

Recurrence statistics for the space of interval exchange maps and the Teichmüller flow on the space of translation surfaces

Romain Aimino^a, Matthew Nicol^b and Mike Todd^c

^a*Departamento de Matemática, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal.*

E-mail: romain.aimino@fc.up.pt; url: http://www.fc.up.pt/pessoas/romain.aimino/

^b*Department of Mathematics, University of Houston, Houston, Texas, USA.*

E-mail: nicol@math.uh.edu; url: http://www.math.uh.edu/~nicol/

^c*Mathematical Institute, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9SS, Scotland.*

E-mail: m.todd@st-andrews.ac.uk; url: http://www.mcs.st-and.ac.uk/~miket/

Received 25 September 2015; revised 15 March 2016; accepted 14 April 2016

Abstract. In this paper we show that the transfer operator of a Rauzy–Veech–Zorich renormalization map acting on a space of quasi-Hölder functions is quasicompact and derive certain statistical recurrence properties for this map and its associated Teichmüller flow. We establish Borel–Cantelli lemmas, Extreme Value statistics and return time statistics for the map and flow. Previous results have established quasicompactness in Hölder or analytic function spaces, for example the work of M. Pollicott and T. Morita. The quasi-Hölder function space is particularly useful for investigating return time statistics. In particular we establish the shrinking target property for nested balls in the setting of Teichmüller flow. Our point of view, approach and terminology derive from the work of M. Pollicott augmented by that of M. Viana.

Résumé. Dans cet article, nous démontrons que l'opérateur de transfert de l'application de renormalisation de Rauzy–Veech–Zorich est quasi-compact sur l'espace des fonctions quasi-Hölder, et nous en déduisons plusieurs propriétés de récurrence statistiques pour cette application et le flot de Teichmüller associé. Nous établissons des lemmes de Borel–Cantelli, des statistiques des valeurs extrêmes et des temps de retour pour l'application et le flot. De précédents résultats ont établi la quasi-compactité dans des espaces de fonctions Hölder ou analytiques, comme par exemple les travaux de M. Pollicott ou de T. Morita. L'espace fonctionnel quasi-Hölder est particulièrement adapté pour analyser les propriétés de récurrence statistiques. En particulier, nous démontrons la propriété des cibles rétrécissantes pour des boules imbriquées dans le cadre du flot de Teichmüller. Notre point de vue, approche et terminologie proviennent du travail de M. Pollicott ainsi que de celui de M. Viana.

MSC: 37A50; 37D40; 60G70

Keywords: Interval exchange map; Teichmüller flow; Rauzy–Veech–Zorich renormalisation map; Transfer operator; Borel–Cantelli lemmas; Extreme Value Laws; Return/hitting time statistics; Quasi-Hölder function space

1. Background and notation

1.1. Dynamical Borel–Cantelli lemmas and other limit laws

Let $T : X \rightarrow X$ be a measure-preserving transformation of a probability space (X, μ) . We assume X is also a metric space equipped with a metric d . Dynamical Borel–Cantelli lemmas concern the following set of questions: suppose (A_n) is a sequence of sets such that $\sum_n \mu(A_n) = \infty$, does $T^n(x) \in A_n$ for infinitely many values of n for μ a.e. $x \in X$? One special example of this is the case where (A_n) is a nested sequence of balls about a point, a setting which is often called the shrinking target problem.

We let $S_n = \sum_{j=0}^{n-1} 1_{A_j} \circ T^j$ and $E_n = \int_X S_n d\mu = \sum_{j=0}^{n-1} \mu(A_j)$. The property $\lim_{n \rightarrow \infty} \frac{S_n(x)}{E_n} = 1$ for μ a.e. $x \in X$ is often called the Strong Borel–Cantelli (SBC) property in contrast to the Borel–Cantelli (BC) property that $S_n(x)$ is unbounded for μ a.e. $x \in X$.

In the setting of uniformly hyperbolic systems pioneering work has been done by W. Philipp [48], Kleinbock and Margulis [37], Chernov and Kleinbock [11] and Dolgopyat [13] (for uniformly partially hyperbolic systems).

More recently dynamical Borel–Cantelli results have been proved for certain non-uniformly hyperbolic systems by example by Kim [36], Gouëzel [21], Gupta et al. [25] and Haydn et al. [27]. These works have also yielded some interesting counterexamples. In the context of flows, Maucourant [43] has proved the analogous Borel Cantelli property for nested balls in the setting of geodesic flows. Athreya [3] gives large deviation and quantitative recurrence results for the Teichmüller geodesic flow.

Related to Borel–Cantelli lemmas are logarithmic laws for the shrinking target problem. These results concern the asymptotic scaling behavior given by the limit

$$\lim_{r \rightarrow 0} \frac{\tau_r(x, y)}{\mu(B_r(y))},$$

where $\tau_r(x, y) = \min\{n : d(T^n x, y) < r\}$ and $B_r(y)$ is a ball of radius r about $y \in X$.

Fundamental work on the ergodic theory of Teichmüller flow and interval exchange transformations include [10, 34, 55, 56, 63]. Of particular relevance to our setting is work of Masur [42], who proved a logarithm type law for Teichmüller geodesic flow on the moduli space of quadratic differentials and work of Galatolo and Kim [19] who obtain Borel–Cantelli like results for generic interval exchange transformations. Marchese [39, 40] also obtained related results on the shrinking target problem for the Rauzy–Veech–Zorich algorithm, with applications to a generalization of the Khinchin theorem for interval exchange transformations. He also obtained logarithmic limit laws for returns for Teichmüller flow on translation surfaces [40, Theorem 1.3]. These quantitative results apply under a logarithmic scaling, unlike our results which apply to the unscaled flow.

Statistical properties of the Teichmüller flow and the Rauzy–Veech–Zorich map have been investigated thoroughly in recent years. Avila, Gouëzel and Yoccoz [5] have shown that the decay of correlations for the flow is exponentially fast for Hölder observables. The corresponding problem for the Rauzy–Veech–Zorich map has been studied by Bufetov and Avila in [9] and [4], where the decay was proven to be exponential as well. The main ingredient of the proof of the latter result was the construction of a Young Tower [60] with an exponential tail of return times. Building upon this fact and work of Melbourne and Nicol [44], Pollicott [49] proved the almost sure invariance principle for Hölder observables, both for the flow and the map. The almost sure invariance principle is a strong reinforcement of the central limit theorem, which was previously established by Bufetov [9], and has several consequences, such as the law of iterated logarithm and the arcsine law. The large deviations principle for Hölder observables follows also directly from the existence of an exponential Young tower and results of Melbourne and Nicol [45].

We also establish recurrence statistics such as Poisson limit laws and Extreme Value Laws (EVLs) for Teichmüller flow, but we leave the detailed description of these properties and results to Section 3. Related work on these distributional limits in a dynamical setting includes [24, 30–32].

1.2. Interval exchange transformations

In this section we synthesize the basic model described by Viana in [57] with the framework developed by Pollicott [49] (see also [47]). Pollicott’s short paper [49] is a very clear account of the Rauzy–Veech–Zorich induction and renormalization from the viewpoint of hyperbolic dynamics. We begin by defining our dynamical systems. This starts with interval exchange transformations, in particular focussing on the formalism described by Viana. We then move to the Rauzy–Veech induction and renormalisation; the Zorich induction and renormalisation; and finally the Morita–Pollicott renormalisation. We will point out the minor differences with Pollicott’s framework as we go along, but broadly speaking, the difference here is that our induced maps are *first* returns. We relate these dynamical systems to the Teichmüller flow on the space of translation surfaces later on.

Following [57, Chapter 1], let $I \subset \mathbb{R}$ be an interval and $\{I_a : a \in \mathcal{A}\}$ a partition of I into intervals indexed by a finite alphabet \mathcal{A} with $d \geq 2$ symbols. An *interval exchange transformation* (IET) is a bijective map $f = f_{(\pi, \lambda)} : I \rightarrow I$ which is a translation of each subinterval I_a , preserves Lebesgue measure and is determined by the following combinatorial and metric data:

- (a) A pair $\pi = (\pi_0, \pi_1)$ of bijections $\pi_\varepsilon : \mathcal{A} \rightarrow \{1, \dots, d\}$ which describe the ordering of the subintervals I_a before and after the action of f :

$$\begin{pmatrix} a_1^0 & a_2^0 & \dots & a_d^0 \\ a_1^1 & a_2^1 & \dots & a_d^1 \end{pmatrix},$$

where $a_j^\varepsilon = \pi_\varepsilon^{-1}(j)$ for $\varepsilon \in \{0, 1\}$ and $j \in \{1, 2, \dots, d\}$.

- (b) A vector $\lambda = (\lambda_a)_{a \in \mathcal{A}}$ of non-negative entries which represent the lengths of the subintervals $(I_a)_{a \in \mathcal{A}}$.

We have a more detailed description of the intervals I_a above which will be useful later: for $\varepsilon \in \{0, 1\}$, let $I_a^{\pi_\varepsilon}$ be the interval of length $\lambda_{\pi_\varepsilon(a)}$ in position $\pi_\varepsilon(a)$ in the interval $[0, \sum_a \lambda_a]$, where “position” means starting at zero and counting to the right.

The transformation $p := \pi_1 \circ \pi_0^{-1}$ is called the *monodromy invariant* of the pair $\pi = (\pi_0, \pi_1)$. As Viana points out, we can always change our pair $\pi = (\pi_0, \pi_1)$ and rearrange the ordering of our lengths so that the resulting data $\pi' = (\pi'_0, \pi'_1)$ and $\lambda' = (\lambda'_a)_{a \in \mathcal{A}}$ represents the same IET as the one above, but with $\pi_0 = \text{id}$. Indeed, this is what is described in Pollicott’s notes: moreover he always assumes that $\sum_a \lambda_a = 1$. However, the setup described here gives a slightly more complicated, but more flexible way for us to describe later dynamics.

The IET can now be described more explicitly as a translation. For $a \in \mathcal{A}$, define

$$w_a := \sum_{\{b: \pi_1(b) < \pi_1(a)\}} \lambda_b - \sum_{\{b: \pi_0(b) < \pi_0(a)\}} \lambda_b.$$

Then

$$f_{(\pi, \lambda)}(x) = x + \sum_a w_a \cdot \mathbb{1}_{I_a}(x).$$

Later it will be useful to think of the translation vector w_a as $\sum_{b \in \mathcal{A}} \mathcal{M}_{ab} \lambda_b$ where the (a, b) entry of the matrix \mathcal{M} is defined by

$$\mathcal{M}_{ab} = \begin{cases} +1 & \text{if } \pi_1(b) < \pi_1(a) \text{ and } \pi_0(b) > \pi_0(a), \\ -1 & \text{if } \pi_1(b) < \pi_1(a) \text{ and } \pi_0(b) < \pi_0(a), \\ 0 & \text{otherwise.} \end{cases}$$

1.3. Rauzy–Veech induction and renormalisation

As is common for families of dynamical systems with parabolic-type behaviour, one way to proceed is to define a good renormalization scheme on the space of parameters. In this setting this was pioneered by Masur and Veech. Given a representative (π, λ) of an IET, for $\varepsilon \in \{0, 1\}$, let $a(\varepsilon)$ denote the last symbol in the expression for π_ε , i.e., $a(\varepsilon) = \pi_\varepsilon^{-1}(d) = a_d^\varepsilon$. Assuming the generic situation where $I_{a(0)}$ and $I_{a(1)}$ have different lengths, we say that

$$(\pi, \lambda) \text{ has } \begin{cases} \text{type 0} & \text{if } \lambda_{a(0)} > \lambda_{a(1)}, \\ \text{type 1} & \text{if } \lambda_{a(0)} < \lambda_{a(1)}. \end{cases}$$

Now set

$$J = \begin{cases} I \setminus f_{(\pi, \lambda)}(I_{a(1)}) & \text{if } (\pi, \lambda) \text{ has type 0,} \\ I \setminus I_{a(0)} & \text{if } (\pi, \lambda) \text{ has type 1.} \end{cases}$$

Then the *Rauzy–Veech induction* $\hat{\mathcal{T}}_0$ is defined as the first return by $f_{(\pi, \lambda)}$ to J . Another way of viewing this, from which we see that we obtain a new IET of the form we started with (although with shorter total length of our intervals), is that $\hat{\mathcal{T}}_0(\pi, \lambda) = (\pi', \lambda')$ where, if (π, λ) is type 0 then

$$\begin{pmatrix} \pi'_0 \\ \pi'_1 \end{pmatrix} = \begin{pmatrix} a_1^0 & \dots & a_{k-1}^0 & a_k^0 & a_{k+1}^0 & \dots & \dots & a(0) \\ a_1^1 & \dots & a_{k-1}^1 & a(0) & a(1) & a_{k+1}^1 & \dots & a_{d-1}^1 \end{pmatrix}$$

and $\lambda' = (\lambda'_a)_{a \in \mathcal{A}}$ for

$$\lambda'_a = \lambda_a \quad \text{for } a \neq a(0), \quad \text{and} \quad \lambda'_{a(0)} = \lambda_{a(0)} - \lambda_{a(1)}.$$

Similarly, if (π, λ) is type 1 then

$$\begin{pmatrix} \pi'_0 \\ \pi'_1 \end{pmatrix} = \begin{pmatrix} a_1^0 & \cdots & a_{k-1}^0 & a(1) & a(0) & a_{k+1}^0 & \cdots & a_{d-1}^0 \\ a_1^1 & \cdots & a_{k-1}^1 & a_k^0 & a_{k+1}^0 & \cdots & \cdots & a(1) \end{pmatrix}$$

and $\lambda' = (\lambda'_a)_{a \in \mathcal{A}}$ for

$$\lambda'_a = \lambda_a \quad \text{for } a \neq a(1), \quad \text{and} \quad \lambda'_{a(1)} = \lambda_{a(1)} - \lambda_{a(0)}.$$

Remark 1.1. This transformation on the set of lengths in \mathbb{R}_+^A can be expressed in terms of a matrix Θ given in (1.9) and (1.10) of [57] and which consists only of 0's and 1's: in fact $\lambda' = \Theta^{-1*}(\lambda)$ where $*$ denotes the transpose. Θ^{-1} is a non-negative matrix.

We are interested in the set of (π, λ) such that $\hat{\mathcal{T}}_0$ is defined for all time. This occurs if and only if (π, λ) satisfies the *Keane condition*, which assumes that

$$f_{(\pi, \lambda)}^n(\partial I_a) \neq \partial I_b \quad \text{for all } n \geq 1 \text{ and } a, b \in \mathcal{A} \text{ with } \pi_0(b) \neq 1,$$

where ∂I_a is the left endpoint of the subinterval I_a . Moreover, if (π, λ) satisfies the Keane condition then $f_{(\pi, \lambda)}$ is minimal (every $f_{(\pi, \lambda)}$ -orbit is dense). A pair $\pi = (\pi_0, \pi_1)$ is called *reducible* if there exists $k \in \{1, \dots, d - 1\}$ such that $\pi_1 \circ \pi_0^{-1}(\{1, \dots, k\}) = \{1, \dots, k\}$. In this case, $f_{(\pi, \lambda)}$ splits into two IETs with simpler combinatorics. If π is not reducible, we say it is *irreducible*. It can be shown that if λ is rationally independent and π is irreducible then (π, λ) satisfies the Keane condition. Keane conjectured that for fixed irreducible π , the map $f_{(\pi, \lambda)}$ was uniquely ergodic for almost-every λ . This conjecture was proved independently by Masur [41] and Veech [55]. The method of proof of Veech was based on a renormalization scheme.

Given a fixed d , as above, we define the Rauzy class $\mathcal{R} = \mathcal{R}(\pi)$ of a pair π as the set of all pairs π' for which there exist $n \geq 0, \lambda$ and λ' with $\hat{\mathcal{T}}_0^n(\pi, \lambda) = (\pi', \lambda')$. They form a partition of the set of all pairs π . Thus we think of $\hat{\mathcal{T}}_0$ acting on sets $\mathcal{R} \times \mathbb{R}_+^A$. For $d = 2$ and $d = 3$ there is a unique Rauzy class, but for $d \geq 4$ there is more than one. Again we refer the reader to [57, Chapter 1] for a nice description of these.

The *Rauzy–Veech renormalization map* \mathcal{T}_0 is simply the transformation $\hat{\mathcal{T}}_0$ renormalised so that the total length of the resulting interval is 1: thus the multiplying factor is

$$\frac{1}{1 - \lambda_{a(1-\varepsilon)}} \quad \text{when } (\pi, \lambda) \text{ is type } \varepsilon.$$

That is $\mathcal{T}_0(\pi, \lambda) = (\pi', \lambda'')$ where $\lambda'' = \frac{\lambda'}{1 - \lambda_{a(1-\varepsilon)}}$. Thus \mathcal{T}_0 acts on the $(d - 1)$ dimensional simplex

$$\Delta = \Delta_{\mathcal{A}} := \{ \lambda = (\lambda_1, \dots, \lambda_d) : \lambda_i > 0, \lambda_1 + \dots + \lambda_d = 1 \}.$$

We define $|\lambda| = \sum_{j=1}^d \lambda_j$, then \mathcal{T}_0 has the form

$$\mathcal{T}_0(\pi, \lambda) = \left(\pi', \frac{\Theta^{-1*}\lambda}{|\Theta^{-1*}\lambda|} \right),$$

where Θ is the matrix defined in Remark 1.1.

Setting

$$\Delta_{\pi, \varepsilon} := \{ \lambda \in \Delta_{\mathcal{A}} : \lambda_{a(\varepsilon)} > \lambda_{a(1-\varepsilon)} \} \quad \text{for } \varepsilon \in \{0, 1\}, \tag{1}$$

$\mathcal{T}_0 : \{\pi\} \times \Delta_{\pi,\varepsilon} \mapsto \{\pi'\} \times \Delta$ is a bijection: a nice Markov property. This also implies that Θ is constant on each $\{\pi\} \times \Delta_{\pi,\varepsilon}$.

As in work of Veech [55] (see also Masur [41]), \mathcal{T}_0 has an absolutely continuous and invariant ergodic measure (acim) μ_0 , which is infinite. \mathcal{T}_0 is not uniformly hyperbolic.

1.4. Zorich induction and renormalisation

Zorich produced accelerated versions of the Rauzy–Veech maps discussed above in order to improve the expansion properties of the system and ultimately to find absolutely continuous invariant probability measures. For this subsection we fix a Rauzy class \mathcal{R} . Now take $\pi = (\pi_0, \pi_1)$ in this class and $\lambda \in \mathbb{R}_+^A$ satisfying the Keane condition. Then for each $k \geq 1$ write $(\pi^k, \lambda^k) = \hat{\mathcal{T}}_0^k(\pi, \lambda)$ and let ε^k denote the type of (π^k, λ^k) and ε denote the type of (π, λ) . Then $n_1 = n_1(\pi, \lambda)$ is defined as the smallest k such that $\varepsilon^k \neq \varepsilon$ and the Zorich induction is defined by

$$\hat{\mathcal{T}}_1(\pi, \lambda) = \hat{\mathcal{T}}_0^{n_1}(\pi, \lambda).$$

Similarly, the Zorich renormalisation $\mathcal{T}_1 : \mathcal{R} \times \Delta \rightarrow \mathcal{R} \times \Delta$ is defined as $\mathcal{T}_1 = \mathcal{T}_0^{n_1}$. This map has a Markov partition into countably many domains. Indeed, let

$$\Delta_{\pi,\varepsilon,n} := \{\lambda \in \Delta_{\pi,\varepsilon} : \varepsilon^1 = \dots = \varepsilon^{n-1} = \varepsilon \neq \varepsilon^n\}.$$

Then for each $\pi \in \mathcal{R}$, $\mathcal{T}_1 : \{\pi\} \times \Delta_{\pi,\varepsilon,n} \mapsto \{\pi^n\} \times \Delta_{\pi^n,1-\varepsilon}$ is a bijection. Moreover,

$$\lambda^n = c_n \Theta^{-n*}(\lambda),$$

where $c_n > 0$ and Θ^{-n*} depends only on π, ε, n . Let also $\Delta_\varepsilon = \bigcup_{\pi \in \mathcal{R}} \Delta_{\pi,\varepsilon}$ and $\Delta_{1-\varepsilon} = \bigcup_{\pi \in \mathcal{R}} \Delta_{\pi,1-\varepsilon}$.

Theorem 1.2 (Zorich). *For a given Rauzy class \mathcal{R} , \mathcal{T}_1 has an absolutely continuous invariant probability measure μ_1 . Moreover, for $\varepsilon \in \{0, 1\}$,*

$$\mathcal{T}_1^2 : \Delta_\varepsilon \rightarrow \Delta_\varepsilon$$

is mixing with respect to the restriction to Δ_ε of the measure $2\mu_1$. Similarly

$$\mathcal{T}_1^2 : \Delta_{1-\varepsilon} \rightarrow \Delta_{1-\varepsilon}$$

is mixing with respect to $2\mu_1$ restricted to $\Delta_{1-\varepsilon}$.

As already noted above, $\mathcal{T}_1(\Delta_\varepsilon) = \Delta_{1-\varepsilon}$, so the absolutely continuous invariant probability measure (acip) μ_1 is not mixing, but has two cyclic classes.

1.5. Morita–Pollicott renormalisation

A common approach (see [5,47,49]) is to consider a map \mathcal{T}_2 derived from \mathcal{T}_1 further by inducing by first return times on an element of a dynamical partition with compact closure in the parameter space. \mathcal{T}_2 has the advantage that it is a multidimensional piecewise expanding map. The setup in Pollicott [49] is slightly different to that outlined here, but for most practical purposes, it is identical.

Recalling the definition of $\Delta_{\pi,0}, \Delta_{\pi,1}$ from (1), let

$$\mathcal{P} = \{\{\pi\} \times \Delta_{\pi,0}, \{\pi\} \times \Delta_{\pi,1} : \pi \in \mathcal{R}\}$$

be the usual finite partition of $\mathcal{R} \times \Delta$ and define for $n \geq 1$

$$\mathcal{P}_n := \bigvee_{k=0}^{n-1} \mathcal{T}_1^{-k} \mathcal{P}.$$

Pollicott’s approach is to choose an $n_B > 1$ and a partition element $B \in \mathcal{P}_{n_B}$ such that B has compact closure \bar{B} contained in the open simplex $\mathbb{R} \times \Delta$. In this case, B is the image of an inverse branch of $\mathcal{T}_1^{n_B}$ which is a strict contraction for the Hilbert metric (see also [57, Corollary 1.21]). Define $n_2(\pi, \lambda)$ to be the first return time of $(\pi, \lambda) \in B$ to B under \mathcal{T}_1 , i.e.

$$n_2(\pi, \lambda) = \inf\{k > 0 : \mathcal{T}_1^k(\pi, \lambda) \in B\}.$$

Then $\mathcal{T}_2 : B \rightarrow B$ is defined as the induced first return time map under \mathcal{T}_1 ,

$$\mathcal{T}_2(\pi, \lambda) = \mathcal{T}_1^{n_2(\lambda, \pi)}(\pi, \lambda).$$

Remark 1.3. Note that for each element $(\pi, \lambda) \in \mathcal{R} \times \Delta$, with λ satisfying the Keane condition, we can find such a B containing (π, λ) .

The set B has a natural countable partition $\mathcal{Q} = \{B_i\}_{i \in \mathcal{I}}$ into sets on which $n_2(\pi, \lambda)$ is constant. The map $\mathcal{T}_2 : B_i \rightarrow B$ is a diffeomorphism for each $i \in \mathcal{I}$ [47, Lemma 3.1]. B has a naturally defined \mathcal{T}_2 -invariant measure, namely $\mu_2 := \frac{\mu_1|_B}{\mu_1(B)}$. The density h_B of μ_2 with respect to Lebesgue measure on B is strictly positive [49, Lemma 2.3] and analytic [49, Corollary 5.1.1]. Let $\mathcal{Q}_n := \bigvee_{k=0}^{n-1} \mathcal{T}_2^{-k} \mathcal{Q}$.

We have the following expansion and distortion properties.

Proposition 1.4 ([49, Lemma 2.2]). There exist $C > 1, \theta > 1$ and D_1, D_2 such that for any $n \geq 1$ and any x, y in the same element of $\mathcal{Q} \in \mathcal{Q}_n$:

- (1) $d(\mathcal{T}_2^n x, \mathcal{T}_2^n y) \geq C\theta^n d(x, y)$;
- (2) $|\log \frac{\text{Jac}(\mathcal{T}_2^n)(x)}{\text{Jac}(\mathcal{T}_2^n)(y)}| \leq D_1 d(\mathcal{T}_2^n x, \mathcal{T}_2^n y)$;
- (3) $\frac{1}{D_2} \leq \mu_2(A) |\text{Jac}(\mathcal{T}_2^n)(x)| \leq D_2$ for all $x \in A \in \mathcal{Q}_n$;
- (4) $\mathcal{T}_2^n : \mathcal{Q} \rightarrow B$ is a diffeomorphism.

Remark 1.5. Since there exists $c > 0$ such that $c^{-1} \leq h_B \leq c$, we can also state the above point (3) using Lebesgue measure m instead of μ_2 (or more accurately, the product of the counting measure on \mathcal{R} and Lebesgue measure on Δ , even though we will always refer to this measure as Lebesgue).

1.6. Gibbs–Markov maps and their transfer operators

The previous subsection motivates a more in depth study of the following class of maps.

Let (Y, d) be a compact metric space endowed with a probability measure m with full support. Let $T : Y \rightarrow Y$ be a non-singular measurable map.

We will say that T is a Gibbs–Markov map if there exists a countable measurable partition $\mathcal{Q} = \{Y_i\}_{i \in \mathcal{I}}$ of Y such that, if we denote by $\mathcal{Q}_n = \bigvee_{k=0}^{n-1} T^{-k} \mathcal{Q}$ the dynamical partition of T^n and by $\text{Jac}(T^n)$ the Jacobian of T^n with respect to m (i.e. $m(T^n A) = \int_A \text{Jac}(T^n) dm$ for every subset $A \subset Y$ on which T^n is injective), we have

- (1) $T^n : \mathcal{Q} \rightarrow Y$ is a bimeasurable bijection;
- (2) $d(T^n x, T^n y) \geq C\theta^n d(x, y)$;
- (3) $|\log \frac{\text{Jac}(T^n)(x)}{\text{Jac}(T^n)(y)}| \leq Dd(T^n x, T^n y)$;

for all $n \geq 1$, all $\mathcal{Q} \in \mathcal{Q}_n$ and all $x, y \in \mathcal{Q}$, where $C, D > 0$ and $\theta > 1$ depend only on the map T .

It is well known such maps admit a spectral gap for their transfer operators on the space of Hölder functions. We will study spectral properties on a larger space which contains discontinuous functions, namely the quasi-Hölder space, introduced by Keller [35] and Saussol [51]. We recall the relevant definitions and properties, and refer to the aforementioned references for more details.

Let $\varepsilon_0 > 0, 0 < \alpha < 1$ and $f : Y \rightarrow \mathbb{R}$ lie in $L_m^1(Y)$. We define the oscillation of f on a Borel subset $S \subset Y$ by

$$\text{osc}(f, S) = \text{ess sup}_S f - \text{ess inf}_S f.$$

We define

$$|f|_\alpha := \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int_Y \text{osc}(f, B_\varepsilon(x)) dm(x)$$

and let $V_\alpha(Y) := \{f \in L^1_m(Y, \mathbb{R}) : |f|_\alpha < \infty\}$. This space is strictly larger than the space of Hölder functions of exponent α on Y and in particular contains characteristic functions of some measurable sets. If we define the norm $\|\cdot\|_\alpha := |\cdot|_\alpha + \|\cdot\|_{L^1_m}$ then $V_\alpha(Y)$ is a Banach space. Since Y is compact, the space $V_\alpha(Y)$ is compactly embedded in $L^1_m(Y)$. Furthermore, $V_\alpha(Y)$ embeds continuously into $L^\infty_m(Y)$ and is a Banach algebra satisfying $|fg|_\alpha \leq |f|_\alpha \|g\|_\infty + \|f\|_\infty |g|_\alpha$ for all $f, g \in V_\alpha(Y)$.

Note also that while $\|\cdot\|_\alpha$ depends on the choice of ε_0 , the space $V_\alpha(Y)$ does not, and two different ε_0 give rise to two equivalent norms on V_α .

Let P denote the transfer operator of T with respect to m . This is the L^1 adjoint of T with respect to L^∞ , i.e. $\int_Y P\phi\psi dm = \int_Y \phi\psi \circ T dm$ for all $\phi \in L^1_m(Y)$ and $\psi \in L^\infty_m(Y)$.

The operator P has the form

$$P\phi(x) = \sum_{i \in \mathcal{I}} \frac{\phi(x_i)}{\text{Jac}(T)(x_i)},$$

where $x_i \in Y_i$ satisfies $Tx_i = x$.

We will prove that the transfer operator P of a Gibbs–Markov map T is quasi-compact and admits a spectral gap on $V_\alpha(Y)$, from which it will follow exponential decay of correlations for T , for observables in $V_\alpha(Y)$. Our main tool will be a Lasota–Yorke type inequality (Lemma 1.7) and Hennion’s theorem [29]. We refer to Baladi [7] for a systematic exposition of this approach.

The next technical lemma will also prove useful later. In order to state it, we need some more notations. For $Q \in \mathcal{Q}_n$, denote $I_{n,Q} : Y \rightarrow Q$ the inverse branch of the restriction of T^n to Q . The transfer operator P^n of T^n has the form

$$P^n\phi(x) = \sum_{Q \in \mathcal{Q}_n} g_n(I_{n,Q}x)\phi(I_{n,Q}x),$$

where $g_n = \frac{1}{\text{Jac}(T^n)}$.

Denote by $M_{n,Q}$ the operator defined on $L^1_m(Y)$ by

$$M_{n,Q}\phi(x) = g_n(I_{n,Q}x)\phi(I_{n,Q}x).$$

Lemma 1.6. *There exists $C > 0$ such that for any $n \geq 1$, $Q \in \mathcal{Q}_n$ and $\phi \in V_\alpha(Y)$, we have $\|M_{n,Q}\phi\|_{L^1_m} = \int_Q |\phi| dm$ and*

$$\int_Y \text{osc}(M_{n,Q}\phi, B_\varepsilon(x)) dm(x) \leq C \int_Q \text{osc}(\phi, B_{c_{n,Q}\varepsilon}(x)) dm(x) + C\varepsilon \int_Q |\phi| dm,$$

where $c_{n,Q}$ is the Lipschitz constant of $I_{n,Q} : Y \rightarrow Q$.

Proof. The relation $\int_Y |M_{n,Q}\phi| dm = \int_Q |\phi| dm$ follows from a change of variables.

Observe that $\text{osc}(M_{n,Q}\phi, B_\varepsilon(x)) = \text{osc}(g_n\phi, I_{n,Q}B_\varepsilon(x))$. Using [51, Proposition 3.2(iii)], we have for all $x \in Y$,

$$\begin{aligned} \text{osc}(M_{n,Q}\phi, B_\varepsilon(x)) &\leq \text{osc}(\phi, I_{n,Q}B_\varepsilon(x)) \text{ess sup}_{I_{n,Q}B_\varepsilon(x)} g_n \\ &\quad + \text{osc}(g_n, I_{n,Q}B_\varepsilon(x)) \text{ess inf}_{I_{n,Q}B_\varepsilon(x)} |\phi|. \end{aligned}$$

By the distortion control of assumption (3), we have $\text{ess sup}_{I_n, Q B_\varepsilon(x)} g_n \leq C g_n(I_n, Qx)$ and, since $|e^t - e^s| \leq |t - s|e^t$ for any $t, s \in \mathbb{R}$,

$$\begin{aligned} \text{osc}(g_n, I_n, Q B_\varepsilon(x)) &\leq \text{ess sup}_{y, z \in I_n, Q B_\varepsilon(x)} \left| \exp \log g_n(y) - \exp \log g_n(z) \right| \\ &\leq \text{ess sup}_{y, z \in I_n, Q B_\varepsilon(x)} \left| \log \frac{g_n(y)}{g_n(z)} \right| g_n(y) \\ &\leq D g_n(I_n, Qx) \text{ess sup}_{y, z \in I_n, Q B_\varepsilon(x)} d(T^n y, T^n z) \\ &\leq C g_n(I_n, Qx) \varepsilon \end{aligned}$$

for some constant $C > 0$. We also have $\text{osc}(\phi, I_n, Q B_\varepsilon(x)) \leq \text{osc}(\phi, B_{c_n, Q\varepsilon}(I_n, Qx))$ and $\text{ess inf}_{I_n, Q B_\varepsilon(x)} |\phi| \leq |\phi(I_n, Qx)|$ for almost every $x \in Y$. Putting together all the above estimates, we get for almost every x ,

$$\text{osc}(M_n, Q\phi, B_\varepsilon(x)) \leq C \text{osc}(\phi, B_{c_n, Q\varepsilon}(I_n, Qx)) g_n(I_n, Qx) + C\varepsilon |\phi(I_n, Qx)| g_n(I_n, Qx).$$

After integration over Y , a change of variables finishes the proof. □

With this lemma, we can prove a Lasota–Yorke type inequality for T .

Lemma 1.7. *If ε_0 is sufficiently small then there exist $0 < \eta < 1$ and $C, D > 0$ such that if $\phi \in V_\alpha(Y)$ then for all $n \geq 0$,*

$$\|P^n \phi\|_\alpha \leq C \eta^n \|\phi\|_\alpha + D \int_Y |\phi| dm.$$

Proof. Since P^n is a contraction on $L_m^1(Y)$ (see for instance Baladi [7]), it is sufficient to estimate $|P^n \phi|_\alpha$. We will next apply Lemma 1.6 to this operator, first noting that by assumption (2), $c_n, Q \leq C\theta^{-n} \leq C$, where $\theta > 1$. Writing $P^n = \sum_{Q \in \mathcal{Q}_n} M_n, Q$ and summing all the relations from Lemma 1.6, [51, Proposition 3.2(i)] then implies that

$$\begin{aligned} \int_Y \text{osc}(P^n \phi, B_\varepsilon(x)) dm(x) &\leq C \int_Y \text{osc}(\phi, B_{C\theta^{-n}\varepsilon}(x)) dm(x) + C\varepsilon \|\phi\|_{L_m^1} \\ &\leq C\varepsilon^\alpha (\theta^{-\alpha n} |\phi|_\alpha + \varepsilon_0^{1-\alpha} \|\phi\|_{L_m^1}) \end{aligned}$$

for all $0 < \varepsilon \leq \frac{\varepsilon_0}{C} = \varepsilon_1$, so that $C\theta^{-n}\varepsilon \leq \varepsilon_0$ and the bound

$$\int_Y \text{osc}(\phi, B_{C\theta^{-n}\varepsilon}(x)) dm(x) \leq C\varepsilon^\alpha \theta^{-\alpha n} |\phi|_\alpha$$

holds.

This shows $\|P^n \phi\|_{\alpha, \varepsilon_1} \leq C\theta^{-\alpha n} \|\phi\|_{\alpha, \varepsilon_0} + C\|\phi\|_{L_m^1}$, where we put the subscript ε_0 or ε_1 in the notation for the quasi-Hölder norm to emphasize the fact it was defined using either ε_0 or ε_1 , and concludes the proof since the two norms $\|\cdot\|_{\alpha, \varepsilon_0}$ and $\|\cdot\|_{\alpha, \varepsilon_1}$ are equivalent. □

Classical arguments then allow us to prove exponential decay of correlations in the quasi-Hölder norm:

Proposition 1.8. *There exists an unique absolutely continuous probability measure μ which is T -invariant, and its density h belongs to $V_\alpha(Y)$. Furthermore, we have*

- (a) $\|P^n \phi - (\int_Y \phi dm)h\|_\alpha \leq C\theta^n \|\phi\|_\alpha$;
- (b) $|\int_Y \phi \psi \circ T^n d\mu - \int_Y \phi d\mu \int_Y \psi d\mu| \leq C\theta^n \|\phi\|_\alpha \|\psi\|_{L_\mu^1}$;

for all $n \geq 1$, for all $\phi \in V_\alpha$ and $\psi \in L^1(\mu)$, for some constants $C > 0$ and $\theta < 1$ which depend only on the map T .

Proof. Lemma 1.7 implies by Hennion’s theorem [29] that P is quasi-compact and has an essential spectral radius strictly less than 1 when acting on the space $V_\alpha(Y)$. To prove (a), it is then sufficient to prove that 1 is a simple eigenvalue of P , and that there is no other eigenvalue on the unit circle. Let then $\phi \in V_\alpha$ be an eigenvector of P for the eigenvalue $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. From standard results, see for instance Aaronson [1], we know that P has an essential spectral radius strictly less than 1 when acting on the space of Lipschitz functions. This shows that ϕ is itself Lipschitz continuous, and then ϕ is a multiple of h and $\lambda = 1$.

We now prove point (b):

$$\begin{aligned} \int_Y \phi \psi \circ T^n d\mu - \int_Y \phi d\mu \int_Y \psi d\mu &= \int_Y \phi h \psi \circ T^n dm - \int_Y \phi d\mu \int_Y \psi d\mu \\ &= \int_Y \left(P^n(\phi h) - \left(\int_Y \phi h dm \right) h \right) \psi dm. \end{aligned}$$

Then, $|\int_Y \phi \psi \circ T^n d\mu - \int_Y \phi d\mu \int_Y \psi d\mu| \leq \|P^n(\phi h) - (\int_Y \phi h dm)h\|_{L_m^\infty} \|\psi\|_{L_m^1}$. By (a), we have that $\|P^n(\phi h) - (\int_Y \phi h dm)h\|_{L_m^\infty} \leq C\theta^n \|\phi\|_\alpha$ since $V_\alpha(Y)$ embeds into L_m^∞ and is a Banach algebra. On the other hand, $\|\psi\|_{L_m^1} \leq c^{-1} \|\psi\|_{L_\mu^1}$ where $c = \inf h$ is strictly positive by Lemma 4.4.1 in [1]. This proves (b). \square

2. Borel–Cantelli lemmas

2.1. Borel–Cantelli lemmas for Gibbs–Markov maps

We first investigate Borel–Cantelli lemmas for the map \mathcal{T}_2 . From Proposition 1.4, we know \mathcal{T}_2 is a Gibbs–Markov map, so we will present general results for this class of maps.

Our result for Gibbs–Markov maps is a fairly straightforward consequence of earlier work (see for example [36, Theorem 2.1], [25, Proposition 2.6]) and the description of their transfer operators we give in the previous subsection.

Proposition 2.1. *Let T be a Gibbs–Markov map on the compact metric space (Y, d) , as in the previous subsection, with absolutely continuous invariant measure μ . Let $\{\phi_n\}$ be a sequence of positive functions on Y such that there exists a constant $K > 0$ with $\|\phi_n\|_\alpha \leq K$ for all n . Let $E_n = \sum_{j=1}^n \mu(\phi_j)$ and suppose E_n is unbounded. Then*

$$\lim_{n \rightarrow \infty} \frac{1}{E_n} \sum_{j=1}^n \phi_j \circ T^j(x) \rightarrow 1$$

for μ a.e. $x \in Y$.

The proof of this proposition, given below, is an easy consequence of a Gal–Koksma type law. We formulate this law as a proposition of W. Schmidt [52,53] as stated by Sprindzuk [54]:

Proposition 2.2. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and let $f_k(\omega)$ ($k = 1, 2, \dots$) be a sequence of non-negative μ measurable functions and g_k, h_k be sequences of real numbers such that $0 \leq g_k \leq h_k \leq 1$ ($k = 1, 2, \dots$). Suppose there exists $C > 0$ such that*

$$\int \left(\sum_{m < k \leq n} (f_k(\omega) - g_k) \right)^2 d\mu \leq C \sum_{m < k \leq n} h_k \tag{*}$$

for arbitrary integers $m < n$. Then for any $\varepsilon > 0$

$$\sum_{1 \leq k \leq n} f_k(\omega) = \sum_{1 \leq k \leq n} g_k + O(\theta^{1/2}(n) \log^{3/2+\varepsilon} \theta(n))$$

for μ a.e. $\omega \in \Omega$, where $\theta(n) = \sum_{1 \leq k \leq n} h_k$.

Proof of Proposition 2.1. In Proposition 2.2 take $f_k = \phi_k \circ T^k$, $g_k = h_k = \mu(\phi_k)$ and, using part (b) of Proposition 1.8, calculate

$$\begin{aligned} & \left| \sum_{i=m}^n \sum_{j=i+1}^n \int \phi_j \circ T^j \phi_i \circ T^i d\mu - \mu(\phi_j)\mu(\phi_i) \right| \\ &= \left| \sum_{i=m}^n \sum_{j=i+1}^n \int \phi_j \circ T^{j-i} \phi_i - \mu(\phi_j)\mu(\phi_i) \right| \\ &\leq \sum_{i=m}^n \sum_{j=i+1}^n C_1 \theta^{j-i} \|\phi_j\|_\alpha \|\phi_i\|_{L^1_\mu} \\ &\leq C_2 \sum_{i=m}^n \|\phi_i\|_{L^1_\mu}. \end{aligned}$$

The result follows immediately from Proposition 2.2. □

Remark 2.3. For any measurable set $A \subset Y$, we have $\|\mathbb{1}_A\|_\alpha \leq m(A) + \sup_{0 < \varepsilon \leq \varepsilon_0} \frac{m(B_\varepsilon(\partial A))}{\varepsilon^\alpha}$. Hence, any sequence of sets (A_n) such that for some $0 < \alpha \leq 1$,

$$\sup_n \sup_{0 < \varepsilon \leq \varepsilon_0} \frac{m(B_\varepsilon(\partial A_n))}{\varepsilon^\alpha} < \infty$$

and $\sum_n \mu(A_n) = \infty$ will satisfy the strong Borel–Cantelli property. In particular, the sequence does not need to be decreasing.

Remark 2.4. As a direct consequence, we get for the Morita–Pollicott renormalization map $\mathcal{T}_2 : B \rightarrow B$ the strong Borel–Cantelli for any sequence of positive functions (f_n) on B bounded in the space $V_\alpha(B)$ for some $0 < \alpha \leq 1$, with $\sum_n \int f_n d\mu_2 = \infty$. Indeed, by Proposition 1.4, this map is Gibbs–Markov with respect to the partition $\mathcal{Q} = \{B_i\}_{i \in \mathcal{I}}$. This applies in particular to any sequences of balls $(B_{r_n}(p_n))$ with $\sum_n \mu_2(B_{r_n}(p_n)) = \infty$, since such sequences satisfy the condition of Remark 2.3 for $\alpha = 1$.

2.2. Borel–Cantelli lemmas for a class of non-uniformly expanding maps

We now turn to investigate Borel–Cantelli lemmas for the Rauzy–Veech–Zorich renormalization map \mathcal{T}_1 .

Remark 2.5. Note that by Haydn et al. [27, Theorem 6.1] or by Galatolo [18, Lemma 6, Lemma 7] if $\{U_n\}$ is a sequence of balls in $\Delta_{\pi, \varepsilon}$, $\varepsilon \in \{0, 1\}$, satisfying $\mu_1(U_n) \geq \frac{C}{n}$ then $\mathcal{T}_1^{2n}(p) \in U_n$ i.o. for μ_1 a.e. $p \in \Delta_{\pi, \varepsilon}$ since $(\mathcal{T}_1^2, \mathcal{R} \times \Delta, \mu_1)$ has exponential decay of correlations for Lipschitz functions [4]. We are interested in obtaining quantitative rates for this almost sure result.

We first proceed to identify a class of maps containing \mathcal{T}_1 for which such results hold.

Let (X, d) be a bounded, locally compact and separable metric space, with a Borel finite positive measure m . Let $T : X \rightarrow X$ be a non-singular transformation for which m is ergodic.

Suppose there exists a compact subset $Y \subset X$ with $m(Y) > 0$ (without loss of generality, we can assume $m(Y) = 1$) and a countable measurable partition $\mathcal{Q} = \{Y_i\}_{i \in \mathcal{I}}$ of Y such that the first return time

$$r(y) = \inf\{n \geq 1 : T^n y \in Y\}$$

of T to Y is constant on each Y_i , and the first return map $\widehat{T} = T^r : Y \rightarrow Y$ is Gibbs–Markov with respect to the partition \mathcal{Q} . We also assume the first return time is integrable with respect to m : $\int_Y r dm < \infty$.

We will refer to such systems as non-uniformly expanding maps, even though more general definitions exist in the literature.

Under these assumptions, there exists a unique absolutely continuous with respect to m probability measure μ which is T -invariant, and the system (X, T, μ) is ergodic. The existence follows directly from the existence of such a measure for the first return map \widehat{T} and the integrability of r , while the uniqueness is ensured by [1, Theorem 1.5.6].

We will deduce a strong Borel–Cantelli property for decreasing sequences of functions supported in Y from our result for Gibbs–Markov maps and the following result of Kim [36, Theorem 3.1]:

Theorem 2.6. *Let (X, T, μ) be an ergodic measure-preserving transformation, and let $T_E : E \rightarrow E$ be the first return map to a set E of positive μ -measure. Let (f_n) be a decreasing sequence of non-negative functions supported in E such that $\sum_n \int f_n d\mu = \infty$. If every subsequence (f_{n_k}) with $\sum_k \int f_{n_k} d\mu = \infty$ is strong Borel–Cantelli with respect to T_E , then (f_n) is strong Borel–Cantelli with respect to T .*

As an immediate corollary of Proposition 2.1 and Theorem 2.6, we get:

Theorem 2.7. *Let (X, T, μ) be a non-uniformly expanding system as described above, with induced set Y . Then any sequence (f_n) of positive functions, supported in Y , bounded in $V_\alpha(Y)$ for some $0 < \alpha \leq 1$, with $\sum_n \int_Y f_n d\mu = \infty$, satisfies the strong Borel–Cantelli property.*

As seen in Section 1.5, the Rauzy–Veech–Zorich renormalization map is a non-uniformly expanding map, with induced set B . Since by Remark 1.3, for any $p^* = (\pi, \lambda)$ satisfying the Keane condition, we can find a good induced set B that contains p , we obtain:

Theorem 2.8. *Let $U_n \subset \mathcal{R} \times \Delta$ be a decreasing sequence of balls, shrinking to a point p^* which satisfies the Keane condition, such that $E_n := \sum_{j=1}^n \mu_1(U_j)$ diverges. Then, for μ_1 almost every $p \in \mathcal{R} \times \Delta$*

$$\frac{1}{E_n} \sum_{j=1}^n \mathbb{1}_{U_j} \circ \mathcal{T}_1^j(p) \rightarrow 1.$$

Proof. Set $f_n = \mathbb{1}_{U_n}$. By the discussion above, for n large enough, f_n will be supported in some fixed good induced set B . Since, as in Remark 2.3, (f_n) is bounded in $V_\alpha(B)$, it follows from Theorem 2.7 that (f_n) is strong Borel–Cantelli with respect to \mathcal{T}_1 . \square

Remark 2.9. *This result remains true for any decreasing sequence of sets U_n shrinking to a point p^* as soon as the boundaries of these sets are sufficiently regular to ensure the condition of Remark 2.3 is satisfied.*

We now consider more general, non necessarily decreasing, sequences of functions supported in the induced set Y . We will require additional properties for the non-uniformly expanding system, and we will see later they are satisfied by the Rauzy–Veech–Zorich map.

We set $C_n = \{r = n\} \subset Y$. This set is a disjoint union of elements of \mathcal{Q} : we have $C_n = \bigcup_{i \in \mathcal{I}_n} Y_i$, where $\mathcal{I}_n = \{i \in \mathcal{I} : r|_{Y_i} \equiv n\}$.

Definition 2.10. *Let T be a non-uniformly expanding map. We say T is good if*

- (1) (X, T, μ) is mixing;
- (2) $m(r > n) \leq C\gamma^n$;
- (3) $c_i \leq C\gamma^n$, for all $n \geq 1$ and $i \in \mathcal{I}_n$;

for some $C > 0$ and $\gamma < 1$, where $c_i = c_{1, Y_i}$ is the Lipschitz constant of $I_i = I_{1, Y_i} : Y \rightarrow Y_i$, the inverse branch of \widehat{T} restricted to Y_i .

Note that (X, T, μ) is mixing if and only if $\gcd\{r|_{Y_i} : i \in \mathcal{I}\} = 1$, see e.g. [61].

Under these assumptions, we have the following result for the decay of correlations of (X, T, μ) for observables supported in Y :

Theorem 2.11. *If T is a good non-uniformly expanding map, there exist $0 < \kappa < 1$ and $C > 0$ such that for all $\phi \in V_\alpha(Y)$ and all $\psi \in L^1(\mu)$ supported in Y ,*

$$\left| \int_X \phi \psi \circ T^n d\mu - \int_X \phi d\mu \int_X \psi d\mu \right| \leq C \kappa^n \|\phi\|_\alpha \|\psi\|_{L^1_\mu}.$$

This theorem has the following corollary:

Corollary 2.12. *Let T be a good non-uniformly expanding map. Suppose $\{\phi_n\}$ is a sequence of positive functions with support in Y bounded in $V_\alpha(Y)$ with $E_n := \sum_{j=1}^n \mu(\phi_j)$ divergent. Then*

$$\frac{1}{E_n} \sum_{j=1}^n \phi_j \circ T^j(x) \rightarrow 1$$

for μ a.e. $x \in X$.

Proof. We will use Proposition 2.2. Take $f_k = \phi_k \circ T^k$ and $h_k = g_k = \mu(\phi_k)$. A rearrangement of terms shows that it suffices to show

$$\sum_{i=m}^n \sum_{j=i+1}^n \mu(\phi_j \circ T^{j-i} \phi_i) - \mu_1(\phi_j) \mu(\phi_i) \leq C \sum_{i=m}^n \mu(\phi_i).$$

But $|\mu(\phi_j \circ T^{j-i} \phi_i) - \mu(\phi_j) \mu(\phi_i)| \leq C \kappa^{j-i} \|\phi_i\|_{L^1_\mu}$ which yields the result as $\sum_{j>i} \kappa^{j-i}$ is summable. □

To prove Theorem 2.11, we will use operator renewal theory, in the spirit of Sarig [50] and Gouëzel [20], even though in our situation of exponential tails for the return time, the situation is easier. We will make use of the following proposition:

Proposition 2.13 ([22, Proposition 3.4]). *Let \mathcal{Q} be a Banach space and suppose $(R_n)_{n \geq 1}$ is a sequence of bounded operators on \mathcal{Q} . Assume that $\|R_n\| = O(\theta^n)$ for some $0 < \theta < 1$. Hence $R(z) = \sum R_n z^n$ and $R'(z) = \sum n R_n z^{n-1}$ are well-defined operators on \mathcal{Q} for z in the unit complex disc $\bar{\mathcal{D}}$. Assume 1 is a simple isolated eigenvalue of $R(1)$ and the eigenprojector Π satisfies $\Pi R'(1) \Pi = \gamma \Pi$ for some $\gamma \neq 1$ and that $I - R(z)$ is invertible for all $z \in \bar{\mathcal{D}} \setminus \{1\}$. Let $V_n = \sum_{l=1}^\infty \sum_{k_1+\dots+k_l=n} R_{k_l} \circ \dots \circ R_{k_1}$. Then V_n is a bounded linear operator on \mathcal{Q} and $\|V_n - \frac{1}{\gamma} \Pi\| = O(\kappa^n)$ for some $0 < \kappa < 1$.*

Let L be the transfer operator associated to the non-uniformly expanding map $T : X \rightarrow X$, defined for $\phi \in L^1(m)$ by

$$L\phi(x) = \sum_{Ty=x} \frac{\phi(y)}{\text{Jac}(T)(y)}.$$

Let P be the transfer operator associated to the first return map $\hat{T} : Y \rightarrow Y$. By the results of Section 1.6, this operator admits a spectral gap on the space $V_\alpha(Y)$.

Let $R_n \phi := \mathbb{1}_Y L^n(\mathbb{1}_{C_n} \phi)$ and $V_n \phi := \mathbb{1}_Y P^n(\mathbb{1}_Y \phi)$. The linear operator R_n corresponds to first returns to Y at time n while V_n considers all points starting in Y which have returned to Y at time n , whether first return or not. The following renewal equation holds:

$$V_n = \sum_{l=1}^\infty \sum_{k_1+\dots+k_l=n} R_{k_l} \circ \dots \circ R_{k_1}.$$

We will show these operators satisfy the three required conditions to apply Proposition 2.13. Recall the definition of a good non-uniformly expanding map from Definition 2.10.

Lemma 2.14. *There exists $0 < \theta < 1$ and $C > 0$ such that $\|R_n\| \leq C\theta^n$.*

Proof. We have $R_n\phi = \sum_{i \in \mathcal{I}_n} \frac{\phi(I_i x)}{\text{Jac}(\widehat{T})(I_i x)}$, whence $R_n = \sum_{i \in \mathcal{I}_n} M_{1, Y_i}$. Thus, by Lemma 1.6, we have

$$\begin{aligned} \|R_n\phi\|_{L_m^1} &\leq \sum_{i \in \mathcal{I}_n} \|M_{1, Y_i}\phi\|_{L_m^1} = \sum_{i \in \mathcal{I}_n} \int_{Y_i} |\phi| dm = \int_{C_n} |\phi| dm \\ &\leq m(C_n)\|\phi\|_{L_m^\infty} \\ &\leq Cm(C_n)\|\phi\|_\alpha, \end{aligned}$$

and

$$\begin{aligned} \int \text{osc}(R_n\phi, B_\varepsilon(x)) dm(x) &\leq \sum_{i \in \mathcal{I}_n} \int \text{osc}(M_{1, Y_i}\phi, B_\varepsilon(x)) dm(x) \\ &\leq C \left(\sum_{i \in \mathcal{I}_n} \int_{Y_i} \text{osc}(\phi, B_{c_i\varepsilon}(x)) dm(x) + \varepsilon \sum_{i \in \mathcal{I}_n} \int_{Y_i} |\phi| dm \right) \\ &\leq C \int_{C_n} \text{osc}(\phi, B_{c^{(n)}\varepsilon}(x)) dm(x) + C\varepsilon \int_{C_n} |\phi| dm, \end{aligned}$$

where $c^{(n)} = \sup_{i \in \mathcal{I}_n} c_i$.

We have

$$\begin{aligned} \int_{C_n} \text{osc}(\phi, B_{c^{(n)}\varepsilon}(x)) dm(x) &\leq \int_B \text{osc}(\phi, B_{c^{(n)}\varepsilon}(x)) dm(x) \leq (c^{(n)})^\alpha \varepsilon^\alpha \|\phi\|_\alpha \\ &\leq (c^{(n)})^\alpha \varepsilon^\alpha \|\phi\|_\alpha \end{aligned}$$

and $\int_{C_n} |\phi| dm \leq m(C_n)\|\phi\|_{L_m^\infty} \leq Cm(C_n)\|\phi\|_\alpha$, whence

$$\|R_n\phi\|_\alpha \leq C((c^{(n)})^\alpha + m(C_n))\|\phi\|_\alpha$$

and similarly for $\|R_n\phi\|_\alpha$. Since $c^{(n)}$ and $m(C_n)$ decay exponentially fast by assumption, one obtains that $\|R_n\| = \mathcal{O}(\theta^n)$ for some $0 < \theta < 1$. □

Lemma 2.15. *$R(1)$ admits 1 as a simple isolated eigenvalue, and the corresponding eigenprojector is given by*

$$\Pi\phi = \left(\int_Y \phi dm \right) \frac{h_Y}{\mu(Y)},$$

where h_Y is the restriction to Y of the density h of the measure μ (and then $\frac{h_Y}{\mu(Y)}$ is the density of the absolutely continuous invariant probability for \widehat{T}).

Furthermore, we have $\Pi R'(1)\Pi = \frac{\Pi}{\mu(Y)}$, so that γ in Proposition 2.13 is equal to $\frac{1}{\mu(Y)}$.

Proof. We note that $R(1) = P$ is the transfer operator of the Gibbs–Markov map \widehat{T} . Consequently, 1 is a simple isolated eigenvalue, and the corresponding eigenprojector is given by the desired formula.

We have

$$\Pi R'(1)\Pi\phi = \left(\frac{\int_Y R'(1)h_Y dm}{\mu(Y)} \right) \left(\frac{\int_Y \phi dm}{\mu(Y)} \right) h_Y,$$

whence $\gamma = \frac{\int_Y R'(1)h_Y dm}{\mu(Y)}$.

Since $R_n\phi = \mathbb{1}_Y L^n(\mathbb{1}_{C_n}\phi) = \mathbb{1}_Y P(\mathbb{1}_{C_n}\phi)$ for any function ϕ , we have

$$\begin{aligned} \int_Y R'(1)h_Y dm &= \sum_n n \int_Y P(\mathbb{1}_{C_n}h_Y) dm = \sum_n n \int_{C_n} h_Y dm \\ &= \sum_n n\mu(C_n) = \int_Y r d\mu = 1 \end{aligned}$$

by Kac’s lemma, and we get $\gamma = \frac{1}{\mu(Y)}$. □

It remains to prove the aperiodicity condition:

Lemma 2.16. *For all $z \in \bar{D} \setminus \{1\}$, $I - R(z)$ is invertible on $V_\alpha(Y)$.*

Proof. We first establish a Lasota–Yorke inequality for the operator $R(z)$. Remark that

$$R(z)^k = \sum_{n_1, \dots, n_k \geq 1} z^{n_1 + \dots + n_k} R_{n_k} \circ \dots \circ R_{n_1},$$

and that

$$R_{n_k} \circ \dots \circ R_{n_1} = \sum_{i_1 \in \mathcal{I}_{n_1}, \dots, i_k \in \mathcal{I}_{n_k}} M_{k, Q_{i_1, \dots, i_k}},$$

where $Q_{i_1, \dots, i_k} \in \mathcal{Q}_k$ is defined by $Q_{i_1, \dots, i_k} = Y_{i_1} \cap \widehat{T}^{-1}Y_{i_2} \cap \dots \cap \widehat{T}^{-(k-1)}Y_{i_k}$. Then, summing all the relations from Lemma 1.6 and noticing that $|z| \leq 1$ and $n_1 + \dots + n_k \geq k$, we have $\|R(z)^k\phi\|_{L_m^1} \leq C|z|^k\|\phi\|_{L_m^1}$ and $|R(z)^k\phi|_\alpha \leq C|z|^k(\theta^{-\alpha k}|\phi|_\alpha + \|\phi\|_{L_m^1})$, arguing as in the proof of Lemma 1.7.

This shows that the spectral radius of $R(z)$ is less than $|z|$, while the essential spectral radius of $R(z)$ is strictly less than 1 if $|z| = 1$, by Hennion’s theorem [29]. Thus, the problem reduces to prove that the relation $R(z)\phi = \phi$, with $|z| = 1$ and $\phi \in V_\alpha(Y)$ implies that $z = 1$ or $\phi = 0$.

Let $|z| = 1$ and $\phi \in V_\alpha(Y)$ non-zero satisfying $R(z)\phi = \phi$, that is $P(z^r\phi) = \phi$. By [46, Proposition 1.1], we deduce that $(\frac{\phi}{h_Y}) \circ \widehat{T} = z^r \frac{\phi}{h_Y}$. Since (X, T, μ) is mixing, and hence weakly mixing, by Proposition A.5 (see the Appendix), we get that $z = 1$, concluding the proof. □

Proof of Theorem 2.11. By Lemmas 2.14, 2.15 and 2.16, we can apply Proposition 2.13 and get $\|V_n - \mu(Y)\Pi\| \leq C\kappa^n$, i.e.

$$\left\| V_n\phi - \left(\int_Y \phi dm \right) h_Y \right\|_\alpha \leq C\kappa^n \|\phi\|_\alpha$$

for all $\phi \in V_\alpha(Y)$.

Let $\phi \in V_\alpha(Y)$ and $\psi \in L^1(\mu)$ supported in Y . We have

$$\int_X \phi\psi \circ T^n dm = \int_X \mathbb{1}_Y L^n(\mathbb{1}_Y\phi)\psi dm = \int_Y (V_n\phi)\psi dm.$$

Since

$$\begin{aligned} \left| \int_Y V_n(\phi)\psi dm - \int_Y \phi dm \int_Y \psi d\mu \right| &= \left| \int_Y \left[V_n\phi - \left(\int_Y \phi dm \right) h_Y \right] \psi dm \right| \\ &\leq \left\| V_n\phi - \left(\int_Y \phi dm \right) h_Y \right\|_\alpha \int_Y |\psi| dm \\ &\leq C\kappa^n \|\phi\|_\alpha \|\psi\|_{L_m^1}, \end{aligned}$$

we get

$$\begin{aligned} \left| \int_X \phi \psi \circ T^n dm - \int_Y \phi dm \int_Y \psi d\mu \right| &\leq C\kappa^n \|\phi\|_\alpha \|\psi\|_{L^1_\mu} \\ &\leq C\kappa^n \|\phi\|_\alpha \|\psi\|_{L^1_\mu}, \end{aligned}$$

as $\|\psi\|_{L^1_\mu} \leq \|h_Y^{-1}\|_{L^\infty} \|\psi\|_{L^1_\mu} \leq C\|\psi\|_{L^1_\mu}$, the density of μ being bounded from below on Y .

The theorem follows by taking ϕh_Y for ϕ , using the fact that $\|\phi h_Y\|_\alpha \leq \|h_Y\|_\alpha \|\phi\|_\alpha \leq C\|\phi\|_\alpha$. □

In order to apply Corollary 2.12 to the Rauzy–Veech–Zorich map, we need mixing, so we will rather consider the map $G = \mathcal{T}_1^2$ restricted to Δ_ε , $\varepsilon = 0, 1$, which admits $\tilde{\mu}_1 = 2\mu_1(\cdot \cap \Delta_\varepsilon)$ as an invariant measure. If the good induced set B is included in Δ_ε , then $\mathcal{T}_2 : B \rightarrow B$ is the first return map of G to B , with associated return time $\tilde{n}_2 = \frac{n_2}{2}$. It has been shown by Avila and Bufetov [4] that the measure of the set $\{n_1 = n\}$ decays exponentially fast with n . To apply Corollary 2.12, it remains to prove the condition on the Lipschitz constants:

Lemma 2.17. *The Lipschitz constant $c_i = c_{i, B_i}$ of $I_i : B \rightarrow B_i$ decays exponentially fast with n : there exist $0 < \gamma < 1$ and $C > 0$ such that $c_i \leq C\gamma^n$ for all $n \geq 1$ and all $i \in \mathcal{I}_n$.*

Proof. By Avila–Bufetov [4], $m(C_n)$ decays exponentially fast. The map $I_n : B \rightarrow Y_i$ is a composition of a linear map $\lambda \rightarrow A\lambda$ followed by $A\lambda \rightarrow \frac{A\lambda}{|A\lambda|_1}$. A is a non-negative matrix and $\frac{\lambda_i}{\lambda_j}$ is bounded for all $\lambda = (\lambda_1, \dots, \lambda_d)$ in B . Hence $1 \geq \frac{|A\lambda|}{\|A\lambda\|} > C > 0$ for all $\lambda \in B$ (this is an observation of Avila and Bufetov [4, page 9]). Furthermore $\frac{|A\lambda'|_1}{|A\lambda|_1} < C$ for all λ, λ' in B by Proposition 1.4. Thus the exponential decay of volume implies that at least one direction contracts exponentially under I_n by a factor $\gamma^{1/d}$ and hence all directions do, this implies $L_n \leq C(\gamma^{1/d})^n$. □

We can then conclude:

Theorem 2.18. *Suppose $\{\phi_n\}$ is a sequence of positive functions with support in B , bounded in $V_\alpha(B)$ with $E_n := \sum_{j=1}^n \tilde{\mu}_1(\phi_j)$ divergent. Then*

$$\frac{1}{E_n} \sum_{j=1}^n \phi_j \circ G^j(x) \rightarrow 1$$

for μ_1 a.e. $x \in \Delta_\varepsilon$.

This theorem applies in particular to sequences of characteristic functions of balls included in B .

3. Extreme Value Laws for \mathcal{T}_1 and \mathcal{T}_2

By expressing \mathcal{T}_2 as a multidimensional piecewise expanding map with exponential decay of correlations with respect to a quasi-Hölder norm versus L^1 we are able to apply results on Extreme Value statistics for such systems. Let $\phi : B \rightarrow \mathbb{R} \cup \{+\infty\}$ be a function, strictly maximized at a point $p_0 \in B$, which is sufficiently regular that for large u the set $\{x \in B : \phi(x) > u\}$ corresponds to a topological ball centered at p_0 . Let

$$M_n(x) := \max\{\phi(x), \phi \circ \mathcal{T}_2(x), \dots, \phi \circ \mathcal{T}_2^n(x)\}.$$

The aim is to show that we have a non-degenerate limit law for M_n , which we think of as a random variable. Since almost surely M_n converges to $\phi(p_0)$, since μ_2 is ergodic, for such a law, we need to rescale our variable. To this end, for each t we define scaling constants $u_n(t)$ by $n\mu_2(\phi > u_n(t)) \rightarrow t$. For example, if $\phi(x) = -\log d(x, p_0)$ then

$u_n(t) = d^{-1}[\log C(d) + \log n - \log t]$ where $C(d)$ is the constant giving the volume of the unit ball in d dimensional Euclidean space (if d is the dimension of B). In fact we may always write $u_n(t)$ in the form

$$u_n(t) = u_n^{\mathcal{T}_2}(t) = \frac{g(t)}{a_n} + b_n$$

for some function $g(t)$ and sequence of constants a_n, b_n . In our example $a_n = d, g(t) = \log C(d) - \log t$ and $b_n = \frac{1}{d} \log n$, where d is the dimension of B . We say that we have an *Extreme Value Law* if the variable M_n under scaling by u_n converges to some non-degenerate distribution. For the classical application of these ideas to i.i.d. processes, see [38]. For more recent applications to dynamical systems, as we have here, see for example [12,15,32].

There is a close connection between rare events point processes (REPP), Extreme Value Laws and hitting times. First we describe what we mean by a compound Poisson process. Let \mathcal{R} be the ring of subsets of \mathbb{R}^+ generated by the semi-ring of subsets of form $[a, b)$ so that an element of $J \in \mathcal{R}$ has the form $J = \bigcup_{i=1}^n [a_i, b_i)$.

Definition 3.1. Let X_1, X_2, \dots , be an i.i.d. sequence of random variables with common exponential distribution of mean $\frac{1}{\theta}$. Let D_1, D_2, \dots be another i.i.d. sequence of random variables, independent of X_i and with distribution function η . We say that N is a compound Poisson process of intensity θ and multiplicity distribution function η if for every $J \in \mathcal{R}$

$$N(J) = \int 1_J d\left(\sum_{i=1}^{\infty} D_i \delta_{X_1 + \dots + X_i}\right),$$

where δ_t is the Dirac measure at t . If $P(D_1 = 1) = 1$ then N is the standard Poisson distribution and for every $t > 0$ the random variable $N([0, t))$ has a Poisson distribution of mean θt .

Remark 3.2. In our applications η will follow a geometric distribution of parameter $\theta \in (0, 1]$ and $\pi(k) := P(D_1 = k) = \theta(1 - \theta)^k$ for every integer $k \geq 0$. In this case the random variable follows a Pólya–Aeppli distribution,

$$P(N([0, t)) = k) = e^{-\theta t} \sum_{i=1}^k \theta^i (1 - \theta)^{k-i} \frac{(\theta t)^i}{i!} \binom{k-1}{i-1}.$$

Define $v_n^{\mathcal{T}_2}(t) := \mu_2(\phi > u_n^{\mathcal{T}_2})^{-1}$ so that $v_n^{\mathcal{T}_2}(t) \sim \frac{n}{t}$. If $J = \bigcup_{i=1}^n [a_i, b_i) \in \mathcal{R}$ and $\gamma > 0$, define $\gamma J = \bigcup_{i=1}^n [\gamma a_i, \gamma b_i) \in \mathcal{R}$.

We define the rescaled REPP $N_n^{\mathcal{T}_2}$ as

$$N_n^{\mathcal{T}_2}(J) := \sum_{j \in v_n^{\mathcal{T}_2} J \cap \mathbb{N}_0} 1_{(\phi \circ \mathcal{T}_2^j > u_n^{\mathcal{T}_2})}. \tag{2}$$

EVLs and limit laws for $N_n^{\mathcal{T}_2}$ for \mathcal{T}_2 follow directly from [6, Proposition 3.3]. We state them here:

Proposition 3.3. Suppose that p_0 satisfies the Keane condition. (1) If p_0 is not a periodic point for \mathcal{T}_2 then $\mu_2\{M_n \leq u_n(t)\} \rightarrow e^{-t}$ and the REPP $N_n^{\mathcal{T}_2}$ converges in distribution to a standard Poisson process N of intensity 1.

(2) If p_0 is a repelling periodic point of prime period k then $\mu_2\{M_n \leq u_n(t)\} \rightarrow e^{-\theta t}$ where $\theta = 1 - |\text{Jac}(D\mathcal{T}_2^{-k})(p_0)|$ and the REPP $N_n^{\mathcal{T}_2}$ converges in distribution to a compound Poisson process N with intensity θ and multiplicity distribution function η given by $\eta(j) = \theta(1 - \theta)^j$ for all integers $j \geq 0$.

Now define $u_n^{\mathcal{T}_1}(t)$ to be so that $n\mu_1(\phi > u_n^{\mathcal{T}_1}) \rightarrow t$ as $n \rightarrow \infty$. Then setting $v_n^{\mathcal{T}_1}(t) := \mu_1(\phi > u_n^{\mathcal{T}_1}(t))^{-1}$, we can define the REPP $N_n^{\mathcal{T}_1}$ by changing all the appearances of \mathcal{T}_2 in (2) to \mathcal{T}_1 . We then have the following corollary.

Corollary 3.4. Suppose that p_0 satisfies the Keane condition. (1) If p_0 is not a periodic point for \mathcal{T}_1 then $\mu_1\{M_n \leq u_n^{\mathcal{T}_1}(t)\} \rightarrow e^{-t}$ and the REPP $N_n^{\mathcal{T}_1}$ converges in distribution to a standard Poisson process N of intensity 1.

(2) If p_0 is a repelling periodic point of prime period k then $\mu_1\{M_n \leq u_n(t)\} \rightarrow e^{-\theta t}$ where $\theta = 1 - |\text{Jac}(D\mathcal{T}_1^{-k})(p_0)|$ and the REPP $N_n^{\mathcal{T}_1}$ converges in distribution to a compound Poisson process N with intensity θ and multiplicity distribution function η given by $\eta(j) = \theta(1 - \theta)^j$ for all integers $j \geq 0$.

Observe that p_0 as above, that is the point where ϕ takes its maximum, we can choose our set B to contain p_0 , so that the result in Proposition 3.3 applies to the corresponding first return map \mathcal{T}_2 . The proof that we can always pass from the result on the first return map (i.e., \mathcal{T}_2 here) to the original case (i.e., for \mathcal{T}_1), which is a simple generalisation of the main result in [28], appears in [17]. Note that the second part was already proved in [16].

4. Return and hitting time statistics

In this section we consider a natural notion of recurrence which, as in [15], is analogous to the EVL perspective in the previous section. Suppose $p_0 \in B$ and U_n is a sequence of balls nested at p_0 . Let $\tau_{2,U}(x) := \min\{n \geq 1 : \mathcal{T}_2^n(x) \in U\}$. We say that \mathcal{T}_2 has *hitting time statistics* to $\{U_n\}$ with distribution $H(t)$ if

$$\lim_{n \rightarrow \infty} \mu_2 \left(x \in B : \tau_{2,U_n}(x) \leq \frac{t}{\mu_2(U_n)} \right) = H(t).$$

We say that \mathcal{T}_2 has *return time statistics* to $\{U_n\}$ with distribution $\tilde{H}(t)$ if

$$\lim_{n \rightarrow \infty} \frac{1}{\mu_2(U_n)} \mu_2 \left(x \in U_n : \tau_{2,U_n}(x) \leq \frac{t}{\mu_2(U_n)} \right) = H(t).$$

There is a large body of literature on this topic: we refer the reader to [2,26] and references therein for further information on this notion of asymptotic recurrence.

As in [15] sets of the form $\{x \in B : M_n \leq u_n(t)\}$ can be rewritten as $\{x \in B : \tau_{2,U_n}(x) \leq \frac{t}{\mu_2(U_n)}\}$, hence the basic part of Proposition 3.3 can be written:

Proposition 4.1. *Suppose that p_0 satisfies the Keane condition. (1) If p_0 is not a periodic point for \mathcal{T}_2 then*

$$\lim_{n \rightarrow \infty} \mu_2 \left(x \in U_n : \tau_{2,U_n}(x) \leq \frac{t}{\mu_2(U_n)} \right) = 1 - e^{-t}.$$

(2) If p_0 is a repelling periodic point of prime period k then

$$\lim_{n \rightarrow \infty} \mu_2 \left(x \in U_n : \tau_{2,U_n}(x) \leq \frac{t}{\mu_2(U_n)} \right) = 1 - e^{-\theta t},$$

where $\theta = 1 - |\text{Jac}(D\mathcal{T}_2^{-k})(p_0)|$

For typical points this was originally proved in [8, Theorem 2.1] and in the periodic case this follows by [16, Corollary 4], but the full dichotomy, covering *all* points, comes from [6]. To convert the results of this proposition from hitting time statistics to return time statistics, we use the main result of [26] which shows that these limits then become $(1 - \theta) + \theta(1 - e^{-\theta t})$ (where we take $\theta = 1$ in the non-periodic case, so nothing changes).

To convert the laws for \mathcal{T}_2 to \mathcal{T}_1 we use [28, Theorem 10.3]. To set up some of the notation here, we suppose that r_U is the first return time to U for the original dynamics, and $r_{Y,U}$ is the first return time for the speeded up (first return map) dynamics on Y and $\mu_U = \mu|_U / \mu(U)$.

Theorem 4.2. *Let $(X, T\mu)$ be an ergodic probability-preserving system and Y be a measurable set with $\mu(Y) > 0$. Assume that $(H_\ell)_\ell$ is a sequence of measurable sets in Y with $\mu(H_\ell) \rightarrow 0$ as $\ell \rightarrow \infty$ and that \tilde{R} is a any random variable with values in $[0, \infty]$. Then*

$$\mu_Y(H_\ell)r_{Y,H_\ell} \xrightarrow{\mu_{H_\ell}} \tilde{R} \quad \text{as } \ell \rightarrow \infty$$

iff

$$\mu(H_\ell)r_{H_\ell} \xrightarrow{\mu_{H_\ell}} \tilde{R} \quad \text{as } \ell \rightarrow \infty.$$

Here $\xrightarrow{\mu_{H_\ell}}$ means pointwise convergence at all continuity points of \tilde{R} where the LHS is considered w.r.t. μ_{H_ℓ} . So for $\tau_{1,U}(x) := \min\{n \geq 1 : \mathcal{T}_1^n(x) \in U\}$, the above result can be interpreted with $\tau_{2,U}$ on H_ℓ being r_{Y,H_ℓ} and $\tau_{1,U}$ on H_ℓ being r_{H_ℓ} , to obtain the statement of Proposition 4.1 for \mathcal{T}_1, μ_1 . As in the previous section, we may assume that the domain B contains p_0 .

5. The Teichmüller flow on the space of translation surfaces

In this section we relate the dynamical structures we described in Section 1 to the Teichmüller flow on the space of translation surfaces. We do not present any new results in this section. We will first introduce invertible versions $\mathcal{R}_0, \mathcal{R}_1$ and \mathcal{R}_2 of the maps presented in Section 1. The key fact we use is that these maps are first return maps for the flow to adapted cross sections, and give a clearer relation to the translation surfaces, which are represented as points in their phase space.

5.1. Translation surfaces: The zippered rectangle construction

Given an irreducible pair $\pi = (\pi_0, \pi_1)$ and a length vector $\lambda \in \mathbb{R}_+^{\mathcal{A}}$, let T_π^+ denote the subset of vectors $\tau = (\tau_a)_{a \in \mathcal{A}} \in \mathbb{R}^{\mathcal{A}}$ such that

$$\sum_{\pi_0(a) \leq k} \tau_a > 0 \quad \text{and} \quad \sum_{\pi_1(a) \leq k} \tau_a < 0$$

for $1 \leq k \leq d - 1$. We say that τ has *type 0* if the total sum $\sum_{a \in \mathcal{A}} \tau_a$ is positive and *type 1* if the total sum is negative.

Next we will use the matrices \mathcal{M} and intervals $I_a^{\pi_\varepsilon}$ defined in Section 1.2. Then given π and $\tau \in T_\pi^+$ we define the height data by $h := -\mathcal{M}\tau$. One can check that $\tau \in T_\pi^+$ implies that each element h_a for $a \in \mathcal{A}$ is strictly positive. Now given (π, λ, τ) , for each $a \in \mathcal{A}$ we can define the rectangles $R_a^{\pi_0} = I_a^{\pi_0} \times [0, h_a] \subset \mathbb{R}^2$ and $R_a^{\pi_1} = I_a^{\pi_1} \times [0, -h_a] \subset \mathbb{R}^2$. We can then form the *translation surface* $M = M(\pi, \lambda, \tau)$ by identifying the top of each rectangle $R_a^{\pi_0}$ with the bottom of the corresponding rectangle $R_a^{\pi_1}$ and then “zipping up” by making a natural identification of pairs of protruding sides of the rectangles: for more details see [57, Chapter 2.7], [59]. The area of $M(\pi, \lambda, \tau)$ can be defined as $\text{area}(\pi, \lambda, \tau) := \lambda \cdot h = \sum_{a \in \mathcal{A}} \lambda_a h_a$. The structure here can be thought of as a Riemann surface with a non-zero holomorphic 1-form or equivalently, as a flat Riemannian metric on a surface with finitely many singularities of conical type and a parallel unit vector field.

Note that the underlying IET here is a first return map of the vertical flow on the translation surface to the interval $[0, \sum_{a \in \mathcal{A}} \lambda_a]$.

Fix \mathcal{R} a Rauzy class. Let

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}(\mathcal{R}) := \{(\pi, \lambda, \tau) \in \mathcal{R} \times \mathbb{R}_+^{\mathcal{A}} \times T_\pi^+\}.$$

We extend the Rauzy–Veech induction map $\hat{\mathcal{T}}_0$ to a map $\hat{\mathcal{R}}_0$ on $\hat{\mathcal{H}}$ by $\hat{\mathcal{R}}_0(\pi, \lambda, \tau) = (\pi', \lambda', \tau')$, where $(\pi', \lambda') = \hat{\mathcal{T}}_0(\pi, \lambda)$ and $\tau' = \Theta^{-1*}(\tau)$ (recall the description of Θ given in Remark 1.1). The height data h' of (π', λ', τ') can be expressed as $h' = \Theta(h)$. Moreover, setting

$$\mathbb{R}_{\pi,\varepsilon}^{\mathcal{A}} := \{\lambda \in \mathbb{R}_+^{\mathcal{A}} : (\pi, \lambda) \text{ has type } \varepsilon\} \quad \text{and} \quad T_{\pi,\varepsilon} := \{\tau \in T_\pi^+ : \tau \text{ has type } \varepsilon\},$$

it can be shown (see e.g. [57, Chapter 2.7]) that:

Proposition 5.1.

- (a) Θ^{-1*} sends T_π^+ injectively inside $T_{\pi'}^+$.

- (b) (Markov) $\widehat{\mathcal{R}}_0(\{\pi\} \times \mathbb{R}_{\pi, \varepsilon}^A \times T_{\pi}^+) = \{\pi'\} \times \mathbb{R}_{\pi'}^A \times T_{\pi', 1-\varepsilon}$.
- (c) Every (π', λ', τ') such that $\sum_{\alpha \in \mathcal{A}} \tau'_{\alpha} \neq 0$ has a unique preimage by $\widehat{\mathcal{R}}_0$.
- (d) If $\widehat{\mathcal{R}}_0(\pi, \lambda, \tau) = (\pi', \lambda', \tau')$ then the areas of $M(\pi, \lambda, \tau)$ and $M(\pi', \lambda', \tau')$ are equal.

5.2. Teichmüller flow

The Teichmüller flow on $\widehat{\mathcal{H}}$ is defined as the induced action $\mathcal{T} = (\mathcal{T}^t)_{t \in \mathbb{R}} : \widehat{\mathcal{H}} \rightarrow \widehat{\mathcal{H}}$ of the diagonal subgroup

$$\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \quad \text{for } t \in \mathbb{R},$$

given by $\mathcal{T}^t(\pi, \lambda, \tau) = (\pi, e^t \lambda, e^{-t} \tau)$. For $c > 0$ we define

$$\mathcal{H}_c := \{(\pi, \lambda, \tau) \in \widehat{\mathcal{H}} : |\lambda| = c\}.$$

The trajectory of a point in $\widehat{\mathcal{H}}$ hits \mathcal{H}_c precisely once. We are looking for transformations from \mathcal{H}_c back to itself of the form $\widehat{\mathcal{R}}_0 \circ \mathcal{T}^t$ for some t . Noticing that if $(\pi', \lambda') = \widehat{\mathcal{R}}_0(\pi, \lambda)$ and (π, λ) is of type ε , then $|\lambda'| = |\lambda|(1 - \frac{\lambda_{\mathcal{A}(1-\varepsilon)}}{|\lambda|})$, we see that the relevant time t is

$$r_0 = r_0(\pi, \lambda) := -\log\left(1 - \frac{\lambda_{\mathcal{A}(1-\varepsilon)}}{|\lambda|}\right), \quad \text{where } (\pi, \lambda) \text{ is of type } \varepsilon.$$

That is to say, we are interested in the map from \mathcal{H}_c to itself given by

$$\mathcal{R}_0 = \widehat{\mathcal{R}}_0 \circ \mathcal{T}^{r_0} : (\pi, \lambda, \tau) \mapsto \widehat{\mathcal{R}}_0(\pi, e^{r_0} \lambda, e^{-r_0} \tau).$$

From now on we restrict ourselves to

$$\mathcal{H} = \mathcal{H}_1.$$

Then we observe that the map above can actually be interpreted as an extension of the Rauzy–Veech *renormalisation* map \mathcal{T}_0 since $\mathcal{R}_0(\pi, \lambda, \tau) = (\pi', \lambda'', \tau'') = (\mathcal{T}_0(\pi, \lambda), \tau'')$ where

$$(\pi', \lambda', \tau') = \widehat{\mathcal{R}}_0(\pi, \lambda, \tau), \quad \lambda'' = \frac{\lambda'}{1 - \lambda_{\mathcal{A}(1-\varepsilon)}}, \quad \tau'' = \tau'(1 - \lambda_{\mathcal{A}(1-\varepsilon)}).$$

The next result is [57, Corollary 2.24] and [57, Lemma 4.3].

Proposition 5.2. $\mathcal{R}_0 : \mathcal{H} \rightarrow \mathcal{H}$ is an (almost everywhere) invertible Markov map and preserves the area of the corresponding translation surfaces. The standard volume form $m_{\mathcal{H}} = d\pi d\lambda_1 d\tau$, where $d\lambda_1$ is the Lebesgue measure induced on $\Delta_{\mathcal{A}}$ and $d\tau$ is the Lebesgue measure on T_{π}^+ , is invariant under \mathcal{R}_0 .

From now on, we will only consider translation surfaces of area 1, i.e. elements of the set

$$\widehat{\mathcal{H}}_{(1)} := \{(\pi, \lambda, \tau) \in \widehat{\mathcal{H}} : \text{area}(\pi, \lambda, \tau) = 1\}.$$

This set is invariant under both the Teichmüller flow $\mathcal{T} = (\mathcal{T}^t)_{t \in \mathbb{R}}$ and the invertible Rauzy–Veech induction $\widehat{\mathcal{R}}_0$. We also set $\mathcal{H}_{(1)} := \widehat{\mathcal{H}}_{(1)} \cap \mathcal{H}$, which is invariant under the invertible Rauzy–Veech renormalization map \mathcal{R}_0 .

We consider the pre-stratum obtained as the quotient of the fundamental domain $\{(\pi, \lambda, \tau) \in \widehat{\mathcal{H}}_{(1)} : 0 \leq \log |\lambda| \leq r_0(\pi, \lambda)\}$ by the equivalence relation

$$\mathcal{T}^{r_0(\pi, \lambda)}(\pi, \lambda, \tau) \sim \mathcal{R}_0(\pi, \lambda, \tau) \quad \text{for all } (\pi, \lambda, \tau) \in \mathcal{H}_{(1)}.$$

Since \mathcal{R}_0 commutes with the flow, the latter induces a flow $\mathcal{T} = (\mathcal{T}^t)_{t \in \mathbb{R}}$ on the pre-stratum, that we also call Teichmüller flow.

The map $\mathcal{R}_0 : \mathcal{H}_{(1)} \rightarrow \mathcal{H}_{(1)}$ is then naturally identified with the Poincaré return map of this flow to the cross section $\mathcal{H}_{(1)}$. The volume form $m_{\mathcal{H}}$ induces a volume form $m_{\mathcal{H}_{(1)}}$ on $\mathcal{H}_{(1)}$ which is still invariant under \mathcal{R}_0 . The key fact is that $m_{\mathcal{H}_{(1)}}$ gives finite mass to $\mathcal{H}_{(1)}$, a fact which was demonstrated by Veech [55].

5.3. Recoded Teichmüller flow and inducing

The moves described above mean that \mathcal{R}_0 can now be interpreted as the first return map of the Teichmüller flow to $\mathcal{H}_{(1)}$, and indeed it is convenient for us to redefine the flow as a suspension flow which is locally defined by $\mathcal{T}^t(\pi, \lambda, \tau, s) = (\pi, \lambda, \tau, t + s)$ on the space

$$\mathcal{H}_{(1)}^{r_0} := \{(\pi, \lambda, \tau, s) \in \mathcal{H}_{(1)} \times \mathbb{R} : 0 \leq s \leq r_0(\pi, \lambda)\} / \sim,$$

where $(\pi, \lambda, \tau, r_0(\pi, \lambda)) \sim (\pi', \lambda'', \tau'', 0)$ and $\mathcal{R}_0(\pi, \lambda, \tau) = (\pi', \lambda'', \tau'')$. We refer to r_0 as the *roof function* for this suspension flow.

A key fact in Proposition 5.1(b) is that given $(\pi, \lambda, \tau) \in \mathcal{H}_{(1)}$, if (π, λ) is of type ε , then τ' is of type $1 - \varepsilon$. So if the first k iterates $(\pi^j, \lambda^j, \tau^j)$ for $j = 1, \dots, k$ of \mathcal{R}_0 do not change the type of (π^j, λ^j) , then the types of (π^j, λ^j) and τ^j are different (ε and $1 - \varepsilon$) for $j \in \{1, \dots, k\}$. So the first time k that the types of (π^k, λ^k) and τ^k are the same is the first time that (π^k, λ^k) changes type. That is, exactly $n_1(\pi, \lambda)$. Therefore, setting $\mathcal{Z} := \mathcal{Z}_0 \cup \mathcal{Z}_1$, where for $\varepsilon \in \{0, 1\}$,

$$\mathcal{Z}_\varepsilon := \{(\pi, \lambda, \tau) \in \mathcal{H}_{(1)} : (\pi, \lambda) \text{ and } \tau \text{ both have type } \varepsilon\},$$

we define $\mathcal{R}_1 : \mathcal{Z} \rightarrow \mathcal{Z}$ as the first return map by \mathcal{R}_0 to \mathcal{Z} . (We can do this with $\hat{\mathcal{R}}_1$ on $\hat{\mathcal{H}}$ too.) This map can be seen as an extension of the Rauzy–Veech–Zorich renormalisation map for the same reasons as for \mathcal{R}_0 : if $\mathcal{R}_1(\pi, \lambda, \tau) = (\pi', \lambda', \tau')$, then $\mathcal{T}_1(\pi, \lambda) = (\pi', \lambda')$. Thus we can produce a new description of our Teichmüller flow.

We omit the description of this since we go straight to the description given by taking an adapted induced set $B_{\mathcal{H}_{(1)}} \subset \mathcal{Z}$ and the first return map \mathcal{R}_2 to $B_{\mathcal{H}_{(1)}}$ by \mathcal{T} . This map will also be the first return map of \mathcal{R}_0 to $B_{\mathcal{H}_{(1)}}$. The choice of B in Section 1.5 was made in order to ensure uniform expansion for the first return map. Since we are now dealing with an invertible map, we will also need uniform contraction in the stable direction. We follow the construction of [5], and choose a good set B , which is the image of an inverse branch of \mathcal{T}_0 . We refer to [5, Section 4.1.3] for the precise definition of B . This set can be written as $B = \{\pi\} \times \{\frac{\Theta^* \lambda}{|\Theta^* \lambda|} : \lambda \in \Delta_{\mathcal{A}}\}$, where Θ is a finite product of the matrices mentioned in Remark 1.1.

We then set $B_{\mathcal{H}_{(1)}} = (B \times T_B^+) \cap \mathcal{H}_{(1)}$, where T_B^+ is defined by the relation $\Theta^* T_B^+ = T_\pi$, and we consider the first return map \mathcal{R}_2 of \mathcal{R}_0 to $B_{\mathcal{H}_{(1)}}$. This map can be written as a skew product over the first return map \mathcal{T}_2 of \mathcal{T}_0 to the set B , i.e. $\mathcal{R}_2(\pi, \lambda, \tau) = (\pi', \lambda', \tau')$, where $(\pi', \lambda') = \mathcal{T}_2(\pi, \lambda)$, and τ' depends on π, λ and τ .

The map \mathcal{R}_2 preserves the renormalised restriction $m_{B_{\mathcal{H}_{(1)}}}$ of $m_{\mathcal{H}_{(1)}}$ to $B_{\mathcal{H}_{(1)}}$. By [5, Lemma 4.3], this map is a hyperbolic skew product over the uniformly expanding Markov map \mathcal{T}_2 , in the sense of [5, Definition 2.5], and henceforth it admits exponential decay of correlations for Lipschitz observables: there exists $C > 0$ and $0 < \alpha < 1$ such that

$$\left| \int \phi \psi \circ \mathcal{R}_2^n dm_{B_{\mathcal{H}_{(1)}}} - \int \phi dm_{B_{\mathcal{H}_{(1)}}} \int \psi dm_{B_{\mathcal{H}_{(1)}}} \right| \leq C \alpha^n \|\phi\|_{\text{Lip}} \|\psi\|_{\text{Lip}}$$

for all $\phi, \psi \in \text{Lip}$, see Young [60].

Since \mathcal{R}_0 is the Poincaré return map of the flow \mathcal{T} to the section $\mathcal{H}_{(1)}$, the map \mathcal{R}_2 is itself the Poincaré return map of \mathcal{T} to the section $B_{\mathcal{H}_{(1)}}$. This gives a roof function $r_2 : B_{\mathcal{H}_{(1)}} \rightarrow \mathbb{R}_+$ defined almost everywhere. Clearly, the roof function depends only on (π, λ) , so we can reduce it to a roof function $r_2 : B \rightarrow \mathbb{R}_+$. We define the suspension

$$B_{\mathcal{H}_{(1)}}^{r_2} := \{(\pi, \lambda, \tau, s) \in B_{\mathcal{H}_{(1)}} \times \mathbb{R} : 0 \leq s \leq r_2(\pi, \lambda)\} / \sim,$$

where $(\pi, \lambda, \tau, r_2(\pi, \lambda)) \sim (\pi', \lambda'', \tau'', 0)$ and $\mathcal{R}_2(\pi, \lambda, \tau) = (\pi', \lambda'', \tau'')$. Again, we can redefine the flow \mathcal{T} as a suspension flow on $B_{\mathcal{H}_{(1)}}^{r_2}$ given by $\mathcal{T}^t(\pi, \lambda, \tau, s) = (\pi, \lambda, \tau, t + s)$, which preserves the measure $\mu_{\mathcal{T}} = \frac{(m_{B_{\mathcal{H}_{(1)}}} \times m)|_{B_{\mathcal{H}_{(1)}}^{r_2}}}{(m_{B_{\mathcal{H}_{(1)}}} \times m)(B_{\mathcal{H}_{(1)}}^{r_2})}$ where m is the Lebesgue measure on \mathbb{R} .

We now revert to a form which matches Pollicott’s [49] notes as well as corresponds to our sections above. Since the roof function depends only on (π, λ) , we can project into a semi-flow by removing the τ parameter: then the actual flow can be reconstructed as the natural extension of what we have produced. Namely, we let

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

where $(\pi, \lambda, r_2(\pi, \lambda)) \sim (\pi', \lambda'', 0)$ and $\mathcal{T}_2(\pi, \lambda) = (\pi', \lambda'')$. Clearly \mathcal{T}_2 is still a first return map to B . Later we will simplify notation further and write simply $x = (\pi, \lambda)$.

The notation we use for the semi-flow is $\mathcal{F}_t : B^{r_2} \rightarrow B^{r_2}$, defined locally by $\mathcal{F}_t(x, u) = (x, u + t)$, with the relevant identifications i.e. $(x, r_2(\pi, \lambda)) \sim (\mathcal{T}_2(x), 0)$.

The semi-flow $\mathcal{F} = \{\mathcal{F}_t\}_{t \in \mathbb{R}}$ preserves the acip $\mu_{\mathcal{F}}$ given by

$$\mu_{\mathcal{F}} = \frac{(\mu_2 \times m)|_{B^{r_2}}}{(\mu_2 \times m)(B^{r_2})} = \frac{(\mu_2 \times m)|_{B^{r_2}}}{\int r_2 d\mu_2},$$

where μ_2 is the acip for \mathcal{T}_2 and m is the Lebesgue measure on \mathbb{R} .

Remark 5.3. Since \mathcal{T}_2 is a first return map for \mathcal{T}_0 , which in turn is a first return map for our Teichmüller semi-flow, any small ball in B^{r_2} is isomorphic to the corresponding ball in B^{r_0} (B^{r_0} being defined similarly to B^{r_2} , with r_0 as the roof function). More precisely, this is true if our ball is contained in a strip $\{(x, t) : x \in B_k, 0 \leq t \leq r_2(x)\}$ for some k . Recall that $\mathcal{Q} = \{B_i\}_{i \in \mathcal{I}}$ is the natural partition of the map \mathcal{T}_2 defined in Section 1.5.

6. Statistical properties of the Teichmüller flow

In this section we extend our Borel–Cantelli Lemmas and EVLs to the Teichmüller flow.

6.1. Borel–Cantelli lemmas for the semi-flow

Here we will use ideas from the proof of [25, Theorem 2], primarily Step 1 of that proof. The main (obvious) difference is that we are dealing with continuous time.

Given a family of sets $U = (U_s)_{s \geq 0}$ set $\psi = (\psi_s)_{s \geq 0}$ where $\psi_s := \mathbb{1}_{U_s}$ and $E_t(U) = E_t(\psi) = \int_0^t (\int \psi_s d\mu_{\mathcal{F}}) ds$. We say that U is a family of *shrinking sets* if $s_1 < s_2$ implies $U_{s_2} \subset U_{s_1}$. In this section we will prove that if $U = (U_s)_{s \geq 0}$ is a family of shrinking sets with some monotonicity condition and $\lim_{t \rightarrow \infty} E_t(U) = \infty$ then

$$\lim_{t \rightarrow \infty} \frac{1}{E_t(U)} \int_0^t \mathbb{1}_{U_s} \circ \mathcal{F}_s(x, u) ds = 1 \quad \text{for } \mu_{\mathcal{F}}\text{-a.e. } (x, u) \in B^{r_2}.$$

This result is contained in Theorem 6.3; in particular, the smoothness condition is given there. We prove in the following subsection that this condition is indeed satisfied for a natural family of sets, namely nested balls.

Recall that B is partitioned (almost everywhere) into sets $\{B_k\}_k$. For $i \in \mathbb{N}_0$, define

$$B_k^i := \{(x, t) \in B_k \times \mathbb{R}_+ : i \leq t < \min\{i + 1, r_2(x)\}\}.$$

So we can write $B^{r_2} = \bigcup_k \bigcup_i B_k^i$ almost everywhere. We will restrict our Borel–Cantelli lemmas to these sets B_k^i , which will be sufficient to prove the general case. Indeed, we define the restricted indicator function

$$\psi_{B_k^i, s} := \mathbb{1}_{U_s \cap B_k^i}$$

and first study the recurrence properties of the family $\psi_{B_k^i} = (\psi_{B_k^i, s})_{s \geq 0}$. We do this by inducing, for which we need the right time scale. Since $\mu_{\mathcal{F}}$ is ergodic and $\int r_2 d\mu_2 < \infty$, we immediately obtain the following lemma where $\bar{r}_2 := \int r_2 d\mu_2$.

Lemma 6.1. For each $\varepsilon > 0$ there exists $T \geq 0$ and a set $X_{\varepsilon,T} \subset B^{r_2}$ such that $(x, u) \in X_{\varepsilon,T}$ and $t \geq T$ implies

$$\left| \frac{t}{\#\{s \in [0, t) : \mathcal{F}_s(x, u) \in B\}} - \bar{r}_2 \right| < \varepsilon.$$

Moreover, $\mu_{\mathcal{F}}(X_{\varepsilon,T}) \rightarrow 1$ as $T \rightarrow \infty$.

Now, for each $\varepsilon \in \mathbb{R}$, we define the induced function on $x \in B$

$$\bar{\psi}_{n, B_k^i, \varepsilon}(x) := \int_0^{r_2(x)} (\mathbb{1}_{U_n(\bar{r}_2+\varepsilon)+s} \cdot \mathbb{1}_{B_k^i}) \circ \mathcal{F}_s(x, 0) ds, \tag{3}$$

and denote the family as $\bar{\psi}_{B_k^i, \varepsilon} = (\bar{\psi}_{n, B_k^i, \varepsilon})_n$. Note that $\int \bar{\psi}_{n, B_k^i, \varepsilon}(x) d\mu_2 \leq \bar{r}_2 \mu_{\mathcal{F}}(U_n(\bar{r}_2+\varepsilon) \cap B_k^i)$ as $\mu_{\mathcal{F}} = \frac{1}{\bar{r}_2}(\mu_2 \times m)|_{B^{r_2}}$. We will be able to compare the long-term behaviour of this function with different values of ε , and compare them all to the long-term behaviour of the flow. This is necessary as we sample at discrete times, and the nested balls are shrinking in continuous time.

We will use the following lemma, which is [25, Lemma 4.2].

Lemma 6.2. Suppose that $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is decreasing and $\sum_{i=0}^{\infty} g(i) = \infty$. Then,

(a) for all $\varepsilon > 0$ and all $n \geq 0$,

$$\frac{\int_0^{(1+\varepsilon)n} g(t) dt}{\int_0^n g(t) dt} \leq 1 + \varepsilon;$$

(b) $\lim_{n \rightarrow \infty} \frac{\int_0^n g(t) dt}{\sum_{j=0}^{n-1} g(j)} = 1$.

Theorem 6.3. Suppose that B_k^i is such that

$$\lim_{t \rightarrow \infty} \int_0^t \mu_{\mathcal{F}}(U_s \cap B_k^i) ds = \infty,$$

i.e. $\lim_{t \rightarrow \infty} E_t(\psi_{B_k^i}) = \infty$. If there exists $K > 0$ and $0 < \alpha \leq 1$ such that $\|\bar{\psi}_{n, B_k^i, 0}\|_{\alpha} < K$ for all $n \in \mathbb{N}_0$, then

$$\lim_{t \rightarrow \infty} \frac{1}{E_t(\psi_{B_k^i})} \int_0^t \mathbb{1}_{U_s \cap B_k^i} \circ \mathcal{F}_s(x, u) ds = 1 \quad \text{for } \mu_{\mathcal{F}}\text{-a.e. } (x, u) \in B^{r_2}.$$

Proof. We use the idea of Step 1 of the proof of [25, Theorem 2]. We will show

$$\lim_{t \rightarrow \infty} \frac{1}{E_t(\psi_{B_k^i})} \int_0^t \mathbb{1}_{U_s \cap B_k^i} \circ \mathcal{F}_s(x, 0) ds = 1 \quad \text{for } \mu_2\text{-a.e. } (x, 0) \in B,$$

as then the proof for $\mu_{\mathcal{F}}$ -a.e. $(x, u) \in B^{r_2}$ follows.

We already know from Proposition 2.1 that for μ_2 -a.e. $x \in B$,

$$\frac{\sum_{j=0}^{n-1} \bar{\psi}_{j, B_k^i, 0}(\mathcal{T}_2^j x)}{E_n(\bar{\psi}_{B_k^i, 0})} \rightarrow 1 \quad \text{as } n \rightarrow \infty,$$

where $E_n(\bar{\psi}_{B_k^i, 0}) := \sum_{j=0}^{n-1} \mu_2(\bar{\psi}_{j, B_k^i, 0})$. Moreover the fact that $E_n(\bar{\psi}_{B_k^i, 0}) \rightarrow \infty$ is equivalent to the divergence assumption in the statement of our theorem, as the sets are shrinking. Lemma 6.2 controls the effect of this perturbation in the limit when we switch on the ε parameter in one of the occurrences of $\psi_{n, B_k^i, \varepsilon}$ above which deals with

the shrinking of the balls during the flow between returns to the base. Note that $\mu_2(\bar{\psi}_{n, B_k^i, \varepsilon}) \leq \mu_2(\bar{\psi}_{n, B_k^i, -\varepsilon})$ and $\lim_{\varepsilon \rightarrow 0} \frac{\mu_2(\bar{\psi}_{j, B_k^i, \varepsilon})}{\mu_2(\bar{\psi}_{j, B_k^i, -\varepsilon})} \rightarrow 1$ uniformly in n . Thus

$$\lim_{\varepsilon \rightarrow 0} \frac{E_{q(t,x)}(\bar{\psi}_{B_k^i, +\varepsilon})}{E_{q(t,x)}(\bar{\psi}_{B_k^i, -\varepsilon})} \rightarrow 1$$

uniformly in n . We will use these observations to squeeze $\frac{\int_0^{r_2^{q(t,x)}(x)} \psi_{B_k^i, s} \circ \mathcal{F}_s(x, 0) ds}{E_{q(t,x)}(\bar{\psi}_{B_k^i, 0})}$ between two corresponding convergent scaled Birkhoff sums.

Given $x \in B$, define $q(t, x)$ as the integer for which

$$r_2^{q(n,x)}(x) \leq t < r_2^{q(n,x)+1}(x),$$

where $r_2^m(x) = r_2(x) + r_2(\mathcal{T}_2 x) + \dots + r_2(\mathcal{T}_2^{m-1} x)$. Observe that since the difference of the integral of $\mathbb{1}_{U_s \cap B_k^i} \circ \mathcal{F}_s(x, \cdot)$ between times $r_2^{q(t,x)}(x)$ and t is made up by at most one passage through B_k^i which integrates to at most the length of B_k^i in the vertical direction, i.e., 1, we have

$$\int_0^t \psi_{B_k^i, s} \circ \mathcal{F}_s(x, 0) ds - \int_0^{r_2^{q(t,x)}(x)} \psi_{B_k^i, s} \circ \mathcal{F}_s(x, 0) ds \leq 1.$$

Hence this difference is uniformly bounded independently of x and t .

$$\text{Thus } \sum_{j=0}^{q(t,x)-1} \bar{\psi}_{j, B_k^i, \varepsilon}(\mathcal{T}_2^j x) \leq \int_0^{r_2^{q(t,x)}(x)} \psi_{B_k^i, s} \circ \mathcal{F}_s(x, 0) \leq \sum_{j=0}^{q(t,x)-1} \bar{\psi}_{j, B_k^i, -\varepsilon}(\mathcal{T}_2^j x) + 1.$$

So by Lemma 6.1, for all small $\varepsilon > 0$,

$$\begin{aligned} & \left(\frac{\sum_{j=0}^{q(t,x)-1} \bar{\psi}_{j, B_k^i, \varepsilon}(\mathcal{T}_2^j x)}{E_{q(t,x)}(\bar{\psi}_{B_k^i, \varepsilon})} \right) \left(\frac{E_{q(t,x)}(\bar{\psi}_{B_k^i, \varepsilon})}{E_{q(t,x)}(\bar{\psi}_{B_k^i, 0})} \right) \\ & \leq \frac{\int_0^{r_2^{q(t,x)}(x)} \psi_{B_k^i, s} \circ \mathcal{F}_s(x, 0) ds}{E_{q(t,x)}(\bar{\psi}_{B_k^i, 0})} \\ & \leq \left(\frac{\sum_{j=0}^{q(t,x)-1} \bar{\psi}_{j, B_k^i, -\varepsilon}(\mathcal{T}_2^j x) + 1}{E_{q(t,x)}(\bar{\psi}_{B_k^i, -\varepsilon})} \right) \left(\frac{E_{q(t,x)}(\bar{\psi}_{B_k^i, -\varepsilon})}{E_{q(t,x)}(\bar{\psi}_{B_k^i, 0})} \right). \end{aligned}$$

Then Lemmas 6.1, 6.2 and the fact that

$$\lim_{\varepsilon \rightarrow 0} \frac{E_{q(t,x)}(\bar{\psi}_{B_k^i, +\varepsilon})}{E_{q(t,x)}(\bar{\psi}_{B_k^i, -\varepsilon})} \rightarrow 1$$

imply that

$$\lim_{t \rightarrow \infty} \frac{\int_0^{r_2^{q(t,x)}(x)} \psi_{B_k^i, s} \circ \mathcal{F}_s(x, 0) ds}{E_{q(t,x)}(\bar{\psi}_{B_k^i, 0})} = \lim_{t \rightarrow \infty} \frac{\int_0^t \psi_{B_k^i, s} \circ \mathcal{F}_s(x, 0) ds}{E_{q(t,x)}(\bar{\psi}_{B_k^i, 0})} = 1.$$

To complete the proof of the proposition, as in Step 2 of the proof of [25, Theorem 2], we show that

$$\lim_{n \rightarrow \infty} \frac{E_n(\psi_{B_k^i})}{E_{\lfloor n/\bar{r}_2 \rfloor}(\bar{\psi}_{B_k^i, 0})} = 1.$$

Notice that this is the one part where our proof is easier than theirs since the flow is a first return to the base (this also accounts for the fact that Step 3 of that proof is unnecessary here).

By Lemma 6.1, $q(n, x) \sim \lfloor \frac{n}{r_2} \rfloor$. Hence

$$\begin{aligned} E_{\lfloor n/\bar{r}_2 \rfloor}(\bar{\psi}_{B_k^i, 0}) &= \sum_{j=0}^{\lfloor \frac{n}{r_2} \rfloor - 1} \int_B \bar{\psi}_{j, B_k^i, 0}(y) d\mu_2(y) \\ &= \sum_{j=0}^{\lfloor \frac{n}{r_2} \rfloor - 1} \int_B \int_0^{r_2(y)} \psi_{B_k^i, j\bar{r}_2+s} \circ \mathcal{F}_s(y, 0) ds d\mu_2(y) \\ &\sim \sum_{j=0}^{\lfloor \frac{n}{r_2} \rfloor - 1} \mu_{\mathcal{F}}(U_j(\bar{r}_2) \cap B_k^i). \end{aligned}$$

Applying Lemma 6.2 with a speeded up time variable, we obtain $\sum_{j=0}^{\lfloor \frac{n}{r_2} \rfloor - 1} \mu_{\mathcal{F}}(U_j(\bar{r}_2) \cap B_k^i) \sim \int_0^{\frac{n}{r_2}} \mu_{\mathcal{F}}(U_{s\bar{r}_2} \cap B_k^i) \bar{r}_2 ds$, so a change of variables then gives $E_{\lfloor n/\bar{r}_2 \rfloor}(\bar{\psi}_{B_k^i, 0}) \sim E_n(\psi_{B_k^i})$, thus completing the proof. \square

6.2. An application of Theorem 6.3

One of the challenges in proving Borel–Cantelli lemmas when moving from the discrete system to the flow is that the induced characteristic functions are not, in general, characteristic functions. In this subsection we prove that characteristic functions of balls in the flow space induce observables which are sufficiently regular that we can apply Theorem 6.3 to them. In fact the averaging in the flow direction regularizes functions. If $(z, u) \in B^{r_2}$ we let $B_\eta(z, u)$ denote a ball of radius η about (z, u) in the Euclidean metric $d_1((z, u), (z', u')) = [(u - u')^2 + \sum_{j=1}^d (z_j - z'_j)^2]^{\frac{1}{2}}$. It is clear from our proof below other Euclidean metrics may be used, for example $d_2((z, u), (z', u')) = |u - u'| + \sum_{j=1}^d |z_j - z'_j|$.

Theorem 6.4. *Let $\delta(s)$ be a decreasing sequence. For $\mu_{\mathcal{F}}$ -a.e. $(z, u) \in B^{r_2}$ setting $U_s = B_{\delta(s)}(z, u)$, if*

$$\lim_{t \rightarrow \infty} E_t(U) = \infty$$

then

$$\lim_{t \rightarrow \infty} \frac{1}{E_t(U)} \int_0^t \mathbb{1}_{U_s} \circ \mathcal{F}_s(x, v) ds = 1 \quad \text{for } \mu_{\mathcal{F}}\text{-a.e. } (x, v) \in B^{r_2}.$$

Proof. As before we define

$$\psi_{B_k^i, s} := \mathbb{1}_{U_{\delta(s)} \cap B_k^i},$$

where

$$B_k^i := \{(x, t) \in B_k \times \mathbb{R}_+ : i \leq t < \min\{i + 1, r_2(x)\}\}.$$

For large s the ball $B_{\delta(s)}(z, u)$ lies inside a fixed $B_{k^*}^{i^*}$ for some specific k^*, i^* . Since we have freedom to induce on a set B placed anywhere in Δ we need not worry about (z, u) lying on the boundary of a B_k^i .

For $\gamma > 0$ we also define the induced function

$$\psi_n := \psi_{n, B_{k^*}^{i^*}, \gamma}(x) := \int_0^{r_2(x)} (\mathbb{1}_{U_{n(\bar{r}_2+\gamma)+s}} \cdot \mathbb{1}_{B_{k^*}^{i^*}}) \circ \mathcal{F}(x, s) ds.$$

We have to show that there exists an α and a constant K such that $\|\psi_n\|_\alpha < K$ for all n .

It suffices to show that there exist α, K such that

$$\varepsilon^{-\alpha} \int_B \text{osc}(\psi_n, B_\varepsilon(x)) dx < K$$

for all n .

If $\delta(n(\bar{r}_2)) \leq \varepsilon$ then $\text{osc}(\psi_n, B_\varepsilon(x)) \leq 2\varepsilon$. This is because for each $y \in B_\varepsilon(x)$,

$$\int_0^{r_2(x)} (\mathbb{1}_{U_{(n(\bar{r}_2+\gamma)+s)}} \cdot \mathbb{1}_{B_k^i}) \circ \mathcal{F}(x, s) ds \leq \delta(|n(\bar{r}_2)|) \leq \varepsilon.$$

So we need only consider the supremum over small $\varepsilon < \delta(n(\bar{r}_2))$. The ball $B_{\delta(s)}(z, u) \subset B_k^{i*}$ lies in a $d + 1$ -dimensional Euclidean space. Its projection onto the d -dimensional space B is a ball $B_{\delta(s)}(z)$ in B_k^* . If the distance of $B_\varepsilon(x)$ to $B_{\delta(s)}(z)$ is greater than 2ε then either $B_\varepsilon(x)$ is in the exterior of $B_{\delta(s)}(z)$ or $B_{2\varepsilon}(x) \subset B_{\delta(s)}(z)$. In the first case $\int_B \text{osc}(\psi_n, B_\varepsilon(x)) = 0$ as the flow starting in $B_\varepsilon(x)$ does not meet $B_{\delta(s)}(z, u)$. In the second case i.e. $B_\varepsilon(x)$ is bounded away from the boundary of $B_{\delta(s)}(z)$ by ε , then the two parts of the boundary of $B_{\delta(s)}(z, u)$ which project to $B_\varepsilon(x)$ may be written locally as graphs over $B_\varepsilon(x)$, the ‘‘height’’ functions are given by $s - u = \sqrt{\delta(s) - \sum_{j=1}^d (t_j - z_j)^2}$ and $s - u = -\sqrt{\delta(s) - \sum_{j=1}^d (t_j - z_j)^2}$, respectively, where $t = (t_1, \dots, t_d)$ and $z = (z_1, \dots, z_d)$ are Euclidean coordinates in B . Here we are restricting to t satisfying $\sqrt{\sum_{j=1}^d (t_j - x_j)^2} < \varepsilon$ where $x = (x_1, \dots, x_d)$ is the center of $B_\varepsilon(x)$. Note that for both branches $|\frac{\partial s}{\partial t_i}| = \frac{1}{2}(\delta(s) - \sum_{j=1}^d (t_j - z_j)^2)^{-\frac{1}{2}}(2|t_i - z_i|)$. In particular since t satisfying $\sqrt{\sum_{j=1}^d (t_j - x_j)^2} < \varepsilon$ is bounded from the boundary of $B_{\delta(s)}(z)$ by ε , i.e. $\sqrt{\delta(s) - \sum_{j=1}^d (t_j - z_j)^2} > \varepsilon$ we have $|\frac{\partial s}{\partial t_i}| \leq \frac{C}{\sqrt{\varepsilon}}$ for all i and hence the oscillation of ψ_n over $B_\varepsilon(x)$ is $O(\sqrt{\varepsilon})$. Finally if $B_\varepsilon(x)$ is within 2ε of the boundary of $B_{\delta(s)}(z)$ then the oscillation of ψ_n over $B_\varepsilon(x)$ is $O(1)$ but the μ_2 measure of points x within a 2ε neighborhood of the boundary of $B_{\delta(s)}(z)$ is $O(\varepsilon)$.

Thus taking $\alpha = \frac{1}{2}$ there exists K such that

$$\varepsilon^{-\frac{1}{2}} \int_B \text{osc}(\psi_n, B_\varepsilon(x)) dx < K$$

for all n . □

6.3. Borel–Cantelli lemmas for the Teichmüller flow

In this section, we prove Borel–Cantelli lemmas for the Teichmüller flow \mathcal{T} seen as a suspension flow over the map $\mathcal{R}_2 : B_{\mathcal{H}(1)} \rightarrow B_{\mathcal{H}(1)}$ with roof function r_2 .

We first prove a similar result for the map \mathcal{R}_2 . Recall that this map preserves the measure $m_{B_{\mathcal{H}(1)}}$ and is a skew-product over the map $\mathcal{T}_2 : B \rightarrow B$, which preserves μ_2 . To simplify the notations, we set $\mu := \mu_2$ and $\hat{\mu} := m_{B_{\mathcal{H}(1)}}$.

Proposition 6.5. *Let (U_n) be a decreasing sequence of nested balls centered at a point $(x, \tau) \in B_{\mathcal{H}(1)}$, with $\sum_n \hat{\mu}(U_n) = \infty$. Assume there exist $C > 0$ and $\gamma > 0$ such that $\hat{\mu}(U_n) \geq Cn^{-\gamma}$ and $(\log n)\mu(U_n) \leq C$ for all $n \geq 0$. Then the sequence (U_n) is strong Borel–Cantelli for \mathcal{R}_2 .*

Proof. We follow the proof of [62, Theorem 1.5]. Let $f_k = \mathbb{1}_{U_k} \circ \mathcal{R}_2^k$. We denote by $E(\cdot)$ the expectation operator with respect to $\hat{\mu}$. We trivialize $B_{\mathcal{H}(1)}$ to a product via the natural diffeomorphism $B_{\mathcal{H}(1)} \rightarrow B \times \mathbb{P}T_B^+$, where $\mathbb{P}T_B^+$ is the image of T_B^+ in the projective space $\mathbb{P}\mathbb{R}^A$. Let Π_x and Π_τ be the projections on the factors B and $\mathbb{P}T_B^+$ respectively. We denote by m_1 the Lebesgue measure on each factor, and by m_2 the product Lebesgue measure on $B \times \mathbb{P}T_B^+$. The measure $\hat{\mu}$ has a smooth density with respect to m_2 , which is bounded uniformly from above and below. Let $E(\cdot)$ be the expectation operator with respect to the measure $\hat{\mu}$.

For $i < j$, we calculate

$$\begin{aligned}
 E(f_i f_j) &= \int \mathbb{1}_{U_i} \circ \mathcal{R}_2^i \mathbb{1}_{U_j} \circ \mathcal{R}_2^j d\hat{\mu} \\
 &= \int \mathbb{1}_{U_i} \mathbb{1}_{U_j} \circ \mathcal{R}_2^{j-i} d\hat{\mu} \\
 &\lesssim \int_{U_i} \mathbb{1}_{\Pi_x U_i} \mathbb{1}_{\Pi_x U_j} \circ \Pi_x \circ \mathcal{R}_2^{j-i} dm_2 \\
 &\lesssim m_1(\Pi_x U_i) m_1(\Pi_x U_i \cap \mathcal{T}_2^{-(j-i)} \Pi_x U_j) \\
 &\lesssim m_1(\Pi_x U_i) \mu(\Pi_x U_i \cap \mathcal{T}_2^{-(j-i)} \Pi_x U_j) \\
 &\lesssim m_1(\Pi_x U_i) (\mu(\Pi_x U_i) \mu(\Pi_x U_j) + C\theta^{j-i} \mu(\Pi_x U_j)) \\
 &\lesssim m_1(\Pi_x U_i) (m_1(\Pi_x U_i) m_1(\Pi_x U_j) + C\theta^{j-i} m_1(\Pi_x U_j)) \\
 &\lesssim (m_2(U_i))^{\frac{1}{2}} ((m_2(U_i))^{\frac{1}{2}} (m_2(U_j))^{\frac{1}{2}} + C\theta^{j-i} (m_2(U_j))^{\frac{1}{2}}) \\
 &\lesssim (m_2(U_i))^{\frac{3}{2}} + \theta^{j-i} m_2(U_i).
 \end{aligned}$$

Throughout this calculation, we have used the fact that μ and $\hat{\mu}$ have a density with respect to m_1 and m_2 respectively which are bounded uniformly from above and below, decay of correlations for \mathcal{T}_2 given by Proposition 1.8 and the fact that there exists a constant K such that for all ball U , $m_1(U) \leq K(m_2(U))^{\frac{1}{2}}$.

So, using decay of correlations for \mathcal{R}_2 and Lipschitz observables, we have

$$\begin{aligned}
 \sum_{j=i+1}^n (E(f_i f_j) - E(f_i)E(f_j)) &\leq \left(\sum_{j=i+1}^{i+a \log i} + \sum_{j>i+a \log i} \right) [E(f_i f_j) - E(f_i)E(f_j)] \\
 &\lesssim (\log i) (m_2(U_i))^{\frac{3}{2}} + m_2(U_i) \\
 &\quad + \sum_{j>i+a \log i} \alpha^{j-i} \|\tilde{f}_i\|_{\text{Lip}} \|\tilde{f}_j\|_{\text{Lip}},
 \end{aligned}$$

where a will be chosen later and \tilde{f}_i is a Lipschitz approximation to f_i , satisfying $m_2(|\tilde{f}_i - f_i|) \lesssim \frac{1}{i^2}$ and $\|\tilde{f}_i\|_{\text{Lip}} \lesssim i^\kappa$ for some fixed κ . We are able to satisfy both conditions as $m_2(U_i) \gtrsim i^{-\nu}$ for some $\nu > 0$. We have $(\log i)(m_2(U_i))^{\frac{3}{2}} \lesssim m_2(U_i)$ and for $a > 0$ sufficiently large

$$\sum_{j>i+a \log i} \alpha^{j-i} \|\tilde{f}_i\|_{\text{Lip}} \|\tilde{f}_j\|_{\text{Lip}} \lesssim m_2(U_i).$$

We have thus shown that

$$\sum_{i=m}^n \sum_{j=i+1}^n (E(f_i f_j) - E(f_i)E(f_j)) \lesssim \sum_{i=m}^n E(f_i)$$

which implies the strong Borel–Cantelli property by Proposition 2.2. □

Remark 6.6. Note that the proof above does not use the assumption that the balls are nested, nor that they are balls just that they may be approximated by Lipschitz functions \tilde{f}_i such that $m_2(|\tilde{f}_i - f_i|) \lesssim \frac{1}{i^2}$ and $\|\tilde{f}_i\|_{\text{Lip}} \lesssim i^\kappa$ for some fixed κ .

We now show that the SBC property for the map \mathcal{R}_2 implies the SBC property for nested balls U_t in the full suspension flow.

Theorem 6.7. *Let $U = (U_t)_{t \geq 0}$ be a family of shrinking balls in $B_{\mathcal{H}(1)}^{r_2}$, with $\mu_{\mathcal{T}}(U_t) \gtrsim t^{-\gamma}$ for some $\gamma > 0$ and $\sup_{t \geq 0} (\log t) \mu_{\mathcal{T}}(U_t) < \infty$. Assume that*

$$E_t := E_t(U) = \int_0^t \mu_{\mathcal{T}}(U_s) ds$$

diverges.

Then the family U is strong Borel–Cantelli for the flow: for $\mu_{\mathcal{T}}$ a.e. $p \in B_{\mathcal{H}(1)}^{r_2}$,

$$\frac{1}{E_t(U)} \int_0^t \mathbb{1}_{U_s}(\mathcal{T}^t(p)) ds \rightarrow 1.$$

Proof. Note that the measure on the flow $\mu_{\mathcal{T}}$ is the product of the base measure and Lebesgue measure in the flow direction, so that $d\mu_{\mathcal{T}} = d\hat{\mu} \times dt$ and that the projection Π , say, via flow lines of the balls U_t in the suspension flow is a t -parametrized sequence of nested “balls” C_t in the Poincaré section $B_{\mathcal{H}(1)}$. The dynamics of the return map to $B_{\mathcal{H}(1)}$ is given by the skew-product map $\mathcal{R}_2 : B_{\mathcal{H}(1)} \rightarrow B_{\mathcal{H}(1)}$. The flow (\mathcal{T}^t) is rectifiable in a sufficiently small neighborhood of the balls U_t . Let $\hat{k}(p)$ be the time that $\mathcal{T}^t(p)$ returns to $B_{\mathcal{H}(1)}$ for the k th time under \mathcal{T} , where $p \in B_{\mathcal{H}(1)}$ or $\hat{\mu}$ a.e. $p \in B_{\mathcal{H}(1)}$,

$$\lim_{k \rightarrow \infty} \frac{\hat{k}(p)}{k} = \int_{B_{\mathcal{H}(1)}} r_2 d\hat{\mu} := \bar{r}_2.$$

We fix an integer n and discretize C_t into disjoint sets $C_{t,j}$, $j = 1$ to n , of roughly equal $\hat{\mu}$ measure and define $\tilde{U}_{t,j} := \{q \in U_t : \Pi q \in C_{t,j}\}$. Hence $C_{t,j}$ lie in $B_{\mathcal{H}(1)}$ while $\tilde{U}_{t,j}$ lies in the full suspension flow $B_{\mathcal{H}(1)}^{r_2}$.

We consider two sequences of sets $C_{\alpha,t,j}$ and $C_{\beta,t,j}$ in the suspension flow defined by flow lines through $C_{t,j}$ of constant length $\tau_1(t, j)$ and $\tau_2(t, j)$ such that for each $\tilde{U}_{t,j}$, $C_{\alpha,t,j} \subset \tilde{U}_{t,j} \subset C_{\beta,t,j}$ and moreover for each $j, t > 0$, $\mu_{\mathcal{T}}(C_{\beta,t,j}) - \mu_{\mathcal{T}}(C_{\alpha,t,j}) \leq e(n)\mu_{\mathcal{T}}(\tilde{U}_{t,j})$ where $e(n) \rightarrow 0$ as $n \rightarrow \infty$. We can ensure this as the boundary of $\tilde{U}_{t,j}$ consists of two manifolds, each a smooth graph over $C_{t,j}$. The role of the sequence of sets $C_{\alpha,t,j}$, $C_{\beta,t,j}$ is to provide discretized lower and upper bounds between which we can squeeze the continuous flow.

Hence $\mu_{\mathcal{T}}(\bigcup_j C_{\alpha,t,j}) \leq \mu_{\mathcal{T}}(U_t) \leq \mu_{\mathcal{T}}(\bigcup_j C_{\beta,t,j})$ and $\mu_{\mathcal{T}}(\bigcup_j C_{\beta,t,j}) - \mu_{\mathcal{T}}(\bigcup_j C_{\alpha,t,j}) \leq e(n)\mu_{\mathcal{T}}(U_t)$ where $e(n) \rightarrow 0$ as $n \rightarrow \infty$.

Recall $\hat{k}(p)$ denotes the k th return time to $B_{\mathcal{H}(1)}$ of a point $p \in B_{\mathcal{H}(1)}$ under the flow \mathcal{T} so that $\mathcal{T}^{\hat{k}(p)} = \mathcal{R}_2^k(p)$. By the ergodic theorem given $\varepsilon > 0$ for $\hat{\mu}$ a.e. p there exists $k^*(\varepsilon)(p)$ such that $k(\bar{r}_2 - \varepsilon) \leq \hat{k}(p) \leq k(\bar{r}_2 + \varepsilon)$ for all $k > k^*(\varepsilon)$.

We fix ε and n . We let $[\alpha]$ denote the integer part of the real number α . For each j , the sequences of sets, indexed by k , $(C_{\lceil k(\bar{r}_2 + \varepsilon) \rceil, j})$ and $(C_{\lfloor k(\bar{r}_2 - \varepsilon) \rfloor, j})$ both have the (SBC) property for $\mathcal{R}_2 : B_{\mathcal{H}(1)} \rightarrow B_{\mathcal{H}(1)}$, i.e.

$$\lim_{k \rightarrow \infty} \frac{1}{E(k, j, \varepsilon, +)} \sum_{i=1}^k \mathbb{1}_{C_{\lceil i(\bar{r}_2 + \varepsilon) \rceil, j}} \circ \mathcal{R}_2^i(p) = 1$$

for $\hat{\mu}$ a.e. $p \in B_{\mathcal{H}(1)}$, where $E(k, j, \varepsilon, +) := \sum_{i=1}^k \hat{\mu}(C_{\lceil i(\bar{r}_2 + \varepsilon) \rceil, j})$ and similarly for $(C_{\lfloor k(\bar{r}_2 - \varepsilon) \rfloor, j})$. Indeed, this follows from Proposition 6.5 since $\hat{\mu}(C_{\lceil k(\bar{r}_2 + \varepsilon) \rceil, j}) \sim \mu_{\mathcal{T}}(U_{\lceil k(\bar{r}_2 + \varepsilon) \rceil})^{1-\frac{1}{d}}$ as $k \rightarrow \infty$, for fixed n and ε .

Note that $k(\bar{r}_2 - \varepsilon) \leq \hat{k} \leq k(\bar{r}_2 + \varepsilon)$ and by the Lipschitz regularity of $\hat{\mu}(C_{t,j})$ in t if $k(\bar{r}_2 - \varepsilon) \leq t \leq k(\bar{r}_2 + \varepsilon)$ then $\hat{\mu}(C_{\lceil k(\bar{r}_2 + \varepsilon) \rceil, j}) - \hat{\mu}(C_{\lfloor k(\bar{r}_2 - \varepsilon) \rfloor, j}) \leq \rho(\varepsilon)\hat{\mu}(C_{\lceil k(\bar{r}_2 + \varepsilon) \rceil, j})$ where $\rho(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$.

Furthermore, for sufficiently large t , once $\mathcal{R}_2^k(p)$ enters $C_{t,j}$ its trajectory spends a length of flow time between $\tau_1(\lfloor k(\bar{r}_2 - \varepsilon) \rfloor, j)$ and $\tau_2(\lceil k(\bar{r}_2 + \varepsilon) \rceil, j)$ in the sets $(\tilde{U}_{t,j})$.

Thus for $\hat{\mu}$ a.e. p , (recall n is fixed)

$$\begin{aligned} \sum_{j=1}^n \sum_{i=1}^T \tau_1([i(\bar{r}_2 - \varepsilon)], j) \hat{\mu}(C_{([i(\bar{r}_2 - \varepsilon)], j)}) &\leq \sum_{j=1}^n \int_0^{T\bar{r}_2} \hat{\mu}(\Pi \tilde{U}_{t,j}) 1_{U_{t,j}} \circ \mathcal{T}^t(p) dt \\ &\leq \sum_{j=1}^n \sum_{i=1}^T \tau_1([i(\tau_1 + \varepsilon)], j) \hat{\mu}(C_{([i(\tau_1 + \varepsilon)], j)}). \end{aligned}$$

The sums $L(T, n) := \sum_{j=1}^n \sum_{i=1}^T \tau_1([i(\bar{r}_2 - \varepsilon)], j) \hat{\mu}(C_{([i(\bar{r}_2 - \varepsilon)], j)})$ and $U(T, n) := \sum_{j=1}^n \sum_{i=1}^T \tau_1([i(\tau_1 + \varepsilon)], j) \times \hat{\mu}(C_{([i(\tau_1 + \varepsilon)], j)})$ are Riemann sums, and $\lim_{T \rightarrow \infty} \frac{U(T, n)}{L(T, n)} = \kappa(n)$ where $\kappa(n) \rightarrow 1$ as $n \rightarrow \infty$.

Using a change of variables

$$\sum_{j=1}^n \int_0^{T\bar{r}_2} \hat{\mu}(\Pi \tilde{U}_{t,j}) 1_{U_{t,j}} \circ \mathcal{T}^t(p) dt \sim \frac{1}{\bar{r}_2} \sum_{j=1}^n \int_0^T \hat{\mu}(\Pi \tilde{U}_{t,j}) 1_{U_{t,j}} \circ \mathcal{T}^t(p) dt,$$

where $H(T) \sim G(T)$ means $\lim_{T \rightarrow \infty} \frac{G(T)}{H(T)} = 1$.

Furthermore

$$\left| \frac{\frac{1}{\tau_1} \sum_{j=1}^n \int_0^T \hat{\mu}(\Pi \tilde{U}_{t,j}) 1_{U_{t,j}} \circ \mathcal{T}^t(p) dt}{\int_0^T \nu(U_t) dt} - 1 \right| \leq \kappa_2(n),$$

where $\kappa_2(n) \rightarrow 0$ as $n \rightarrow \infty$.

This proves the SBC property for nested balls in the full suspension flow. □

6.4. Extreme Value Laws for the flow

We have established EVLs for sufficient regular observations on the dynamical system $(\mathcal{T}_2, B, \mu_2)$. We now consider EVLs for the flow $\mathcal{F}_s : B^{r_2} \rightarrow B^{r_2}$. To do this we use [32, Theorem 2.6] which relates Extreme Value Theory for functions on the suspension of a base transformation to the Extreme Value statistics of observations on the base.

We start with some preliminary notation. Let $\bar{r}_2 = \int_B r_2(x) d\mu_2$. Let $\phi : B^{r_2} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a function, strictly maximized at a point $(x_0, u_0) \in B^{r_2}$, which is sufficiently regular that for large r the set $\{(x, u) \in B^{r_2} : \phi((x, u)) > r\}$ corresponds to a topological ball centered at (x_0, u_0) . Let $\bar{\phi}(x) = \sup_{0 \leq u \leq r_2(x)} \phi((x, u))$ and define $u_n(t)$ by the requirement that $n\mu_2\{\bar{\phi} > u_n(t)\} \rightarrow t$. Let $M_T(x, s) := \max\{\phi(F_s(x, u)) : 0 \leq s \leq T\}$. As a consequence of [32, Theorem 2.6]:

Proposition 6.8. *Suppose when we write $u_n(t) = \frac{g(t)}{a_n} + b_n$ the normalizing constants $a_n > 0$ and b_n satisfy:*

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} a_n |b_{[n+\varepsilon n]} - b_n| = 0, \tag{4}$$

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \left| 1 - \frac{a_{[n+\varepsilon n]}}{a_n} \right| = 0. \tag{5}$$

Then:

- (1) If x_0 is not a periodic point for \mathcal{T}_2 then $\mu\{M_T \leq u_{[T/\bar{r}_2]}(t)\} \rightarrow e^{-t}$.
- (2) If x_0 is a repelling periodic point of prime period k then $\mu\{M_T \leq u_{[T/\bar{r}_2]}(t)\} \rightarrow e^{-\theta t}$ where $\theta = 1 - |\text{Jac}(D\mathcal{T}_2^{-k})(p_0)|$.

The Extreme Value result for the Teichmüller flow $\mathcal{T} = (\mathcal{T}^t)_{t \in \mathbb{R}}$ holds from combining [23, Theorem 2.1] with [32, Corollary 2.3] (note that the proof for Gibbs–Markov maps holds in any dimension as long as conformality holds) and [32, Theorem 2.6].

Appendix: Aperiodicity and weak mixing

Let (X, T, μ) be an ergodic measure-preserving dynamical system.

Definition A.1. (X, T, μ) is weakly mixing if $f \circ T = e^{it} f$ for some non-zero $f \in L^2(\mu)$ and $t \in [0, 2\pi)$ implies that $t = 0$ and f is constant.

Remark A.2. This definition is equivalent to the classical one, stating that

$$\frac{1}{n} \sum_{k=0}^{n-1} |\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)| \rightarrow 0$$

for any measurable sets A and B . See [58, Theorem 1.26] in the case where (X, T, μ) is invertible, and [33, Theorem 664] or [14, Theorem 2.36] for a proof of the equivalence valid in any case.

Let $Y \subset X$ be a subset of positive μ -measure. We denote by $\tau(y)$ the first return time of $y \in Y$ to Y :

$$\tau(y) = \min\{n \geq 1 : T^n y \in Y\}.$$

We then define the first return map $\widehat{T} : Y \rightarrow Y$ by $\widehat{T} = T^\tau$. It preserves the normalisation μ_Y of the restriction to Y of the measure μ and is ergodic with respect to it.

Definition A.3. We will say that the first return time is aperiodic if $f \circ \widehat{T} = e^{it\tau} f$ for some non-zero $f \in L^2(\mu_Y)$ and $t \in [0, 2\pi)$ implies that $t = 0$ and f is constant.

Remark A.4. By [46, Proposition 1.1], the relation $f \circ \widehat{T} = e^{it\tau} f$ is equivalent to $\mathcal{L}(e^{it\tau} f) = f$, where \mathcal{L} is the transfer operator of \widehat{T} with respect to the measure μ_Y .

Proposition A.5. The first return time is aperiodic if and only if (X, T, μ) is weakly mixing.

Proof. Suppose first that the first return time is aperiodic and let $f \in L^2(\mu)$ non-zero and $t \in [0, 2\pi)$ such that $f \circ T = e^{it} f$. We easily verify that the restriction f_Y of f to Y satisfies $f_Y \circ \widehat{T} = e^{it\tau} f_Y$:

$$f_Y(\widehat{T}y) = f(T^{\tau(y)}y) = e^{it\tau(y)} f(y) = e^{it\tau(y)} f_Y(y).$$

f_Y is also non identically zero: otherwise, f would vanish on the set $\bigcup_{n \geq 0} T^{-n}Y$, which by ergodicity is equal to X mod μ . Aperiodicity yields that $t = 0$, which means that $f \circ T = f$. Ergodicity implies that f is constant.

Conversely, suppose that (X, T, μ) is weakly mixing and that $f \in L^2(\mu_Y)$ is non identically zero and satisfies $f \circ \widehat{T} = e^{it\tau} f$. We first extend τ on the whole space X as being the first hitting time. By ergodicity, it is well defined μ -a.e. We then define $\tilde{f} \in L^2(\mu)$ by $\tilde{f} = e^{-it\tau} f \circ T^\tau$. Since $T^{\tau(x)}x$ belongs to Y for μ -a.e. $x \in X$ by definition, \tilde{f} is well defined. Our assumption on f implies that \tilde{f} and f coincide on Y , so that it is non identically zero.

Now, we verify that $\tilde{f} \circ T = e^{it} \tilde{f}$. Let $x \in X$ with $\tau(x) > 1$. Since τ is the first hitting time, we have $\tau(Tx) = \tau(x) - 1$. Hence, $\tilde{f}(Tx) = e^{-it\tau(Tx)} f(T^{\tau(Tx)}Tx) = e^{it} e^{-it\tau(x)} f(T^{\tau(x)}x) = e^{it} \tilde{f}(x)$. If $\tau(x) = 1$, which implies $Tx \in Y$, we have by definition of \tilde{f} that $\tilde{f}(x) = e^{-it} f(Tx) = e^{-it} \tilde{f}(Tx)$.

Weak mixing implies that $t = 0$ and \tilde{f} is constant. Since the restriction of \tilde{f} to Y is f , this shows that f is constant, and concludes the proof. \square

Acknowledgements

RA was supported by Conseil Régional Provence-Alpes-Côte d'Azur, by the ANR-project *Perturbations*, by the PICS (Projet International de Coopération Scientifique), Propriétés statistiques des systèmes dynamiques déterministes et aléatoires, with the University of Houston, n. PICS05968 and by the European Advanced Grant Macroscopic Laws

and Dynamical Systems (MALADY) (ERC AdG 246953). Most of this work was done when RA was affiliated to Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France and Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France and Università di Roma (Tor Vergata), Roma, Italy. MN was partially supported by NSF grant DMS 1101315 and by the French CNRS with a *poste d'accueil* position at the Center of Theoretical Physics in Luminy. MT was partially supported by NSF grant DMS 1109587. RA and MN would like to thank Huyi Hu for discussions on quasi-Hölder space and aperiodicity. MN would like to thank Mark Pollicott for interesting and helpful discussions concerning Rauzy–Veech renormalization and Teichmüller flow. The authors wish to thank Sandro Vaienti for helpful remarks, encouragement and many useful discussions concerning this work. They also wish to thank the referees of this paper for important suggestions and comments.

References

- [1] J. Aaronson. *An Introduction to Infinite Ergodic Theory. Mathematical Surveys and Monographs* **50**. AMS, Providence, 1997. [MR1450400](#)
- [2] M. Abadi and A. Galves. Inequalities for the occurrence times of rare events in mixing processes. The state of the art. *Markov Process. Related Fields* **7** (2001) 97–112. [MR1835750](#)
- [3] J. Athreya. Quantitative recurrence and large deviations for Teichmüller geodesic flow. *Geom. Dedicata* **119** (2006) 121–140. [MR2247652](#)
- [4] A. Avila and A. Bufetov. Exponential decay of correlations for the Rauzy–Veech–Zorich induction map. In *Partially Hyperbolic Dynamics, Laminations, and Teichmüller Flow* 203–211. *Fields Inst. Commun.* **51**. Amer. Math. Soc., Providence, RI, 2007. [MR2388696](#)
- [5] A. Avila, S. Gouëzel and J.-C. Yoccoz. Exponential mixing of the Teichmüller flow. *Publ. Math. Inst. Hautes Etudes Sci.* **104** (2006) 143–211. [MR2264836](#)
- [6] H. Ayaç, J. M. Freitas and S. Vaienti. Laws of rare events for deterministic and random dynamical systems. *Trans. Amer. Math. Soc.* **367** (2015) 8229–8278. [MR3391915](#)
- [7] V. Baladi. *Positive Transfer Operators and Decay of Correlations. Advanced Series in Nonlinear Dynamics* **16**. World Scientific, River Edge, 2000. [MR1793194](#)
- [8] H. Bruin, B. Saussol, S. Troubetzkoy and S. Vaienti. Return time statistics via inducing. *Ergodic Theory Dynam. Systems* **23** (2003) 991–1013. [MR1997964](#)
- [9] 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** (2006) 579–623. [MR2220100](#)
- [10] A. I. Bufetov and B. M. Gurevich. Existence and uniqueness of a measure with maximal entropy for the Teichmüller flow on the moduli space of abelian differentials. *Mat. Sb.* **202** (2011) 3–42. [MR2857792](#)
- [11] N. Chernov and D. Kleinbock. Dynamical Borel–Cantelli lemmas for Gibbs measures. *Israel J. Math.* **122** (2001) 1–27. [MR1826488](#)
- [12] P. Collet. Statistics of closest return for some non-uniformly hyperbolic systems. *Ergodic Theory Dynam. Systems* **21** (2001) 401–420. [MR1827111](#)
- [13] D. Dolgopyat. Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.* **356** (2004) 1637–1689. [MR2034323](#)
- [14] M. Einsiedler and T. Ward. *Ergodic Theory with a View Towards Number Theory. Graduate Texts in Mathematics* **259**. Springer, London, 2011. [MR2723325](#)
- [15] A. C. M. Freitas, J. M. Freitas and M. Todd. Hitting time statistics and extreme value theory. *Probab. Theory Related Fields* **147** (2010) 675–710. [MR2639719](#)
- [16] A. C. M. Freitas, J. M. Freitas and M. Todd. The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics. *Comm. Math. Phys.* **321** (2013) 483–527. [MR3063917](#)
- [17] A. C. M. Freitas, J. M. Freitas, M. Todd and S. Vaienti. Rare events for the Manneville–Pomeau map. *Stochastic Process. Appl.* **126** (11) (2016) 3463–3479. [MR3549714](#)
- [18] S. Galatolo. Dimension and hitting time in rapidly mixing systems. *Math. Res. Lett.* **14** (5) (2007) 797–805. [MR2350125](#)
- [19] S. Galatolo and D. Kim. The dynamical Borel–Cantelli lemma and the waiting time problems. *Indag. Math. (N.S.)* **18** (3) (2007) 421–434. [MR2373690](#)
- [20] S. Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.* **139** (2004) 29–65. [MR2041223](#)
- [21] S. Gouëzel. A Borel–Cantelli lemma for intermittent interval maps. *Nonlinearity* **20** (6) (2007) 1491–1497. [MR2327135](#)
- [22] S. Gouëzel. Local limit theorem for nonuniformly partially hyperbolic skew-products and Farey sequences. *Duke Math. J.* **147** (2009) 192–284. [MR2495076](#)
- [23] C. Gupta. Extreme value distributions for some classes of non-uniformly partially hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **30** (2010) 757–771. [MR2643710](#)
- [24] C. Gupta, M. Holland and M. Nicol. Extreme value theory and return time statistics for dispersing billiard maps and flows, Lozi maps and Lorenz-like maps. *Ergodic Theory Dynam. Systems* **31** (2011) 1363–1390. [MR2832250](#)
- [25] C. Gupta, M. Nicol and W. Ott. A Borel–Cantelli lemma for non-uniformly expanding dynamical systems. *Nonlinearity* **23** (2010) 1991–2008. [MR2669635](#)
- [26] N. Haydn, Y. Lacroix and S. Vaienti. Hitting and return time statistics in ergodic dynamical systems. *Ann. Probab.* **33** (2005) 2043–2050. [MR2165587](#)
- [27] N. Haydn, M. Nicol, T. Persson and S. Vaienti. A note on Borel–Cantelli lemmas for non-uniformly hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **33** (2013) 475–498. [MR3035294](#)

- [28] N. T. Haydn, N. Winterberg and R. Zweimüller. Return-time statistics, hitting-time statistics and inducing. In *Ergodic theory, Open Dynamics, and Coherent Structures* 217–227. Springer Proceedings in Mathematics & Statistics **70**. Springer, New York, 2014. MRMR3213501
- [29] H. Hennion. Sur un théorème spectral et son application aux noyaux lipschitziens. *Proc. Amer. Math. Soc.* **118** (1993) 627–634. MR1129880
- [30] M. Hirata. Poisson limit law for Axiom-A diffeomorphisms. *Ergodic Theory Dynam. Systems* **13** (1993) 533–556. MR1245828
- [31] M. Hirata, B. Saussol and S. Vaienti. Statistics of return times: A general framework and new applications. *Comm. Math. Phys.* **206** (1999) 33–55. MR1736991
- [32] M. P. Holland, M. Nicol and A. Török. Extreme value distributions for non-uniformly hyperbolic dynamical systems. *Trans. Amer. Math. Soc.* **364** (2012) 661–688. MR2846347
- [33] S. Kalikow and R. McCutcheon. *An Outline of Ergodic Theory. Cambridge Studies in Advanced Mathematics* **122**. Cambridge University Press, Cambridge, 2010. MR2650005
- [34] M. Keane. Interval exchange transformations. *Math. Z.* **141** (1975) 25–31. MR0357739
- [35] G. Keller. Generalized bounded variation and applications to piecewise monotonic transformations. *Z. Wahrsch. Verw. Gebiete* **69** (1985) 461–478. MR0787608
- [36] D. Kim. The dynamical Borel–Cantelli lemma for interval maps. *Discrete Contin. Dyn. Syst.* **17** (4) (2007) 891–900. MR2276480
- [37] D. Kleinbock and G. Margulis. Logarithm laws for flows on homogeneous spaces. *Invent. Math.* **138** (1999) 451–494. MR1719827
- [38] G. Lindgren, M. R. Leadbetter and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes. Springer Series in Statistics*. Springer-Verlag, New York-Berlin, 1983. MR0691492
- [39] L. Marchese. The Khinchin theorem for interval-exchange transformations. *J. Mod. Dyn.* **5** (2011) 123–183. MR2787600
- [40] L. Marchese. Khinchin type condition for translation surfaces and asymptotic laws for the Teichmüller flow. *Bull. Soc. Math. France* **140** (2013) 485–532. MR3059848
- [41] H. Masur. Interval exchange transformations and measured foliations. *Ann. of Math. (2)* **115** (1982) 169–200. MR0644018
- [42] H. Masur. Logarithm law for geodesics in moduli space. In *Mapping Class Groups and Moduli Spaces of Riemann Surfaces* 229–245. *Contemp. Math.* **150**. Amer. Math. Soc., Providence, RI, 1993. MR1234267
- [43] F. Maucourant. Dynamical Borel–Cantelli lemma for hyperbolic spaces. *Israel J. Math.* **152** (2006) 143–155. MR2214457
- [44] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* **260** (2005) 131–146. MR2175992
- [45] I. Melbourne and M. Nicol. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.* **360** (2008) 6661–6676. MR2434305
- [46] T. Morita. A generalized local limit theorem for Lasota–Yorke transformations. *Osaka J. Math.* **26** (1989) 579–595. MR1021432
- [47] T. Morita. Renormalized Rauzy inductions. *Adv. Stud. Pure Math.* **49** (2007) 263–288.
- [48] W. Philipp. Some metrical theorems in number theory. *Pacific J. Math.* **20** (1967) 109–127. MR0205930
- [49] M. Pollicott. Statistical properties of the Rauzy–Veech–Zorich map. Unpublished notes. Available at <http://homepages.warwick.ac.uk/~masdbl/teichmuller-asip.pdf>.
- [50] O. Sarig. Subexponential decay of correlations. *Invent. Math.* **150** (2002) 629–653. MR1946554
- [51] B. Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math.* **116** (2000) 223–248. MR1759406
- [52] W. Schmidt. A metrical theory in Diophantine approximation. *Canad. J. Math.* **12** (1960) 619–631. MR0118711
- [53] W. Schmidt. Metrical theorems on fractional parts of sequences. *Trans. Amer. Math. Soc.* **110** (1964) 493–518. MR0159802
- [54] V. G. Sprindzuk. *Metric Theory of Diophantine Approximations. Scripta Series in Mathematics*. V. H. Winston and Sons, Washington, D.C., 1979. Translated from the Russian and edited by Richard A. Silverman. With a foreword by Donald J. Newman. MR0548467
- [55] W. A. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math. (2)* **115** (1982) 201–242. MR0644019
- [56] W. A. Veech. The Teichmüller geodesic flow. *Ann. of Math. (2)* **124** (1986) 441–530. MR0866707
- [57] M. Viana. Dynamics of interval exchange maps and Teichmüller flows. IMPA, 2008. Available at <http://w3.impa.br/~viana/out/ietf.pdf>.
- [58] P. Walters. *An Introduction to Ergodic Theory. Graduate Texts in Mathematics* **79**. Springer, New York-Berlin, 1982. MR0648108
- [59] J.-C. Yoccoz. Continued fraction algorithms for interval exchange maps: An introduction. In *Frontiers in Number Theory, Physics, and Geometry. I* 401–435. Springer, Berlin, 2006. MR2261103
- [60] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650. MR1637655
- [61] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999) 153–188. MR1750438
- [62] L. Zhang. Borel–Cantelli lemmas and extreme value theory for geometric Lorenz models. *Nonlinearity* **29** (1) (2016) 232–255. MR3460754
- [63] A. Zorich. Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents. *Ann. Inst. Fourier (Grenoble)* **46** (1996) 325–370. MR1393518