval exchange [20]; to avoid confusion, we shall speak of *Kerckhoff names* of subintervals.

Take an interval exchange (λ, π). A *Kerckhoff naming* on the subintervals of (λ, π) is defined by an arbitrary permutation $i_1, \ldots, i_m$ of the symbols $\{1, \ldots, m\}$. Once such a permutation is given, we assign names $I_{i_1}, \ldots, I_{i_m}$ to the subintervals of (λ, π), from the left to the right (i.e., the subinterval $[0, \lambda_1)$ is named I_{i_1}, the subinterval $[\lambda_1, \lambda_1 + \lambda_2)$ is named I_{i_2} and so forth).

A Kerckhoff naming of the subintervals of (λ, π) induces a naming on the subintervals of $\mathcal{T}(\lambda, \pi)$ in the following way. Assume $\lambda_m < \lambda_{\pi^{-1}m}$ and that the Rauzy operation a was applied to (λ, π) in order to obtain $\mathcal{T}(\lambda, \pi)$. Then the subintervals of $\mathcal{T}(\lambda, \pi)$ are named, from the left to the right, by $I_{i_1}, \ldots, I_{i_{\pi^{-1}m}}, I_{i_m}, I_{i_{\pi^{-1}m+1}}, \ldots, I_{i_{m-1}}$. If $\lambda_m > \lambda_{\pi^{-1}m}$ and the Rauzy operation b was applied, then the subintervals of $\mathcal{T}(\lambda, \pi)$ are just named, as before, by $I_{i_1}, \ldots, I_{i_m}$, from the left to the right. Proceeding inductively, we obtain a naming for any $\mathcal{G}^n(\lambda, \pi)$. Conversely, if we have a Kerckhoff naming of subintervals of (λ, π), then, for any word $w \in W_{\mathcal{A},B}$ compatible with (λ, π), we automatically obtain a Kerckhoff naming on the subintervals of $t_w(\lambda, \pi)$ and $T_w(\lambda, \pi)$.

Let (λ, π) be an interval exchange with a Kerckhoff naming $I_{i_1}, \ldots, I_{i_m}$. If $(\lambda, \pi) \in \Delta^+$ (that is, $\lambda_m > \lambda_{\pi^{-1}m}$), then we say that $I_{i_{\pi^{-1}m}}$ is the subinterval *in the critical position* (we shall also sometimes say "in the a-critical position"). If $(\lambda, \pi) \in \Delta^-$ (that is, $\lambda_m < \lambda_{\pi^{-1}m}$), then we say that I_{i_m} is the subinterval *in the critical position* (we shall also sometimes say "in the b-critical position").

9. Exponential growth

Let $\mathbf{x} \in \overline{\Delta}$, that is, $\mathbf{x} = (\ldots, (\lambda(-n), \pi(-n)), \ldots, (\lambda, \pi))$, where, as usual, $\mathcal{G}(\lambda(-n), \pi(-n)) = (\lambda(1-n), \pi(1-n))$. Define the words $w(n)$ by the relation $(\lambda(-n), \pi(-n)) = t_{w(n)}(\lambda, \pi)$. Set $(\Lambda(-n), \pi(-n)) = T_{w(n)}(\lambda, \pi)$.

Lemma 14. *There exists N such that the following is true. For any $\mathbf{x} \in \overline{\Delta}(\mathcal{R})$, there exist $i_1, i_2 \in \{1, \ldots, m\}$, $i_1 \neq i_2$, such that*

$$\Lambda(-N)_{i_1} \geq 2(\lambda(0)_{i_1}),\ \Lambda(-N)_{i_2} \geq 2(\lambda(0)_{i_2}).$$

Remark 8. In what follows, we only use the existence of one subinterval whose length doubles; note, however, that the proof given below yields two such subintervals.

Lemma 14 immediately implies the following.

Corollary 9. *There exist constants $C > 0, \alpha > 0$, depending on the Rauzy class only, such that for any $w \in W_{\mathcal{A},B}$ we have*

$$||A(w)|| \geq C \exp(\alpha|w|).$$

Proof of Lemma 14. Take a point $\mathbf{x} \in \overline{\Delta}$,

$$\mathbf{x} = (\ldots, (\lambda(-n), \pi(-n)), \ldots, (\lambda, \pi)).$$

Give Kerckhoff names $I_1, \ldots, I_m$ to the subintervals of the exchange (λ, π) from the left to the right, so that the length of I_i is λ_i. We thus automatically obtain a Kerckhoff naming for the subintervals of $((\lambda(-n), \pi(-n))$ for any n. Let I_{j_n} be the critical subinterval for $(\lambda(-n), \pi(-n))$. Consider the infinite sequence

$$I_{j_1} \ldots I_{j_n} \ldots. \tag{32}$$

Note that $j_n \neq j_{n+1}$. A subword $I_{j_k} \ldots I_{j_{k+l}}$ will be called a *simple cycle* if $I_{j_k} = I_{j_{k+l}}$ whereas $I_{j_k}, \ldots, I_{j_{k+l-1}}$ are all distinct. Naturally, $l \leq m$. There are finitely many possible simple cycles. Therefore there exists N, depending only on m, such that for any word of length N in the alphabet $\{I_1, \ldots, I_m\}$, some simple cycle occurs at least m times. Now take the word

$$I_{j_1} \ldots I_{j_N} \tag{33}$$

at the beginning of the sequence (32), and take a simple cycle which occurs m times, say

$$I_{l_1} \dots I_{l_r}. \tag{34}$$

Here, of course, $r \leq m$. Now estimate the nonrenormalized length of the subintervals with names $I_{l_1}, \dots, I_{l_r}$ $(r \leq m)$ at the times $n = 0, -1, \dots, -N$. In the beginning, these are $\lambda_{l_1}, \dots, \lambda_{l_r}$. The key observation is, as usual, that the interval in critical position at a given inverse Zorich step was, at the previous step, added to the previous critical interval. After the first occurrence of the cycle (34), therefore, the (nonnormalized) length of I_{l_1} is at least $\lambda_{l_1} + \lambda_{l_2}$, that of I_{l_2} is at least $\lambda_{l_2} + \lambda_{l_3}$ and so forth. After the second occurrence of (34), the length of I_{l_1} is at least $\lambda_{l_1} + \lambda_{l_2} + \lambda_{l_3}$, that of I_{l_2} is at least $\lambda_{l_2} + \lambda_{l_3} + \lambda_{l_4}$, and so forth. Finally, after the r-th occurrence of (34), the length of I_{l_1} is not less than $\lambda_{l_1} + \lambda_{l_2} + \cdots + \lambda_{l_r}$, that is, not less than $2\lambda_{l_1}$, since $\lambda_{l_1} = \lambda_{l_r}$. The lemma is proven. □

10. Proof of Lemma 5

An informal sketch of the proof of Lemma 5 is presented first. One divides the subintervals into "big" ones and "small" ones: the aim is to obtain one more "big" interval. For this, one must first put a small subinterval into critical position. This is achieved by Lemma 15. In the previous section, we have seen that the total length of the (nonrenormalized) interval grows exponentially with the number of Zorich steps (with an exponent depending on ϵ). When the total length of the interval doubles, we obtain a new "big" subinterval.

10.1. **Putting a small interval into critical position.** Take an interval exchange (λ, π) and name the subintervals $I_1, \dots, I_m$, from the left to the right.

Proposition 13. *Any interval can be put either in the a-critical or in the b-critical position.*

Proof. Assume I_i can be put into the a-critical position. Then all I_j for $j > i$ can be put into the b-critical position. To prove this, take the shortest word w that puts I_i into the a-critical position. Then, in the preimage, all I_j, $j > i$, still stand to the right of I_i, though perhaps in a different order (because an inversion of order between I_i and I_j can only happen once I_i reaches the critical position). Therefore, we can immediately place any of the I_j, $j > i$, into the b-critical position. Similarly, if I_i can be put into the b-critical position, then all $I_{\pi(j)+1}, \dots, I_{\pi m}$ can be put into the a-critical position. Since the permutation π is irreducible, the proposition follows. □

Remark 9. Proposition 13 can also be proved by noting that any Rauzy class has a permutation π such that $\pi(1) = m$, $\pi(m) = 1$. This observation shows even slightly more, namely, that every interval, except I_1 and $I_{\pi^{-1}1}$, can be put *both* into the a-critical and the b-critical positions.

Now pick a positive integer $k \leq m$ and a real $\gamma > 0$. We say that we have a (k, γ)-*big-small decomposition* if the intervals of the exchange are divided into two groups: $I_{i_1}, \dots, I_{i_k}$, each of length at least γ, and the remaining ones (nothing is said about the length of the remaining ones).

Under the Kerckhoff convention, a big-small decomposition of (λ, π) is inherited by all $T_w(\lambda, \pi)$ (one just takes the intervals with the same names).

Lemma 15. *There exists a constant $L(\mathcal{R})$, depending on the Rauzy class only, and, for any $\gamma > 0$, there exists a constant $p(\gamma)$ such that the following is true. Let $(\lambda, \pi) \in \Delta_{k,\gamma}$ with a fixed big-small decomposition. Then there exists $w \in \mathcal{W}_{\mathcal{A},B}$ such that*

(1) $\mathbb{P}(w|(\lambda, \pi)) \geq p(\gamma)$,
(2) $|T_w(\lambda, \pi)| < L(\mathcal{R})$,
(3) *the exchange $T_w(\lambda, \pi)$ has a small interval in critical position.*

Proof. Take the shortest word (in terms of the number of Zorich operations) that puts a small interval into critical position. Among all such words, pick the one that involves the smallest number of Rauzy operations. The length of this word, as well as the number of Rauzy operations involved, only depends on the Rauzy class. At each intermediate Rauzy step, all subintervals following the critical one either in the preimage or in the image must be big; otherwise there would exist a shorter word placing a small interval into critical position. Therefore, by Lemma 9 and Corollary 6, the probability of each Zorich operation involved is bounded from below by a constant that only depends on γ. The lemma is proved. □

10.2. **Completion of the proof of Lemma 5.**

Proof. Take any $x \in \overline{\Delta}$ such that $(\lambda(0), \pi(0)) \in \Delta_{\gamma,k,\epsilon}$. By Lemma 15, place a small interval into the critical position. Now take the first n such that $|\Lambda(-n)| > 2$. From Lemma 14 it follows that $n < K|\log \epsilon|$. Applying Lemma 7 repeatedly and using stationarity, we obtain that the relation $|\Lambda(-n)| < 2M$ holds with positive probability depending only on M.

Consider two cases:

(1) at all steps from 1 to n, only small intervals were added between themselves;
(2) at some step a large interval was added to a small one.

(Note, that since we start with a small interval in critical position, either one or the other case holds (for, in order that a small interval be added to a big interval, a big interval must first be placed into critical position, and for that it must first be added to a small one).)

In the first case, the lengths of all large intervals remain the same, and after renormalization at step n, each large interval has length at least $\gamma/(2M)$. Since $|\Lambda(-n)| > 2$, and we have only m subintervals, there must be another interval of length at least $1/(2mM)$, and Lemma 5 is proved in this case.

In the second case, let n_1 be the first moment at which a big interval is added to a small one. Then $|\Lambda(-n_1)| < 2$, and, since at previous moments only small intervals were added between themselves, we have $k + 1$ intervals of length at least $\gamma/2$, and Lemma 5 is proved completely. □

11. Return times for the Teichmüller flow

We have in fact proven a stronger statement, namely, the following lemma.

Lemma 16. *For any word $\mathbf{q} \in W_{\mathcal{A},B}$ such that all entries of the matrix $A(\mathbf{q})$ are positive, there exist constants $K_0(\mathbf{q}), p(\mathbf{q})$, depending only on $\mathbf{q}$ and such that the following is true. For any $K \geq K_0$ and any $(\lambda, \pi) \in \Delta(\mathcal{R})$,*

$$\mathbb{P}(\exists n : (\lambda(-n), \pi(-n)) \in \Delta(\mathbf{q}), |\Lambda(-n)| < K|(\lambda, \pi)) \geq p(\mathbf{q}).$$

To prove this lemma, it suffices to notice that in the proof of Lemma 5 given above, the addition of a new "big" subinterval involves only a bounded increase in the total length of the interval.

Lemma 16 has the following corollary for the Teichmüller flow on the space of zippered rectangles.

Take an arbitrary word $\mathbf{q} = q_1 \dots q_{2l+1} \in W_{\mathcal{A},B}$ such that all entries of the matrix $A(q_1 \dots q_l)$ are positive and all entries of the matrix $A(\mathbf{q})$ are positive. Consider the cylinder set

$$\overline{\Delta}_{\mathbf{q}} = \{\omega \in \Omega^{\mathbb{Z}}_{\mathcal{A},B}, \omega_{-l} = q_1, \dots, \omega_l = q_{2l+1}\}.$$

Consider the flow P^t as a special flow over $\overline{\Delta}_{\mathbf{q}}$. Denote the roof function of the flow by $\tau_{\mathbf{q}}$. Lemma 16 then implies

Corollary 10. *There exists $\epsilon > 0$ such that*

$$\int_{\overline{\Delta}_{\mathbf{q}}} \exp(\epsilon \tau_{\mathbf{q}}(\omega)) d\mathbb{P}(\omega) < +\infty.$$

Proof of the Corollary. Take $\omega \in \Omega^{\mathbb{Z}}_{\mathcal{A},B}$. As usual, set

$$(\lambda(-n), \pi(-n)) = \Phi^{-1}(\omega_{-n} \dots \omega_0 \omega_1 \dots),$$
$$(\Lambda(-n), \pi(-n)) = T_{\omega_{-n} \dots \omega_{-1} \omega_0}(\lambda(0), \pi(0)).$$

Set $n_{\mathbf{q}}(\omega)$ to be the smallest n such that

$$\omega_{-n} = q_1, \dots, \quad \omega_{-n+2l} = q_{2l+1}.$$

Finally, set $L_{\mathbf{q}}(\omega) = \log |\Lambda(-n_{\mathbf{q}}(\omega))|$.

Remark 10. Informally, $L_{\mathbf{q}}(\omega)$ is the "Teichmüller flow time" (in the negative direction) it takes ω to reach $\overline{\Delta}_{\mathbf{q}}$.

To establish Corollary 10, it suffices to prove

Proposition 14. *There exists $\epsilon > 0$ such that*

$$\int_{\Omega^{\mathbb{Z}}_{\mathcal{A},B}} \exp(\epsilon L_{\mathbf{q}}(\omega)) d\mathbb{P}(\omega) < +\infty.$$

Proof of Proposition 14. Our main tool will be Lemma 16. Take a constant $K > K_0$ such that $1 - p(\mathbf{q}) + \frac{2}{K} < 1$. Define a random time $k_1(\omega)$ to be the first moment n such that $|\Lambda(-n)(\omega)| > K$. Note that the map

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

is invertible (here, as always, σ is the right shift on $\Omega_{\mathcal{A},B}$).

Introduce a function $\eta : \Omega^{\mathbb{Z}}_{\mathcal{A},B} \to \mathbb{N}$ by the formula

$$\eta(\omega) = [\frac{\log |\Lambda(-k_1(\omega))|}{\log K}].$$

In other words, $\eta(\omega) = n$ if

$$K^n \le |\Lambda(-k_1(\omega))| < K^{n+1}.$$

Proposition 15. *There exists a constant C such that the following is true for any $K > K_0$. For any $c_1 \dots c_n \dots \in \Omega^+_{\mathcal{A},B}$, we have*

$$\mathbb{P}(\{\omega : \eta(\omega) = r | \omega_1 = c_1, \dots \omega_n = c_n \dots\}) \le \frac{C}{K^{r-1}}.$$

This immediately follows from Lemma 7.

Proposition 16.

$$\mathbb{P}(\{\omega : \eta(\omega) = 1, \omega_{-k_1(\omega)} \ldots \omega_0 \text{ does not contain the word } \mathbf{q}\}) \le 1 - p(\mathbf{q}).$$

This is a direct corollary of Lemma 16: indeed, the probability of hitting $\overline{\Delta}_{\mathbf{q}}$ in Teichmüller time not exceeding K is at least $p(\mathbf{q})$.

Finally, take a large N and let

$$n_N(\omega) = \min\{n : \eta(\omega) + \cdots + \eta(\tilde{\sigma}^n(\omega)) \le N\}.$$

Note that, by definition,

$$K^N \le |\Lambda(-n_N)(\omega)| \le K^{2N}.$$

Now consider the set

$$\tilde{\Omega}(N) = \{\omega : \omega_{-n_N(\omega)} \ldots \omega_0 \text{ does not contain the word } \mathbf{q}\}.$$

Note that

$$\{\omega : L_{\mathbf{q}}(\omega) > 2N\} \subset \tilde{\Omega}(N).$$

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

$$\mathbb{P}(\tilde{\Omega}(N)) \le r^N.$$

But by the previous two propositions, using the binomial formula, we immediately obtain

$$\mathbb{P}(\tilde{\Omega}) \le C(1 - p(\mathbf{q}) + \frac{2}{K})^N,$$

and, since $1 - p(\mathbf{q}) + \frac{2}{K} < 1$, the proposition follows. □

This proposition admits an equivalent formulation in terms of the norms of renormalization matrices on the space of interval exchange transformations.

More precisely, for $(\lambda, \pi) \in \Delta_{\mathbf{q}}$, $\Phi(\lambda, \pi) = \omega_1 \ldots \omega_n \ldots$, we let $n^{\mathbf{q}}(\lambda, \pi)$ be the smallest $n > 0$ such that $\mathcal{G}^n(\lambda, \pi) \in \Delta_{\mathbf{q}}$, and we set

$$\mathcal{N}(\lambda, \pi) = ||A(\omega_1 \ldots \omega_{n^{\mathbf{q}}(\omega)})||.$$

Since the ratio of the norm of the matrix $A(\omega_1 \ldots \omega_{n^{\mathbf{q}}(\omega)})$ and the exponential of the Teichmüller time elapsing between two successive returns into $\overline{\Delta}_{\mathbf{q}}$ is bounded from above and from below by two constants depending only on $\mathbf{q}$, we obtain

Corollary 11. *There exists $\epsilon > 0$ such that*

$$\int_{\Delta_{\mathbf{q}}} \mathcal{N}(\lambda, \pi)^{\epsilon} d\mathbb{P} < +\infty.$$

Remark 11. The first results on exponential decay for the probabilities of return times were obtained by Jayadev Athreya [30]. In his approach, Athreya used the dynamics of $SL(2, \mathbb{R})$-action, which allowed him to obtain optimal exponents. The argument above is an attempt to recover some of Athreya's theorems using the language of interval exchange transformations; the argument above does not, however, give an optimal exponent.

Avila, Gouëzel, and Yoccoz have recently announced exponential decay of correlations for the Teichmüller flow. One of the steps in their proof is, again, an exponential estimate for return times, which they have obtained independently (Avila [oral communication]). Their exponent is optimal.

12. Estimate of the measure

Lemma 17. *There exists a constant $C(\mathcal{R})$ depending only on the Rauzy class $\mathcal{R}$ such that*

$$\nu(\Delta(\mathcal{R}) \setminus \Delta_\epsilon(\mathcal{R})) < C\epsilon.$$

The proof repeats that of Proposition 13.2 in Veech [1].

Lemma 4 and Corollary 5 therefore imply the following.

Corollary 12. *Let $\mathbf{q} \in W_{\mathcal{A},B}$, $\mathbf{q} = q_1 \dots q_l$ be such that all entries of the matrix $A(\mathbf{q})$ are positive and $\Delta(\mathbf{q}) \subset \Delta^+$. Then there exist constants $C > 0, \alpha > 0$, depending on $\mathbf{q}$ only, such that for any n we have*

$$\mathbb{P}((\lambda,\pi) \in \Delta^+ |\ \mathcal{G}^{2k}(\lambda,\pi) \notin \Delta(\mathbf{q}) \text{ for all } \ k, 1 \le k \le n) \le C\exp(-\alpha\sqrt{n}).$$

Proof. Let $n = r^2$ and denote

$$X(n,\mathbf{q}) = \{(\lambda,\pi) : \mathcal{G}^{2k}(\lambda,\pi) \notin \Delta(\mathbf{q}) \text{ for all } \ k, 1 \le k \le n\}.$$

Take

$$B(n) = \{(\lambda,\pi) : \mathcal{G}^{2k}(\lambda,\pi) \notin \Delta_{\exp(-r)} \text{ for some } \ k, 1 \le k \le n\}.$$

Then, by the previous lemma, $\nu(B(n)) \le Cr^2\exp(-r)$, whereas, by Corollary 5, using stationarity of the measure ν we obtain

$$\nu(X(n,\mathbf{q}) \setminus B(n)) \le (1 - p(\mathbf{q}))^r,$$

and Corollary 12 is proven. □

Remark 12. This result allows us to use the tower method of L.-S. Young [11] and to obtain the decay rate $\exp(-\alpha\sqrt{n})$ for correlations of bounded Hölder functions. For bounded Lipschitz functions, one can also use the method of V. Maume-Deschamps [12] and obtain the uniform rate of decay at the rate $\exp(-\alpha n^{1/2-\epsilon})$. It is not clear to me, however, how to use either of these methods in the invertible case.

13. Inequalities

Let

$$W^+_{\mathcal{A},B} = \{w \in W_{\mathcal{A},B} :\ |w| \text{ is even },\ \Delta(w) \subset \Delta^+\}.$$

Lemma 18. *For any $C_1, C_2 > 0$ there exists $C_3 > 0$ such that the following is true. Suppose $row(A) < C_1$ and $\lambda \in \Delta_{C_2}$. Then*

$$\frac{1}{C_3} \le \frac{|A\lambda|^m}{\Pi_{j=1}^m \sum_{i=1}^m A_{ij}} \le C_3.$$

Proof. Denote $A_j = \sum_{i=1}^m A_{ij}$, so that $|A| = \sum_{j=1}^m A_j$. Then

$$\frac{A_j}{A_k} \le row(A),$$

whence

$$\frac{A_j}{|A|} \ge \frac{1}{m\ row(A)}.$$

Finally, if $\lambda \in \Delta_{C_2}$, then

$$|A\lambda| \ge C_2|A|,$$

which completes the proof. □

Corollary 13. *For any $C_4 > 0$, $C_5 > 0$ there exists $C_6 > 0$ such that the following is true. Suppose $(\lambda, \pi) \in \Delta_{C_4}$. Suppose $w \in \mathcal{W}_{\mathcal{A},B}$ is compatible with (λ, π) and is such that $row(A(w)) < C_5$. Then*

$$\frac{1}{C_6} \leq \frac{\mathbf{m}(C(w))}{\mathbb{P}(w|(\lambda, \pi))} \leq C_6.$$

Corollary 14. *For any $C_7 > 0$, $C_8 > 0$, $C_9 > 0$, there exists $C_{10} > 0$ such that the following is true. Suppose $(\lambda, \pi) \in \Delta_{C_7}$. Suppose $w \in \mathcal{W}_{\mathcal{A},B}$ is compatible with (λ, π) and furthermore satisfies*

$$row(A(w)) < C_8, \ \Delta(w) \subset \Delta_{C_9}.$$

Then

$$\frac{1}{C_{10}} \leq \frac{\mathbb{P}(C(w))}{\mathbb{P}(w|(\lambda, \pi))} \leq C_{10}.$$

Corollary 15. *Let M be such that for any $n > M$ any two vertices in the Rauzy graph can be joined in n steps. Then for any $C_{17} > 0$, $C_{18} > 0$, $C_{19} > 0$, there exists $C_{20} > 0$ such that the following is true. Suppose $(\lambda, \pi) \in \Delta^+ \cap \Delta_{C_{17}}$. Suppose $w \in W^+_{\mathcal{A},B}$ satisfies*

$$row(A(w)) < C_{18}, \ \Delta(w) \subset \Delta^+ \cap \Delta_{C_{19}}.$$

Then for any $n \geq M$, we have

$$\frac{1}{C_{20}} \leq \frac{\mathbb{P}(C(w))}{\mathbb{P}^{(2n)}(w|(\lambda, \pi))} \leq C_{20}.$$

From the definition (8) of the Hilbert metric it easily follows that for any $\lambda, \lambda' \in \Delta_{m-1}$ we have

$$e^{-d(\lambda,\lambda')}\lambda'_i \leq \lambda_i \leq e^{d(\lambda,\lambda')}\lambda'_i. \tag{35}$$

Proposition 17. *Assume $\lambda, \lambda' \in \Delta^+_\pi$. Then*

$$\exp(-md(\lambda, \lambda')) \leq \frac{\rho(\lambda, \pi)}{\rho(\lambda', \pi)} \leq \exp(md(\lambda, \lambda')).$$

Proof. Indeed, there exist linear forms

$$l_i^{(j)}(\lambda) = \sum_{k=1}^{m} a_{ik}^{(j)} \lambda_k,$$

where $a_{ik}^{(j)}$ are nonnegative integers (in fact, either 0 or 1, but we do not need this here) such that

$$\rho(\lambda, \pi) = \sum_{j=1}^{s} \frac{1}{l_1^{(j)}(\lambda) l_2^{(j)}(\lambda) \dots l_m^{(j)}(\lambda)}.$$

Clearly, if for all $i = 1, \dots, m$ and some $\alpha > 0$, we have $\alpha^{-1}\lambda_i \leq \lambda'_i \leq \alpha\lambda_i$, then

$$\alpha^{-m} \leq \frac{\rho(\lambda, \pi)}{\rho(\lambda', \pi)} \leq \alpha^m,$$

and the proposition is proved. □

For similar reasons we have

Proposition 18. *Assume $\lambda, \lambda' \in \Delta_\pi^+$ and let A be an arbitrary matrix with non-negative integer entries. Then*

$$\exp(-md(\lambda, \lambda')) \leq \frac{\rho(A\lambda, \pi)}{\rho(A\lambda', \pi)} \leq \exp(md(\lambda, \lambda')).$$

From these propositions and formula (20) we obtain

Corollary 16. *Let $c \in \mathcal{A}$ be compatible with π. Then for any $\lambda, \lambda' \in \Delta_\pi^+$ we have*

$$\exp(-2md(\lambda, \lambda')) \leq \frac{\mathbb{P}(c|(\lambda, \pi))}{\mathbb{P}(c|(\lambda', \pi))} \leq \exp(2md(\lambda, \lambda')).$$

This corollary implies the following.

Lemma 19. *Let $w \in W_{\mathcal{A},B}^+$ be such that the cylinder $C(w)$ has finite Hilbert diameter. Then for any c compatible with w and any $(\lambda_0, \pi) \in C(w)$ we have*

$$\exp(-2m\ diam C(w)) \leq \frac{\mathbb{P}(c|(\lambda_0, \pi))}{\mathbb{P}(\omega_0 = c|\omega|_{[1,|w|]} = w)} \leq \exp(2m\ diam C(w)).$$

Proof. We have

$$\nu(C(cw)) = \int_{C(w)} \mathbb{P}(c|(\lambda, \pi)) d\nu(\lambda, \pi).$$

Let $d = diam C(w)$. For any $(\lambda, \pi), (\lambda', \pi) \in C(w)$, we have, by Corollary 16,

$$\exp(-2md) \leq \frac{\mathbb{P}(c|(\lambda, \pi))}{\mathbb{P}(c|(\lambda', \pi))} \leq \exp(2md).$$

Fix an arbitrary $(\lambda_0, \pi) \in \Delta_w$.

Then, from the above,

$$\nu(C(w)) P(c|(\lambda_0, \pi)) \exp(-2md) \leq \int_{C(w)} P(c|(\lambda, \pi)) d\nu(\lambda, \pi)$$
$$\leq \nu(C(w)) P(c|(\lambda_0, \pi)) \exp(2md),$$

and, since, by definition, we have

$$\mathbb{P}(\omega_0 = c|\omega|_{[1,|w|]} = w) = \frac{\mathbb{P}(cw)}{\mathbb{P}(w)},$$

the lemma is proved. □

For $N \in \mathbb{N}$ and $A \subset \Delta(\mathcal{R})$, we denote

$$\mathbb{P}^{(N)}(A|(\lambda, \pi)) = \mathbb{P}((\lambda(-N), \pi(-N)) \in A|(\lambda(0), \pi(0)) = (\lambda, \pi));$$

for $w \in W_{\mathcal{A},B}$, we write $\mathbb{P}^{(N)}(w|(\lambda, \pi)) = \mathbb{P}^{(N)}(\Delta(w)|(\lambda, \pi))$.

Lemma 20. *Let M be a number such that for any $N \geq M$ any two vertices of the Rauzy graph can be connected in N steps. For any $\gamma > 0$, $N \geq M$ there exists a constant C_0 depending only on γ and N such that for any word $w \in \mathcal{W}_{\mathcal{A},B}^+$ and any $(\lambda, \pi) \in \Delta_\gamma$,*

$$\mathbb{P}^{(2N)}(w|(\lambda, \pi)) \geq \frac{C_0}{|A(w)\lambda|^m}.$$

Proof. Let $w = w_1 \dots w_{2n}$, and let $w_{2n} = (a, m_1, \pi_1)$.

Let $\pi'_1\pi'_2 \dots \pi'_{2N}$ be a path of length $2N$ between π and π_1 (here $\pi'_1 = \pi$, $\pi'_{2n} = \pi_1$, $\pi_{2k+1} = a\pi_{2k}$, $\pi_{2k+2} = b\pi_{2k+1}$).

Denote $w_{n+2i+1} = (a, 1, \pi_{2i+1})$, $w_{n+2i} = (b, 1, \pi_{2i})$. In other words, the word $= w_{2n+1} \dots w_{2n+2N} \in \mathcal{W}_{\mathcal{A},B}$ is the word corresponding to the path $\pi'_1\pi'_2 \dots \pi'_{2N}$ in the Rauzy graph. Then $w' = w_1 \dots w_{2n+2N}$ is a word compatible with (λ, π). Besides,

$$|A(w_{2n+1}c_{n+2} \dots w_{2n+2N})| < (2N)^{(2N)}.$$

We have

$$P^{(2n)}(w|(\lambda,\pi)) \geq P(w'|(\lambda,\pi)) = \frac{\rho(T_{w'}(\lambda), w'\pi)}{|A(w')\lambda|^m \rho(\lambda,\pi)}.$$

There exists a universal constant C_1 such that $\rho(\lambda', \pi') > C_1$ for any $(\lambda', \pi') \in \Delta^+$ (the density of the invariant measure is bounded from below).

Then, $|A(w')\lambda|^m \leq |A(w')|^m \leq (2N)^{2mN}|A(w)|^m$.

Finally, there exists a C_2 depending on c only such that if $\lambda_i > c$ for all i, then $\rho(\lambda, \pi) > C_2$. □

Combining all of the above, we obtain the result of the lemma.

14. Markov approximation and the Doeblin condition

14.1. **Good cylinders.** Let $\mathbf{q} = q_1 \dots q_l$ be a word such that all entries of the matrix $A(\mathbf{q})$ are positive. Fix $\epsilon > 0$ and let k_0 be such that

$$\mathbb{P}(\Delta(\mathbf{q}) \cap \mathcal{G}^{-2n}\Delta(\mathbf{q})) \geq \epsilon \text{ for } n > k_0. \tag{36}$$

Note that, due to mixing, Corollary 5 implies the following.

Proposition 19. *Let $\mathbf{q} \in W_{\mathcal{A},B}$, $\mathbf{q} = q_1 \dots q_l$ be such that all entries of the matrix $A(\mathbf{q})$ are positive and that $\Delta(\mathbf{q}) \subset \Delta^+$. Then there exist positive constants $K(\mathbf{q}), p(\mathbf{q})$ such that the following is true for any $\epsilon > 0$. Suppose $(\lambda, \pi) \in \Delta_\epsilon \cap \Delta^+$ and set n to be the integer part of $K(\mathbf{q})|\log \epsilon|$. Then*

$$\mathbb{P}((\lambda(-2n), \pi(-2n)) \in \Delta(\mathbf{q}) | (\lambda(0), \pi(0)) = (\lambda, \pi)) \geq p(\mathbf{q}).$$

Take $k \geq k_0$. Let $r = 2(K+1)k + 2M$, where K is the constant from Lemma 4 and M is the connecting constant of the Rauzy graph from Lemma 20.

Let θ, $0 < \theta < 1$ be arbitrary. A word $w = w_1 \dots w_k$ is called *good* if

(1) $\Delta(w) \subset \Delta_{\exp(-k)}$;
(2) the word $\mathbf{q}$ appears at least $\frac{k^\theta}{l}$ times in w (we only count disjoint appearances).

A word $w_1 \dots w_r$ is called good if $w_1 \dots w_k$ is good, a word $w_1 \dots w_{Nr}$ is called good if all words $w_1 \dots w_r$, $w_{r+1} \dots w_{2r}$, $\dots w_{(N-1)r+1} \dots w_{Nr}$ are good, and a word $w_1 \dots w_{Nr+L}$, $L < r$, is good if $w_1 \dots w_{Nr}$ is good and either $L < k$ or $w_{Nr+1} \dots w_{Nr+k}$ is good.

We denote by $\mathbf{G}(N)$ the set of all good words of length N.

Let

$$\Delta(\mathbf{G}(N)) = \bigcup_{w \in \mathbf{G}(N)} \Delta(w)$$

and

$$\Delta(B(N)) = \Delta^+ \setminus \Delta(\mathbf{G}(N)).$$

By Corollary 12, there exist constants C_{31}, C_{32} such that for all r we have

$$\mathbb{P}(\Delta(B(N))) \le C_{31} N \exp(-C_{32} r^{(1-\theta)/2}), \tag{37}$$

and, for any $(\lambda, \pi) \in \Delta(\mathbf{q})$, also

$$\mathbb{P}((\lambda(-1), \pi(-1)) \in \Delta(B(N)) | (\lambda(0), \pi(0)) = (\lambda, \pi)) \le C_{31} N \exp(-C_{32} r^{(1-\theta)/2}). \tag{38}$$

14.2. **Preliminary estimates for the Doeblin condition.** From Corollary 16 we deduce that there exists a constant C_{33} such that for any $(\lambda, \pi), (\lambda', \pi) \in \Delta(\mathbf{q})$, and any word w compatible with $\mathbf{q}$, we have

$$\frac{1}{C_{33}} \le \frac{\mathbb{P}(w|(\lambda, \pi))}{\mathbb{P}(w|(\lambda', \pi))} \le C_{33}.$$

Finally, by Lemma 20, there exists a constant C_{34} such that for any $w \in \mathcal{W}_{\mathcal{A},B}$ and for any $N > M$ we have

$$\frac{1}{C_{34}} \le \frac{\mathbb{P}^{(2N)}(w|(\lambda, \pi))}{\mathbb{P}^{(2N)}(w|(\lambda', \pi))} \le C_{34}.$$

Take an arbitrary point $(\lambda, \pi) \in \Delta_{\mathbf{q}}$. Define a new measure φ on Δ^+. Namely, for a set $A \subset \Delta^+$ put

$$\varphi(A) = \mathbb{P}((\lambda(-2M), \pi(-2M)) \in A) | (\lambda(0), \pi(0)) = (\lambda, \pi)). \tag{39}$$

Lemma 21. *There exists a constant $\alpha > 0$ such that the following is true for any r. Let $\mathcal{C}_1, \mathcal{C}_2 \in \mathbf{G}(r)$. Then*

$$\mathbb{P}(\omega|_{[1,r]} = \mathcal{C}_1, \omega|_{[r+1,2r]} \in \mathbf{G}(r) \ |\omega|_{[2r+1,3r]} = \mathcal{C}_2) \ge \alpha \varphi(\mathcal{C}_1).$$

Indeed, we have the following propositions.

Proposition 20. *There exists a constant p_1 such that the following is true for all r and all $n \ge r$.*

Let $C_2 \in \mathbf{G}(r)$, $(\lambda, \pi) \in \mathcal{C}_2$. Then

$$\mathbb{P}((\lambda(-2n), \pi(-2n)) \in \Delta(\mathbf{q}) | (\lambda(0), \pi(0)) = (\lambda, \pi)) \ge p_1.$$

This follows from the definition of a good cylinder and Corollary 5.

Proposition 21. *There exists a constant p_2 such that the following is true for all k:*

$$\mathbb{P}(\omega|_{[1,r]} \in \mathbf{G}(r), \omega|_{[2M+1,l+2M+1]} = \mathbf{q} \ |\omega|_{[r+1,r+l+1]} = \mathbf{q}) \ge p_2.$$

This follows from the estimates (37), (38) on the measure of bad cylinders and from Proposition 19.

Proposition 22. *There exists a constant p_3 such that the following is true for all r. Let $c_1 \dots c_n \dots \in \Delta(\mathbf{q})$. Then*

$$\mathbb{P}(\omega|_{[1,r]} = \mathcal{C}_1 | \omega_{r+2M+1} = c_1, \omega_{r+2M+2} = c_2, \dots) \ge p_3 \varphi(C_1).$$

This follows directly from Lemma 20.

The three propositions imply Lemma 21.

14.3. **Approximation by a Markov measure.** We define a new measure $\mathbf{p}_{r,\theta}$ on the set $\mathbf{G}(r^2)$ of good cylinders of length r^2.

Let $\mathcal{C} = c_1 \dots c_{r^2}$ be an (r,θ)-good cylinder. Set $\mathcal{C}_i = c_{ir+1} \dots c_{(i+1)r}$. Define

$$\mathbf{p}_{r,\theta}(\mathbf{C}) = \mathbb{P}(\omega|_{[1,r]} = \mathcal{C}_1 | \omega|_{[r+1,2r]} = \mathcal{C}_2)\mathbb{P}(\omega|_{[r+1,2r]} = \mathcal{C}_2|\omega|_{[2r+1,3r]} = \mathcal{C}_3)\dots \\ \mathbb{P}(\omega|_{[r^2-r+1,r^2]} = \mathcal{C}_r).$$

If D is not a good cylinder, then $\mathbf{p}_{r,\theta}(D) = 0$.

Normalize to get a probability measure:

$$\mathbf{P}_{r,\theta}(\mathcal{C}) = \frac{\mathbf{p}_{r,\theta}(\mathcal{C})}{\sum_{\mathcal{D}\in\mathbf{G}(r^2)} \mathbf{p}_{r,\theta}(\mathcal{D})}.$$

$\mathbf{P}_{r,\theta}$ is a Markov measure of memory r (in general, nonhomogeneous), as is shown by the following well-known lemma [14].

Lemma 22. *For any k, $0 < k < r$, we have*

$$\mathbf{P}_{r,\theta}(\omega|_{[kr+1,(k+1)r]} = \mathcal{C}_k|\omega|_{[(k+1)r+1,r^2]}) = \mathcal{C}_{k+1}\dots\mathcal{C}_r) \\ = \mathbf{P}_{r,\theta}(\omega|_{[kr+1,(k+1)r]} = \mathcal{C}_k|\omega|_{[(k+1)r+1,(k+2)r]}) = \mathcal{C}_{k+1}).$$

From the Hölder property for the transition probability, we have

Proposition 23. *There exist constants C_{41}, C_{42} such that the following is true for any r.*

Let $c_1 \dots c_n \dots \in \Omega_{\mathcal{A},B}$ and assume $c_{n+1}\dots c_{n+r} \in \mathbf{G}(r)$. Then

$$\exp(-C_{41}\exp(-C_{42}k^\theta)) \\ \le \frac{P(\omega_1 = c_1,\dots,\omega_n = c_n|\omega_{n+1} = c_{n+1},\dots,\omega_{n+r} = c_{n+r})}{\mathbb{P}(\omega_1 = c_1,\dots,\omega_n = c_n|\omega_{n+1} = c_{n+1},\dots,\omega_{n+i} = c_{n+i},\dots)} \\ \le \exp(C_{41}\exp(-C_{42}k^\theta)).$$

Corollary 17. *There exist constants C_{43}, C_{44} such that the following is true for any r. Let $A \in \mathcal{F}_n$, let $c_{n+1}\dots c_{n+i}\dots \in \Omega_{\mathcal{A}}$, and assume $c_{n+1}\dots c_{n+r} \in \mathbf{G}(r)$. Then*

$$\exp(-C_{43}\exp(-C_{44}k^\theta)) \le \frac{\mathbb{P}(A|\omega_{n+1} = c_{n+1},\dots,\omega_{n+r} = c_{n+r})}{\mathbb{P}(A|\omega_{n+1} = c_{n+1},\dots,\omega_{n+i} = c_{n+i},\dots)} \\ \le \exp(C_{43}\exp(-C_{44}k^\theta)).$$

Applying l times, we obtain

Lemma 23. *There exist constants $C_{45}, C_{46}, C_{47}, C_{48}$ such that the following is true for any r. Let $c_1\dots c_{r^2} \in \mathbf{G}(r^2)$. Then for any l, $1 \le l \le r$, we have*

$$\exp(-C_{45}l\exp(-C_{46}k^\theta)) \\ \le \frac{\mathbb{P}(\omega_1 = c_1,\dots,\omega_{lr} = c_{lr}|\omega_{lr+1} = c_{lr+1},\dots,\omega_{r^2} = c_{r^2})}{\mathbf{p}_{r,\theta}(\omega_1 = c_1,\dots,\omega_{lr} = c_{lr}|\omega_{lr+1} = c_{lr+1},\dots,\omega_{r^2} = c_{r^2})} \\ \le \exp(C_{45}l\exp(-C_{46}k^\theta))$$

and

$$\exp(-C_{47}l\exp(-C_{48}k^\theta)) \le \frac{\mathbb{P}(\omega_1 = c_1,\dots,\omega_{lr} = c_{lr})}{\mathbf{p}_{r,\theta}(\omega_1 = c_1,\dots,\omega_{lr} = c_{lr})} \le \exp(C_{47}l\exp(-C_{48}k^\theta)).$$

Summing over cylinders of length lr, we obtain the following.

Corollary 18. *There exist constants C_{49}, C_{50} such that the following is true for any r. Let $c_1 \dots c_{r^2} \in \mathbf{G}(r^2)$. Then for any l, $1 \le l \le r$, and any $A \in \mathcal{F}_{lr}$, we have*

$$\begin{aligned}&\exp(-C_{49} l \exp(-C_{50} k^\theta)) \\ &\quad \le \frac{\mathbb{P}(A \cap \mathbf{G}(lr) | \omega_{lr+1} = c_{lr+1}, \dots, \omega_{r^2} = c_{r^2})}{\mathbf{p}_{r,\theta}(A | \omega_{lr+1} = c_{lr+1}, \dots, \omega_{r^2} = c_{r^2})} \\ &\quad \le \exp(C_{49} l \exp(-C_{50} k^\theta))\end{aligned}$$

and

$$\exp(-C_{49} l \exp(-C_{50} k^\theta)) \le \frac{\mathbb{P}(A \cap \mathbf{G}(lr))}{\mathbf{p}_{r,\theta}(A)} \le \exp(C_{49} l \exp(-C_{50} k^\theta)).$$

Using (37), we can estimate the total mass of the measure $\mathbf{p}_{r,\theta}$.

Corollary 19. *There exist constants C_{51}, C_{52} such that for any r we have*

$$\mathbf{p}_{r,\theta}(\mathbf{G}(r^2)) \ge \exp(-C_{51} r \exp(-C_{52} k^{(1-\theta)/2})).$$

We now have normalized versions of previous statements.

Corollary 20. *There exist constants $C_{53}, C_{54}, C_{55}, C_{56}$ such that the following is true for any r. Let $c_1 \dots c_{r^2} \in \mathbf{G}(r^2)$. Then for any l, $1 \le l \le r$, and any $A \in \mathcal{F}_{lr}$, we have*

$$\begin{aligned}&\exp(-C_{53} l \exp(-C_{54} k^\theta) - C_{55} r \exp(-C_{56} k^{(1-\theta)/2})) \\ &\quad \le \frac{\mathbb{P}(A \cap \mathbf{G}(lr) | \omega_{lr+1} = c_{lr+1}, \dots, \omega_{r^2} = c_{r^2})}{\mathbf{P}_{r,\theta}(A | \omega_{lr+1} = c_{lr+1}, \dots, \omega_{r^2} = c_{r^2})} \\ &\quad \le \exp(C_{53} l \exp(-C_{54} k^\theta) + C_{55} r \exp(-C_{56} k^{(1-\theta)/2})\end{aligned}$$

and

$$\begin{aligned}&\exp(-C_{53} l \exp(-C_{54} k^\theta) - C_{55} r \exp(-C_{56} k^{(1-\theta)/2})) \\ &\quad \le \frac{\mathbb{P}(A \cap \mathbf{G}(lr))}{\mathbf{P}_{r,\theta}(A)} \\ &\quad \le \exp(C_{53} l \exp(-C_{54} k^\theta) + C_{55} r \exp(-C_{56} k^{(1-\theta)/2}).\end{aligned}$$

Using the Markov approximation, we can estimate conditional measure of good cylinders for the measure $\mathbb{P}$.

Corollary 21. *There exist constants $C_{57}, C_{58}, C_{59}, C_{60}$ such that the following is true for any r. Let $c_1 \dots c_{r^2} \in \mathbf{G}(r^2)$. Then for any l, $1 \le l \le r$, we have*

$$\begin{aligned}&\mathbb{P}((\omega_1 \dots \omega_{lr}) \in \mathbf{G}(lr) | \omega_{lr+1} = c_{lr+1}, \dots, \omega_{r^2} = c_{r^2}) \\ &\qquad \ge \exp(-C_{57} l \exp(-C_{58} k^\theta) - C_{59} r \exp(-C_{60} k^{(1-\theta)/2})).\end{aligned}$$

Proof. Indeed,

$$\mathbf{P}_{r,\theta}((\omega_1 \dots \omega_{lr}) \in \mathbf{G}(lr) | \omega_{lr+1} = c_{lr+1}, \dots, \omega_{r^2} = c_{r^2}) = 1. \qquad \square$$

14.4. **Doeblin Condition.**

Proposition 24. *There exists C_{61} such that the following holds for any r. For any $\mathcal{C}_1 \subset \Delta(\mathbf{q})$, $C_2 \subset \Delta_{\mathbf{q}}$, and any $\mathcal{C}_3 \in \mathbf{G}(r)$, we have either*

$$\frac{1}{C_{61}} \leq \frac{\mathbf{p}_{r,\theta}(C_3|C_2)}{\mathbf{p}_{r,\theta}(C_3|C_1)} \leq C_{61}$$

or $\mathbf{p}_{r,\theta}(C_3|C_2) = \mathbf{p}_{r,\theta}(C_3|C_1) = 0$.

Considering n-step transition probabilities, we obtain

Proposition 25. *There exists a constant C_{62} such that the following holds for any r. For any $\mathcal{C}_1 \subset \Delta(\mathbf{q})$, $C_2 \subset \Delta_{\mathbf{q}}$, any $\mathcal{C}_3 \in \mathbf{G}(r)$, and any $n \geq M$, we have*

$$\frac{1}{C_{62}} \leq \frac{\mathbf{p}_{r,\theta}(\omega|_{[1,r]} = C_1|\omega_{[2n+r,2n+2r]} = C_2)}{\mathbf{p}_{r,\theta}(\omega|_{[1,r]} = C_1|\omega_{[2n+r,2n+2r]} = C_3)} \leq C_{62}.$$

Now, mixing Proposition 19 and Proposition 20, and the definition of a good cylinder implies that

Proposition 26. *There exists a constant C_{63} such that the following holds for any r. For any $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3 \in \mathbf{G}(r)$ we have*

$$\frac{1}{C_{63}} \leq \frac{\mathbf{p}_{r,\theta}(\omega|_{[1,r]} = C_1|\omega_{[2r,3r]} = C_2)}{\mathbf{p}_{r,\theta}(\omega|_{[1,r]} = C_1|\omega_{[2r,3r]} = C_3)} \leq C_{63}.$$

Now let $c_1 \dots c_{r^2} \in \mathbf{G}(r^2)$. Denote $\mathcal{C}_i = c_{ir+1} \dots c_{(i+1)r}$. Lemma 21, together with the above estimates, implies the following.

Corollary 22. *There exist constants C_{71}, C_{72} such that the following is true. For any l, $1 \leq l \leq r$, we have*

$$\begin{aligned}\mathbb{P}(\omega|_{[1,lr]} \in & \mathbf{G}(lr), \omega|_{[lr+1,(l+1)r]} = \mathcal{C}_l, \\ & \omega|_{[(l+1)r+1,(l+2)r]} \in \mathbf{G}(r)|\omega_{(l+2)r+1,(l+3)r]}) = \mathcal{C}_3) \geq C_{71} \times \varphi(\mathcal{C}_l)\end{aligned}$$

and

$$\begin{aligned}\mathbf{P}_{r,\theta}(\omega|_{[1,lr]} \in & \mathbf{G}(lr), \omega|_{[lr+1,(l+1)r]} = \mathcal{C}_l, \\ & \omega|_{[(l+1)r+1,(l+2)r]} \in \mathbf{G}(r)|\omega_{(l+2)r+1,(l+3)r]}) = \mathcal{C}_3) \geq C_{72} \times \varphi(\mathcal{C}_l).\end{aligned}$$

This is the Doeblin Condition for the measure $\mathbf{P}_{r,\theta}$ (see [13], [14], [22]). The Doeblin Condition implies that there exist constants C_{73}, C_{74} such that for any $\mathcal{C}_1, \mathcal{C}_2 \in \mathbf{G}(r)$, we have

$$\begin{aligned}\exp(-C_{73}\exp(-C_{74}r)) \leq & \frac{\mathbf{P}_{r,\theta}(\omega|_{[1,r]} = \mathcal{C}_1|\omega|_{[r^2,r^2+r]} = \mathcal{C}_2)}{\mathbf{P}_{r,\theta}(\mathcal{C}_1)} \\ & \leq \exp(C_{73}\exp(-C_{74}r)),\end{aligned}$$

whence we obtain

Proposition 27. *There exist constants $C_{81}, C_{82}, C_{83}, C_{84}$ such that the following is true for any r:*

$$\begin{aligned}&\exp(-C_{81}(\exp(-C_{82}r) + \exp(-C_{83}r^{\theta}) + \exp(-C_{84}r^{(1-\theta)/2}))) \\ &\leq \frac{\mathbb{P}(\omega|_{[1,r]} = \mathcal{C}_1|\omega|_{[r+1,r^2]} \in \mathbf{G}(r^2 - r), \omega|_{[r^2,r^2+r]} = \mathcal{C}_2)}{\mathbb{P}(\mathcal{C}_1)} \\ &\leq \exp(C_{81}(\exp(-C_{82}r) + \exp(-C_{83}r^{\theta}) + \exp(-C_{84}r^{(1-\theta)/2}))).\end{aligned}$$

Moreover, in view of mixing, Proposition 19, and Proposition 20, the same estimate, up to a constant, takes place for any $n \geq r^2$.

Proposition 28. *There exist constants $C_{85}, C_{86}, C_{87}, C_{88}$ such that the following is true for all r and all $n \geq r^2$:*

$$\begin{aligned}
&\exp(-C_{85}(\exp(-C_{86}r) + \exp(-C_{87}r^{\theta}) + \exp(-C_{88}r^{(1-\theta)/2}))) \\
&\quad \leq \frac{\mathbb{P}(\omega|_{[1,r]} = \mathcal{C}_1 | \omega|_{[r+1,n]} \in \mathbf{G}(n-r), \omega|_{[n,n+r]} = \mathcal{C}_2)}{\mathbb{P}(\mathcal{C}_1)} \\
&\quad \leq \exp(C_{85}(\exp(-C_{86}r) + \exp(-C_{87}r^{\theta}) + \exp(-C_{88}r^{(1-\theta)/2}))).
\end{aligned}$$

15. Approximation of Hölder functions and completion of the Proof of Theorems 4, 7, and 8

We shall prove the decay of correlations for a slightly more general class of functions on $\Delta(\mathcal{R})$ than Hölder functions. (We shall need this slightly more general class in the proof of the Central Limit Theorem.)

Namely, we shall only require that a function be Hölder in restriction to cylinders of some given length, and we shall also allow a moderate growth of the Hölder constant at infinity.

Formally, we say that a function $\phi : \Delta(\mathcal{R}) \to \mathbb{R}$ is *weakly l, α-Hölder* if the following holds. Let k be a positive integer, and let $w \in W_{\mathcal{A},B}$, $|w| \leq l$ be such that $\Delta(w) \subset \Delta_{\exp(-k)}$. Then there exists a constant $C(\phi)$ such that for any $(\lambda, \pi), (\lambda', \pi) \in \Delta(w)$, we have

$$|\phi(\lambda, \pi) - \phi(\lambda', \pi)| \leq Ckd(\lambda, \lambda')^{\alpha}.$$

The smallest such C for a given ϕ will be denoted $C_{l,\alpha}^{weak}(\phi)$. Clearly, if ϕ is Hölder with exponent α, then it is also weakly l, α-Hölder for any l and $C_{l,\alpha}^{weak}(\phi) \leq C_{\alpha}(\phi)$.

Recall that $\mathcal{B}_n$ is the σ-algebra of sets of the form $\mathcal{G}^{-n}(A)$, $A \subset \Delta(\mathcal{R})$.

To prove the decay of correlations, it suffices to estimate the L_2-norm of $E(\phi|\mathcal{B}_{2n})$ for a given *weakly l, α-Hölder ϕ*.

It will be convenient to assume that $\phi \geq 1$ (by linearity, it suffices to consider that case).

Proposition 29. *Let $\theta \in \mathbb{R}$, $0 < \theta < 1$. Let $p > 2$ and $\alpha > 0$. There exist constants C_{91}, C_{92}, C_{93} such that the following is true for any r and any $n \geq r^2$.*

Let $l \leq r$. Let $\phi \in L_p(\Delta(\mathcal{R})^+, \nu)$ be weakly l, α-Hölder and satisfy $\phi \geq 1$.

Then $\phi = \phi_1 + \phi_2 + \phi_3$ where

(1) *$\phi_1 \geq 1$ on $\mathbf{G}(n)$ and $\phi_1 = \phi_2 = 0$ on $\Delta(B(n))$,*
(2) *for any $(\lambda, \pi) \in \mathbf{G}(n)$, we have*

$$|\frac{E(\phi_1|\mathcal{F}_n)(\lambda, \pi)}{E(\phi_1)} - 1| \leq \exp(-C_{91}(r^{(1-\theta)/2} + r^{\theta})),$$

(3) *for $(\lambda, \pi) \in G(n)$, we have $|\phi_2| \leq C_{l,\alpha}^{weak}(\phi) \exp(-C_{92}r^{\theta})$,*
(4) *$||\phi_3||_{L_2} \leq \exp(-C_{93}r^{(1-\theta)/2})||\phi||_{L_p}$.*

Proof. For any good word $w = w_1 \ldots w_{n+r}$, consider its beginning $w_1 \ldots w_r$ and choose a point $x_{w_1 \ldots w_r} \in \Delta(w_1 \ldots w_r)$.

Denote by $\chi_{\Delta(w)}$ the characteristic function of $\Delta(w)$ and set

$$\phi_1 = \sum_{w \in \mathbf{G}(n+r)} \phi(x_{w_1 \ldots w_r}) \chi_{\Delta(w)}.$$

Proposition 28 yields the required properties of ϕ_1 (note that we sum over all good words of length $n+r$ in order to be able to apply the Proposition).

We set $\phi_2 = (\phi - \phi_1)\chi_{G(n+r)}$ and $\phi_3 = \phi\chi_{\Delta(B(n+r))}$. The estimate for ϕ_2 is satisfied by the definition of a Hölder function.

Finally, we have

$$||\phi_3||_{L_2}^2 = E(|\phi\chi_{\Delta(B(n))}|^2),$$

whence, by Hölder's inequality, using the estimate (37), we obtain the desired estimate for ϕ_3, and the Proposition is proved completely. □

Proposition 29 with $\theta = 1/3$ yields Theorem 4.

We now complete the proof of Theorem 7.

For a word $w \in W_{\mathcal{A},B}$, $|w| = 2n+1$, $w = w_1 \ldots w_{2n+1}$, denote $C^{[-n,n]}(w) = \{\omega \in \Omega_{\mathcal{A},B}^{\mathbb{Z}} : \{\omega_{-n} = w_1, \ldots, \omega_n = w_{2n+1}\}$ and set $\overline{\Delta}(w) = \overline{\Phi}^{-1} C^{[-n,n]}(w)$. Denote by $\mathcal{B}_{[-n,n]}$ the sigma-algebra generated by $\overline{\Delta}(w)$ for all $w \in W_{\mathcal{A},B}$.

Also, for $\epsilon > 0$, denote

$$\overline{\Delta}_\epsilon = \{(\lambda, h, a, \pi) \in \overline{\Delta}(\mathcal{R}) : \lambda \in \Delta_\epsilon\}.$$

Again, we shall prove the theorem for a slightly larger class of functions.

We say that a function $\phi : \overline{\Delta}(\mathcal{R}) \to \mathbb{R}$ is *weakly l, α-Hölder* if the following holds. Let k be a positive integer, and let $w \in W_{\mathcal{A},B}$, $|w| \le 2l+1$ be such that $\overline{\Delta}(w) \subset \Delta_{\exp(-k)}$. Then there exists a constant $C(\phi)$ such that for any (λ, h, a, π), $(\lambda', h', a', \pi) \in \overline{\Delta}(w)$, we have

$$|\phi(\lambda, h, a, \pi) - \phi(\lambda', h', a', \pi)| \le Ckd((\lambda, h, a, \pi), (\lambda', h', a', \pi))^\alpha.$$

The smallest such C for a given ϕ will be denoted $C_{l,\alpha}^{weak}(\phi)$. Clearly, if ϕ is Hölder with exponent α, then it is also *weakly l, α-Hölder* for any l and $C_{l,\alpha}^{weak}(\phi) \le C_\alpha(\phi)$.

Denote by $\overline{G}(2n+1)$ the union of all $\overline{\Delta}(w)$ for good w, by $\overline{B}(2n+1)$ the complement of $\overline{G}(2n+1)$.

Proposition 30. *Let $\theta \in \mathbb{R}$, $0 < \theta < 1$. Let $p > 2$ and $\alpha > 0$. There exist constants C_{101}, C_{102}, such that the following is true for any r and any $n \ge r^2$.*

Let $l \le r$. Let $\phi \in L_p(\overline{\Delta}(\mathcal{R}) \cdot \overline{\nu})$ be weakly l, α-Hölder and satisfy $\phi \ge 1$. Then there exist functions ϕ_1, ϕ_2, ϕ_3 such that

(1) $\phi = \phi_1 + \phi_2 + \phi_3$,
(2) ϕ_1 *is $\mathcal{B}_{[-n,n]}$-measurable and supported on $\overline{G}(2n+1)$,*
(3) $|\phi_2| \le C_{101} C_\alpha(\phi) \exp(-r^{(1-\theta)/2} + r^\theta)$,
(4) $|\phi_3|_{L_2} \le C_{102} \exp(-r^{(1-\theta)/2}) ||\phi||_{L_p}$.

For any good w, $|w| = 2n+1$, take an arbitrary point x_w in $\overline{\Delta}(w)$. Set

$$\phi_1 = \sum_{w \in G(2n+1)} \phi(x_w) \chi_{\overline{\Delta}(w)},$$

$$\phi_2 = (\phi - \phi_1) \cdot \chi_{G(2n+1)},$$

$$\phi_3 = \phi \cdot \chi_{B(2n+1)},$$

and the Proposition is proved.

Proposition 30 with $\theta = 1/3$ yields Theorem 7.

It remains to establish the Central Limit Theorem for the flow P^t. Consider the special function $\tilde{\tau}$ of the flow P^t over the transformation $\mathcal{F}$. Note that $\tilde{\tau}(\lambda, h, a, \pi)$ only depends on (λ, π). Consider the restriction of $\tilde{\tau}$ on a cylinder of the form $\Delta(w_1)$, $w_1 \in \mathcal{A}$. Then there exist distinct $j(1), \ldots, j(l) \in \{1, \ldots, m\}$ such that

$$\tilde{\tau}(\lambda, \pi) = \log(\lambda_{j(1)} + \lambda_{(j(2)} + \cdots + \lambda_{j(l)}),$$

which shows that the function $\tilde{\tau}$, restricted to an arbitrary $\Delta(w_1)$, is Lipshitz with respect to the Hilbert metric on $\Delta(\mathcal{R})$.

Now for a Hölder ϕ consider the function

$$\tilde{\phi}(x) = \int_0^{\tilde{\tau}(x)} \phi(P^t x).$$

For any $k > 1$, if $(\lambda, \pi) \in \Delta_{\exp(-k)}$, then, by definition, $\tilde{\tau}(\lambda, \pi) \leq k$. Therefore, if ϕ is Hölder of exponent α, then $\tilde{\phi}$ is weakly $1, \alpha$-Hölder.

It is easy to see that $\tilde{\tau}(\lambda, \pi) \in L_r(\Delta(\mathcal{R}), \nu)$ for any $r > 1$, whence, if $\phi \in L_p(\Omega_0(\mathcal{R}), \mu_{\mathcal{R}})$ for some $p > 2$, then there exists $p' > 2$ such that the function

$$\tilde{\phi}(x) = \int_0^{\tilde{\tau}(x)} \phi(P^t x)$$

satisfies $\tilde{\phi} \in L_{p'}(\mathcal{Y}^{\pm}, \overline{\nu})$.

Therefore, the Theorem of Melbourne and Török [15] implies Theorem 8, the Central Limit Theorem for the flow P^t.

Acknowledgements

I am deeply grateful to Yakov G. Sinai, who introduced me to interval exchange transformations, explained to me the method of Markov approximations, and encouraged me in every possible way as the work progressed (more importantly, when it did not).

I am deeply grateful to Alexander Eskin, who suggested to me the problem of the decay of correlations for the induction map. I am deeply grateful to Giovanni Forni who introduced me to Teichmüller theory. I am deeply grateful to Anton Zorich for our many discussions on the behaviour of the induction map. I am deeply grateful to Tullio Ceccherini-Silberstein for our discussions and his friendly encouragement. I am deeply grateful to Corinna Ulcigrai and Pavel Batchourine for their suggestions that have been of invaluable help to me. I am deeply grateful to Jayadev Athreya, Valdo Durrleman, Charles L. Fefferman, Boris M. Gurevich, Carlangelo Liverani, Michael Ludkovski, Ian Melbourne, Klaus Schmidt, and Andrei Török for useful discussions. I am deeply grateful to the referee for many helpful suggestions on improving the presentation. Part of this work was done at the Erwin Schrödinger Institute in Vienna, at the Institute of Mathematics “Guido Castelnuovo” of the University of Rome “La Sapienza”, and at the CIRM-IML in Marseille. I am deeply grateful to these institutions for their hospitality.

References

[1] William Veech, Gauss measures for transformations on the space of interval exchange maps, Annals of Mathematics, 15(1982), 201–242. MR0644019 (83g:28036b)

[2] W.Veech, Interval exchange transformations. J. Analyse Math. 33 (1978), 222–272. MR0516048 (80e:28034)

[3] Veech, William A. Projective Swiss cheeses and uniquely ergodic interval exchange transformations. Ergodic theory and dynamical systems, I (College Park, Md., 1979–1980), pp. 113–193, Progr. Math., 10, Birkhäuser, Boston, Mass., 1981. MR0633764 (83g:28036a)

[4] Anton Zorich, Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents. Ann. Inst. Fourier (Grenoble) 46 (1996), no. 2, 325–370. MR1393518 (97f:58081)

[5] Oseledets, V. I. The spectrum of ergodic automorphisms. (Russian) Dokl. Akad. Nauk SSSR 168 (1966), 1009–1011. MR0199347 (33:7494)

[6] G.Rauzy, Échanges d'intervalles et transformations induites. Acta Arith. 34, (1979), no. 4, 315–328. MR0543205 (82m:10076)

[7] M. I. Gordin, The central limit theorem for stationary processes, Dokl. Akad. Nauk SSSR 188 (1969), 739-741. MR0251785 (40:5012)

[8] Carlangelo Liverani, Central limit theorem for deterministic systems. International Conference on Dynamical Systems (Montevideo, 1995), 56-75. MR1460797 (98k:28025)

[9] M. Keane, Interval exchange transformations, Math. Zeitschrift, 141(1975), 25–31. MR0357739 (50:10207)

[10] Giovanni Forni, Deviation of ergodic averages for area-preserving flows on surfaces of higher genus. Ann. of Math. (2), 155 (2002), no. 1, 1–103. MR1888794 (2003g:37009)

[11] Lai-Sang Young, Recurrence times and rates of mixing. Israel Journal of Mathematics, 110(1999), 153-188. MR1750438 (2001j:37062)

[12] Véronique Maume-Deschamps, Projective metrics and mixing properties on towers, Trans. Amer. Math. Soc., 353 (2001), no.8, 3371-3389. MR1828610 (2002a:37004)

[13] Sinai, Ya. G. Gibbs measures in ergodic theory. Uspehi Matematicheskih Nauk 27 (1972), no. 4(166), 21–64. MR0399421 (53:3265)

[14] Bunimovich, L. A.; Sinai, Ya. G. Statistical properties of Lorentz gas with periodic configuration of scatterers. Comm. Math. Phys. 78 (1980/81), no. 4, 479–497. MR0606459 (82m:82007)

[15] I. Melbourne, A. Török, Statistical limit theorems for suspension flows, Israel Journal of Mathematics, 144 (2004), 191–209. MR2121540 (2006c:37005)

[16] Kolmogorov, A. N. A local limit theorem for classical Markov chains. Izvestiya Akad. Nauk SSSR. Ser. Mat. 13, (1949). 281–300. MR0031216 (11:119c)

[17] Marcelo Viana, Stochastic Dynamics of Deterministic Systems, Brazilian Mathematics Colloquium 1997, IMPA; online at www.impa.br/ ˜viana.

[18] Caroline Series, The modular surface and continued fractions. J.London Math.Soc.(2) 31 (1985), no.1, 69-80. MR0810563 (87c:58094)

[19] Caroline Series, Geometrical Markov coding of geodesics on surfaces of constant negative curvature. Ergodic theory and dynamical systems 6 (1986), no. 4, 601-625. MR0873435 (88k:58130)

[20] Kerckhoff, S. P. Simplicial systems for interval exchange maps and measured foliations. Ergodic Theory Dynam. Systems 5 (1985), no. 2, 257–271. MR0796753 (87g:58075)

[21] H. Masur, Interval exchange transformations and measured foliations. Ann. of Math. (2) 115 (1982), no. 1, 169–200. MR0644018 (83e:28012)

[22] Doob, J. L. Stochastic processes. John Wiley and Sons, New York, 1990. MR1038526 (91d:60002)

[23] V.A. Rokhlin, Exact endomorphisms of a Lebesgue space, Izv. Akad. Nauk SSSR Ser. Mat. 25 (1961) 499-530. MR0143873 (26:1423)

[24] V.A. Rokhlin, New progress in the theory of transformations with invariant measure. Uspehi Mat. Nauk, 15 no. 4, 3– 26. MR0132155 (24:A2002)

[25] Ya.G. Sinai, Topics in ergodic theory, Princeton University Press, 1994. MR1258087 (95j:28017)

[26] M.Kontsevich, A.Zorich, Connected components of the moduli spaces of Abelian differentials with prescribed singularities, Inventiones Mathematicae, 153(2003), no.3, 631-678. MR2000471 (2005b:32030)

[27] I.P. Cornfeld, Ya. G. Sinai, S.V. Fomin, Ergodic theory, "Nauka", Moscow, 1980. MR0610981 (83a:28017)

[28] Hubbard, John; Masur, Howard, Quadratic differentials and foliations. Acta Math. 142 (1979), no. 3-4, 221–274. MR0523212 (80h:30047)

[29] M.Kontsevich, Lyapunov exponents and Hodge theory, "Mathematical Beauty of Physics", Saclay, 1996. MR1490861 (99b:58147)

[30] J. Athreya, Quantitative recurrence and large deviations for Teichmüller geodesic flows, preprint.

Department of Mathematics, Princeton University, Princeton, New Jersey 08544

Current address: (Until June 30, 2006) Department of Mathematics, The University of Chicago, 5734 South University Avenue, Chicago, Illinois 60637; (starting July 1, 2006) Department of Mathematics, Rice University, MS 136, 6100 Main Street, Houston, Texas 77251-1892

E-mail address: bufetov@math.rice.edu