STATISTICAL PROPERTIES OF THE RAUZY-VEECH-ZORICH MAP

ROMAIN AIMINO AND MARK POLLICOTT

Abstract. In this note we survey some very basic statistical properties of the Rauzy-Veech map and the Zorich acceleration. Our aim is to give a particularly thermodynamic perspective of well known results.

1. Introduction

In this note we will consider the Rauzy-Veech-Zorich renormalization map for interval exchange maps. The special case of interval exchange transformations on two intervals simply corresponds to rotations on the unit circle, and in this case the corresponding renormalization map reduces to the usual Farey map, and its acceleration to the continued fraction transformation. Thus, one might naturally view interval exchange maps on $m \geq 3$ intervals as generalizations of circle rotations; and the renormalization map as a generalization of the classical continued fraction transformation. It was shown by Masur and Veech that their original renormalization map $\mathcal{T}_0$ possesses an absolutely continuous ergodic invariant measure, and Zorich showed that for the accelerated version $\mathcal{T}_1$ there is a finite invariant measure.

A number of interesting statistical results already have already been established for the renormalization map, and related transformations (e.g., Central Limit Theorems and other Limit Theorems cf. [2], [4], [21]). The first aim of this paper is to present an alternative approach to some of these results, and to give some simple generalizations. Indeed, for dynamical systems in general there is a potential hierarchy of statistical properties that one may establish for such maps, beginning with ergodicity; central limits theorems; functional central limit theorems, and finally almost sure invariance principles. In this paper we will re-derive the central limit theorem, the stronger functional central limit theorem, and establish the almost sure invariance principle, from which the others then follow. A basic technique, familiar from other non-uniformly hyperbolic settings, is to induce a hyperbolic map $\mathcal{T}_2$ on a smaller set B in the domain of $\mathcal{T}_1$. In particular, statistical properties are typically easier to establish for $\mathcal{T}_2$, and these can then be "lifted" to the map $\mathcal{T}_1^2$. There is a well known application of related results to Teichmüller flows for abelian differentials, which can be modeled in terms of suspended flows over these maps (and their natural extensions).

One of the interesting applications of the (accelerated) Rauzy-Veech-Zorich map is to the theory of Teichmüller flows. In particular, a suspension semi-flow for the

The first author was partially supported by FCT fellowship SFRH/BPD/123630/2016, by FCT project PTDC/MAT-PUR/28177/2017 and by CMUP, which is financed by national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., under the project with reference UIDB/00144/2020. The second author was partially supported by ERC-Advanced Grant 833802-Resonances and EPSRC grant EP/T001674/1.

(accelerated) Rauzy-Veech-Zorich map corresponds to a well known model for the Teichmüller flow.

Theorem 1.1. *The transformations $\mathcal{T}_1$ and $\mathcal{T}_2$ satisfy the functional central limit theorem with respect to the natural absolutely continuous invariant probability measure for Hölder continuous observables. In particular, they satisfy the law of the iterated logarithm and the arcsine law for Hölder continuous observables.*

The second aim of this paper is to describe a "zeta function" associated to $\mathcal{T}_2$. This is defined by analogy with the Ruelle zeta function for Axiom A diffeomorphisms. The poles of these zeta functions (and the residues of associated complex functions) encapsulate dynamical information about the maps. Moreover, when these invariants vanish then the zeta function takes a particularly trivial form.

We will initially follow Morita in studying a transfer operator associated to $\mathcal{T}_2$ acting on Lipschitz (or, more generally, Hölder) continuous functions [21]. This allows us to apply the method of Mackey and Tyran Kamiǹski [13, 14], to give a simple and direct proof of the (Functional) Central Limit Theorem, and the method of Philipp-Stout [23], as developed in the dynamical context by Melbourne and Nicol [19], to show the almost everywhere invariance principles. Subsequently, we will consider a transfer operator associated to $\mathcal{T}_2$ on a smaller space of analytic functions and study the complex function $d(z,s)$ of two variables formally defined by

$$d(z,s) = \exp\left(-\sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{\mathcal{T}_2^n x = x} |\det(D\mathcal{T}_2^n)(x)|^{-s}\right), \quad z, s \in \mathbb{C},$$

in terms of the periodic points $\mathcal{T}_2^n x = x$ and the weights $|\det(D\mathcal{T}_2^n)(x)|$.

In particular, we can apply a powerful approach of Ruelle [25] (cf also Mayer [16, 17] for particularly readable account in specific cases related to continued fractions) based on Fredholm determinants to show such functions have a meromorphic extension, and we can give an alternative expression for (the sum of the Lyapunov exponents):

$$\Lambda = \int \log|\det(D\mathcal{T}_2)(x)| d\mu_2(x)$$

for the Kontsevich-Zorich cocycle, where μ_2 is the unique absolutely continuous invariant probability measure for $\mathcal{T}_2$.

Theorem 1.2. *The function $d(z,s)$ is analytic on $\mathbb{C}^2$. We can write*

$$\Lambda = \frac{\frac{\partial d(1,s)}{\partial s}|_{s=1}}{\frac{\partial d(z,1)}{\partial z}|_{z=1}}.$$

The methods in this note will work for other multidimensional continued fraction type algorithms, for which the (accelerated) Rauzy-Veech-Zorich algorithm forms a topical example.

In section 2, we recall results on interval exchanges and their renormalizations. In section 3, we introduce the transfer operator on Hölder continuous functions and recall the results of Morita on its spectra. In section 4, we prove the statistical properties for the induced map $\mathcal{T}_2$. In section 5, we derive the statistical properties for the Zorich map $\mathcal{T}_1$. In section 6, we study the transfer operator on the smaller space of analytic functions, and in section 7, we use these results to study

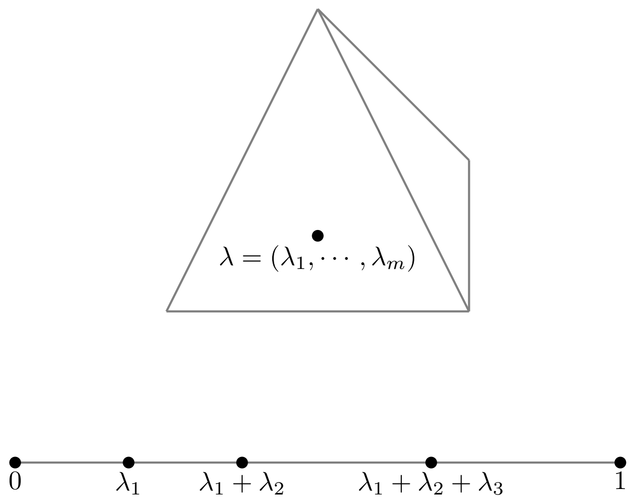

FIGURE 1. A partition of the unit interval corresponds to a point in a simplex

Lyapunov exponents and $d(z, s)$. Finally, in section 8, we describe the connection to Teichmüller flows and in the last section we speculate on the connection to pressure.

2. Interval exchange transformation

In this section we recall some of the basic constructions. We refer the reader to the excellent surveys [32] and [35] for further details.

Interval exchange transformations $T : [0, 1] \to [0, 1]$ are orientation preserving piecewise isometries of the unit interval. In the case of two intervals, this corresponds to a rotation of the circle, i.e., a translation of the interval (modulo one). More generally, assume that I is partitioned into m intervals $I_1, \cdots, I_m$ of lengths $\lambda_1, \cdots, \lambda_m$, respectively, upon each of which T acts isometrically. We can represent this partition as a vector λ in the standard $(m-1)$-dimensional simplex

$$\Delta = \{\lambda = (\lambda_1, \cdots, \lambda_m) : 0 < \lambda_1, \cdots, \lambda_m < 1 \text{ and } \lambda_1 + \cdots + \lambda_m = 1\}.$$

Thus the transformation T is completely determined by these lengths, and by order of the images of the original intervals. This latter information is encapsulated by a permutation π on $\{1, \cdots, m\}$. In particular, every interval exchange transformation corresponds to a pair (λ, π), where $\lambda \in \Delta$ and π is a permutation. Moreover, corresponding to the natural assumption that T doesn't contain an invariant subsystem, we say that π is irreducible if there is no $1 \leq l < m$ such that $\pi(\{1, \cdots, l\}) = \{1, \cdots, l\}$. We will always assume from now on that π is irreducible.

The classical Keane Conjecture (proved by Masur and Veech, independently) states that the transformation T is uniquely ergodic for almost all $\lambda \in \Delta$. The method of proof lead to the development of an important renormalization scheme on such transformations, which we will briefly describe.

2.1. **The Rauzy class of permutations.** Given a permutation π, let us denote by $k = \pi^{-1}(n)$ (i.e., $\pi(k) = n$). A key idea of Rauzy was to replace the permutation

π by one of two new permutations: either

$$a\pi(j) := \begin{cases} \pi(j) & \text{if } 1 \le j \le k \\ \pi(m) & \text{if } j = k+1 \\ \pi(j-1) & \text{if } k+2 \le j \le m \end{cases} \quad \text{or } b\pi(j) := \begin{cases} \pi(j) & \text{if } 1 \le \pi(j) \le \pi(m) \\ \pi(j)+1 & \text{if } \pi(m) < \pi(j) < n \\ \pi(m)+1 & \text{if } j = k \end{cases}$$

If we start from a given permutation we do not necessarily get all permutations by these two operations. This leads to the following definition.

Definition 2.1. *Given a permutation π the Rauzy class $\mathcal{R}$ consists of all permutations that can be derived from π by repeatedly applying these two operations.*

It can be shown that belonging to the same Rauzy class is an equivalence relation. The irreducible permutations are a union of a finite number of Rauzy classes.

Example 2.2. *($n = 4$) The irreducible permutation $\pi_0 = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{smallmatrix}\right)$ lies in a Rauzy class of 7 permutations. These are illustrated in the following diagram, where an arrow labeled by a goes from π to $a\pi$ (and an arrow labeled by b goes from π to $b\pi$).*

$$\begin{array}{ccccc}
a \circlearrowright \left(\begin{smallmatrix} 1&2&3&4 \\ 3&2&4&1 \end{smallmatrix}\right) & & & & \left(\begin{smallmatrix} 1&2&3&4 \\ 4&2&1&3 \end{smallmatrix}\right) \circlearrowright b \\
\uparrow b & \searrow b & & a \swarrow & \uparrow a \\
\left(\begin{smallmatrix} 1&2&3&4 \\ 2&4&3&1 \end{smallmatrix}\right) & \leftarrow_b & \left(\begin{smallmatrix} 1&2&3&4 \\ 4&3&2&1 \end{smallmatrix}\right) & \rightarrow_a & \left(\begin{smallmatrix} 1&2&3&4 \\ 4&1&3&2 \end{smallmatrix}\right) \\
\updownarrow a & & & & \updownarrow b \\
\left(\begin{smallmatrix} 1&2&3&4 \\ 2&4&1&3 \end{smallmatrix}\right) & & & & \left(\begin{smallmatrix} 1&2&3&4 \\ 3&1&4&2 \end{smallmatrix}\right) \\
\circlearrowright b & & & & \circlearrowright a
\end{array}$$

Similarly, one can look at the Rauzy class of $\pi_0 = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{smallmatrix}\right)$ described by the following diagram.

$$\begin{array}{ccccccccc}
a \circlearrowright \left(\begin{smallmatrix} 1&2&3&4 \\ 2&3&4&1 \end{smallmatrix}\right) & \rightarrow_b & \left(\begin{smallmatrix} 1&2&3&4 \\ 3&4&2&1 \end{smallmatrix}\right) & \leftrightarrow_a & \left(\begin{smallmatrix} 1&2&3&4 \\ 3&4&1&2 \end{smallmatrix}\right) & \leftrightarrow_b & \left(\begin{smallmatrix} 1&2&3&4 \\ 4&3&1&2 \end{smallmatrix}\right) & \leftarrow_a & \left(\begin{smallmatrix} 1&2&3&4 \\ 4&1&2&3 \end{smallmatrix}\right) \circlearrowright b \\
b \nwarrow \quad \swarrow b & & & & & & a \searrow \quad \nearrow a & & \\
\left(\begin{smallmatrix} 1&2&3&4 \\ 4&2&3&1 \end{smallmatrix}\right) & & & & & & \left(\begin{smallmatrix} 1&2&3&4 \\ 4&2&3&1 \end{smallmatrix}\right) & & \\
a \swarrow \quad \nwarrow a & & & & & & b \nearrow \quad \searrow b & & \\
b \circlearrowright \left(\begin{smallmatrix} 1&2&3&4 \\ 4&1&2&3 \end{smallmatrix}\right) & \rightarrow_a & \left(\begin{smallmatrix} 1&2&3&4 \\ 4&3&1&2 \end{smallmatrix}\right) & \leftrightarrow_b & \left(\begin{smallmatrix} 1&2&3&4 \\ 3&4&1&2 \end{smallmatrix}\right) & \leftrightarrow_a & \left(\begin{smallmatrix} 1&2&3&4 \\ 3&4&2&1 \end{smallmatrix}\right) & \leftarrow_b & \left(\begin{smallmatrix} 1&2&3&4 \\ 2&3&4&1 \end{smallmatrix}\right) \circlearrowright a
\end{array}$$

We notice a symmetry with respect to the centre of the diagram.

There are excellent descriptions of this procedure in [32] to which we refer the interested reader.

2.2. **The Rauzy-Veech renormalization $\mathcal{T}_0$.** Consider some given $1 \le k \le m$. We can then apply one of the following two operations on the vector $\lambda = (\lambda_1, \cdots, \lambda_m)$, to produce a new vector $\lambda' = (\lambda'_1, \cdots, \lambda'_m)$: Either

Case I ($\lambda_m > \lambda_k$): Let $\lambda \mapsto \lambda' = (\lambda_1, \cdots, \lambda_{m-1}, \lambda_m - \lambda_k)$; or

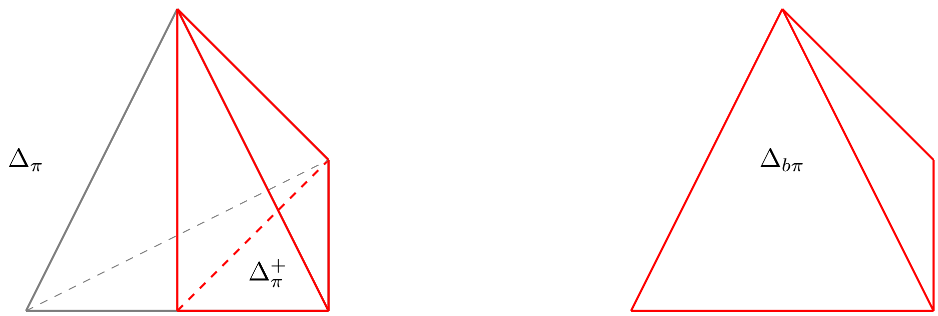

FIGURE 2. The image of half of each copy of the simplex gets mapped to a copy of the simplex

Case II ($\lambda_k > \lambda_m$)*:* Let $\lambda \mapsto \lambda' = (\lambda_1, \cdots, \lambda_{k-1}, \lambda_k - \lambda_m, \lambda_m, \lambda_{k+1}, \cdots, \lambda_{m-1})$.

Firstly, we would like to make a particular choice of case such that vector λ' is strictly positive. The case $\lambda_k = \lambda_m$ is therefore ambiguous, but atypical, and shall be ignored. Secondly, we observe that the definition of λ' is such that it does not lie in the simplex Δ. However, this will soon be corrected by rescaling.

We can define a map $\mathcal{T}_0$ from $\Delta \times \mathcal{R}$ to itself (modulo some codimension one planes, as described above, on which it is ambiguously defined). This will be a renormalization map, in the sense that it associates a new interval exchange map to an old one (with the same number of intervals, m). To be more precise, given $\pi \in \mathcal{R}$ we denote

$$\Delta_\pi^+ = \{(\lambda, \pi) \in \Delta \times \{\pi\} : \lambda_m > \lambda_{\pi^{-1}m}\} \text{ and}$$
$$\Delta_\pi^- = \{(\lambda, \pi) \in \Delta \times \{\pi\} : \lambda_m < \lambda_{\pi^{-1}m}\}.$$

We can define a transformation $\mathcal{T}_0 : \Delta \times \mathcal{R} \to \Delta \times \mathcal{R}$ a.e. by

$$\mathcal{T}_0(\lambda, \pi) = \left(\frac{\lambda'}{\|\lambda'\|_1}, \pi'\right) = \begin{cases} \left(\frac{(\lambda_1, \cdots, \lambda_{m-1}, \lambda_m - \lambda_k)}{1-\lambda_k}, a\pi\right) & \text{if } \lambda \in \Delta_\pi^+ \\ \left(\frac{(\lambda_1, \cdots, \lambda_{k-1}, \lambda_k - \lambda_m, \lambda_m, \lambda_{k+1}, \cdots, \lambda_{m-1})}{1-\lambda_m}, b\pi\right) & \text{if } \lambda \in \Delta_\pi^- \end{cases}$$

with $k = \pi^{-1}(m)$, where we divide by $\|\lambda'\|_1 = \sum_i \lambda_i'$ so as to rescale the image vectors to lie on the simplex Δ.

Example 2.3. *(Example 2.2 revisited) Let* $\underline{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$*. We can again consider the Rauzy class* $\mathcal{R}$ *of* $\pi = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{smallmatrix}\right)$ *as described above. We can then consider, say, the restriction of the map to the simplex labelled by* $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{smallmatrix}\right)$*. Since* $k = \pi^{-1}(4) = 3$ *we have that*

$$\mathcal{T}_0\left((\lambda_1, \lambda_2, \lambda_3, \lambda_4), \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{smallmatrix}\right)\right)$$
$$= \begin{cases} \left(\left(\frac{\lambda_1}{1-\lambda_3}, \frac{\lambda_2}{1-\lambda_3}, \frac{\lambda_3}{1-\lambda_3}, \frac{\lambda_4-\lambda_3}{1-\lambda_3}\right), \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{smallmatrix}\right)\right) & \textit{if } \lambda_4 > \lambda_3 \\ \left(\left(\frac{\lambda_1}{1-\lambda_4}, \frac{\lambda_2}{1-\lambda_4}, \frac{\lambda_3-\lambda_4}{1-\lambda_4}, \frac{\lambda_4}{1-\lambda_4}\right), \left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{smallmatrix}\right)\right) & \textit{if } \lambda_4 < \lambda_3. \end{cases}$$

Unfortunately, these transformations aren't uniformly hyperbolic, as one can readily see since some of the boundaries of the simplicies remain fixed (e.g., the side $\lambda_3 = 0$ in the simplex). This will be partly remedied by replacing $\mathcal{T}_0$ by maps which are "more hyperbolic".

2.3. **The Zorich accelerated remormalization $\mathcal{T}_1$.** Following Zorich, one can consider a map $\mathcal{T}_1 : \Delta \times \mathcal{R} \to \Delta \times \mathcal{R}$ defined a.e. by $\mathcal{T}_1(\lambda, \pi) = \mathcal{T}_0^{n(\lambda,\pi)}(\lambda, \pi)$ where

$$n(\lambda, \pi) = \inf\{k > 0 : \mathcal{T}_0^k(\lambda, \pi) \in \Delta^{\pm} \times \mathcal{R} \text{ where } \lambda \in \Delta^{\mp}\}$$

and where we denote $\Delta^+ = \bigcup_{\pi \in \mathcal{R}} \Delta_\pi^+$ and $\Delta^- = \bigcup_{\pi \in \mathcal{R}} \Delta_\pi^-$.

The following elegant result was proved by Zorich.

Proposition 2.4 (Zorich)**.** *The transformation $\mathcal{T}_1$ preserves a finite absolutely continuous invariant measure μ_1 (i.e., $\mu_1(\Delta \times \mathcal{R}) < +\infty$). Moreover, the restriction $\mathcal{T}_1^2 : \Delta^+ \to \Delta^+$ is ergodic (and $\mathcal{T}_1^2 : \Delta^- \to \Delta^-$ is ergodic).*

Previously, Masur and Veech had shown the existence of a sigma finite $\mathcal{T}_0$-invariant measure μ_0, which can be easily recovered from μ_1.

However, to gain more control over the distortion properties of the transformations one can induce on a smaller set, so as to get a transformation which has even stronger properties.

2.4. **The induced map $\mathcal{T}_2$ on a smaller set.** Let $\mathcal{P} = \{\Delta_\pi^+, \Delta_\pi^- : \pi \in \mathcal{R}\}$ be the natural finite partition of $\Delta \times \mathcal{R}$ then we can define the refinements

$$\mathcal{P}_n := \vee_{k=0}^{n-1} \mathcal{T}_1^{-k} \mathcal{P} = \{P_{i_1} \cap \mathcal{T}_1^{-1} P_{i_2} \cap \cdots \cap \mathcal{T}_1^{-(n-1)} P_{i_{n-1}} : P_j \in \mathcal{P}\}$$

for any $n \geq 1$. Following a now standard approach we can can choose $n_0 > 1$ and $B \in \mathcal{P}_{n_0}$, say, to be any image of an inverse branch of $\mathcal{T}^{n_0}$ which is a contraction. [1]

Finally, we can then consider the induced map $\mathcal{T}_2 : B \to B$ defined by $\mathcal{T}_2(\lambda, \pi) = \mathcal{T}_1^{\widehat{n}(\lambda,\pi)}(\lambda, \pi)$ where

$$\widehat{n}(\lambda, \pi) = \inf\{k > 0 : \mathcal{T}_1^k(\lambda, \pi) \in B\}$$

is the first return time to B. The following is immediate from the observation that the composition of projective transformation remains projective, see Morita [21, Lemma 3.1].

Lemma 2.5. *The induced map $\mathcal{T}_2 : B \to B$ is a piecewise projective expanding map of the general form*

$$(\lambda_1, \cdots, \lambda_n) \mapsto \left(\frac{\sum_{j=1}^d a_{1j}\lambda_j}{\sum_{i,j=1}^d a_{ij}\lambda_j}, \cdots, \frac{\sum_{j=1}^d a_{dj}\lambda_j}{\sum_{i,j=1}^d a_{ij}\lambda_{ij}} \right)$$

on each piece of the partition of smoothness of $\mathcal{T}_2$.

We are now in a position to use familiar techniques for the study of hyperbolic maps.

---

[1] All of these transformations are projective, i.e., matrices act linearly on vectors, followed by normalizing. Such a transformation is contracting in the projective metric when the simplex is mapped strictly inside itself, which happens when the matrix is strictly positive.

3. Transfer operators

Let ω denote the natural volume form on B. We can formally define a linear map $\mathcal{L} : L^1(B,\omega) \to L^1(B,\omega)$ associated to $\mathcal{T}_2 : B \to B$ by the identity

$$\int_B \mathcal{L}f(x)g(x)d\omega(x) = \int_B f(x)g(\mathcal{T}_2 x)d\omega(x), \text{ where } f \in L^1(B), g \in L^\infty(B)$$

and we denote $x = (\underline{\lambda}, \pi) \in B$. (The existence of such a $\mathcal{L}f \in L^1(B)$ follows immediately from the Riesz representation theorem). Moreover, we can use the change of variables formula to formally write:

$$\mathcal{L}^k f(x) = \sum_{y \in \mathcal{T}_2{}^{-k} x} \frac{f(y)}{|\mathrm{Jac}(\mathcal{T}_2^k)(y)|} \text{ a.e..}$$

In fact, a simple calculation, see Veech [29, Proposition 5.2], shows:

Lemma 3.1. *Let A be the matrix such that $y = \frac{Ax}{\|Ax\|_1}$. We can write the Jacobian as $\mathrm{Jac}(\mathcal{T}_2^k)(y) = \|Ax\|_1^m$.*

From this explicit formula for the Jacobian one easily sees that $\mathcal{L}(C^0(B)) \subset C^0(B)$. In order to get stronger results on $\mathcal{T}_2$, we need to consider the operator acting on smaller Banach spaces than $C^0(B)$. In section 6, we will consider the operator acting on analytic functions. However, for the present we shall follow the more classical approach of studying the operator acting on Hölder continuous functions.

Given $\beta > 0$ and a function $w : B \to \mathbb{C}$, we define $\|w\|_\beta = \|w\|_\infty + |w|_\beta$ where

$$|w|_\beta = \sup_{x \neq y} \frac{|w(x) - w(y)|}{\|x - y\|^\beta}$$

and let $C^\beta(B) = \{w : B \to \mathbb{C} : \|w\| < \infty\}$. When $\beta = 1$ these are simply the Lipschitz functions. The next result can be used to show that $\mathcal{L}$ preserves Hölder functions. Let $\mathcal{Q}$ be the partition of smoothness of $\mathcal{T}_2$, and let $\mathcal{Q}_k = \vee_{i=0}^{k-1} \mathcal{T}_2^{-i}\mathcal{Q}$. The following result is basically due to Morita [21]:

Lemma 3.2.
1. *There exists $C > 0$ and $\Theta > 1$ such that for any $n \geq 1$ and x, y in the same element of $\mathcal{Q}_n$ we have*
$$\|\mathcal{T}_2{}^n x - \mathcal{T}_2{}^n y\| \geq C\Theta^n \|x - y\|.$$
2. *There exists $C > 0$ such that for any $n \geq 1$ and x, y lie in the same element of $\mathcal{Q}_n$ we have*
$$\left|\log\left(\frac{\mathrm{Jac}(\mathcal{T}_2^n)(x)}{\mathrm{Jac}(\mathcal{T}_2^n)(y)}\right)\right| \leq C\|\mathcal{T}_2{}^n x - \mathcal{T}_2{}^n y\|.$$
3. *There exists $D > 1$ such that for any $A \in \mathcal{Q}_n$ and any $x \in A$ we can estimate*
$$\frac{1}{D} \leq \omega(A)\,|\mathrm{Jac}(\mathcal{T}_2^n)(x)| \leq D.$$

Proof. These results are based on the basic observation that the first return map $\mathcal{T}_2 : B \to B$ must be of the form $\mathcal{T}_2(x) = \mathcal{T}_1^{\widehat{n}(\lambda,\pi)}(x) = \mathcal{T}_1^{\widehat{n}(\lambda,\pi)-n_0} \circ \mathcal{T}_1^{n_0}(x)$, where $\mathcal{T}_1^{\widehat{n}(\lambda,\pi)-n}$ does not contract distances and $\mathcal{T}_1^{n_0}$ definitely expands them. Full details can be found in [21, Lemma 3.4]. □

Corollary 3.3. *The operator $\mathcal{L}$ preserves the space of Hölder functions, i.e., $\mathcal{L} : C^\beta(B) \to C^\beta(B)$ is well defined.*

Many of the statistical results for $\mathcal{T}_2$ are related to the existence of a spectral gap for $\mathcal{L}$. In the case of the operator acting on analytic functions is essentially automatic since the operator is compact (as we will see later). However, in the present context of Hölder continuous functions it remains true.

Lemma 3.4. *The value 1 is a simple eigenvalue with a positive eigenfunction $\rho > 0$. The rest of the spectrum is contained in a disk of radius τ strictly smaller than 1.*

Proof. The proof follows a classical approach [22]. Given $g \in C^\beta(B)$, we can estimate for each $x \in B$ that

$$|(\mathcal{L}^n g)(x)| \leq \|g\|_\infty \left(\sum_{\mathcal{T}_2^n y = x} \frac{1}{\mathrm{Jac}(\mathcal{T}_2^n)(y)} \right) \leq D\|g\|_\infty$$

by part (3) of Lemma 3.2. Thus $\|\mathcal{L}^n g\|_\infty \leq D\|g\|_\infty$. Similarly, in the special case $g = 1$ we can see that $D^{-1} \leq \mathcal{L}^n(1)(x) \leq D$, for all $x \in B$.

Given $x_1, x_2 \in B$, assume that $y_i \in (\mathcal{T}_2^n)^{-1} x_i$ $(i = 1, 2)$ are chosen in the same inverse branch. With this convention, we write that

$$\begin{aligned} &(\mathcal{L}^n g)(x_1) - (\mathcal{L}^n g)(x_2) \\ &= \sum_{\mathcal{T}_2^n y_i = x_i} \left(\frac{1}{\mathrm{Jac}(\mathcal{T}_2^n)(y_1)} - \frac{1}{\mathrm{Jac}(\mathcal{T}_2^n)(y_2)} \right) g(y_1) + \sum_{\mathcal{T}_2^n y_2 = x_2} \frac{(g(y_1) - g(y_2))}{\mathrm{Jac}(\mathcal{T}_2^n)(y_2)}. \end{aligned}$$

Note that by part (3) of Lemma 3.2, we have

$$D^{-2} \leq \frac{\mathrm{Jac}(\mathcal{T}_2^n)(y_1)}{\mathrm{Jac}(\mathcal{T}_2^n)(y_2)} \leq D^2,$$

and hence, we can write

$$\begin{aligned} \left| \frac{1}{\mathrm{Jac}(\mathcal{T}_2^n)(y_1)} - \frac{1}{\mathrm{Jac}(\mathcal{T}_2^n)(y_2)} \right| &\leq \frac{D^2}{\mathrm{Jac}(\mathcal{T}_2^n)(y_2)} \left| \log \left(\frac{\mathrm{Jac}(\mathcal{T}_2^n)(y_2)}{\mathrm{Jac}(\mathcal{T}_2^n)(y_1)} \right) \right| \\ &\leq \frac{D^2 C}{\mathrm{Jac}(\mathcal{T}_2^n)(y_2)} \|x_1 - x_2\|. \end{aligned}$$

Thus we can bound

$$\begin{aligned} &|(\mathcal{L}^n g)(x_1) - (\mathcal{L}^n g)(x_2)| \\ &\leq \sum_{\mathcal{T}_2^n y_2 = x_2} \frac{1}{|\mathrm{Jac}(\mathcal{T}_2^n)(y_2)|} \left(D^2 C \|g\|_\infty + \Theta^{-n} \|g\|_\beta \right) \|x_1 - x_2\| \\ &\leq D \left(D^2 C \|g\|_\infty + \frac{|g|_\beta}{\Theta^n} \right) \|x_1 - x_2\|. \end{aligned}$$

(This gives the well known Doeblin-Fortet, Marinescu-Tulcea or Lasota-Yorke inequality for $\mathcal{L}$: there exists $C > 0$ such that $\|\mathcal{L}^n g\|_\beta \leq C \left(\|g\|_\infty + \Theta^{-n} \|g\|_\beta \right)$ for all $n \geq 0$ and all $g \in C^\beta(B)$.)

In particular, the family $\{\frac{1}{N} \sum_{n=0}^{N-1} \mathcal{L}^n 1\}_{N=1}^\infty$ is equicontinuous and bounded, and thus has a uniform accumulation point $\rho \in C^\beta(B)$, say, where $D^{-1} \leq \rho(x) \leq D$, for all $x \in B$. Clearly, $\mathcal{L}\rho = \rho$ is a positive eigenfunction for the eigenvalue 1.

Let $d\mu_2(x) = \rho(x)d\omega(x)$ be the corresponding invariant probability measure. To see that 1 is a simple eigenvalue, assume that $\mathcal{L}\rho' = \rho'$, and then choose the largest $\epsilon > 0$ that the eigenfunction $\rho_\epsilon := \rho + \epsilon\rho' \geq 0$. Since we can find $x \in B$ with $\rho_\epsilon(x) = 0$, it then follows from $\mathcal{L}\rho_\epsilon = \rho_\epsilon$ that $\rho_\epsilon(y) = 0$, for all $y \in \mathcal{T}_2^{-1}x$. Proceeding inductively, we see that $\rho_\epsilon(y)$ vanishes on the dense set $y \in \cup_{n=0}^{\infty}\mathcal{T}_2^{-n}x$, and thus $\rho' = \epsilon\rho$, i.e., 1 is a simple eigenvalue. We can define $\widehat{\mathcal{L}} : C^\beta(B) \to C^\beta(B)$ by

$$\widehat{\mathcal{L}}w(x) = \frac{1}{\rho(x)}\mathcal{L}(w\rho)(x).$$

Then $\widehat{\mathcal{L}}1 = 1$ (and $\widehat{\mathcal{L}}^*\mu_2 = \mu_2$) and again the Doeblin-Fortet inequality holds for $\widehat{\mathcal{L}}$, i.e., $\|\widehat{\mathcal{L}}^n w\|_\beta \leq C\|w\|_\infty + \Theta^{-n}\|w\|_\beta$. Moreover, since for any positive $w \in C^\beta(B)$ we have $\sup w \geq \sup \widehat{\mathcal{L}}w \geq \sup \widehat{\mathcal{L}}^2 w \geq \cdots$ we can deduce from the equicontinuity that there is a unique limit in the uniform norm which, using that $\widehat{\mathcal{L}}1 = 1$, we conclude must be the constant $\int w d\mu_2$, i.e., $\widehat{\mathcal{L}}^n w \to \int w d\mu_2$ as $n \to +\infty$, see [22, Theorem 2.2].

Finally, to show that the rest of the spectrum of $\mathcal{L}$ is contained strictly within the unit disc it suffices to show the same for $\widehat{\mathcal{L}}$ and, more particularly, $\widehat{\mathcal{L}} : C^\beta(B)/\mathbb{C} \to C^\beta(B)/\mathbb{C}$ has spectral radius strictly smaller than 1. However, the convergence of $\widehat{\mathcal{L}}^n w$ implies that $\|\widehat{\mathcal{L}}^n w + \mathbb{C}\|_\infty \to 0$ as $n \to +\infty$ and thus two applications of the Marinescu-Tulcea inequality gives

$$\begin{aligned}\|\widehat{\mathcal{L}}^{2n}w\|_\beta &\leq C\left(\|\widehat{\mathcal{L}}^n w + \mathbb{C}\|_\infty + \Theta^{-n}\|\widehat{\mathcal{L}}^n w\|_\beta\right) + \Theta^{-n}\|\widehat{\mathcal{L}}^n w\|_\beta \\ &\leq C\left(\|\widehat{\mathcal{L}}^n w + \mathbb{C}\|_\infty + \Theta^{-n}(C+1)\left(C\|w\|_\infty + \|w\|_\beta\Theta^{-n}\right)\right) \\ &< 1\end{aligned}$$

for large enough $n \geq 0$, uniformly on the unit ball of $C^\beta(B)/\mathbb{C}$. The result follows from the spectral radius theorem. □

As usual, the probability measure μ_2 which is the eigenprojection associated to 1 (i.e., $\widehat{\mathcal{L}}\mu_2 = \mu_2$) is the unique absolutely continuous $\mathcal{T}_2$-invariant probability measure on B. In particular, μ_2 is the renormalized restriction of μ_1 to B.

Corollary 3.5. *(1) The transformation $\mathcal{T}_2 : B \to B$ is exponentially mixing on Hölder functions, i.e., there exists $0 < \tau < 1$ and $C > 0$ such that for all $F \in L^\infty(B)$ and $G \in C^\beta(B)$ with $\int F d\mu_2 = \int G d\mu_2 = 0$,*

$$\left|\int F \circ \mathcal{T}_2^n . G d\mu_2 - \int F d\mu_2 \int G d\mu_2\right| \leq C\tau^n \|F\|_{L^1(\mu_2)}\|G\|_\beta \text{ for all } n \geq 0.$$

(2) For μ_2-almost all $x = (\lambda, \pi) \in B$ we have that

$$\frac{1}{N}\sum_{n=0}^{N-1} F({\mathcal{T}_2}^n(x,\lambda)) = \int F d\mu_2 + O\left(\frac{\log N}{\sqrt{N}}\right).$$

Proof. For the first part, we can write

$$\int F \circ \mathcal{T}_2^n . G d\mu_2 - \int F d\mu_2 \int G d\mu_2 = \int \left(\mathcal{L}^n(G\rho) - \left(\int G d\mu_2\right)\rho\right) F\, d\omega.$$

Thus, $\left|\int F \circ \mathcal{T}_2^n . G d\mu_2 \int F d\mu_2 \int G d\mu_2\right| \leq \|\mathcal{L}^n(G\rho) - \left(\int G d\mu_2\right)\rho\|_\infty \|F\|_{L^1(\omega)}$. By Lemma 3.4, $\|\mathcal{L}^n(G\rho) - \left(\int G d\mu_2\right)\rho\|_\infty \leq C\tau^n\|G\|_\beta$, since $C^\beta(B)$ embeds into

$L^\infty(B)$, is a Banach algebra and $\rho \in C^\beta(B)$. On the other hand, $\|F\|_{L^1(\omega)} \leq c^{-1}\|F\|_{L^1(\mu_2)}$, where $c = \inf \rho$ is strictly positive.

The second part follows immediately from the first part by a standard spectral result [10]. □

4. Statistical properties for $\mathcal{T}_2$

Let $d\mu_2(x) = \rho(x)d\omega(x)$ be the unique absolutely $\mathcal{T}_2$-invariant probability measure on B given by Proposition 2.4. This measure μ_2 is ergodic (cf. [4] or, alternatively, by part (1) of Corollary 3.5) and so we can apply the Birkhoff ergodic theorem which gives that for any $f \in L^1(X, \mu_2)$ and for μ_2-a.e. $x \in B$ we have that

$$\frac{1}{n}\sum_{j=0}^{n-1} f(\mathcal{T}_2^j x) \to \int f d\mu_2, \text{ as } n \to +\infty,$$

pointwise and in L^1. In this section we want to discuss various generalizations of this basic property.

4.1. The Central Limit Theorem and Functional Central Limit Theorem.

A classical result for expanding dynamical systems is the Central Limit Theorem, and the stronger Functional Central Limit Theorem.

Definition 4.1. *We say that $\mathcal{T}_2$ satisfies the Functional Central Limit Theorem whenever for a Hölder continuous function $h \in C^\beta(B,\mathbb{R})$ with $\int h d\mu_2 = 0$ (not equal to a coboundary) there exists $\sigma > 0$ such that for $0 \leq t \leq 1$,*

$$w_n(t) = \frac{1}{\sigma\sqrt{n}}\left(\sum_{j=0}^{[nt]-1} h \circ \mathcal{T}_2^j + (nt - [nt])h \circ \mathcal{T}_2^{[nt]}\right)$$

converges weakly to the Wiener measure on $C([0,1],\mathbb{R})$.

This is sometimes called a weak invariance principle, in reference to the topology of covergence.

The Central Limit Theorem could be deduced directly from the spectral results on $\widehat{\mathcal{L}}$ in the previous section, but, with no additional work we can deduce the stronger Functional Central Limit Theorem.

Proposition 4.2. *The Functional Central Limit Theorem holds for $\mathcal{T}_2$.*

Proof. By a quite general result of Mackey and Tyran-Kamiǹska [13, 14] (cf. also [28]) if $h_0 \in L^2(B, \mu_2)$ satisfies $\int h_0 d\mu_2 = 0$ and $\widehat{\mathcal{L}}h_0 = 0$, and

$$\sum_{n=1}^{\infty} \frac{1}{n^{3/2}} \sqrt{\int \left(\sum_{k=0}^{n-1} \widehat{\mathcal{L}}^k h_0\right)^2 d\mu_2} < \infty,$$

then setting $\sigma^2 = \int |h_0|^2 d\mu_2$ gives

$$w_n^0(t) = \frac{1}{\sqrt{n}} \sum_{j=0}^{[nt]-1} h_0 \circ \mathcal{T}_2^j \to \sigma w(t), \text{ for } t \in [0,1].$$

(i.e., the Functional Central Limit Theorem for h_0). More generally, given a Hölder continuous function h with $\int h d\mu_2 = 0$, we recall from Lemma 3.4 that there exists $0 < \tau < 1$ such that $\|\widehat{\mathcal{L}}^n h\|_\beta = O(\theta^n)$, and therefore $u = \sum_{n=1}^\infty \widehat{\mathcal{L}}^n h$ converges in

$C^\beta(B)$. Let $u = \sum_{n=1}^{\infty} \widehat{\mathcal{L}}^n h$ and set $h_0 := h - u \circ T + u$ then $\widehat{\mathcal{L}}(h_0) = \widehat{\mathcal{L}}h - u + \widehat{\mathcal{L}}u = 0$. Since h and h_0 are cohomologous we can bound $|w_n(t) - w_n^0(t)| \leq 2\|u\|_\infty/\sqrt{n}$ and thus deduce the Functional Central Limit Theorem for h. If $\sigma^2 = 0$, then we would have $h_0 \equiv 0$, and so h would be equal to a coboundary, which is not the case by assumption. □

The following are standard corollaries for Hölder continuous functions f using the Continuous Mapping Theorem [8, 9] beginning with the central limit theorem.

Corollary 4.3 (Central Limit Theorem)**.** *For $y \in \mathbb{R}$ we have that*

$$\lim_{n\to+\infty} \mu_2 \left\{ x \in B : \frac{1}{\sqrt{n}} \sum_{j=1}^{n} f(\mathcal{T}_2^j x) \leq y \right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt$$

The Central Limit Theorem (and much more besides) has already been proved by Butetov [4] and Morita [21]. The approach of Bufetov involved studying the rate of mixing of $\mathcal{T}_2$; and the method of Morita involved perturbation theory of the transfer operator.

The following are other standard corollaries [8, 9].

Corollary 4.4. *For $y \geq 0$ we have that*

$$\lim_{n\to+\infty} \mu_2 \left\{ x \in B : \frac{1}{\sqrt{n}} \max_{1\leq k\leq n} \sum_{j=1}^{k} f(\mathcal{T}_2 x) \leq y \right\} = \frac{\sqrt{2}}{\sqrt{\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt - 1.$$

Corollary 4.5 (Arcsine Law)**.** *For $0 \leq y \leq 1$ we have that*

$$\lim_{n\to+\infty} \mu_2 \left\{ x \in B : \frac{N_n(x)}{n} \leq y \right\} = \frac{2}{\sqrt{\pi}} \sin^{-1} \sqrt{y}$$

where $N_n(x) = Card\left\{1 \leq k \leq n \ : \sum_{j=1}^{k} f(\mathcal{T}_2^j x) > 0\right\}$.

Corollary 4.6 (Law of the iterated logarithm)**.** *For μ_2-a.e. $x \in B$ we have*

$$\limsup_{n\to+\infty} \frac{\sum_{j=1}^{n} f(\mathcal{T}_2^j x)}{\sigma\sqrt{2n \log\log n}} = 1.$$

Remark 4.7. *There are a number of other statistical results which could be considered. For example, Morita has shown that there is a local limit theorem and Berry-Esseen estimates for $\mathcal{T}_2$. We could also consider Edgeworth expansions, following Fernando and Liverani* [6].

4.2. **Almost Sure Invariance Principles.** With only a little further work, we next establish a class of stronger results, from which the preceeding (and several others) can easily be deduced.

Given a Hölder continuous function $f : B \to \mathbb{R}$ with $\int f d\mu_2 = 0$ we can associate the summation $f^n(x) := \sum_{i=0}^{n-1} f(\mathcal{T}^i x)$, for each $n \geq 1$.

Definition 4.8. *We say that $\mathcal{T}_2 : B \to B$ satisfies the Almost Sure Invariance Principle relative to Hölder continuous functions and the measure μ_2 if for any such function $f : B \to \mathbb{R}$ with $\int f d\mu_2 = 0$ not equal to a coboundary, there exists a sequence of random variables $\{S_n\}$, possibly on a larger probability space, equal in distribution under μ_2 with $\{f^n\}$ and there exists $\epsilon > 0$ such that $S_n = W_n + O(n^{\frac{1}{2}-\epsilon})$ as $n \to +\infty$, where $\{W_t\}_{t\geq 0}$ is a Brownian motion with variance $\sigma^2 > 0$.*

The following result is a strengthening of Proposition 4.2.

Theorem 4.9 (Almost sure invariance principle for $\mathcal{T}_2$)**.** *The transformation $\mathcal{T}_2 : B \to B$ satisfies the Almost Sure Invariance Principle.*

Proof. The standard approach is to deduce this from an application of a result of Philipp and Stout [23] (cf. [19] for a dynamical reformulation). In particular, we only need to establish that the hypotheses there hold. More precisely, given a β-Hölder function $f : B \to \mathbb{R}$ with $\int f d\mu_2 = 0$ we observe that:

(1) $f \in L^{2+\delta}(B)$, for any $\delta > 0$ (since v is automatically bounded);

(2) for any $n \geq 1$,
$$\int |f^n|^2 d\mu_2 = n\sigma^2 + O(1)$$
(by expanding the Left Hand Side and bounding the cross terms using Part (1) of Corollary 3.5), see [19, Proof of Corollary 2.3] for more details;

(3) for any $k \geq 0$,
$$\begin{aligned} E\left(|f - E(f| \vee_{i=0}^{k-1} \mathcal{T}_2^{-i}\mathcal{Q})|^{2+\delta})| \vee_{i=0}^{k-1} \mathcal{T}_2^{-i}\mathcal{Q}\right) &\leq \|f - E(f| \vee_{i=0}^{k-1} \mathcal{T}_2^{-i}\mathcal{Q}))\|_\infty^{2+\delta} \\ &\leq (\|f\|_\beta \sup_{a \in \mathcal{Q}_k} \operatorname{diam}(a))^{2+\delta} \\ &\leq (\|f\|_\beta \Theta^{-k})^{2+\delta}, \end{aligned}$$
(where, as usual, $E(\cdot| \vee_{i=0}^{k-1} \mathcal{T}_2^{-i}\mathcal{Q}) = \sum_{a \in \mathcal{T}_2^{-i}\mathcal{Q}} \frac{1}{\mu(a)} \int_a (\cdot) d\mu)$; and, finally,

(4) given any $A_1 \in \vee_{i=0}^{k-1}\mathcal{T}_2^{-i}\mathcal{Q}$ and any Borel measurable set $A_2 \subset B$, and for any $n, k \geq 0$, we can bound
$$\begin{aligned} &\left|\mu_2(A_1 \cap \mathcal{T}_2^{-(k+n)}A_2) - \mu_2(A_1)\mu_2(A_2)\right| \\ &= \left|\int \chi_{A_1}(\chi_{A_2} \circ \mathcal{T}_2^{k+n})d\mu_2 - \int \chi_{A_1} d\mu_2 . \int \chi_{A_2} d\mu_2\right| \\ &= \left|\int (\widehat{\mathcal{L}}^n \chi_{A_1})(\chi_{A_2} \circ \mathcal{T}_2^k)d\mu_2 - \int \widehat{\mathcal{L}}^n \chi_{A_1} d\mu_2 \int \chi_{A_2} \circ \mathcal{T}_2^k d\mu_2\right| \\ &= \left|\int \left[\widehat{\mathcal{L}}^k \chi_{A_1} - \int \widehat{\mathcal{L}}^k \chi_{A_1} d\mu_2\right] (\chi_{A_2} \circ \mathcal{T}_2^n) d\mu_2\right| \\ &= \left|\int \widehat{\mathcal{L}}^n \left[\widehat{\mathcal{L}}^k \chi_{A_1} - \int \widehat{\mathcal{L}}^k \chi_{A_1} d\mu_2\right] \chi_{A_2} d\mu_2\right| \\ &\leq \left(\int \left|\widehat{\mathcal{L}}^n \left[\widehat{\mathcal{L}}^k \chi_{A_1} - \int \widehat{\mathcal{L}}^k \chi_{A_1} d\mu_2\right]\right|^2 d\mu_2\right)^{\frac{1}{2}} \left(\int \chi_{A_2}^2 d\mu_2\right)^{\frac{1}{2}} \\ &\leq C\tau^n \|\widehat{\mathcal{L}}^k \chi_{A_1}\|_\beta \mu_2(A_2)^{\frac{1}{2}}, \end{aligned}$$
for some $C > 0$, using the Cauchy-Schwartz inequality, that $\widehat{\mathcal{L}}^* \mu_2 = \mu_2$ and (again) that $0 < \tau < 1$ is a bound on the modulus of the second eigenvalue of $\widehat{\mathcal{L}}$. Finally, we can observe that $\|\widehat{\mathcal{L}}^k \chi_{A_1}\|_\beta \leq D\mu(A_1)$, as in the proof of [19, Lemma 2.4], and so the bound can be taken to be $C\tau^n$.

We can then apply Theorem 7.1 in [23] (cf. Theorem A.1 in [19]) to deduce that the Almost Sure Invariance Principle holds for $\mathcal{T}_2$. □

There is an immediate application of the preceding analysis to return times for $\mathcal{T}_2$. Given any Borel set A we denote by $r_A : A \to \mathbb{N}$ the first return time to A, i.e., $r_A(x) = \inf\{n \geq 1 : \mathcal{T}_1^n x \in A\}$. In particular, the value defined inductively by $r_A^{(n)}(x) = r_A^{(n-1)}(x) + r_A(\mathcal{T}_2^{r_A^{(n-1)}(x)})$ is the nth return time. Using Birkhoff's theorem and Kac's theorem on return times we have that

$$\lim_{n\to+\infty} \frac{r_A^{(n)}(x)}{n} = \frac{1}{\mu_1(A)} \text{ for } \mu\text{-a.e. } x \in B.$$

For the particular choice $A = B$ we can consider the function $r_B(x) = \widehat{n}(x)$ and by Kac's theorem $\int r_B d\mu_2 = 1/\mu_1(B)$. It is easy to see that the variance is non-zero and thus this leads, for example, to the following corollary:

Corollary 4.10. *There exists $\sigma > 0$ such that*

$$\lim_{N\to+\infty} \mu_2 \left\{ x : \frac{1}{N} r_B^{(N)}(x) - \frac{1}{\mu_1(B)} \leq y \right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt$$

for $y \in \mathbb{R}$.

Remark 4.11. *Finer results about recurrence properties and the statistical behavior of return times for $\mathcal{T}_1$ and $\mathcal{T}_2$ can also be deduced from the spectral gap (Lemma 3.4), see Aimino, Nicol and Todd [1].*

5. Statistical properties for $\mathcal{T}_1$

The statistical properties of $\mathcal{T}_2$ described above can be used to establish analogous results for the original Zorich map $\mathcal{T}_1 : \Delta\times\mathcal{R} \to \Delta\times\mathcal{R}$, with respect to μ_1, by viewing it as a suspension. More precisely, we can associate to the map $\mathcal{T}_2 : B \to B$ and the return time $\widehat{n} : B \to \mathbb{Z}^+$ a suspension space

$$B^{\widehat{n}} := \{(x, k) \in B \times \mathbb{Z} : 0 \leq k \leq \widehat{n}(x) - 1\}/\sim$$

where we identify $(\lambda, \pi; \widehat{n}(x))$ and $(\mathcal{T}_2(\lambda, \pi); 0)$. We can also define the natural map $\mathcal{T}_2^{\widehat{n}} : B^{\widehat{n}} \to B^{\widehat{n}}$ on this suspension space by

$$\mathcal{T}_2^{\widehat{n}}(x, k) = \begin{cases} (x, k+1) & \text{if } 0 \leq k \leq \widehat{n}(x) - 2 \\ (\mathcal{T}_2 x, 0) & \text{if } k = \widehat{n}(x) - 1. \end{cases}$$

There is a natural $\mathcal{T}_2^{\widehat{n}}$-invariant measure $d\mu_2 \times d\mathbb{N}/\int \widehat{n} d\mu_2$, where $d\mathbb{N}$ corresponds to the usual counting measure. The following result is standard.

Lemma 5.1. *The map $\Psi : B^{\widehat{n}} \to \Delta \times \mathcal{R}$ defined by $\Psi(x, k) = {\mathcal{T}_1}^k(x)$ is:*

1. *a semi-conjugacy, i.e., $\mathcal{T}_1 \circ \Psi = \Psi \circ \mathcal{T}_2^{\widehat{n}}$, and*
2. *an isomorphism (with respect to $d\mu_2 \times d\mathbb{N}/\int \widehat{n} d\mu_2$ and $d\mu_1$).*

We can deduce the almost sure invariance principle for the Zorich map $\mathcal{T}_1 : \Delta \times \mathcal{R} \to \Delta \times \mathcal{R}$, by applying a result given in a paper of Melbourne and Nicol [19] (which is formulated from the results of Melbourne and Török [18]), and whose proof is made precise by Korepanov [11]. The other statistical properties follow as a direct consequence.

The main technical condition we require is the following:

Lemma 5.2. *For any $\delta > 0$ we have that*

$$\sum_{k=1}^{\infty} \mu_2 \{x = (\lambda, \pi) \in B : \widehat{n}(x) = k\} k^{2+\delta} < +\infty.$$

Proof. By an estimate of Avila-Bufetov [2, Lemma 1], there exists $C > 0$ and $0 < \theta < 1$ such

$$\mu_2 \{x \in B : \widehat{n}(\underline{\lambda}, \pi) \geq k\} \leq C\theta^k, \text{ for all } k \geq 1.$$

Thus $\sum_{k=1}^{\infty} \mu_2 \{x \in B : \widehat{n}(\underline{\lambda}, \pi) = k\} k^{2+\delta} \leq C \sum_{k=1}^{\infty} \theta^k k^{2+\delta} < +\infty$. □

We now describe a general class of function for which the results will be established. Let $f : \Delta \times \mathcal{R} \to \mathbb{R}$ be Hölder continuous and satisfy $\int f d\mu_1 = 0$. We can associate to f a function $\overline{f} : B \to \mathbb{R}$ defined μ_2-a.e. by

$$\overline{f}(x) = \sum_{l=0}^{\widehat{n}(x)-1} f(\mathcal{T}_1^l x).$$

In particular, we have that $\int \overline{f} d\mu_2 = 0$. A key property is that Birkhoff sums of $\overline{f}$ with respect to $\mathcal{T}_2$ constitute a subsequence of Birkhoff sums of f with respect to $\mathcal{T}_1$. Thus, to obtain statistical properties for the latter, it is enough to prove them for the former, and to have some control on the gaps between two consecutive terms of the subsequence. This is the approach followed in [18, 11]. If, in the interests of expediency, we make the hypothesis that the function $\overline{f} : B \to \mathbb{R}$ is Hölder continuous, then we can lift the results for $\mathcal{T}_2$ in Theorem 4.9 (with respect to $\overline{f}$) to those for $\mathcal{T}_1$ (with respect to f). More generally, we can assume that f is Hölder continuous and the associated function $\overline{f}$ satisfies a weaker "local Hölder" condition that if x and y belong to the same element of $\mathcal{Q}$ with $\widehat{n}(x) = \widehat{n}(y) = n$, say, then $|\overline{f}(x) - \overline{f}(y)| \lesssim n\|\overline{f}\|_\beta \|x - y\|^\beta$. However, following [19] we can then consider the slightly larger Banach space $\mathcal{B}$ with respect to the norm

$$\|h\|_{\mathcal{B}} = \sup_{A \in \mathcal{Q}} \sup_{x \in A} \frac{|f(x)|}{\widehat{n}(A)} + \sup_{A \in Q} \sup_{\substack{x,y \in A \\ x \neq y}} \frac{1}{\widehat{n}(A)} \frac{|h(x) - h(y)|}{\|x - y\|^\beta},$$

for which the proofs of Lemma 3.4 and Theorem 4.9 readily generalize.

To extend the almost sure invariance principle from $\mathcal{T}_2$ to $\mathcal{T}_1$ we need first to check the hypotheses of the theorem of Melbourne and Török [18]. This will prove the almost sure invariance principle for $\mathcal{T}_1$ and the renormalized restriction of μ_1 to B, and we can then use the result of Korepanov [11, Theorem 3.7] to conclude the results for $(\mathcal{T}_1, \mu_1)$. In particular,

(1) by the Lemma 5.2, we can choose $\delta > 0$ so that $\widehat{n} \in L^{2+\delta}(B, \mu_2)$, and
(2) by the analogue of part (2) of Corollary 3.5 we have that

$$\frac{1}{N} \sum_{i=0}^{N-1} \widehat{n}({\mathcal{T}_2}^i x) = \int \widehat{n} d\mu + O\left(\frac{1}{N^{1-\epsilon}}\right), \ \mu_2\text{-a.e. } x \in B.$$

In particular, we can now conclude that the almost sure invariance principle holds for $\mathcal{T}_1$ with variance $\widehat{\sigma}^2 = \sigma^2 / \int \widehat{n} d\mu_2$.

Theorem 5.3 (Almost sure invariance principle for $\mathcal{T}_1$)**.** *The almost sure invariance principle holds for $\mathcal{T}_1$ and μ_1.*

Remark 5.4. *It can be interesting to precise the error rates in the almost sure invariance principle above. Even if the result of Philipp and Stout [23] used to prove Theorem 4.9 does not provide very insightful bounds, it is possible, using different methods, to prove that, for $\mathcal{T}_2$ and $\mathcal{T}_1$, we have $S_n = W_n + o(n^\lambda)$ for every $\lambda > 0$, see Korepanov [12].*

This theorem has several consequences for Hölder continuous functions f, including the the analogues of Proposition 4.2 and Corollaries 4.3 - 4.6. for $\mathcal{T}_1$.

More precisely, we have the following results.

Proposition 5.5. *The Functional Central Limit Theorem holds for $\mathcal{T}_1$.*

This completes the proof of Theorem 1.1.

Corollary 5.6 (Central Limit Theorem)**.** *For $y \in \mathbb{R}$ we have that*

$$\lim_{n\to+\infty} \mu_1 \left\{ x \in B : \frac{1}{\sqrt{n}} \sum_{j=1}^{n} f(\mathcal{T}_1^j x) \le y \right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt$$

Corollary 5.7. *For $y \ge 0$ we have that*

$$\lim_{n\to+\infty} \mu_1 \left\{ x \in B : \frac{1}{\sqrt{n}} \max_{1\le k\le n} \sum_{j=1}^{k} f(\mathcal{T}_1^j x) \le y \right\} = \frac{\sqrt{2}}{\sqrt{\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt - 1$$

Corollary 5.8 (Arcsine Law)**.** *For $0 \le y \le 1$ we have that*

$$\lim_{n\to+\infty} \mu_1 \left\{ x \in B : \frac{N_n(x)}{n} \le y \right\} = \frac{2}{\sqrt{\pi}} \sin^{-1} \sqrt{y}$$

where $N_n(x) = Card \left\{ 1 \le k \le n \ : \ \sum_{j=1}^{k} f(\mathcal{T}_1^j x) > 0 \right\}$.

Corollary 5.9 (Law of the iterated logarithm)**.** *For μ-a.e. $x \in B$ we have*

$$\limsup_{n\to+\infty} \frac{\sum_{j=1}^{n} f(\mathcal{T}_1^j x)}{\sigma\sqrt{2n \log\log n}} = 1.$$

From the structure of the map $\mathcal{T}_1$, one can deduce many other interesting statistical properties. For instance, using Lemmata 3.2 and 5.2, we can obtain a local large deviations principle, thanks to Melbourne and Nicol [20, Theorem 2.1] (see also Rey-Bellet and Young [24, Theorem B]):

Theorem 5.10 (Local large deviations principle for $\mathcal{T}_1$)**.** *For any Hölder continuous function $f : \Delta \times \mathcal{R} \to \mathbb{R}$ not equal to a coboundary such that $\int f d\mu_1 = 0$, there exists $\epsilon_0 > 0$ and a rate function $c : (-\epsilon_0, \epsilon_0) \to \mathbb{R}$ continuous, strictly convex, vanishing only at 0, such that for every $0 < \epsilon < \epsilon_0$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mu_1(f^n > n\epsilon) = -c(\epsilon).$$

6. Transfer operators and analytic functions

To take advantage of the transformation $\mathcal{T}_2$ being piecewise analytic, we can also consider the transfer operator acting on a space of analytic functions. This will prove useful in the proof of Theorem 1.2. Let us denote $\underline{\lambda} = (\lambda_1, \cdots, \lambda_m), \xi = (\xi_1, \cdots, \xi_m) \in \mathbb{R}^m$. For sufficiently small $\epsilon > 0$ we denote by

$$B_\epsilon^{\mathbb{R}} = \left\{ \underline{\lambda} \in \mathbb{R}^m : \sum_{j=1}^{m} \lambda_j = 1 \text{ and } |\lambda - B| < \epsilon \right\}$$

an ϵ-neighbourhood of B in the (hyperplane containing the) simplex and consider a simple complexification of the form

$$B_\epsilon^{\mathbb{C}} = \left\{ \underline{\lambda} + i\underline{\xi} \in \mathbb{C}^m \,:\, |\underline{\lambda} - B| < \epsilon, \sum_{j=1}^m \lambda_j = 1, \sum_{j=1}^m \xi_j = 0 \text{ and } |\xi_j| \le \epsilon \right\}.$$

Let $\mathcal{T}_2 : B_\epsilon^{\mathbb{C}} \to \mathbb{C}^n$ also denote the analytic extension from B to $B_{\mathbb{C}}$ provided $\epsilon > 0$ is sufficiently small.

In order to show that $\mathcal{L}$ preserves a space of analytic functions on this space we can use the following simple lemma.

Lemma 6.1. *Providing $\epsilon > 0$ is sufficiently small we have that $\overline{\mathcal{T}_2^{-1} B_\epsilon^{\mathbb{C}}} \subset \mathrm{int}(B_\epsilon^{\mathbb{C}})$. Moreover, for $x = \underline{\lambda} + i\underline{\xi} \in B_\epsilon^{\mathbb{C}}$ we have that*

$$\sup_{x \in B_\epsilon^{\mathbb{C}}} \left| \sum_{\mathcal{T}_2^{-1} y = x} \frac{1}{(\sum_i (Ay)_i)^m} \right| < +\infty$$

Proof. Since the inverse branches of $\mathcal{T}_2 : B \to B$ are uniformly contracting, we can choose $\epsilon > 0$ sufficiently small and $0 < \theta < 1$ such that $\mathcal{T}_2^{-1} B_\epsilon^{\mathbb{R}} \subset B_{\theta\epsilon}^{\mathbb{R}}$. We can show that their complexifications have a similar property with respect to $B_{\mathbb{C}}$. To begin, observe that the linear action of any of the positive matrices A corresponding to an inverse branch of $\mathcal{T}_2$ act on both the real and imaginary coordinates independently, and the complexification of the linear action is again a linear action:

$$(\lambda_1, \cdots, \lambda_m) + i(\xi_1, \cdots, \xi_m) \mapsto A(\lambda_1, \cdots, \lambda_m) + iA(\xi_1, \cdots, \xi_m).$$

The image under the projective action comes from dividing by $\sum_j (A\underline{\lambda})_j + i \sum_j (A\underline{\xi})_j$ (i.e., the complexification of $\|A\underline{\lambda}\|$) to get:

$$\frac{A\underline{\lambda} + iA\underline{\xi}}{\sum_j (A\underline{\lambda})_j + i \sum_j (A\underline{\xi})_j} = \frac{A\underline{\lambda}}{\sum_j (A\underline{\lambda})_j} - \left(\frac{A\underline{\lambda} \frac{(\sum_j (A\underline{\xi})_j)^2}{(\sum_j (A\underline{\lambda})_j)} - A\underline{\xi} (\sum_j (A\underline{\xi})_j)}{\left(\sum_j (A\underline{\lambda})_j\right) \left(\sum_j (A\underline{\lambda})_j + \frac{(\sum_j (A\underline{\xi})_j)^2}{(\sum_j (A\underline{\lambda})_j)} \right)} \right) + i \frac{\left(A\underline{\xi} - A\underline{\lambda} \frac{\sum_j (A\underline{\xi})_j}{\sum_j (A\underline{\lambda})_j} \right)}{\sum_j (A\underline{\lambda})_j + \frac{(\sum_j (A\underline{\xi})_j)^2}{(\sum_j (A\underline{\lambda})_j)}}.$$

In particular, for $\theta' = (1+\theta)/2$ and $\epsilon > 0$ sufficiently small we can deduce that $\mathcal{T}_2^{-1} B_\epsilon^{\mathbb{C}} \subset B_{\theta'\epsilon}^{\mathbb{C}}$. This completes the proof of the first part of the lemma.

For the second part of the lemma, we first observe that uniformly in $\underline{\lambda} + i\underline{\xi} \in B_\epsilon^{\mathbb{C}}$ we have

$$\begin{aligned} \frac{1}{(\sum_j (A\underline{\lambda})_j + i \sum_j (A\underline{\xi})_j)^m} &= \frac{1}{(\sum_j (A\underline{\lambda})_j)^m} \frac{1}{\left(1 + i \frac{(\sum_j (A\underline{\xi})_j)^m}{(\sum_j (A\underline{\lambda})_j)^m}\right)} \\ &= \left(\frac{1}{(\sum_j (A\underline{\lambda})_j)^m} \right) (1 + O(\epsilon)). \end{aligned} \tag{6.1}$$

However, from the formula of the transfer operator, we know that, as in the proof of Lemma 3.4, for $x \in B$,

$$\sup_{x \in B} \left| \sum_{\mathcal{T}_2^{-1} y = x} \frac{1}{\left(\sum_i (Ay)_i\right)^m} \right| < +\infty. \tag{6.2}$$

Comparing (6.1) and (6.2) completes the proof. □

We can consider the Banach space $H(B_\epsilon^{\mathbb{C}})$ of analytic functions $f : B_\epsilon^{\mathbb{C}} \to \mathbb{C}$ with a continuous extension to the closure of $B_\epsilon^{\mathbb{C}}$ endowed wiith supremum norm $\|f\| = \sup_{B_\epsilon^{\mathbb{C}}} |f(z)|$. We can apply Lemma 6.1 to deduce that the operator $\mathcal{L} : H(B_\epsilon^{\mathbb{C}}) \to H(B_\epsilon^{\mathbb{C}})$ is well defined. In particular, that the series expression for $\mathcal{L}w(x)$ converges to an analytic function for $x \in B_\epsilon^{\mathbb{C}}$ merely follows by complex differentiation under summation sign.

This leads to the following definition and result.

Definition 6.2. *Any bounded linear operator $L : B \to B$ on a Banach space B with norm $\|\cdot\|$ is called nuclear (of order α) if there exist:*

(i) vectors $u_n \in B$ (with $\|u_n\| = 1$);
(ii) bounded linear functionals $l_n \in B^*$ (with $\|l_n\| = 1$); and
(iii) a sequence (ρ_n) of complex numbers such that $\sum_{n=0}^{\infty} |\rho_n|^\alpha < +\infty$, with

$$L(v) = \sum_{n=0}^{\infty} \rho_n l_n(v) u_n, \qquad \text{for all } v \in B.$$

We say that L has order zero, if property holds for any $\alpha > 0$.

In particular, a nuclear operator is automatically a compact operator, for which the non-zero eigenvalues are of finite multiplicity (and the eigenspaces and dual spaces are of finite multiplicity).

Proposition 6.3. *The operator $\mathcal{L} : H(B_\epsilon^{\mathbb{C}}) \to H(B_\epsilon^{\mathbb{C}})$ is nuclear (of order zero).*

Proof. The proof follows the same lines as that in [16, 17], see also [25]. We denote by $C^\omega(B_\epsilon^{\mathbb{C}})$ the Fréchet space of analytic functions on $B_\epsilon^{\mathbb{C}}$, endowed with the compact-open topology. We observe that $\mathcal{L} : H(B_\epsilon^{\mathbb{C}}) \to C^\omega(B_\epsilon^{\mathbb{C}})$ is a bounded linear operator and recall that the space $C^\omega(B_\epsilon^{\mathbb{C}})$ is nuclear [7]. In particular, if we compose $\mathcal{L}$ with the continuous inclusion $H(B_\epsilon^{\mathbb{C}}) \hookrightarrow C^\omega(B_\epsilon^{\mathbb{C}})$, we conclude that the operator $\mathcal{L}$ is nuclear (or order zero) [7] (cf. [17], proof of Lemma 3). □

Many of the statistical results for $\mathcal{T}_2$ described in the previous sections are related to the existence of a spectral gap for $\mathcal{L}$. In the present analytic context this is essentially automatic since the operator is compact. Moreover, one can apply an approach of Mayer ([17], p.12) to recover that the value 1 is a simple eigenvalue of maximal modulus, and that eigenfunction ρ is real analytic.

We can recover the following:

Corollary 6.4. *The invariant density of $\mathcal{T}_2$ (and thus $\mathcal{T}_1$) is real analytic.*

Remark 6.5. *Zorich* [34, Theorem 1] *actually proved that the invariant density is, when restricted to a subset of the form Δ_π^+ or Δ_π^-, a function which is rational, positive and homogeneous of degree $-m$ on $\mathbb{R}^m$.*

We can again define $\widehat{\mathcal{L}} : C^\omega(B) \to C^\omega(B)$ by

$$\widehat{\mathcal{L}}w(x) = \frac{1}{\rho(x)}\mathcal{L}(w\rho)(x).$$

then $\widehat{\mathcal{L}}1 = 1$ and $\widehat{\mathcal{L}}^*\mu = \mu$.

7. Zeta functions and Lyapunov Exponents

We now turn to the proof of Theorem 1.2. Recall that we can write the sum Λ of the Lyapunov exponents of $\mathcal{T}_2$ as

$$\Lambda = \int_B \log|\det D\mathcal{T}_2(x)| d\mu_2(x).$$

We shall describe an approach to the Lyapunov exponents using complex functions. The connection between zeta functions and both the standard and multi-dimensional continued fraction transformations was explored by Mayer in [16] (cf. also [17]). We also refer the reader to the monograph of Baladi [3] for an account of the theory of dynamical zeta functions and determinants for hyperbolic maps.

Definition 7.1. *We can associate to $\mathcal{T}_2$ a complex function $d(z,s)$ in two variables defined by*

$$d(z,s) = \exp\left(-\sum_{n=1}^{\infty}\frac{z^n}{n}\sum_{\mathcal{T}_2^n x = x}|\det(D\mathcal{T}_2^n)(x)|^{-s}\right)$$

where we interpret the periodic points as points in the disjoint union. This converges for $|z|$ and $Re(s)$ sufficiently small.

The function $d(z,s)$ can be viewed as the reciprocal of a zeta function (in the sense of Ruelle).

The main technical result on such functions is the following.

Proposition 7.2.
1. *If $|s|$ is sufficiently small, then $d(z,s)$ is an entire function in z;*
2. *Moreover, if we expand $d(z,s) = 1 + \sum_{n=1}^{\infty} a_n(s)z^n$, then there exists $c > 0$ such that $|a_n| = O(e^{-cn^{1+1/(m-1)}})$;*
3. *The zeros z_0 for $d(z,1)$ correspond to eigenvalues $\lambda = 1/z_0$. In particular, 1 is the zero of smallest modulus; and*
4. *We can write*

$$\frac{\frac{\partial d(1,s)}{\partial s}|_{s=1}}{\frac{\partial d(z,1)}{\partial z}|_{z=1}} = \int \log|\det(D\mathcal{T}_2)(x)| d\mu_2(x).$$

Proof. This follows from the method of Ruelle [25] and Grothendieck [7]. The only additional feature is that the operator has infinitely many inverse branches but, as in [16, 17], this presents no additional complications to the proof. □

This gives an alternative expression for Lyapunov exponent in terms of the fixed points of powers of $\mathcal{T}_2$.

Corollary 7.3. *We can write Λ in terms of rapidly convergent series*

$$\Lambda = \frac{\sum_{n=1}^{\infty} c_n}{\sum_{n=1}^{\infty} b_n}$$

where

(1) b_n *and* c_n *are explicit values (given below) using fixed points of powers of* $\mathcal{T}_2$*; and*
(2) $|b_n| = O(e^{-cn^{1+1/(m-1)}})$ *and* $|c_n| = O(e^{-cn^{1+1/(m-1)}})$.

Proof. By Proposition 7.2 we can write

$$\Lambda = \frac{\frac{\partial d(1,s)}{\partial s}|_{s=1}}{\frac{\partial d(z,1)}{\partial z}|_{z=1}} = \frac{\sum_{n=1}^{\infty} a_n'(1)}{\sum_{n=1}^{\infty} n a_n(1)}.$$

Using the expansion $\exp(z) = 1 + \sum_{l=1}^{\infty} z^l/l!$ we can write that for $Re(s)$ sufficiently large and $|z|$ sufficiently small

$$\begin{aligned} d(z,s) &= 1 + \sum_{l=1}^{\infty} \frac{1}{l!} \left(- \sum_{k=1}^{\infty} \frac{z^k}{k} \sum_{\mathcal{T}_2^k x = x} |\det(D\mathcal{T}_2^k)(x)|^{-s} \right)^l \\ &= 1 + \sum_{n=1}^{\infty} z^n \left(\sum_{k_1+\cdots+k_l=n} \frac{(-1)^l}{l!} \prod_{i=1}^{l} \left(\frac{1}{k_i} \sum_{\mathcal{T}_2^{k_i} x = x} |\det(D\mathcal{T}_2^{k_i})(x)|^{-s} \right) \right) \end{aligned}$$

by grouping together terms with the same power of z. Thus by

$$a_n(s) = \left(\sum_{k_1+\cdots+k_l=n} \frac{(-1)^l}{l!} \prod_{i=1}^{l} \left(\frac{1}{k_i} \sum_{\mathcal{T}_2^{k_i} x = x} |\det(D\mathcal{T}_2^{k_i})(x)|^{-s} \right) \right)$$

and thus by part (2) of Proposition 7.2

$$b_n = n a_n(1) = n \left(\sum_{k_1+\cdots+k_l=n} \frac{(-1)^l}{l!} \prod_{i=1}^{l} \left(\frac{1}{k_i} \sum_{\mathcal{T}_2^{k_i} x = x} |\det(D\mathcal{T}_2^{k_i})(x)|^{-1} \right) \right)$$

and

$$c_n = a_n'(1) = \frac{d}{ds}|_{s=1} \left(\sum_{k_1+\cdots+k_l=n} \frac{(-1)^l}{l!} \prod_{i=1}^{l} \left(\frac{1}{k_i} \sum_{\mathcal{T}_2^{k_i} x = x} |\det(D\mathcal{T}_2^{k_i})(x)|^{-s} \right) \right).$$

The bounds on b_n come directly from the bounds on $a_n(1)$ in part (2) of Proposition 7.2.

Using the bounds on $a_n(s)$ in part (2) of Proposition 7.2 applied to s small neighbourhood of $s = 1$ we get bounds on $c_n = a_n'(1)$ using Cauchy's theorem, i.e., for small enough $\epsilon > 0$ we let

$$|a_n'(0)| \leq \frac{1}{2\pi} \left| \int_{|\xi|=\epsilon} a_n(\xi) \xi^{-2} d\xi \right| = O(e^{-cn^{1+1/(m-1)}})$$

and so the bounds on $|a_n(\cdot)|$ also serve to bound c_n. □

By the estimate in Part (2) of Proposition 7.2 we see that for each fixed t the function $d(z,t)$ is an entire function of order 1 in z. In particular, if $\{z_n(t)\}$ are poles of $d(z,t)$ then by the Hadamard Weierstrauss theorem the function $d(z,t)$ takes the form

$$d(z,t) = e^{A(t)z+B(t)} \prod_n \left(1 - \frac{z}{z_n(t)} \right) e^{\frac{z}{z_n(t)}}$$

where $A(t), B(t) \in \mathbb{C}$ and each $z_n(t)$ depend analytically on t by the Implicit Function Theorem.

Remark 7.4. *Following Zorich [34], we can also consider the largest Lyapunov exponent θ_1 for these transformations. Let E_{ij} $(1 \leq i,j \leq m)$ denote the $m \times m$ matrix with entries 1 on the diagonal and in the (i,j)th place and 0 otherwise, and let P_π denote the permutation matrix associated to π. Consider the matrices*

$$A(\pi, a) = (I + I_{\pi^{-1}m,m}).P(\tau^{\pi^{-1}(m)}) \text{ and } A(\pi, a) = E + I_{m,\pi^{-1}m}.$$

We then define a matrix valued function $B(\lambda, \mu)$ on $\cup_{\pi \in \mathcal{R}} \Delta_\pi^+ \cup \Delta_\pi^-$ by

$$B(\lambda, \pi) = A(\lambda, \pi) \left(A\mathcal{T}_0(\lambda, \pi)\right) \cdots \left(A\mathcal{T}_0^{\widehat{n}(\lambda,\pi)-1}(\lambda, \pi)\right).$$

The general definition for the (leading) Lyapunov exponent for this matrix is

$$\theta_1 = \inf_{n \geq 1} \left\{ \frac{1}{n} \int \log \|B(\lambda, \pi) B\mathcal{T}_1(\lambda, \pi) \cdots B\mathcal{T}_1^n(\lambda, \pi)\| d\mu_1 \right\}.$$

Zorich ([34], Theorem 4) proved the following elegant result: The Lyapunov exponent can be written

$$\begin{aligned} \theta_1 &= -\sum_{\pi \in \mathcal{R}} \int_{\Delta_\pi^\pm} \left|\log(1-\lambda_m) - \log(1-\lambda_{\pi^{-1}m})\right| d\mu_1(\lambda) \\ &= \frac{1}{m} \sum_{\pi \in \mathcal{R}} \int_{\Delta_\pi^\pm} \log|\det D\mathcal{T}_1| d\mu_1. \end{aligned}$$

To complete this section, we briefly consider a related complex function. We can formally define

$$\eta(z) = -\sum_{n=1}^\infty \frac{z^n}{n} \sum_{\mathcal{T}_2^n x = x} \frac{\log|\det(D\mathcal{T}_2^n)(x)|}{|\det(D\mathcal{T}_2^n)(x)|}, \quad z \in \mathbb{C}.$$

In particular, we observe that since $\eta(z) = \frac{\partial \log d(z,t)}{\partial t}|_{t=1}$ then by part (1) of Proposition 7.2 we see that $\eta(z)$ is meromorphic in the entire complex plane and we can write

$$\eta(z) = B'(1) + \sum_n \frac{\frac{z z_n'(1)}{z_n(1)}}{(z_n(1) - z)} + z\left(A'(1) + \frac{z_n'(1)}{[z_n(1)]^2}\right),$$

for which the poles are $\{z_n\}$ and the residues are $\mu_n := \frac{z_n'(1)}{z_n(1)}$ $(n \geq 1)$. Moreover, by part (3) of Proposition 7.2 the poles also correspond to derivatives of the eigenvalues of the associated transfer operator. This gives a connection to the approach to resonances considered by Ruelle in the context of Axiom A diffeomorphisms and is suggestive of an analogous interpretation.

Finally, we conclude with the following curiousity.

Proposition 7.5. *Assume that $\mu_n = 0$ for every $n \geq 1$ then $\eta(z) = 0$ for all $z \in \mathbb{C}$.*

Of course, the conclusion of the Proposition is equivalent to $\sum_{\mathcal{T}_2^n x = x} \frac{\log|\det(D\mathcal{T}_2^n)(x)|}{|\det(D\mathcal{T}_2^n)(x)|} = 0$ for each $n \geq 1$.

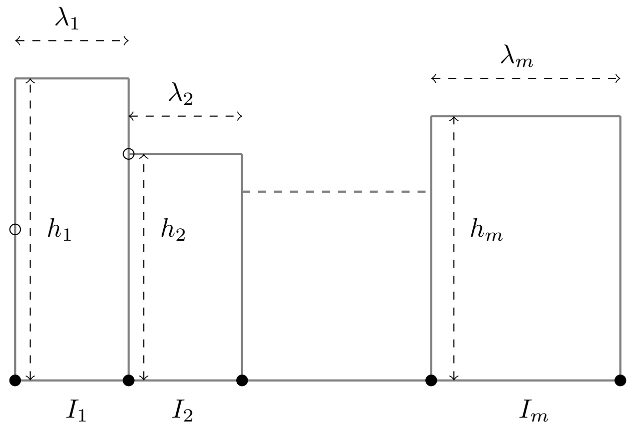

FIGURE 3. A zippered rectangle

8. A glimpse into Teichmüller flows

Thus far we have only considered the case of discrete transformations ($\mathcal{T}_1$, $\mathcal{T}_2$, etc.), but not the case for continuous flows. For completeness, we briefly describe in this section a small piece of the relationship with Teichmüller flows and suggest a connection with the preceding statistical results for $\mathcal{T}_1$ and $\mathcal{T}_2$. We begin by recalling a well known connection between flat surfaces (or translation surfaces) and interval exchange transformations, although we will keep our description brief and informal and the refer the reader to one of the several excellent surveys in this area, such as Veech [30], Viana [33, Chapter 2], Bufetov [4, Section 1.6] or Zorich [34, Section 5] to name a few.

There is a close connection between interval exchange maps and flat metrics on surfaces. A particularly convenient presentation of a flat surface is as a union of m rectangles in the plane based on the intervals I_i and of height l_i, for $i = 1, \cdots, m$. Thus the information we need to reconstruct the flat torus begins with

(a) The lengths λ_i of the intervals I_i ($i = 1, \cdots, m$);
(b) The heights h_i of the rectangles ($i = 1, \cdots, m$).

Since we will assume that the surface has unit area we can write that $\lambda_1 h_1 + \cdots + \lambda_m h_m = 1$. In addition in order to attach the tops of the rectangles back to their bottoms in the correct order we need:

(c) The permutation π on $\{1, \cdots, m\}$ which tells the change in order in which we reattach the tops of the rectangles.

In addition, to define the flow and invariant measure it is convenient to introduce two other coordinates (which obviously depend on those above):

(d) $a_0, \cdots, a_m$, which are actually dependent on the other variables by $h_i - a_i = h_{\pi^{-1}(\pi(i)+1)} - a_{\pi^{-1}(\pi(i)+1)-1}$ for $i = 1, \cdots, m-1$, with the convention $a_0 = a_{m+1} = 0$; and
(e) $\delta_i = a_{i-1} - a_i$, for $i = 1, \cdots, m$

and the heights of other singularities (which lie in the sides of the rectangles).

This construction is usually called a zippered rectangle. Let $\Omega_{\mathcal{R}}$ denote the space of all unit area (zippered) rectangles. There is a natural volume $d\lambda_1 \cdots d\lambda_m d\delta_1 \cdots d\delta_m$

on $\Omega_{\mathcal{R}}$. Let μ denote the normalized measure. A version of the Teichmüller flow $T_t : \Omega_{\mathcal{R}} \to \Omega_{\mathcal{R}}$ is defined locally by $T_t(\lambda, h, a, \pi) = (e^t\lambda, e^{-t}h, e^{-t}a, \pi)$ (i.e, flattening the rectangles from above) and this preserves the volume. There is a natural projection from $\Omega_{\mathcal{R}}$ to the moduli space of flat metrics $\mathcal{M}$ and the corresponding semi-conjugate flow $S_t : \mathcal{M} \to \mathcal{M}$ is the *Teichmüller flow*. However, to emphasize the connection to our previous discussion we will persist with the model flow T_t. We can consider the cross section

$$\mathcal{Y} = \left\{ (\lambda, h, a, \pi) \in \Omega_{\mathcal{R}} : \sum_{i=1}^{m} \lambda_i = 1 \right\}$$

to the flow T_t. Under the natural identification on $\Omega_{\mathcal{R}}$ corresponding to different presentations of surfaces as rectangles: the return time function to $\mathcal{Y}$ corresponds to the natural extension of the map $\mathcal{T}_0$ and the return time function is simply $r(\lambda, \pi) = \log(1 - \min\{\lambda_m, \lambda_{\pi^{-1}m}\})$. This shows that the properties of the flow T_t are closely related to those of the maps related to the Rauzy-Veech map.

In particular, the Teichmüller flow T_t is a finite-to-one factor of the natural extension of the suspended semi-flow associated to the map $\mathcal{T}_0$ and the function r, i.e., let

$$(\Delta \times \mathcal{R})^r = \left\{ (\underbrace{\lambda, \pi}_{=:x}, u) \in \Delta \times \mathcal{R} \times \mathbb{R} : 0 \leq u \leq r(\lambda, \pi) \right\}$$

where we identify $(x, r(x)) = (\mathcal{T}_0(x), 0)$ and we define the semi-flow

$$(\mathcal{T}_0)_t^r : (\Delta \times \mathcal{R})^r \to (\Delta \times \mathcal{R})^r$$

locally by $(\mathcal{T}_0)_t^r(x, u) = (x, u + t)$, subject to the identifications.

Since inducing on $B \subset \Delta$ (as described in the discrete case) gives the map $\mathcal{T}_2 : B \to B$, we can also represent this semi flow as a suspension semiflow over $\mathcal{T}_2 : B \to B$ with respect to a related function $r_2 : B \to \mathbb{R}$, i.e., let

$$B^{r_2} = \{(x, u) \in B \times \mathbb{R} : 0 \leq u \leq r_2(x)\} / \sim$$

where we identify $(x, r_2(x)) = (\mathcal{T}_2(x), 0)$ and we define $(\mathcal{T}_2)_t^{r_2} : B^{r_2} \to B^{r_2}$ locally by $(\mathcal{T}_2)_t^{r_2}(x, u) = (x, u + t)$, subject to the identifications.

The following lemma was established by Bufetov [4].

Lemma 8.1. *(1) $r_2 \in L^\gamma(B, \mu_2)$, for every $\gamma > 1$; and*
(2) if $F : \Omega_{\mathcal{R}} \to \mathbb{R}$ is Hölder and $f : B \to \mathbb{R}$ is defined by $f(x) := \int_0^{r_2(x)} F(S_t x)dt$ then there exists $\delta > 0$ such that $f \in L^{2+\delta}(B, \mu_2)$.

Since r_2 is integrable, the Teichmüller flow preserves the probability measure μ_{r_2} defined by $d\mu_{r_2} = \left(\int_B r_2 d\mu_2\right)^{-1} d\mu_2 \times ds$.

We now recall the continuous analogue of the Almost Sure Invariance Principle.

Definition 8.2. *A flow $\psi_t : X \to X$ with invariant probability measre μ is said to satisfy the Almost Sure Invariance Principle with respect to a probability measure ν if for a Hölder function $\Phi : X \to \mathbb{R}$ not equal to a coboundary such that $\int \Phi d\mu = 0$ there is a $\epsilon > 0$ and a random variable $\{S_t\}_{t \geq 0}$ and a Brownian motion B with variance σ^2 such that $\left\{\int_0^t \Phi(\psi_s)ds\right\}_{t \geq 0}$, seen as a random process defined on (X, ν), is equal in distribution to random variables $\{S_t\}_{t \geq 0}$ and $S_t = B_t + O(t^{1/2-\epsilon})$.*

The result for Teichmüller flows corresponding to Theorem 4.9 is the following.

Theorem 8.3. *The Teichmüller flow satisfies the almost sure invariance principle with respect to the probability measure μ_2 seen as a measure on B^{r_2} supported on $B \times \{0\}$.*

Proof. It suffices to show the result for the associated semi-flow (the result for the natural extension requiring a standard argument involving changing functions by a coboundary, see for instance [19, Lemma 3.2]). Let $\Phi : B^{r_2} \to B^{r_2}$ be a Hölder function with $\phi(x) = \int_0^{r(x)} \Phi((\mathcal{T}_2)_t^{r_2} x))dt$.

(i) $\overline{r}_2 \in L^{2+\beta}(B, \mu_2)$, for some $\beta > 1$ (by part (1) of Lemma 8.1)
(ii) $\phi \in L^{2+\delta}(B, \mu_2)$, for some $\delta > 0$ (by part (2) of Lemma 8.1); and
(iii) $\mathcal{T}_2 : B \to B$ satisfies the Almost Sure Invariance Principle (by Theorem 4.9)

The Teichmüller flow then satisfies the Almost Sure Invariance Principle by the results of [18]. □

Even though the Almost Sure Invariance Principle has been proven only for the measure μ_2, and not for the measure μ_{r_2}, this is enough to deduce the following corollaries for Hölder continuous functions Φ, see Denker and Philipp [5].

Corollary 8.4 (Central Limit Theorem)**.** *For $y \in \mathbb{R}$ we have that*

$$\lim_{T\to+\infty} \mu_{r_2} \left\{ (x,s) \; : \; \frac{1}{\sqrt{T}} \int_0^T \Phi((\mathcal{T}_2)_t^{r_2}(x,s))dt \le y \right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt$$

The central limit theorem for Teichmüller flows was proved by Bufetov [4].

Corollary 8.5. *For $y \ge 0$ we have that*

$$\lim_{T\to+\infty} \mu_{r_2} \left\{ (x,s) \; : \; \frac{1}{\sqrt{T}} \max_{1\le t\le T} \int_0^t \Phi((\mathcal{T}_2)_t^{r_2}(x,s))dt \le y \right\} = \frac{\sqrt{2}}{\sqrt{\pi}\sigma} \int_{-\infty}^{y} e^{-t^2/2\sigma^2} dt - 1$$

Corollary 8.6 (Arcsine Law)**.** *For $0 \le y \le 1$ we have that*

$$\lim_{T\to+\infty} \mu_{r_2} \left\{ (x,s) \; : \; \frac{N_T(x)}{T} \le y \right\} = \frac{2}{\sqrt{\pi}} \sin^{-1} \sqrt{y}$$

where $N_T(x,s) = Leb \left\{ 0 \le t \le T \;\; : \; \int_0^t \Phi((\mathcal{T}_2)_t^{r_2}(x,s))dt > 0 \right\}$

Corollary 8.7 (Law of the iterated logarithm)**.** *For μ_{r_2}-a.e. (x,s) we have*

$$\limsup_{T\to+\infty} \frac{\int_0^T \Phi((\mathcal{T}_2)_t^{r_2}(x,s))dt}{\sigma\sqrt{2T\log\log T}} = 1.$$

Remark 8.8. *If Korepanov's results [11] can be extended to suspended flows, then it would be possible to prove Theorem 8.3 for the invariant measure μ_{r_2}, and then, using arguments from Melbourne and Nicol [19], to pass to the natural extension, thus obtaining the Almost Sure Invariance Principle and all its corollaries for the original (invertible) Teichmüller flow defined on $\Omega_{\mathcal{R}}$.*

9. Comments on pressure

It is natural to relate these statistical properties to classical ideas on pressure. To accommodate the complication of having a countable-to-one map $\mathcal{T}_2 : B \to B$ (and also an unbounded return time $\widehat{n} : B \to \mathbb{Z}^+$ when we look at the tower to reconstruct $\mathcal{T}_1$) it is convenient to work with the Gurevich pressure (as developed by Sarig [26], by analogy with the more familiar Gurevich entropy for countable subshifts of finite type). Let us consider a fairly general formulation of these results.

Recall that $\mathcal{Q}$ is the partition of smoothness of $\mathcal{T}_2$, and that $\mathcal{Q}_n = \vee_{i=0}^{n-1} \mathcal{T}_2^{-i}\mathcal{Q}$ is its nth level refinement. For $x, y \in B$, we denote by

$$s(x, y) = \inf \{n \geq 0 \,:\, x \text{ and } y \text{ belong to two different elements of } \mathcal{Q}_n\}$$

their separation time with respect to $\mathcal{T}_2$. Assume that $\phi : B \to \mathbb{R}$ is (locally) Hölder continuous, in the sense that there exists $0 < \theta < 1$ and $A > 0$ such that $V_n(\theta) \leq A\theta^n$ for $n \geq 0$, where

$$V_n(\phi) := \sup\{|\phi(x) - \phi(y)| \,:\, x, y \in B, s(x, y) \geq n\}$$

(i.e., the variation of the function over elements of the nth level refinement of the partition associated with $\mathcal{T}_2$.) In particular, Lipschitz functions satisfy these conditions. On the other hand, more generally this condition doesn't require ϕ to be bounded, say.

Definition 9.1. *To define the (Gurevich) pressure can fix any element $A \in \mathcal{Q}_{n_0}$, for chosen value n_0. We then define*

$$P(\phi) := \lim_{n\to\infty} \frac{1}{n} \log \left(\sum_{\mathcal{T}_2^n x = x \in A} e^{\phi^n(x)} \right),$$

where $\phi^n(x) := \phi(x) + \phi(\mathcal{T}_2 x) + \cdots + \phi(\mathcal{T}_2^{n-1} x)$.

Under very modest mixing conditions (i..e., the "Big Images Property" which applies in the case of $\mathcal{T}_2$) we can see that the definition is independent of the choice of n_0 and A. However, in general some additional assumptions are required to ensure that the pressure is finite.

One would anticipate that the properties of $P(\phi)$ would be useful in further studies of the properties of these maps and flows.

Remark 9.2. *Sarig [27] has results which suggest that the map $t \mapsto P(\phi + t\psi)$ is analytic for suitable Hölder continuous functions ϕ, ψ whenever the pressure is finite and $t \in (-\epsilon, \epsilon)$. This is based on spectral properties of a suitable transfer operator.*

References

[1] R. Aimino, M. Nicol, M. Todd, *Recurrence statistics for the space of Interval Exchange maps and the Teichmüller flow on the space of translation surfaces*, Ann. Inst. H. Poincaré Probab. Statist., **53**, 1371-1401, (2017)

[2] A. Avila, A. Bufetov, *Exponential decay of correlations for the Rauzy-Veech-Zorich induction map*, in *Partially Hyperbolic Dynamics, Laminations, and Teichmüller Flow*, Fields Inst. Commun. **51**, 203-211, Amer. Math. Soc., Providence, RI, (2007)

[3] V. Baladi, *Dynamical Zeta Functions and Dynamical Determinants for Hyperbolic Maps - A Functional Approach*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge/A Series of Modern Surveys in Mathematics, **68**, Springer International Publishing, 2018)

[4] A. Bufetov, *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*, J. Amer. Math. Soc., **19**, 579-623, (2006)

[5] M. Denker, W. Philipp, *Approximation by Brownian motion for Gibbs measures and flows under a function*, Ergodic Theory Dynam. Systems, **4**, 541-552, (1984)

[6] K. Fernando, C. Liverani, *Edgeworth expansions for weakly dependent random variables*, to appear in Ann. Inst. H. Poincaré Probab. Statist., (2020)

[7] A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc., **16**, Amer. Math. Soc., (1955)
[8] P. Hall, C. Heyde, *Martingale limit theory and its application*, Academic Press, New York, (1980)
[9] C. Heyde, *Invariance principles in statistics*, International Statistical Review, **49**, 143-152, (1981)
[10] A. G. Kachurovskiĭ, *Rates of convergence in ergodic theorems*, Russian Math. Surveys, **51**, 653-703, (1996)
[11] A. Korepanov, *Equidistribution for Nonuniformly Expanding Dynamical Systems, and Application to the Almost Sure Invariance Principle*, Comm. Math. Phys., **359**, 1123-1138, (2018)
[12] A. Korepanov, *Rates in Almost Sure Invariance Principle for Dynamical Systems with Some Hyperbolicity*, Comm. Math. Phys., **363**, 173-190, (2018)
[13] M. Mackey, M. Tyran Kamińska, *Deterministic Brownian motion: The effects of perturbing a dynamical system by a chaotic semi-dynamical system*, Physics reports, **422**, 167-222, (2006)
[14] M. Mackey, M. Tyran Kamińska, *Central limit theorems for non-invertible measure preserving maps*, Colloq. Math., **110**, 167-191, (2008)
[15] H. Masur *Interval exchange transformations and measured foliations*, Annals of Math., **115**, 169-200, (1982)
[16] D. Mayer, *On a ζ function related to the continued fraction transformations*, Bull. Soc. Math. France, **104**, 195-203, (1976)
[17] D. Mayer, *Approach to equilibrium for locally expanding maps in* $\mathbb{R}^k$, Comm. Math. Phys., **95**, 1-15, (1984)
[18] I. Melbourne, A. Török, *Statistical limit theorems for suspension flows*, Israel J. Math., **144**, 191-209, (2004)
[19] I. Melbourne, M. Nicol, *Almost sure invariance principle for nonuniforlmy hyperbolic systems*, Comm. Math. Phys., **260**, 131-145, (2005)
[20] I. Melbourne, M. Nicol, *Large deviations for nonuniformly hyperbolic systems*, Trans. Amer. Math. Soc., **360**, 6661-6676, (2008)
[21] T. Morita, *Renormalized Rauzy inductions*, Advanced Studies in Pure Mathematics, **43**, 1-25, (2005)
[22] W. Parry, M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic systems*, Astérisque, **187-188**, 1-268, (1990)
[23] W. Philipp, W. Stout, *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables*, Mem. Amer. Math. Soc., **161**, Amer. Math. Soc, Providence, RI, (1975)
[24] L. Rey-Bellet, L.-S. Young, *Large deviations in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems, **28**, 587-612, (2008)
[25] D. Ruelle, *Zeta-functions for expanding maps and Anosov flows*, Invent. Math., **34**, 231-242, (1976)
[26] O. Sarig, Thermodynamic formalism for countable Markov shifts, Ergodic Theory Dynam. Systems, **19**, 1565-1593, (1999)
[27] O. Sarig, Phase transitions for countable Markov shifts, Comm. Math. Phys., **217**, 555-577, (2001)
[28] M. Tyran Kamińska, *An invariance principle for maps with polynomial decay of correlations*, Comm. Math. Phys., **260**, 1-15, (2005)
[29] W. Veech, *Interval exchange transformations*, JAMA, **33**, 222-272, (1978)
[30] W. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Annals of Math., **115**, 201-242, (1982)
[31] W. Veech, *The Teichmüller geodesic flow*, Annals of Math., **124**, 441-530, (1986)
[32] J.-C. Yoccoz, *Continued Fraction Algorithms for Interval Exchange Maps: an Introduction*, in Frontiers in Number Theory, Physics, and Geometry I, Springer, 401-435, (2006)
[33] M. Viana, *Dynamics of interval exchange maps and Teichm"uller flows*, IMPA, (2008), available at http://w3.impa.br/~viana/out/ietf.pdf
[34] A. Zorich, *Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents*, Ann. Inst. Fourier, **46**, 325-370, (1996)
[35] A. Zorich, *Flat surfaces*, Frontiers in number theory, physics, and geometry I, Springer, Berlin, pp 437-583, (2006)