Vaibhav Gadre

Partial sums of excursions along random geodesics and volume asymptotics for thin parts of moduli spaces of quadratic differentials

Abstract. For a non-uniform lattice in SL(2, ℝ), we consider excursions of a random geodesic in cusp neighborhoods of the quotient finite area hyperbolic surface or orbifold. We prove a strong law for a certain partial sum involving these excursions. This generalizes a theorem of Diamond and Vaaler [9] for continued fractions. In the Teichmüller setting, we consider invariant measures for the SL(2, ℝ) action on the moduli spaces of quadratic differentials. By the work of Eskin and Mirzakhani [12], these measures are supported on affine invariant submanifolds of a stratum of quadratic differentials. For a Teichmüller geodesic random with respect to an SL(2, ℝ)-invariant measure, we study its excursions in thin parts of the associated submanifold. Under a regularity hypothesis for the invariant measure, we prove similar strong laws for certain partial sums involving these excursions. The limits in these laws are related to the volume asymptotic of the thin parts. By Siegel–Veech theory, these are given by Siegel–Veech constants. As a direct consequence, we show that the word metric of mapping classes that approximate a Teichmüller geodesic ray that is random with respect to the Masur–Veech measure, grows faster than $T \log T$.

Keywords. Teichmüller theory, moduli of Riemann surfaces

1. Introduction

This paper provides a specific analogy between non-uniform lattices in SL(2, ℝ) and mapping class groups. The analogy is established from the point of view of cusp excursions of random geodesics. For non-uniform lattices in SL(2, ℝ), we consider excursions of a random geodesic in horoball neighborhoods of the cusps of the quotient hyperbolic surface. For an SL(2, ℝ) orbit closure in a stratum of a moduli space of quadratic differentials we consider excursions of random Teichmüller geodesics in a thin part of the orbit closure.

1.1. *Non-uniform lattices in* SL(2, ℝ)

Let Γ be a non-uniform lattice in SL(2, ℝ). The quotient $X = \Gamma \backslash \mathbb{H}^2$ is a complete finite area surface/orbifold with finitely many cusps $c_1, \ldots, c_J$. Let $X_{\text{cusps}} \subset X$ be the union

V. Gadre: School of Mathematics and Statistics, University of Glasgow,
University Place, Glasgow G12 8SQ, UK; e-mail: gadre.vaibhav@gmail.com

of disjoint horoball neighborhoods of the cusps. The set X_{cusps} lifts to a countable collection $\mathcal{H}$ of disjoint horoballs in $\mathbb{H}^2$. The complement $X \setminus X_{\text{cusps}}$ is a compact set X_{thick} called the *thick part* of X. The lift $\widetilde{X}_{\text{thick}}$, which we call the thick part in $\mathbb{H}^2$, is the complement of the union of horoballs in $\mathcal{H}$.

The Liouville measure is a natural $\text{SL}(2, \mathbb{R})$-invariant measure on the unit tangent bundle $T^1\mathbb{H}^2$. In the upper half-plane model it is given by

$$d\ell = \frac{dx\, dy\, d\theta}{2\pi y^2}.$$

The hyperbolic geodesic flow on T^1X is given by the action of the diagonal subgroup. The measure $d\ell$ descends to a flow-invariant measure on T^1X. We will continue to call it *Liouville measure*. The conditional measure on the unit tangent circle at any point is the pullback via the visual map of the standard Lebesgue measure on $\partial\mathbb{H}^2 = S^1$.

Since the number of horoballs is countable, $d\ell$-almost every geodesic ray is recurrent to X_{thick}. By the ergodicity of the geodesic flow and the fellow traveling property of hyperbolic geodesics, Leb-almost every geodesic ray from a fixed base-point $x_0 \in X$ ventures into X_{cusps} infinitely often. In particular, a geodesic ray γ in $\mathbb{H}^2$ whose endpoint r in S^1 is Leb-typical enters and leaves infinitely many horoballs. By analyzing the collection $\mathcal{H}$, Sullivan [28] showed that the lim sup of the maximum depth that γ achieves in X_{cusps} is asymptotically $(1/2)\log T$, where T is the time along γ. It is convenient to assume $x_0 \in X_{\text{thick}}$, which can be achieved by making X_{cusps} smaller if necessary.

To set up notation, let $\gamma(x_0, r)$ be the geodesic ray from x_0 to $r \in S^1$. We denote by $\gamma_T(x_0, r)$ the point on γ at distance T from x_0. When the context is clear we will use just γ and γ_T. Let $\pi : \mathbb{H}^2 \to \widetilde{X}_{\text{thick}}$ be the closest point projection. Let $N = N(T)$ be the number of horoballs that γ intersects up to γ_T. We enumerate this collection $\mathcal{H}_{\gamma,T}$ of horoballs as $\{H_1, \ldots, H_N\}$ in the order of increasing time. For all $k < N$, γ enters and exits H_k; H_N may be an exception if $\gamma_T \in H_N$. Let d_{thick} be the path metric on $\widetilde{X}_{\text{thick}}$. For a horoball H that γ enters and exits, the *complete excursion* $E(\gamma, H)$ is defined as the d_{thick}-distance between the entry and exit points. If $\gamma_T \in H_N$ then the *partial excursion* $E(\gamma, H_N)$ is the d_{thick}-distance between the entry point for H_N and $\pi(\gamma_T)$.

The *total excursion* $E(\gamma, T)$ till time T, as defined in [15], is

$$E(\gamma, T) = \sum_{k \leq N} E(\gamma, H_k).$$

Along Leb-typical geodesic rays, [15, Proposition 5.4] shows that $E(\gamma, T)/T \to \infty$. We prove here:

Theorem 1.2. *For Lebesgue almost every r in S^1,*

$$\lim_{T \to \infty} \frac{E(\gamma, T) - \max_{k \leq N} E(\gamma, H_k)}{T \log T} = \frac{2}{\pi} \frac{\ell(T^1 X_{\text{cusp}})}{\ell(T^1 X)}.$$

Continued fractions. Let $r \in [0, 1]$. The continued fraction expansion is given by writing r as

$$r = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{+ \cdots}}},$$

where each a_i is in $\mathbb{N}$. When r is irrational the expansion is infinite. We denote r as $[a_1, a_2, \dots]$.

Theorem 1.3 (Diamond–Vaaler [9]). *For Leb-almost every $r \in [0, 1]$,*

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} a_i - \max_{i \leq n} a_i}{n \log n} = \frac{1}{\log 2}.$$

We will derive Theorem 1.3 from Theorem 1.2 in the special case when X is the modular surface $\mathrm{SL}(2, \mathbb{Z}) \backslash \mathbb{H}^2$. Excursions of geodesic rays into the maximal cusp neighborhood in X are related to coefficients in the continued fraction expansion of the point at infinity. Diamond–Vaaler used techniques specific to the symbolic dynamics (Gauss map) in the theory of continued fractions. Theorem 1.2 relies on more general features, viz. asymptotic for $\ell(X_{\mathrm{cusps}})$ and exponential mixing of the geodesic flow. These features hold for the Teichmüller geodesic flow in the analogous setting of quadratic differentials (Theorem 1.9).

Word metric along random geodesics. As a direct implication of Theorem 1.2, we analyze how the word metric grows along random geodesics.

For distinct points $x, y \in \mathbb{H}^2$ let $\gamma(x, y)$ be the hyperbolic geodesic between them. The *projected path* is defined by $p(x, y) = \pi(\gamma(x, y))$. Let $L(x, y)$ be the d_{thick}-length of $p(x, y)$, where d_{thick} is the path metric on the thick part. The quantity $L(x_0, \gamma_T) - E(\gamma, T)$ is the time spent by γ in the interior of X_{thick}. By the ergodicity of the geodesic flow, $L(x_0, \gamma_T) - E(\gamma, T)$ grows linearly in T. Hence, a direct consequence of Theorem 1.2 is:

Theorem 1.4. *For Leb-almost every $r \in S^1$,*

$$\lim_{T \to \infty} \frac{L(x_0, \gamma_T) - \max_{k \leq N} E(\gamma, H_k)}{T \log T} = \frac{2}{\pi} \frac{\ell(T^1 X_{\mathrm{cusp}})}{\ell(T^1 X)}.$$

By [15, Lemma 5.1], the projected path $p(x_0, \gamma_T)$ is a quasi-geodesic in $(\widetilde{X}_{\mathrm{thick}}, d_{\mathrm{thick}})$. A closer look at the proof of [15, Lemma 5.1] shows that $L(x_0, \gamma_T) - d_{\mathrm{thick}}(x_0, \gamma_T)$ grows at most linearly in N. In Lemma 3.8, we show that N grows linearly in T. As a direct consequence of these linear bounds and Theorem 1.4 we get:

Theorem 1.5. *For Leb-almost every $r \in S^1$,*

$$\lim_{T \to \infty} \frac{d_{\mathrm{thick}}(x_0, \gamma_T) - \max_{k \leq N} E(\gamma, H_k)}{T \log T} = \frac{2}{\pi} \frac{\ell(T^1 X_{\mathrm{cusp}})}{\ell(T^1 X)}.$$

A base-point $x_0 \in \mathbb{H}^2$ is said to be *generic* if the stabilizer of x_0 in Γ is trivial. The Γ-orbit of x_0 is called a *lattice*. When x_0 is generic, each lattice point corresponds to a unique group element. Each point γ_T has at least one closest lattice point which we denote by $h_T x_0$. In fact, this closest point is unique for almost all points along γ.

The group Γ is finitely generated. A finite choice of generators for Γ defines a proper word metric d_Γ on Γ. Different choices of generators produce quasi-isometric metrics. Let $d_\Gamma(1, h_T)$ be the word length where $h_T x_0$ is the lattice point closest to γ_T.

The group Γ acts cocompactly on $\widetilde{X}_{\text{thick}}$, so by the Švarc–Milnor lemma, (Γ, d_Γ) is quasi-isometric to $(\widetilde{X}_{\text{thick}}, d_{\text{thick}})$. Thus, as a direct consequence of Theorem 1.5 we get:

Theorem 1.6. *There exists a constant $M_1 > 0$ that depends on the word metric such that for Leb-almost every $r \in S^1$,*

$$M_1 T \log T < d_\Gamma(1, h_T)$$

for T sufficiently large depending on r.

In fact, if the contribution from the largest excursion is removed, then the word metric grows like $T \log T$ up to uniform multiplicative and additive constants. Theorem 1.6 should be thought of as a refinement of [15, Proposition 5.6] which states that along a Leb-generic geodesic ray, $d_\Gamma(1, h_T)/T \to \infty$ as $T \to \infty$.

1.7. Moduli spaces of quadratic differentials

Let S be a hyperbolic surface of finite type. S is *non-sporadic* if it is not a sphere with at most four punctures or boundary components, or a torus with at most one puncture or boundary component. In the sporadic examples, the Teichmüller space is either trivial or isometric to $\mathbb{H}^2$ and the mapping class group is a non-uniform lattice in $\text{SL}(2, \mathbb{R})$. This reduces us to the previous case.

The *Teichmüller space* $\mathcal{T}(S)$ is the space of marked conformal structures on S, or alternatively, marked hyperbolic metrics on S. The *mapping class group* $\text{Mod}(S)$ is the group of orientation preserving diffeomorphisms of S modulo isotopy. $\text{Mod}(S)$ acts on $\mathcal{T}(S)$ by changing the marking. The quotient $\mathcal{M}(S) = \text{Mod}(S)\backslash\mathcal{T}(S)$ is the *moduli space of Riemann surfaces*.

The *Teichmüller metric* is given by

$$d_{\mathcal{T}}(X, Y) = \tfrac{1}{2} \inf_f \log K(f),$$

where the infimum is over all quasi-conformal maps $f : X \to Y$, and $K(f)$ is the quasi-conformal constant of f. The group $\text{Mod}(S)$ acts by isometries of $d_{\mathcal{T}}$. Thus, the metric descends to a metric on $\mathcal{M}(S)$. The *thin part* of $\mathcal{T}(S)$ consists of hyperbolic surfaces that contain a simple closed curve with small hyperbolic length. The thin part of $\mathcal{T}(S)$ is obviously $\text{Mod}(S)$-invariant. The quotient $\text{Mod}(S)\backslash\mathcal{T}(S) \subset \mathcal{M}(S)$ is the thin part of the moduli space of Riemann surfaces.

For a Riemann surface X, let $\mathcal{Q}(X)$ be the set of meromorphic quadratic differentials on X with simple poles at the punctures. If $(k_1, \ldots, k_r)$ are the multiplicities of the zeros

then $k_1 + \cdots + k_r = 2g - 2 + n$, where n is the number of punctures. A quadratic differential is equivalent to a half-translation structure on S, i.e. it defines charts from S to $\mathbb{C}$ with transition functions of the form $\pm z + c$. The resulting flat metric has a cone singularity with angle $(k + 2)\pi$ at a k-order zero (or with $k = -1$ for a simple pole). A quadratic differential is *unit area* if the corresponding singular flat metric has area 1. The space $\mathcal{Q}$ of unit area quadratic differentials can be identified with the unit cotangent bundle to $\mathcal{T}(S)$ [16]. The space $\mathcal{Q}$ is stratified by the multiplicity of its zeros; we denote the stratum with multiplicities $\alpha = (k_1, \ldots, k_r)$ by $\mathcal{Q}(\alpha)$. Each stratum is $\mathrm{Mod}(S)$-invariant. We will continue to denote the moduli space by $\mathcal{Q}(\alpha)$.

The periods/holonomies of a fixed basis for the homology of S relative to the singularities give local co-ordinates on $\mathcal{Q}(\alpha)$. The natural volume form μ_{hol} in these co-ordinates, called the *Masur–Veech measure*, can be thought of as an analog of the Liouville measure. It is invariant under $\mathrm{Mod}(S)$-action and descends to a finite measure on the moduli space (see [21], [29]).

The affine action of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{C} = \mathbb{R}^2$ preserves the transitions $z \to \pm z + c$. This defines an $\mathrm{SL}(2, \mathbb{R})$ action on each stratum $\mathcal{Q}(\alpha)$. The action of the diagonal part defines the Teichmüller geodesic flow. The compact part $\mathrm{SO}(2, \mathbb{R})$ leaves the underlying conformal structure unchanged. This gives an isometric embedding $\mathbb{H}^2 = \mathrm{SL}(2, \mathbb{R})/\mathrm{SO}(2, \mathbb{R}) \to \mathcal{T}(S)$. These are called *Teichmüller discs* and they foliate $\mathcal{T}(S)$. We denote by $\mathrm{SL}(2, \mathbb{R})(q)$ the orbit and by $\mathbb{D}(q)$ the Teichmüller disc.

In the flat metric defined by q, a *saddle connection* is a geodesic segment in the q-metric that connects a pair of (same or distinct) singularities. For $\epsilon > 0$, the *ϵ-thin part* $\mathcal{Q}(\alpha)_\epsilon$ is the set of $q \in \mathcal{Q}(\alpha)$ such that some saddle connection has q-length squared less than ϵ.

The points q' in $\mathrm{SL}(2, \mathbb{R})(q)$ for which a particular saddle connection β has q'-length squared shorter than ϵ projects to a horoball H in $\mathbb{D}(q)$. The *point at infinity* for H is given by the direction in which β is vertical. For saddle connections that are parallel for q, the ratios of their holonomies are constant on $\mathrm{SL}(2, \mathbb{R})(q)$. Thus, in a collection of parallel saddle connections, the shortest holonomy will be used to define the horoball.

For a μ_{hol}-typical q, the horoballs in $\mathbb{D}(q)$ do not form a packing; their interiors have intersections. Every point $X \in \mathbb{D}(q)$ is contained in at most finitely many horoballs. But there need not be an upper bound over $\mathbb{D}(q)$ for this number. This makes it hard to estimate the d_{thick}-distance between the entry and exit points of a geodesic in $\mathcal{Q}(\alpha)_\epsilon$, which would have been a natural analog for an excursion in this context. Instead, we consider excursions in individual horoballs.

The excursions in horoballs in $\mathbb{D}(q)$ of a geodesic γ till time T are defined as follows. For a horoball H that γ enters and exits, the *complete excursion* $E(\gamma, H)$ is defined as the distance along ∂H between the entry and exit points. If $\gamma_T \in H$ then the *partial excursion* $E(\gamma, H)$ is defined as the distance along ∂H between the entry point of γ and $\pi_H(\gamma_T)$ where $\pi_H : H \to \partial H$ is the closest point projection. Let $H_1, \ldots, H_{N(T)}$ denote all the horoballs that γ intersects till time T. The *total excursion* $E(\gamma, T)$ is defined as

$$E(\gamma, T) = \sum_{k \leq N(T)} E(\gamma, H_k).$$

When horoballs have disjoint interiors, these definitions coincide with the definitions in Section 1.1. With intersections, the excursions $E(\gamma, H_k)$ can be simultaneous and there can be several partial excursions when γ_T is in the thin part.

The main theorem we prove is:

Theorem 1.8. *There exists a constant $c(\alpha) > 0$ such that for μ_{hol}-almost every $q \in \mathcal{Q}(\alpha)$, the Teichmüller geodesic ray γ that q determines satisfies*

$$\lim_{T \to \infty} \frac{E(\gamma, T) - \max_{k \leq N(T)} E(\gamma, H_k)}{T \log T} = 2\epsilon c(\alpha).$$

The constant $c(\alpha)$ is the Siegel–Veech constant associated to $\mathcal{Q}(\alpha)$ (see Section 4.5 for the definition).

Theorem 1.8 is a special case of a more general theorem, viz. Theorem 1.9 for $\mathrm{SL}(2, \mathbb{R})$-invariant measures and $\mathrm{SL}(2, \mathbb{R})$-invariant loci of saddle connection holonomies.

$\mathrm{SL}(2, \mathbb{R})$ *orbit closures and invariant measures.* Recently, Eskin and Mirzakhani [12, Theorem 1.4] showed that ergodic $\mathrm{SL}(2, \mathbb{R})$-invariant probability measures are of Lebesgue class and supported on invariant complex submanifolds in $\mathcal{Q}(\alpha)$. These manifolds are affine in the sense that they are given by linear equations in the holonomy coordinates. Furthermore, Eskin, Mirzakhani and Mohammadi [13, Theorem 2.1] showed that all $\mathrm{SL}(2, \mathbb{R})$ orbit closures are affine invariant submanifolds. See [12, Section 1] for more details. More recently, Filip [14] showed that these submanifolds are algebraic subvarieties.

Let μ be an ergodic $\mathrm{SL}(2, \mathbb{R})$-invariant probability measure supported on an affine invariant submanifold $\mathcal{N}$. For $\epsilon > 0$, the ϵ-*thin part* $\mathcal{N}_\epsilon$ is the subset of $q \in \mathcal{N}$ such that some saddle connection has q-length squared less than ϵ. Saddle connections β_1, β_2 are $\mathcal{N}$-*parallel* if they are parallel for an open set of quadratic differentials in $\mathcal{N}$. We assume that μ satisfies the following *regularity* condition. For $\epsilon, \kappa > 0$, let $\mathcal{N}_{\epsilon,\kappa}$ be the subset of $q \in \mathcal{N}$ that have at least one saddle connection β_1 with $\ell_q^2(\beta_1) \leq \epsilon$ and a saddle connection β_2 not $\mathcal{N}$-parallel to β_1 with $\ell_q^2(\beta_2) \leq \kappa$. We assume that there exists $m_1 > 0$ such that for ϵ, κ small enough

$$\mu(\mathcal{N}_{\epsilon,\kappa}) \leq m_1 \epsilon \kappa.$$

For μ_{hol}, [23, Section 10, Claim (7)] proves the regularity above. A weaker regularity for any $\mathrm{SL}(2, \mathbb{R})$-invariant measure is proved in [4, Theorem 1.2].

$\mathrm{SL}(2, \mathbb{R})$-*invariant loci.* For $q \in \mathcal{N}$, let $V(q) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ be an assignment of a nonempty subset (with multiplicity) of holonomies of saddle connections on q. We require that the assignment varies linearly under the $\mathrm{SL}(2, \mathbb{R})$ action, i.e. $V(gq) = gV(q)$ for all $g \in \mathrm{SL}(2, \mathbb{R})$. Such a V will be called an $\mathrm{SL}(2, \mathbb{R})$-*invariant locus*. Let $c(V, \mu)$ be the Siegel–Veech constant associated to V and μ (see Section 4.5 for the definition). We assume V satisfies $c(V, \mu) > 0$.

For $R \geq 1$, the ϵ/R-*thin part* of $\mathcal{N}$ *corresponding to* V is the subset of $q \in \mathcal{N}$ for which some saddle connection with holonomy in $V(q)$ has q-length squared less than ϵ/R. We denote this set by $\mathcal{N}(V)_{\epsilon/R}$. The regularity condition and the Siegel–Veech formula (4.6) can be used to prove the volume asymptotic

$$\lim_{R \to \infty} \frac{\mu(\mathcal{N}(V)_{\epsilon/R})}{\pi \epsilon/R} = c(V, \mu).$$

See [11, Section 7] for the main ideas.

Here, for μ-almost every $q \in \mathcal{N}$, we consider excursions of the Teichmüller geodesic ray determined by q in horoballs in $\mathbb{D}(q)$ for saddle connections with holonomy in V. Let $E_V(\gamma, T)$ be the sum till time T of excursions of γ in horoballs for saddle connections with holonomy in V. Let N_V be the number of such excursions of γ till time T. We prove:

Theorem 1.9. *Let μ be a regular* $\mathrm{SL}(2, \mathbb{R})$-*invariant measure supported on an affine invariant submanifold $\mathcal{N}$. For μ-almost every $q \in \mathcal{N}$, the Teichmüller geodesic ray γ that q determines satisfies*

$$\lim_{T \to \infty} \frac{E_V(\gamma, T) - \max_{k \leq N_V} E(\gamma, H_k)}{T \log T} = 2\epsilon c(V, \mu),$$

where $c(V, \mu)$ is the Siegel–Veech constant associated to V and μ.

Theorem 1.8 is a special case of Theorem 1.9 when $\mathcal{N} = \mathcal{Q}(\alpha)$, $\mu = \mu_{\mathrm{hol}}$ and V is the $\mathrm{SL}(2, \mathbb{R})$-invariant locus of holonomies of all saddle connections.

Configurations with cylinders. For the analog of Theorem 1.6, we state a special case of Theorem 1.9. For completeness we give some background.

For a connected component of $\mathcal{Q}(\alpha)$, a *configuration* $\mathcal{C}$ of saddle connections is a geometric type of maximal collections of homologous saddle connections on a translation or half-translation surface in it. Here, the homology is the appropriate relative homology (see [24] for details). The homology condition implies that the saddle connections in a configuration are parallel. For holomorphic 1-forms their holonomies coincide. For quadratic differentials their holonomies can take two values which differ by a factor of 2. The saddle connections with the smaller holonomy will be called the *small saddle connections* in $\mathcal{C}$. It was shown in [11] and [24] that in a μ_{hol}-typical degeneration all saddle connections in some configuration shrink to length zero. A configuration $\mathcal{C}$ gives an $\mathrm{SL}(2, \mathbb{R})$-invariant locus $V_{\mathcal{C}}$ for $\mathcal{Q}(\alpha)$. It follows that $c(V, \mu_{\mathrm{hol}}) > 0$.

A special subset of configurations corresponds to metric cylinders. A *metric cylinder* is an embedded cylinder that is a union of freely homotopic closed trajectories of q such that the boundary components are a concatenation of saddle connections. If some of the saddle connections in a configuration $\mathcal{C}$ bound a metric cylinder, we call $\mathcal{C}$ a *configuration with cylinders*. Masur and Zorich [24] showed that each boundary component of such cylinders has exactly one or two saddle connections in $\mathcal{C}$. The q-length of the core curve is equal to the boundary saddle connection or twice the length of one of the boundary saddle connections depending on the case.

Given $\mathcal{C}$, the *$\mathcal{C}$-thin part* of $SL(2, \mathbb{R})(q)$ consists of all points for which the length squared of the small saddle connections in $\mathcal{C}$ is less than ϵ. Its projection to $\mathbb{D}(q)$ is a horoball. The point at infinity for the horoball is the direction in which the saddle connections in $\mathcal{C}$ are vertical.

If ϵ is sufficiently small compared to the q-area of a cylinder then the core curve is also short in the hyperbolic metric. For $0 < \sigma < 1$ small enough depending on the orbit closure, we can specialize further to configurations with cylinders in which some cylinder has area at least σ. Such a restriction gives a horoball "packing": any point in $\mathbb{D}(q)$ is contained in at most $3g-3+n$ horoballs. By construction, the packing is $\mathrm{Mod}(S)$-equivariant. Masur [22] showed that in each Teichmüller disc the packing satisfies Sullivan's criteria and used it to prove the lower bound in the log law: a Leb-typical geodesic ray in every Teichmüller disc is recurrent to the thick part with lim sup of the maximum depth in $\mathcal{T}(S)_\epsilon$ asymptotically of size $(1/2)\log T$.

Let V be the subset of holonomies of saddle connections forming configurations with cylinders such that some cylinder has area at least σ. It follows from the Siegel–Veech machinery and [22, Proposition 2.5] that if σ is small enough depending on $\mathcal{Q}(\alpha)$ then $c_{\mathrm{cyl}_\sigma} > 0$, where c_{cyl_σ} is the Siegel–Veech constant for $\mathcal{Q}(\alpha)$ for configurations with cylinders such that some cylinder is of area at least σ. For translation surfaces, there are finer results due to Vorobets, for example [30, Theorem 1.8]. For a geodesic $\gamma \in \mathcal{Q}(\alpha)$, let $E_{\mathrm{cyl}_\sigma}(\gamma, T)$ be the sum till time T of excursions of γ in horoballs for such configurations. Let $N_{\mathrm{cyl}_\sigma}(T)$ be the number of such excursions till time T. As a special case of Theorem 1.9 we have

Theorem 1.10. *For μ_{hol}-almost every $q \in \mathcal{Q}(\alpha)$, the Teichmüller geodesic γ that q determines satisfies*

$$\lim_{T \to \infty} \frac{E_{\mathrm{cyl}_\sigma}(\gamma, T) - \max_{k \leq N_{\mathrm{cyl}_\sigma}} E(\gamma, H_k)}{T \log T} = 2\epsilon c_{\mathrm{cyl}_\sigma}(\alpha),$$

where $c_{\mathrm{cyl}_\sigma}(\alpha)$ is the Seigel-Veech constant for $\mathcal{Q}(\alpha)$ for configurations with cylinders such that some cylinder is of area at least σ.

Using Theorem 1.10 for the principal stratum, we can prove a lower bound on word-metric growth along typical Teichmüller geodesics. Along a Teichmüller geodesic, the twisting in the core curve of a metric cylinder is, up to a uniform multiplicative constant, A/ϵ times the excursion, where A is the q-area of the cylinder [15, Proposition 2.7]. By Mumford compactness, the closure of $\mathcal{M}(S) \setminus \mathcal{M}(S)_\epsilon$ is compact. Hence, by the Švarc–Milnor lemma, $\mathrm{Mod}(S)$ is quasi-isometric to the thick part $\mathcal{T}(S) \setminus \mathcal{T}(S)_\epsilon$. With a base-point X_0 in the thick part, the orbit $\mathrm{Mod}(S)X_0$ will be called a *Teichmüller lattice*. If γ is recurrent to the thick part then along recurrence times γ_T, there is a lattice point $h_T X_0$ closest to γ_T. The distance between γ_T and $h_T X_0$ is bounded by the diameter of $\mathcal{M}(S) \setminus \mathcal{M}(S)_\epsilon$. Because of compactness, this diameter is finite. As shown in [15, Proposition 3.11], along a recurrent Teichmüller geodesic γ the total excursion $E(\gamma, T)$ in the Masur collection gives a coarse lower bound on the word metric of the approximating group elements h_T, i.e. there exist constants $a_1, a_2 > 0$ such that

$$d_{\mathrm{Mod}(S)}(1, h_T) \geq a_1 E(\gamma, T) - a_2.$$

Thus, as a direct consequence of Theorem 1.10 in the principal stratum we get

Theorem 1.11. *There exists a constant $M_2 > 0$ depending on the word metric such that for almost every $q \in \mathcal{Q}(X_0)$ the approximating group elements h_T along the Teichmüller geodesic γ that q determines satisfy*

$$M_2 T \log T < d_{\mathrm{Mod}(S)}(1, h_T)$$

for all T sufficiently large depending on q.

Suppose q_1, q_2 have the same uniquely ergodic measured foliation as their vertical foliation. For such quadratic differentials, Masur [20, Theorem 2] showed that the geodesic rays that they determine are positively asymptotic. For any base-point $X_0 \in \mathcal{T}(S)$, a μ_{hol}-typical $q \in \mathcal{Q}(X_0)$ has a uniquely ergodic vertical foliation. Thus, Theorem 1.11 implies

Theorem 1.12. *Let $X_0 \in \mathcal{T}(S)$ be a base-point. There is a constant $M_3 > 0$ such that for Leb-almost every $q \in Q(X_0)$ the approximating group elements h_T along the Teichmüller geodesic γ that q determines satisfy*

$$M_3 T \log T < d_{\mathrm{Mod}(S)}(1, h_T)$$

for all T sufficiently large depending on q.

Strategy of proof

The main idea is to approximate the sum of excursions till T by an integral over time of a function defined over $T^1 X_{\mathrm{cusps}}$ or $\mathcal{N}(V)_\epsilon$. This function is not L^1. Analyzing the largest excursion, we prove that if it exceeds $T(\log T)^c$ for some $1 > c > 1/2$, then it is the unique excursion that exceeds this threshold. This follows from a Borel–Cantelli argument which requires quasi-independence of excursions. We use mixing of the geodesic flow to establish quasi-independence (Proposition 3.2). By removing the largest excursion from the sum we get a quantity that can be approximated by a suitable T-dependent truncation of the above function. This truncation is L^1. The leading term of its L^1-norm is a constant times $\log T$. The constant is in terms of the proportional volume of the cusp neighborhoods. To deduce the main theorems, we apply an effective ergodic theorem to the truncation. This shows that the integral over $[0, T]$ of the truncation is equal to T times the L^1-norm of the truncation with an error term which is $o(T \log T)$. To prove the effective ergodic theorem (Theorem 2.4), we use a decay of correlations for the geodesic flow. This decay of correlations is independently due to Moore and Ratner [26], [27] in the context of non-uniform lattices. For quadratic differentials, the decay is due to Avila–Resende [5] for the Masur–Veech measure, and Avila–Gouëzel [2] for general $\mathrm{SL}(2, \mathbb{R})$-invariant measures.

In the setting of quadratic differentials, matters are complicated by the fact that a half-translation surface can have several non-parallel short saddle connections. While this number is finite for any given half-translation surface, there is no upper bound for

it over the $SL(2, \mathbb{R})$ orbit closure. This means that a Teichmüller geodesic can do several excursions simultaneously and typically it does so. We impose a regularity condition on the $SL(2, \mathbb{R})$-invariant measure, namely quasi-independence for two non-parallel saddle connections to be simultaneously short. To prove that the truncation is L^1, the main technical work leverages two facts: the above quasi-independence and a bound due to Eskin and Masur [10] for the number of short saddle connections in terms of length of shortest saddle connection. We also show that the leading term of its L^1-norm is asymptotically a constant times $\log T$. The constant is related to the asymptotic of volumes of thin parts. By Siegel–Veech theory, these are the associated Siegel–Veech constants.

2. Ergodic theory

This section develops the ergodic-theoretic tools which will be used later. In particular, the main goal is to derive the effective ergodic theorem (Theorem 2.4) which gives a uniform rate of convergence in the ergodic theorem simultaneously for a sequence of non-negative functions that satisfy a certain decay of correlations.

Let $(X, \mathcal{B}, \ell)$ be a probability measure space. Let g_t be a measure preserving flow on X such that g_t is exponentially mixing. More precisely, we assume that an appropriate subspace of $L^2(X)$ satisfies the following decay of correlations: if f_1 and f_2 are functions in the subspace then $\int_X f_1 \, d\ell = \int_X f_2 \, d\ell = 0$ and there exist constants $K, \rho > 0$ such that

$$\left| \int_X f_1(g_s x) f_2(g_t x) \, d\ell \right| \leq K|t - s|e^{-\rho|t-s|} \|f_1\|_{L^2} \|f_2\|_{L^2}. \tag{2.1}$$

For $f \in L^1(X)$, let $I(f) = \int_X f \, d\ell$. We denote by $\mathcal{L}$ the subspace in $L^2(X)$ of functions f such that the decay of correlations (2.1) is satisfied when f_1 and f_2 are both set to $f - I(f)$.

Lemma 2.2. *Any function $f \in \mathcal{L}$ with $I(f) = 0$ satisfies*

$$\int_X \left(\int_0^T f(g_t x) \, dt \right)^2 d\ell \leq 2KT \|f\|_{L^2}^2. \tag{2.3}$$

Proof. Observe that

$$\int_X \left(\int_0^T f(g_t x) \, dt \right)^2 d\ell = \int_X \left(\int_0^T \int_0^T f(g_s x) f(g_t x) \, ds \, dt \right) d\ell$$

$$= \int_0^T \int_0^T \left(\int_X f(g_s x) f(g_t x) d\ell \right) ds \, dt$$

$$\leq \int_0^T \int_0^T K|t - s|e^{-\rho|t-s|} \|f\|_{L^2}^2 \, ds \, dt$$

where we have used (2.1) in the last inequality. A direct computation shows

$$\int_0^T \int_0^T K|t - s|e^{-\rho|t-s|} \|f\|_{L^2}^2 \, ds \, dt = K\|f\|_{L^2}^2 \left(\frac{T}{\rho^2}(1 + e^{-\rho T}) + \frac{2}{\rho^3}(-1 + e^{-\rho T}) \right)$$

$$\leq 2KT \|f\|_{L^2}^2. \qquad \square$$

Suppose $n : \mathbb{R} \to \mathbb{N}$ is a function that is constant on each interval $[2^k, 2^{k+1})$.

Theorem 2.4. *For any $c > 1/2$, $m > 1$ and any sequence of non-negative functions $f_j \in \mathcal{L}$, almost every x satisfies*

$$\frac{1}{m} T \| f_n \|_{L^1} - T^{1/2} (\log T)^c (\| f_n \|_{L^2}^2 - \| f_n \|_{L^1}^2)^{1/2}$$

$$\leq \int_0^T f_n(g_t x) \, dt \leq mT \| f_n \|_{L^1} + T^{1/2} (\log T)^c (\| f_n \|_{L^2}^2 - \| f_n \|_{L^1}^2)^{1/2}$$

for all T large enough depending on x and m and where $n = n(T)$.

Proof. Given $f \in \mathcal{L}$ define
$$F(x) = f(x) - I(f).$$
Then $I(F) = 0$ and so by Lemma 2.2,

$$\int_X \left(\int_0^T F(g_t x) \, dt \right)^2 d\ell \leq 2KT \| F \|_{L^2}^2$$

for all T. By Chebyshev's inequality, for any positive function $r(T, F)$ we have

$$\ell\left(\left\{ x : \left(\int_0^T F(g_t x) \, dt \right)^2 \geq r(T, F) \right\} \right) \leq \frac{2KT \| F \|_{L^2}^2}{r(T, F)}. \tag{2.5}$$

Let $c > 1/2$ and set $r(T, F) = T (\log T)^{2c} \| F \|_{L^2}^2$ in (2.5). Then we get

$$\ell\left(\left\{ x : \left(\int_0^T F(g_t x) \, dt \right)^2 \geq T (\log T)^{2c} \| F \|_{L^2}^2 \right\} \right) \leq \frac{2K}{(\log T)^{2c}}. \tag{2.6}$$

Starting from our sequence f_j, let F_j be the sequence of functions given by

$$F_j(x) = f_j(x) - I(f_j).$$

The estimate (2.6) above is satisfied by all functions F_j and in particular by F_n where $n = n(T)$. Fix $r = 1/a$ for some positive integer $a > 1$. Observe that for $T_k = 2^{rk}$,

$$\sum_{k=1}^{\infty} \frac{2K}{(\log T_k)^{2c}} = \sum_{k=1}^{\infty} \frac{2K}{(rk)^{2c}} < \infty.$$

Hence by the Borel–Cantelli lemma, almost every x satisfies

$$\left(\int_0^{T_k} F_n(g_t x) \, dt \right)^2 \leq T_k (\log T_k)^{2c} \| F_n \|_{L^2}^2 \tag{2.7}$$

for all k large enough depending on x and a. Similarly, if we set $r(T, F) = (T/2^r)(\log(T/2^r))^{2c} \| F \|_{L^2}^2$ and shift $n(T)$ to $n(T/2)$, the same reasoning by using the Borel–Cantelli lemma implies that almost every x satisfies

$$\left(\int_0^{T_{k+1}} F_n(g_t x) \, dt \right)^2 \leq T_k (\log T_k)^{2c} \| F_n \|_{L^2}^2 \tag{2.8}$$

for all k large enough depending on x and a. Hence a full measure set of x satisfies both (2.7) and (2.8). Noting that $F_n(g_t x) = f_n(g_t x) - I(f_n)$ and $\|F_n\|_{L^2}^2 = \|f_n\|_{L^2}^2 - I(f_n)^2$, we can rewrite the above estimates as

$$\left| \int_0^{T_k} f_n(g_t x)\, dt - T_k I(f_n) \right| \le T_k^{1/2} (\log T_k)^c (\|f_n\|_{L^2} - I(f_n)^2)^{1/2},$$

$$\left| \int_0^{T_{k+1}} f_n(g_t x)\, dt - T_{k+1} I(f_n) \right| \le T_k^{1/2} (\log T_k)^c (\|f_n\|_{L^2} - I(f_n)^2)^{1/2}.$$

Over the intermediate times $T_k < T < T_{k+1}$ the number n does not vary. So the function f_n being considered remains the same. Now we use the assumption that f_n is a non-negative function to get an estimate as above for the intermediate times. Since f_n is non-negative, the time integral of f_n is non-decreasing. In particular,

$$\int_0^{T_k} f_n(g_t x)\, dt \le \int_0^T f_n(g_t x)\, dt \le \int_0^{T_{k+1}} f_n(g_t x)\, dt.$$

Observe that

$$T_k I(f_n) - T_k^{1/2}(\log T_k)^c (\|f_n\|_{L^2}^2 - I(f_n)^2)^{1/2}$$
$$\ge \frac{1}{2^r} T I(f_n) - T^{1/2}(\log T)^c (\|f_n\|_{L^2}^2 - I(f_n)^2)^{1/2} \quad (2.9)$$

and

$$T_{k+1} I(f_n) + T_k^{1/2}(\log T_k)^c (\|f_n\|_{L^2}^2 - I(f_n)^2)^{1/2}$$
$$\le 2^r T I(f_n) + T^{1/2}(\log T)^c (\|f_n\|_{L^2}^2 - I(f_n)^2)^{1/2}. \quad (2.10)$$

Finally, note $I(f_n) = \|f_n\|_{L^1}$. The left hand side of (2.9) is the lower bound in (2.7) and the left hand side in (2.10) is the upper bound in (2.8). Thus we get

$$\frac{1}{2^r} T I(f_n) - T^{1/2}(\log T)^c (\|f_n\|_{L^2}^2 - \|f_n\|_{L^1}^2)^{1/2}$$
$$\le \int_0^T f_n(g_t x)\, dt \le 2^r T I(f_n) + T^{1/2}(\log T)^c (\|f_n\|_{L^2}^2 - \|f_n\|_{L^1}^2)^{1/2}.$$

The theorem follows by choosing a so large that $r = 1/a$ satisfies $2^r < m$. $\qquad\square$

We also prove a variant of Lemma 2.2 which we will need later for quasi-independence of excursions.

Lemma 2.11. *For any $S_1 < S_2 < T$ and every non-negative function $f \in \mathcal{L}$,*

$$\int_X \left(\int_{S_1}^{S_2} f(g_s x)\, ds \int_{S_2}^T f(g_t x)\, dt \right) d\ell$$
$$< (S_2 - S_1)(T - S_2)\|f\|_{L^1}^2 + \frac{5K}{\rho}(\|f\|_{L^2}^2 - I(f)^2) \quad (2.12)$$

where K, c are the constants in the decay of correlations (2.1).

Proof. For any $f \in \mathcal{L}$, define

$$F(x) = f(x) - I(f).$$

Then $I(F) = 0$ and so F satisfies the decay of correlations (2.1). Note that

$$\int_X F(g_s x) F(g_t x) \, d\ell = \int_X f(g_s x) f(g_t x) \, d\ell - I(f)^2$$

and $\|F\|_{L^2}^2 = \|f\|_{L^2}^2 - I(f)^2$. It follows that

$$\left| \int_X f(g_s x) f(g_t x) \, d\ell - I(f)^2 \right| \leq K|t - s|e^{-\rho|t-s|}(\|f\|_{L^2}^2 - I(f)^2),$$

which implies

$$\left| \int_X f(g_s x) f(g_t x) \, d\ell \right| \leq I(f)^2 + K|t - s|e^{-\rho|t-s|}(\|f\|_{L^2}^2 - I(f)^2).$$

For non-negative functions f this implies

$$\int_X \left(\int_{S_1}^{S_2} f(g_s x) \, ds \int_{S_2}^{T} f(g_t x) \, dt \right) d\ell = \int_{S_2}^{T} \int_{S_1}^{S_2} \left(\int_X f(g_s x) f(g_t x) \, d\ell \right) ds \, dt$$

$$\leq \int_{S_2}^{T} \int_{S_1}^{S_2} [I(f)^2 + K|t - s|e^{-\rho|t-s|}(\|f\|_{L^2}^2 - I(f)^2)] \, ds \, dt$$

$$< (S_2 - S_1)(T - S_2)I(f)^2 + \frac{5K}{\rho}(\|f\|_{L^2}^2 - I(f)^2)$$

where the last inequality follows from a direct computation. $\qquad\square$

3. Partial sums of excursions for non-uniform lattices in $\mathrm{SL}(2, \mathbb{R})$

The Liouville measure ℓ on $T^1\mathbb{H}^2$ is invariant under the $\mathrm{SL}(2, \mathbb{R})$ action and descends to a flow-invariant measure on $T^1 X = \Gamma \backslash T^1\mathbb{H}^2$. To get a probability measure ℓ on $T^1 X$ we normalize by passing to

$$d\ell \to \frac{1}{2\pi |\chi(X)|} d\ell,$$

where $\chi(X)$ is the Euler characteristic of X. For notational simplicity we continue to call the probability measure $d\ell$.

The geodesic flow on $T^1 X$ is given by the action of the diagonal subgroup of $\mathrm{SL}(2, \mathbb{R})$. It is ergodic with respect to ℓ. In fact, it is known to be exponentially mixing. As shown in [26], $\mathrm{SO}(2, \mathbb{R})$-invariant L^2-functions on $T^1 X$ satisfy the following decay of correlations for the diagonal flow: There exist constants $K, \rho > 0$ such that for any pair f_1, f_2 of $\mathrm{SO}(2, \mathbb{R})$-invariant L^2-functions on $T^1 X$ with $\int_{T^1 X} f_1 \, d\ell = \int_{T^1 X} f_2 \, d\ell = 0$,

$$\left| \int_{T^1 X} f_1(x) f_2(g_t x) \, d\ell \right| \leq K t e^{-\rho t} \|f_1\|_{L^2} \|f_2\|_{L^2}. \tag{3.1}$$

See also [27, Theorem 2], [25, Corollary 2.1]. In particular, the lifts to T^1X of L^2-functions on X are $SO(2, \mathbb{R})$-invariant by default, so the above decay of correlations applies to them.

For $R \geq 1$, let Y_R be the subset of the horoballs $\mathcal{H}$ consisting of those points which are at least $\log R$ away from the boundary of the horoballs in the hyperbolic metric, i.e.

$$Y_R := \bigcup_{H \in \mathcal{H}} \{x \in H : d(x, \partial H) \geq \log R\}.$$

Let $X_R \subset X$ be the quotient of $\Gamma \backslash Y_R$. In particular, $X_1 = X_{\text{cusp}}$. We will write T^1Y for the restriction of the unit tangent bundle to any subset $Y \subset X$. An elementary calculation shows that

$$\ell(T^1X_R) = \frac{1}{R} \frac{\ell(T^1X_{\text{cusp}})}{\ell(T^1X)} = \frac{C_X}{R}$$

where to simplify notation, henceforth we set $\ell(T^1X_{\text{cusp}})/\ell(T^1X)$ to be C_X. Let χ_R be the characteristic function of T^1X_R and let

$$\phi_R = \chi_{R/2} - \chi_R.$$

Note that

$$\|\phi_R\|_{L^1} = C_X/R \quad \text{and} \quad \|\phi_R\|_{L^2} = \sqrt{C_X}/\sqrt{R}.$$

During an excursion of size at least R, a geodesic γ must cross $T^1X_{R/2} \backslash T^1X_R$ twice during a complete excursion and at least once during a partial excursion. By basic hyperbolic geometry, the geodesic spends time greater than $\log 2$ each time it crosses $T^1X_{R/2} \backslash T^1X_R$.

The next proposition allows us to show that along Leb-almost every geodesic ray, for all times T large enough there is at most a single "large" excursion. The proposition is a continuous time refinement of [9, Lemma 2.1] and the proof uses Lemma 2.11.

Proposition 3.2. *For any $c > 1/2$ and for ℓ-almost every $v \in T^1X$ there exists $T(v)$ such that for all $T > T(v)$,*

$$E(\gamma, H_i) \geq T(\log T)^c$$

for at most a single $H_i \in H_{\gamma, T}$, where γ is the geodesic ray with $v(\gamma_0) = v$.

For the rest of the discussion, let $T_n = 2^n$. Proposition 3.2 follows from

Proposition 3.3. *For any $c > 1/2$ and for ℓ-almost every $v \in T^1X$ there exists a non-negative integer $n(v)$ such that for all $n > n(v)$,*

$$E(\gamma, H_i) \geq T_{n-1}(\log T_{n-1})^c$$

for at most a single $H_i \in H_{\gamma, T_n}$, where γ is the geodesic ray with $v(\gamma_0) = v$.

Proof. Let $\lambda = \log T_{n-1} + c \log \log T_{n-1}$. By basic hyperbolic geometry, the time it takes a geodesic to go from the boundary of a horoball to X_R where $R = T_{n-1}(\log T_{n-1})^c$ is bounded between λ and $\log(T_{n-1}(\log T_{n-1})^c + \sqrt{T_{n-1}^2(\log T_{n-1})^{2c} - 1}) < \lambda + \log 2$. Similarly, let $\lambda' = \lambda - \log 2$. Then λ' is a lower bound on the time it takes a geodesic to go from the boundary of a horoball to $X_{R/2}$ where $R = T_{n-1}(\log T_{n-1})^c$.

For positive integers $j \leq \lfloor T_n/\lambda \rfloor$, let $S_j = j\lambda$. Let

$$V_{n,k} = \left\{ v \in T^1X : \int_{S_k}^{S_{k+2}+\lambda'} \phi_R(v(\gamma_s))\, ds > \log 2 \text{ and } \int_{S_{k+2}+\lambda'}^{T_n} \phi_R(v(\gamma_t))\, dt > \log 2 \right\}.$$

By applying Chebyshev's inequality to the estimate in Lemma 2.11 for the function ϕ_R we get

$$
\begin{aligned}
\ell(V_{n,k}) &\leq \frac{(S_{k+2} + \lambda' - S_K)(T_n - S_{k+2} - \lambda')}{(\log 2)^2} \frac{C_X^2}{R^2} + \frac{1}{(\log 2)^2} \frac{5KC_X}{\rho R}\left[1 - \frac{C_X}{R}\right] \\
&< \frac{6\lambda C_X^2}{(\log 2)^2 T_{n-1}(\log T_{n-1})^{2c}} + \frac{5KC_X}{\rho(\log 2)^2 T_{n-1}(\log T_{n-1})^c}.
\end{aligned}
$$

In the last inequality we have used $T_n - S_{k+2} - \lambda' < T_n = 2T_{n-1}$.

Let W be the set of $v \in T^1X$ such that the corresponding geodesic γ has two excursions $E(\gamma, H_i)$ and $E(\gamma, H_j)$ till time T_n satisfying $E(\gamma, H_i) \geq T_{n-1}(\log T_{n-1})^c$ and $E(\gamma, H_j) \geq T_{n-1}(\log T_{n-1})^c$. Let S be the time at which the first big excursion $E(\gamma, H_i)$ begins. Let k be such that $S_k \leq S < S_{k+1}$. Because of our choice of λ the first excursion $E(\gamma, H_i)$ must end after S_{k+2}. Thus, the second big excursion $E(\gamma, H_j)$ also has to begin after S_{k+2}. Because of the choice of λ' the geodesic cannot cross $T^1X_{R/2} \setminus T^1X_R$ during $E(\gamma, H_j)$ before $S_{k+2} + \lambda'$. Thus, $v \in V_{n,k}$. Let

$$V_n = \bigcup_{k=0}^{\lfloor T_n/\lambda \rfloor - 2} V_{n,k}.$$

Using the upper bound on $\ell(V_{n,k})$ we get

$$
\begin{aligned}
\ell(V_n) \leq \sum_{k=1}^{\lfloor T_n/\lambda \rfloor} \ell(V_{n,k}) &< \frac{6T_n C_X^2}{(\log 2)^2 T_{n-1}(\log T_{n-1})^{2c}} + \frac{5KT_n C_X}{\rho(\log 2)^2 \lambda T_{n-1}(\log T_{n-1})^c} \\
&< \frac{b_1 T_n}{T_{n-1}(\log T_{n-1})^{2c}} + \frac{b_2 T_n}{T_{n-1}(\log T_{n-1})^{1+c}} \\
&\leq \frac{2b_1}{(\log T_{n-1})^{2c}} + \frac{2b_2}{(\log T_{n-1})^{1+c}}
\end{aligned}
$$

for some constants $b_1, b_2 > 0$. Since $c > 1/2$ it follows that $\sum_n \ell(V_n) < \infty$. Proposition 3.3 then follows by the Borel–Cantelli lemma. $\qquad \square$

Proof of Proposition 3.2. Let n be such that $T_{n-1} < T \leq T_n$. Since $E(\gamma, H) \geq T(\log T)^c$ implies $E(\gamma, H) > T_{n-1}(\log T_{n-1})^c$, Proposition 3.2 follows from Proposition 3.3. $\qquad \square$

Remark 3.4. It is important to observe that Proposition 3.3 holds under the weaker condition that there is a constant $A > 1$ such that for R sufficiently large

$$1/(AR) < \ell(T^1 X_R) < A/R.$$

This observation will be of relevance for a similar proposition in the setting of quadratic differentials.

Recall that x_0 is a base-point and $T_{x_0}^1 X$ can be identified with S^1.

Corollary 3.5. *For any $c > 1/2$ and Leb-almost every $r \in S^1$ there is $T(r)$ such that if $T > T(r)$ then*

$$E(\gamma, H_i) \geq T(\log T)^c$$

for at most a single $H_i \in H_{\gamma, T}$, where γ is the geodesic ray from x_0 to r.

Proof. It follows from Proposition 3.2 that the conclusion is true for generic base-points. Suppose γ_0 and γ_1 are geodesic rays from distinct base-points x_0 and x_1 converging to the same point r at infinity and let H be a horoball. Let π_H be the closest point projection to H and let $a = d_{\partial H}(\pi_H x_0, \pi_H x_1)$. Then we have the crude bound

$$E(\gamma_0, H) - 2ae^{-\tau} - 2 \leq E(\gamma_1, H) \leq E(\gamma, H) + 2ae^{-\tau} + 2,$$

where τ is the minimum of $d(x_0, H)$ and $d(x_1, H)$. So for H that is far enough the excursions by γ_0 and γ_1 are the same up to a uniform additive constant. This implies the corollary. $\square$

Define the function $\psi : X \to \mathbb{R}$ by

$$\psi(x) = \begin{cases} 0 & \text{if } x \in X_{\text{thick}}, \\ (2/\pi)e^{d(x, \partial X_{\text{thick}})} & \text{otherwise.} \end{cases}$$

Let $\Psi : T^1 X \to \mathbb{R}$ be the lift of ψ to $T^1 X$. By definition, the function Ψ is $SO(2, \mathbb{R})$-invariant.

Suppose a geodesic ray γ has a complete excursion in a horoball H entering and exiting H at times T_1 and T_2 respectively. We relate the time integral of Ψ between T_1 and T_2 to $E(\gamma, H)$.

Conjugating if necessary, we may assume that H is the horoball $\{z : \text{Im}(z) \geq 1\}$ in the upper half-space model. The function ψ then takes the form $\psi(z) = (2/\pi)\text{Im}(z)$ on H and 0 otherwise. Conjugating further by $z \mapsto z + a$ for some $a \in \mathbb{R}$ we may assume that γ is the geodesic with endpoints $\{A, -A\} \in \mathbb{R}$ for some $A \geq 1$. We parameterize γ with unit speed from A to $-A$. Then, as complex numbers, $\gamma(T_1) = \sqrt{A^2 - 1} + i$ and $\gamma(T_2) = -\sqrt{A^2 - 1} + i$. Note that $E(\gamma, H) = 2\sqrt{A^2 - 1}$. Let $\theta_0 = \text{Arg}(\gamma(T_1))$. Then $\text{Arg}(\gamma(T_2)) = \pi - \theta_0$. By elementary integration, it follows that

$$\int_{T_1}^{T_2} \Psi(v(\gamma_t)) \, dt = \int_{\theta_0}^{\pi - \theta_0} \frac{2A}{\pi} \, d\theta = \frac{2A}{\pi}(\pi - 2\theta_0) = 2A - \frac{4A\theta_0}{\pi}$$

Since $\sin \theta_0 = 1/A$, it follows that

$$E(\gamma, H) - 2 < \int_{T_1}^{T_2} \Psi(v(\gamma_t)) \, dt < E(\gamma, H) + 2, \qquad (3.6)$$

i.e. for complete excursions the time integral of Ψ is $E(\gamma, H)$ up to a uniform additive error.

To analyze a partial excursion, we parameterize γ by $Ae^{i\theta}$ between θ_0 and $\pi - \theta_0$. Suppose $\theta \in (\theta_0, \pi - \theta_0)$. Then the partial excursion till $Ae^{i\theta}$ is given by

$$E(\gamma, H) = A(\cos \theta_0 - \cos \theta).$$

On the other hand, the time integral is given by

$$\int \Psi(v(\gamma_t)) \, dt = \frac{2A}{\pi}(\theta - \theta_0),$$

where $T_3 < T$ are the times such that $\gamma_{T_3} = Ae^{i\theta_0}$ and $\gamma_T = Ae^{i\theta}$. The difference of the right hand sides above is symmetric about $\pi/2$ so we may restrict to $(\theta_0, \pi/2]$. Then the difference has a crude upper bound given by

$$A\left(\frac{2}{\pi}(\theta - \theta_0) - (\cos \theta_0 - \cos \theta)\right) \leq \frac{2A}{\pi}\theta \leq A \sin \theta.$$

Note that $\Psi(v(\gamma_\theta)) = (2/\pi)\mathrm{Im}(Ae^{i\theta}) = (2/\pi)A \sin \theta$. It follows that for a partial excursion

$$E(\gamma, H) - \frac{\pi}{2}\Psi(v(\gamma_T)) < \int_{T_3}^{T} \Psi(v(\gamma_t)) \, dt < E(\gamma, H) + \frac{\pi}{2}\Psi(v(\gamma_T)), \qquad (3.7)$$

where $T_3 < T$ is the time at which the excursion begins.

Recall that χ_R is the characteristic function of $T^1 X_R$. We define truncations of Ψ by

$$\Psi_R(v) = \Psi(v)(\chi_1(v) - \chi_R(v)),$$

where χ_1 is the characteristic function of $T^1 X_1 = T^1 X_{\mathrm{cusp}}$. Note that while Ψ is not L^1, the truncations Ψ_R satisfy

$$\|\Psi_R\|_{L^1} = \frac{2C_X}{\pi} \log R \quad \text{and} \quad \|\Psi_R\|_{L^2} = \frac{2\sqrt{C_X}}{\pi}\sqrt{R}.$$

Inequalities (3.6) and (3.7) show that partial sum of excursions (minus the largest excursion), i.e. the numerator in Theorem 1.2, is estimated by the time integral of a suitable truncation of Ψ up to an additive error that is linear in the number N of excursions. The next lemma shows that N grows linearly in T.

Lemma 3.8. *There is a constant $\eta > 0$ such that for Leb-almost every $r \in S^1$,*

$$\lim_{T \to \infty} \frac{N}{T} = \eta.$$

Proof of Lemma 3.8. One can follow an approach similar to Schmidt's theorem in the theory of Diophantine approximation [1, Theorem 1.1 with $k = 1$]. For completeness, we give a weaker but more direct proof below.

For $v \in T^1(X \setminus X_{\text{thick}})$, let γ be the geodesic such that $v(\gamma_0) = v$, i.e. the geodesic whose unit tangent vector at $t = 0$ is v. For almost every v, the geodesic γ intersects $\partial X_{\text{thick}}$ both in the forward and backward directions. Let $T_b < 0$ and $T_f > 0$ be the first instances of these intersections, i.e. the first instances backward and forward along γ when it intersects the boundary of the horoball containing $\pi(v)$. Set

$$\xi(v) = \frac{1}{T_f - T_b}.$$

This defines a non-negative function $\xi : T^1(X \setminus X_{\text{thick}}) \to \mathbb{R}_{\geq 0}$ which we extend by setting it zero outside. It is straightforward to see that for almost every $r \in S^1$,

$$\int_0^T \xi(v(\gamma_t))\, dt = N.$$

We claim that $\xi \in L^1(T^1X)$. Let $R_k = 1 + 1/2^k$ and consider $T^1 X_{R_k} \setminus T^1 X_{R_{k-1}}$. Since $\log(1 + 1/2^k) = 1/2^k - 1/2^{2k+1} + $ higher order terms, there exists a constant $b_3 > 0$ such that

$$\ell(T^1 X_{R_{k-1}} \setminus T^1 X_{R_k}) < b_3/2^k$$

for all k large enough. By basic hyperbolic geometry, for any $v \in T^1 X_{R_k} \setminus T^1 X_{R_{k-1}}$,

$$\xi(v) \leq \frac{1}{2\log(1 + 1/2^k + \sqrt{(1 + 1/2^k)^2 - 1})}$$
$$< \frac{1}{2\log(1 + \sqrt{1/2^k} + 1/2^k)} < b_4 2^{k/2}$$

for some constant $b_4 > 0$ and all k large enough. This gives the bound

$$\int_{T^1 X_{R_k} \setminus T^1 X_{R_{k-1}}} \xi\, d\ell < \frac{b_3 b_4}{2^{k/2}},$$

which proves that $\xi \in L^1$. The lemma follows by applying the ergodic theorem to ξ. □

Proof of Theorem 1.2. While not necessary, for notational simplicity we set the constants $c > 1/2$ in Proposition 3.2 and in Theorem 2.4 to be equal. Consider the sequence of functions Ψ_{2^k}. For T such that $2^k \leq T < 2^{k+1}$ we set $n(T) = \lfloor k + c \log_2 k \rfloor$ where $\lfloor\ \rfloor$ is the greatest integer function. By Theorem 2.4 applied to the sequence of functions Ψ_{2^n}, we know that for any $c > 1/2$ and $m > 1$, ℓ-almost every $v = v(\gamma_0) \in T^1X$ satisfies the lower bound

$$\frac{1}{m} T \|\Psi_{2^n}\|_{L^1} - T^{1/2}(\log T)^c (\|\Psi_{2^n}\|_{L^2}^2 - \|\Psi_{2^n}\|_{L^1}^2)^{1/2} \leq \int_0^T \Psi_{2^n}(v(\gamma_t))\, dt$$

and the upper bound

$$\int_0^T \Psi_{2^n}(v(\gamma_t))\, dt \leq mT \|\Psi_{2^n}\|_{L^1} + T^{1/2}(\log T)^c (\|\Psi_{2^n}\|_{L^2}^2 - \|\Psi_{2^n}\|_{L^1}^2)^{1/2}$$

for all T large enough depending on v and m. Let $r = T(\log T)^c/2^n$. Then $1 < r < 3$. Substituting the L^1- and L^2-norms of Ψ_{2^n} in $T^{1/2}(\log T)^c(\|\Psi_{2^n}\|_{L^2}^2 - \|\Psi_{2^n}\|_{L^1}^2)^{1/2}$, we get the upper bound

$$T^{1/2}(\log T)^c \left(\frac{4C_X}{r\pi^2} T(\log T)^c - \frac{4C_X^2}{\pi^2}(\log T + c\log\log T - \log r)^2 \right)^{1/2}$$

$$\leq \frac{2\sqrt{C_X}}{\pi\sqrt{r}} T(\log T)^{3c/2}.$$

Thus the preceding pair of inequalities become

$$\frac{2C_X}{m\pi} T(\log T + c\log\log T - \log r) - \frac{2\sqrt{C_X}}{\pi\sqrt{r}} T(\log T)^{3c/2} \leq \int_0^T \Psi_{2^n}(v(\gamma_t))\,dt, \qquad (3.9)$$

$$\int_0^T \Psi_{2^n}(v(\gamma_t))\,dt \leq \frac{2mC_X}{\pi} T(\log T + c\log\log T - \log r) + \frac{2\sqrt{C_X}}{\pi\sqrt{r}} T(\log T)^{3c/2}. \quad (3.10)$$

We choose $c < 2/3$. Let U_m be the full measure set in T^1X satisfying (3.9) and (3.10). Consider the countable intersection

$$U = \bigcap_{a \in \mathbb{N}} U_{1+1/a}.$$

Then U has full measure and for v in U the constraint $c < 2/3$ implies

$$\lim_{T \to \infty} \frac{1}{T\log T} \int_0^T \Psi_{2^n}(v(\gamma_t))\,dt = \frac{2C_X}{\pi}. \qquad (3.11)$$

By the same reasoning as in the proof of Corollary 3.5 the above limit is true for any base-point x_0 and for Leb-almost every $r \in S^1$.

It remains to relate the time integral of Ψ_{2^n} to partial sums of excursions. Enumerate the horoballs in $\mathcal{H}_{\gamma,T}$ as $H_1, \ldots, H_N$ in the order of increasing time. For a complete excursion $E(\gamma, H_i)$ satisfying $E(\gamma, H_i) \leq T(\log T)^c$, the difference between $E(\gamma, H_i)$ and the time integral of Ψ_{2^n} between the entry and exit points for H_i is at most 2 by (3.6). This implies that the additive error contributed by complete excursions less than the threshold is bounded above by $2N$.

If $E(\gamma, H_i) > T(\log T)^c$ then by Corollary 3.5 it is the unique excursion that exceeds the threshold. Let $T_1 < T_2$ be the entry and exit times for H_i. In this case, note that

$$\int_{T_1}^{T_2} \Psi_{2^n}(v(\gamma_t))\,dt < 2T(\log T)^c.$$

If there is a partial excursion then let $T_3 < T$ be the time at which γ enters H_N. If the partial excursion exceeds the threshold then notice that

$$\int_{T_3}^T \Psi_{2^n}(v(\gamma_t))\,dt < 2T(\log T)^c.$$

Otherwise, by (3.7) the difference between the time integral and the partial excursion is at most $(\pi/2)\Psi(v(\gamma_T))$, which in turn is bounded above by $T(\log T)^c$.

Using the estimates above and also (3.6) and (3.7) we get

$$\int_0^T \Psi_{2^n}(v(\gamma_t))\,dt - 2N - 4T(\log T)^c < E(\gamma, T) - \max_{1 \le k \le N} E(\gamma, H_k)$$

$$< \int_0^T \Psi_{2^n}(v(\gamma_t))\,dt + 2N + 4T(\log T)^c. \quad (3.12)$$

Theorem 1.2 then follows from putting together (3.11), (3.12) and Lemma 3.8. $\qquad \square$

Proof of the Diamond–Vaaler Theorem 1.3. For the modular surface $X = \mathrm{SL}(2, \mathbb{Z}) \backslash \mathbb{H}^2$ the lift to $\mathbb{H}^2$ of the largest embedded cusp neighborhood in X is the well-known Ford packing: in the upper half-space model, we get horoballs resting at rational points, the Euclidean radius of the horoball with the point at infinity p/q in reduced form being $1/(2q^2)$.

With the cusp neighborhood fixed as above, Theorem 1.2 for X states that for any base-point x_0 and for Leb-almost every $r \in S^1$,

$$\lim_{T \to \infty} \frac{E(\gamma, T) - \max_{k \le N} E(\gamma, H_k)}{T \log T} = \frac{6}{\pi^2}, \quad (3.13)$$

where γ is the geodesic ray from x_0 to r. To derive the Diamond–Vaaler result (Theorem 1.3) from the above limit, we relate excursions to continued fraction coefficients of r, and time T along the geodesic to the number n of continued fraction coefficients.

In the upper half-space model, for $r \in [0, 1]$ irrational, let $[a_1, a_2, \ldots]$ be the infinite continued fraction expansion of r. Let $p_n/q_n = [a_1, \ldots, a_n]$ be the n-th convergent of r and let H_n' be the horoball with p_n/q_n as the point at infinity. We first consider vertical geodesics: for $r \in [0, 1]$ let γ' be the vertical geodesic ray from $(r, i) \in \mathbb{H}^2$ to $(r, 0) \in S^1 = \mathbb{R} \cup \infty$.

The ray γ' has excursions in horoballs that are given by rational approximations of r satisfying $|r - p/q| \le 1/(2q^2)$. By a classical theorem for continued fractions, such rationals are a subset of the convergents p_n/q_n. If $a_n \ge 2$ then $a_n - 1 < E(\gamma', H_n') < a_n + 1$. However, if $a_n = 1$ then γ' may or may not intersect H_n' and we set $E(\gamma', H_n') = 0$ if it does not. In any case, excursions of γ' are equal to the coefficients up to a uniform additive error, and hence we get

$$\sum_{k=1}^n E(\gamma', H_k') - n \le \sum_{k=1}^n a_k \le \sum_{k=1}^n E(\gamma', H_k') + n. \quad (3.14)$$

By classical theory of continued fractions [8, Proposition 4.8.2(4)], for Leb-almost every r,

$$\lim_{n \to \infty} \frac{\log q_n}{n} = \frac{\pi^2}{12 \log 2}.$$

Since $p_n/q_n \to r$, the same limit is true for $(\log p_n)/n$. Up to a transposition of columns, the matrix Q_n with columns $[p_{n-1}, q_{n-1}]^t$ and $[p_n, q_n]^t$ is in $\mathrm{SL}(2, \mathbb{Z})$. The hyperbolic translation length of the matrix up to a uniform additive error is $2\log(\mathrm{trace})$. By the above discussion, $\log(\mathrm{trace})$ is $\log q_n$ up to a uniform additive error. So let $2\log q_n = T_n$. Recall that Q_n acts on the upper half-plane by Möbius transformations. Geometrically

$Q_n(r, i)$ is the orbit point closest to γ'_{T_n} with the distance between them bounded above by the diameter of X_{thick}, i.e. uniformly bounded from above. This implies that along the sequence of times $2 \log q_n = T_n$ the limit n/T_n is $(6 \log 2)/\pi^2$. It should be pointed out that the number N of horoballs that γ' actually intersects till T_n is less than or equal to n, and in fact N/T_n will have a different limit as $T_n \to \infty$.

The geodesic ray γ from x_0 to r and the vertical ray γ' are asymptotic. Set

$$a = \max_{H \in \mathcal{H}} d_{\partial H}(\pi_H x_0, \pi_H(r, i))$$

where π_H is the closest point projection to H. Then we have the crude bound

$$E(\gamma, H) - 2ae^{-\tau} - 2 < E(\gamma', H) < E(\gamma, H) + 2ae^{-\tau} + 2$$

where τ is the minimum of $d(x_0, H)$ and $d((r, i), H)$. Let d be the distance between horocycles with r at infinity that pass through x_0 and (r, i) respectively. Then depending on the case we get the crude bound

$$\begin{aligned} E(\gamma, T_n \pm d) &- 2(a+1)(n + 2\eta(T_n + d)) \\ &\leq \sum_{k \leq n} E(\gamma', H'_k) \leq E(\gamma, T_n \pm d) + 2(a+1)(n + 2\eta(T_n + d)). \end{aligned}$$

The estimate above implies that

$$\begin{aligned} \lim_{n \to \infty} &\frac{\sum_{k \leq n} E(\gamma', H'_k) - \max_{k \leq n} E(\gamma', H'_k)}{T \log T} \\ &= \lim_{T_n \to \infty} \frac{E(\gamma, T_n \pm d) - \max_{k \leq N} E(\gamma, H_k)}{T_n \log T_n} = \frac{6}{\pi^2} \end{aligned}$$

where the second equality follows from the fact that passing to T_n instead of $T_n \pm d$ in the numerator introduces an additive error of at most e^d. Finally, note that $a_k - 1 \leq E(\gamma', H'_k) \leq a_k + 1$, and so for Leb-almost every $r \in [0, 1]$,

$$\begin{aligned} \lim_{n \to \infty} &\frac{\sum_{k=1}^{n} a_k - \max_{k \leq n} a_k}{n \log n} \\ &= \lim_{n \to \infty} \frac{\sum_{k \leq n} E(\gamma', H'_k) - \max_{k \leq n} E(\gamma', H'_k)}{T_n \log T_n} \frac{T_n \log T_n}{n \log n} = \frac{6}{\pi^2} \frac{\pi^2}{6 \log 2} = \frac{1}{\log 2}, \end{aligned}$$

finishing the proof of Theorem 1.3. □

4. Partial sums along random Teichmüller geodesics in a stratum of quadratic differentials

4.1. Preliminaries from Teichmüller theory

Let S be a hyperbolic surface of finite type, i.e. a surface of finite area which may have boundary components or punctures. We say such a surface S is *sporadic* if it is a sphere with at most four punctures or boundary components, or a torus with at most one puncture

or boundary component. We shall be primarily interested in non-sporadic surfaces, as in the sporadic cases the Teichmüller spaces are either trivial, or isometric to $\mathbb{H}^2$, which reduces us to the case of a non-uniform lattice in $SL(2, \mathbb{R})$.

Let S be a non-sporadic surface which has no boundary components, but may have punctures. The Teichmüller space $\mathcal{T}(S)$ is the space of marked conformal structures on S. Alternatively, by uniformization, it is the space of marked hyperbolic metrics on S. We shall consider $\mathcal{T}(S)$ together with the Teichmüller metric

$$d_{\mathcal{T}}(X, Y) = \tfrac{1}{2} \inf_f \log K(f)$$

where the infimum is taken over all quasiconformal maps $f: X \rightarrow Y$, and $K(f)$ is the quasiconformal constant for the map f. The mapping class group $\mathrm{Mod}(S)$ acts by isometries on $\mathcal{T}(S)$. Let $\mathcal{T}(S)_\epsilon$ be the *thin part* of the Teichmüller space, i.e. all surfaces which contain a curve of hyperbolic length at most ϵ. Let $\mathcal{M}(S)$ be the moduli space $\mathrm{Mod}(S) \backslash \mathcal{T}(S)$. The thin part $\mathcal{T}(S)$ is $\mathrm{Mod}(S)$-invariant. The thin part $\mathcal{M}(S)_\epsilon$ of the moduli space is the quotient $\mathrm{Mod}(S) \backslash \mathcal{T}(S)_\epsilon$.

Let $\mathcal{Q}(X)$ be the unit area meromorphic quadratic differentials on X with simple poles at all the punctures of X. If $(k_1, \ldots, k_r)$ are the multiplicities of the zeros of a quadratic differential q then $k_1 + \cdots + k_r = 4g - 4 + 2m$ where m is the number of punctures of X. By contour integration, a quadratic differential q defines a half-translation structure on S, i.e. it defines charts from S to $\mathbb{C}$ with transition functions of the form $z \mapsto \pm z + c$. The resulting flat metric has a cone singularity with cone angle $(k + 2)\pi$ at a zero of q of order k (or with $k = -1$ at a simple pole). A quadratic differential is *unit area* if the corresponding flat metric has unit area. The space $\mathcal{Q}$ of unit area quadratic differentials can be identified with the unit cotangent bundle to $\mathcal{T}(S)$ [16]. We let $\pi : \mathcal{Q} \rightarrow \mathcal{T}(S)$ be the projection which sends a quadratic differential to its underlying Riemann surface. The space $\mathcal{Q}$ is stratified by the multiplicities of the zeros: we denote the strata with multiplicities $\alpha = (k_1, \ldots, k_r)$ by $\mathcal{Q}(\alpha)$. For each stratum, the number of connected components is bounded [18], [7]. To simplify notation, we continue to write $\mathcal{Q}(\alpha)$ for a connected component of $\mathcal{Q}(\alpha)$.

For any $q \in \mathcal{Q}(\alpha)$ there is a canonical ramified double cover such that the lift of q is the square of a holomorphic 1-form ω and (X, q) is a quotient of the double cover with respect to hyper-elliptic involution. Fixing a basis for the anti-invariant (with respect to hyper-elliptic involution) part of the homology of the double cover relative to the singularities, the holonomies (periods) given by integrating ω over the basis define local co-ordinates on $\mathcal{Q}(\alpha)$. The natural volume form in these co-ordinates defines the *Masur–Veech measure*. Alternatively, it is known as the *holonomy measure*. We shall denote it by μ_{hol}. The measure μ_{hol} is $\mathrm{Mod}(S)$-invariant. So it descends to a measure on $\mathrm{Mod}(S) \backslash \mathcal{Q}(\alpha)$, the corresponding stratum of the moduli space of quadratic differentials. We continue to denote it by $\mathcal{Q}(\alpha)$. The μ_{hol}-volume of $\mathcal{Q}(\alpha)$ is finite [21], [29].

The affine action of $SL(2, \mathbb{R})$ on the charts of $\mathbb{C} = \mathbb{R}^2$ preserves the glueing by half-translations. This defines an action of $SL(2, \mathbb{R})$ on $\mathcal{Q}(\alpha)$. The orbits $SL(2, \mathbb{R})(q)$ foliate $\mathcal{Q}(\alpha)$. The compact part $SO(2, \mathbb{R})$ acts by rotations of $\mathbb{R}^2$. Hence, it preserves the conformal structure. The action of the diagonal subgroup defines the Teichmüller

geodesic flow. It shrinks the leaves of the vertical foliation for q and stretches the leaves of the horizontal foliation for q by the same factor. It follows from the definition that μ_{hol} is SL$(2, \mathbb{R})$-invariant.

Since SO$(2, \mathbb{R})$ preserves the conformal structure, we get an isometrically embedded $\mathbb{H}^2 = $ SL$(2, \mathbb{R})/$SO$(2, \mathbb{R})$ in $\mathcal{T}(S)$. This is called a *Teichmüller disc*, and we will denote the Teichmüller disc determined by q as $\mathbb{D}(q)$. The Teichmüller metric restricted to $\mathbb{D}(q)$ is isometric to that of the hyperbolic plane of constant curvature -4.

In the flat metric defined by a quadratic differential q, a *saddle connection* is a geodesic segment that connects a pair of (same or distinct) singularities of q. The ϵ-*thin part* $\mathcal{Q}(\alpha)_\epsilon$ of $\mathcal{Q}(\alpha)$ is the subset of q such that $\ell_q^2(\beta) \leq \epsilon$ for some saddle connection β.

The points q' in SL$(2, \mathbb{R})(q)$ where $\ell_{q'}^2(\beta) \leq \epsilon$ project to a horoball in $\mathbb{D}(q)$. The point at infinity of the horoball is given by the direction in which β is vertical. When two saddle connections β_1 and β_2 are parallel, the proportion $[\ell_{q'}(\beta_1) : \ell_{q'}(\beta_2)]$ as a function of q' is constant. Hence, the horoball is determined by the saddle connection with the shortest holonomy in a collection of parallel saddle connections. Typically, the intersection $\mathcal{Q}(\alpha)_\epsilon \cap \mathbb{D}(q)$ is a complicated collection of horoballs in $\mathbb{D}(q)$ with their interiors having intersections. Every point in $\mathbb{D}(q)$ is contained in at most finitely many horoballs. But there need not be a uniform upper bound over $\mathbb{D}(q)$ for this number. This makes it hard to estimate the d_{thick}-distance between the entry and exit points of a geodesic in $\mathcal{Q}(\alpha)_\epsilon \cap \mathbb{D}(q)$, which would have been a natural analog of an excursion in this context. Instead, we consider excursions in individual horoballs.

The excursions in horoballs in $\mathbb{D}(q)$ of a geodesic γ till time T are defined as follows. For a horoball H that γ enters and exits, the complete excursion $E(\gamma, H)$ is defined as the distance along ∂H between the entry and exit points. If $\gamma_T \in H$ then the partial excursion $E(\gamma, H)$ is defined as the distance along ∂H between the entry point of γ and $\pi_H(\gamma_T)$ where $\pi_H : H \to \partial H$ is the closest point projection. Let $H_1, \ldots, H_{N(T)}$ denote all the horoballs that γ intersects till time T. The total excursion $E(\gamma, T)$ is defined as

$$E(\gamma, T) = \sum_{k \leq N(T)} E(\gamma, H_k).$$

When horoballs have disjoint interiors, these definitions coincide with the definitions in Section 1.1. With intersections, the excursions $E(\gamma, H_k)$ can be simultaneous and there can be several partial excursions when γ_T is in the thin part.

4.2. SL$(2, \mathbb{R})$ *orbit closures and invariant measures*

Recently, Eskin and Mirzakhani [12, Theorem 1.4] showed that ergodic SL$(2, \mathbb{R})$-invariant probability measures are of Lebesgue class and are supported on invariant complex submanifolds in $\mathcal{Q}(\alpha)$. These manifolds are affine in the sense that in the holonomy coordinates on $\mathcal{Q}(\alpha)$ they are cut out by linear equations. Going further, Eskin, Mirzakhani and Mohammadi [13, Theorem 2.1] showed that all SL$(2, \mathbb{R})$ orbit closures are affine invariant submanifolds. See [12, Section 1] for more details. More recently, Filip [14] showed that these submanifolds are in fact algebraic subvarieties.

4.3. Thin parts and regularity for invariant measures

Let μ be an ergodic $SL(2, \mathbb{R})$-invariant probability supported on an affine invariant sub-manifold $\mathcal{N} \subset \mathcal{Q}(\alpha)$. For $\epsilon > 0$, we define the ϵ-*thin part* of $\mathcal{N}$ as follows:

$$\mathcal{N}_\epsilon = \{q \in \mathcal{N} : \ell_q^2(\beta) \leq \epsilon \text{ for some saddle connection } \beta\}.$$

Saddle connections β_1, β_2 are $\mathcal{N}$-*parallel* if they are parallel for an open subset of quadratic differentials in $\mathcal{N}$ [31, Definition 4.6].

Regularity. For $\epsilon, \kappa > 0$ small enough, let $\mathcal{N}_{\epsilon,\kappa}$ be the subset of $q \in \mathcal{N}$ such that there is a pair of saddle connections β_1, β_2 not $\mathcal{N}$-parallel such that $\ell_q^2(\beta_1) \leq \epsilon$ and $\ell_q^2(\beta_2) \leq \kappa$. The measure μ is said to be *regular* if there exists a constant $m_1 > 0$ such that for $\epsilon, \kappa > 0$ small enough we have

$$\mu(\mathcal{N}_{\epsilon,\kappa}) \leq m_1 \epsilon \kappa. \tag{4.4}$$

Masur and Smillie [23, Section 10, Claim (7)] showed that the holonomy measure μ_{hol} is regular. Avila, Matheus and Yoccoz [4, Theorem 1.2] proved a weaker version of regularity for any $SL(2, \mathbb{R})$-invariant measure.

4.5. $SL(2, \mathbb{R})$-invariant loci, Siegel–Veech transform and volume asymptotic

For $q \in \mathcal{N}$, let $V(q) \subset \mathbb{R}^2 \setminus \{(0, 0)\}$ be an assignment of a non-empty subset of holonomies of saddle connections on q. We require that the assignment varies linearly under the $SL(2, \mathbb{R})$ action, i.e. $V(gq) = gV(q)$ for all $g \in SL(2, \mathbb{R})$. As observed in [10], such an assignment satisfies conditions (B) and C_μ mentioned in that paper for any $SL(2, \mathbb{R})$-invariant measure μ. Such a V will be called an $SL(2, \mathbb{R})$-*invariant locus*.

Let f be a smooth function on $\mathbb{R}^2$ with compact support. The *Siegel–Veech transform* associated to an $SL(2, \mathbb{R})$-invariant locus V is defined as

$$\widehat{f}(q) = \sum_{v \in V(q)} f(v).$$

Veech showed that $f \in L^1(\mathcal{N}, \mu)$ and proved the *Siegel–Veech formula*

$$\int_{\mathcal{N}} \widehat{f} \, d\mu = c(V, \mu) \int_{\mathbb{R}^2} f \, dx \, dy \tag{4.6}$$

where the constant $c(V, \mu)$ does not depend on f. The constant $c(V, \mu)$ is called the *Siegel–Veech constant* associated to V and μ. We assume that the assignment V is such that $c(V, \mu) > 0$.

For $R \geq 1$, the ϵ/R-*thin part of $\mathcal{N}$ corresponding to V* is the set of q with some saddle connection with holonomy in V having q-length squared less than ϵ/R. We denote this set by $\mathcal{N}(V)_{\epsilon/R}$.

Let $f_{\epsilon/R}$ be the characteristic function of the ball $B((0, 0), \sqrt{\epsilon/R})$. While $f_{\epsilon/R}$ is not smooth, the Siegel–Veech formula extends to such characteristic functions. The regularity

condition (4.4) and the Siegel–Veech formula (4.6) applied to $f_{\epsilon/R}$ can be used to prove the volume asymptotic

$$\lim_{R \to \infty} \frac{\mu(\mathcal{N}(V)_{\epsilon/R})}{\pi \epsilon / R} = c(V, \mu). \tag{4.7}$$

See [11, Section 7] for the main ideas.

4.8. Exponential mixing of Teichmüller flow:

It is known that the Teichmüller flow is exponentially mixing. For the Masur–Veech measure, the decay of correlations (3.1) for $\mathrm{SO}(2, \mathbb{R})$-invariant L^2-functions is due to Avila–Gouëzel–Yoccoz [3] for holomorphic 1-forms and due to Avila–Resende [5] for quadratic differentials. For general $\mathrm{SL}(2, \mathbb{R})$-invariant measures this is due to Avila–Gouëzel [2]. Since the functions we consider are pullbacks from Teichmüller discs, they are $\mathrm{SO}(2, \mathbb{R})$-invariant. Hence, the decay of correlations applies to them.

5. Proofs of Theorem 1.9

5.1. The simplest case

We first prove Theorem 1.9 in the case when $V(q)$ is the set of holonomies of *all* saddle connections on q. This allows us to convey the key ideas while getting into fewer subtleties. We denote the corresponding Siegel–Veech constant simply by $c(\mu)$.

Let $q \in \mathcal{N}_\epsilon$. Consider short saddle connections in q. If some saddle connections are parallel we choose the one with the smallest holonomy. Suppose that the lengths of these short saddle connections are $\ell_q^2(\beta_1) = \epsilon/R_1, \ldots, \ell_q^2(\beta_k) = \epsilon/R_k$ with $R_1 \geq \cdots \geq R_k \geq 1$. We define

$$\Psi(q) = \frac{2}{\pi} R_1 \quad \text{and} \quad \overline{\Psi}(q) = \frac{2}{\pi}(R_1 + \cdots + R_k).$$

Obviously $\overline{\Psi}(q) \geq \Psi(q)$ for all q. At first glance, the function Ψ above is similar to the function Ψ defined in the context of non-uniform lattices. However, here there can be simultaneous excursions. Hence, the sum over all excursions between successive entry and exit times $T_1 < T_2$ for $\mathcal{N}_\epsilon$ can satisfy

$$\sum_{H:\gamma[T_1,T_2] \cap H \neq \emptyset} E(\gamma, H) \gg \int_{T_1}^{T_2} \Psi(v(\gamma_t)) \, dt.$$

This discrepancy is rectified by using the larger function $\overline{\Psi}$. The key point is to estimate the difference in the L^1- and L^2-norms of the truncations of Ψ and $\overline{\Psi}$ in terms of the depth in $\mathcal{N}_\epsilon$ of the truncations. This will enable us to show that the above discrepancy does not happen too often.

For times T when γ_T is in the thick part, it would be interesting to relate the integral over time of $\overline{\Psi}$ to $d_{\text{thick}}(\gamma_0, \gamma_T)$. In analogy with the observations preceding Theorem 1.5, we expect their ratio to be asymptotic to a constant $M > 0$ that depends only on $\mathcal{N}$.

Let χ_R denote the characteristic function of $\mathcal{N}_{\epsilon/R}$ and define the truncation

$$\Psi_R = (\chi_1 - \chi_R)\Psi.$$

Lemma 5.2.

$$\lim_{R\to\infty} \frac{\|\Psi_R\|_{L^1}}{\log R} = 2\epsilon c(\mu), \qquad \lim_{R\to\infty} \frac{\|\Psi_R\|_{L^2}}{\sqrt{R}} = 2\frac{\sqrt{\epsilon c(\mu)}}{\sqrt{\pi}}.$$

Proof. It follows from (4.7) that for any $A > 1$ there is R_0 such that for all $R > R_0$,

$$\frac{1}{A}\frac{\pi\epsilon c(\mu)}{R} < \mu(\mathcal{N}_{\epsilon/R}) < A\frac{\pi\epsilon c(\mu)}{R}.$$

Fix $r > 0$ and for any positive integer k consider $\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}$. If k is large enough that $2^{(k-1)r} > R_0$ then the measure of the above set satisfies

$$\frac{\pi\epsilon c(\mu)}{2^{kr}}\frac{2^r - A^2}{A} < \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) < \frac{\pi\epsilon c(\mu)}{2^{kr}}\frac{2^r A^2 - 1}{A}.$$

Given r, we choose A close to 1 such that

$$\frac{2^r - 1}{2^r} < \frac{2^r - A^2}{A} < \frac{2^r A^2 - 1}{A} < 2^r(2^r - 1). \tag{5.3}$$

Let n be the largest integer such that $2^{nr} \leq R$. The L^1-norm of Ψ_R can be estimated by

$$\frac{2}{\pi}\sum_{k=1}^{n} 2^{(k-1)r}\mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) < \|\Psi_R\|_{L^1} < \frac{2}{\pi}\sum_{k=1}^{n+1} 2^{kr}\mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}).$$

Let n_0 be the smallest integer such that $2^{n_0 r} \geq R_0$. We assume that $R \gg R_0$. The summation in the lower bound on the left satisfies

$$\sum_{k=0}^{n} 2^{(k-1)r}\mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) > \sum_{k=1}^{n_0-1} 2^{(k-1)r}\mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}})$$

$$+ \sum_{k=n_0}^{n} 2^{(k-1)r}\frac{\pi\epsilon c(\mu)}{2^{kr}}\frac{2^r - 1}{2^r}.$$

The right hand side of the above inequality simplifies to

$$\sum_{k=1}^{n_0-1} 2^{(k-1)r}\mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) + \frac{(n - n_0)\pi\epsilon c(\mu)}{2^r}\frac{2^r - 1}{2^r}.$$

As R becomes large, the second term dominates and since $(n - n_0)/\log R \to 1/(r\log 2)$, the above expression simplifies to

$$\frac{2}{2^{2r}\log 2}\frac{2^r - 1}{r}\epsilon c(\mu) < \lim_{R\to\infty}\frac{\|\Psi_R\|_{L^1}}{\log R},$$

which as $r \to 0$ implies

$$2\epsilon c(\mu) \leq \lim_{R \to \infty} \frac{\|\Psi_R\|_{L^1}}{\log R}.$$

Similarly the summation for the upper bound on $\|\Psi_R\|_{L^1}$ gives

$$\frac{2}{2^r \log 2} \frac{2^r - 1}{r} \epsilon c(\mu) > \lim_{R \to \infty} \frac{\|\Psi_R\|_{L^1}}{\log R},$$

which as $r \to 0$ implies

$$2\epsilon c(\mu) \geq \lim_{R \to \infty} \frac{\|\Psi_R\|_{L^1}}{\log R},$$

proving the lemma for the L^1-norm.

In a similar way, the square of the L^2-norm of Ψ_R can be estimated by

$$\frac{4}{\pi^2} \sum_{k=1}^{n} 2^{2(k-1)r} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) < \|\Psi_R\|_{L^2}^2 < \frac{4}{\pi^2} \sum_{k=1}^{n+1} 2^{2kr} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}).$$

The summation in the lower bound satisfies

$$\sum_{k=1}^{n} 2^{2(k-1)r} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) > \sum_{k=1}^{n_0-1} 2^{2(k-1)r} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}})$$
$$+ \sum_{k=n_0}^{n} 2^{2(k-1)r} \frac{\pi \epsilon c(\mu)}{2^{kr}} \frac{2^r - 1}{2^r}.$$

The right hand side above is equal to

$$\sum_{k=1}^{n_0-1} 2^{2(k-1)r} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) + \frac{2^{(n+1)r} - 2^{n_0 r}}{2^r - 1} \frac{\pi \epsilon c(\mu)}{2^{2r}} \frac{2^r - 1}{2^r},$$

which is greater than

$$\sum_{k=1}^{n_0-1} 2^{2(k-1)r} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) + \frac{\pi \epsilon c(\mu) R}{2^{3r}} - \frac{2^{n_0 r} \pi \epsilon c(\mu)}{2^{3r}}.$$

As R becomes large, the term containing R dominates, and letting $r \to 0$ we get the lower bound

$$\frac{2\sqrt{\epsilon c(\mu)}}{\sqrt{\pi}} \leq \lim_{R \to \infty} \frac{\|\Psi_R\|_{L^2}}{\sqrt{R}}.$$

Similarly the summation in the upper bound satisfies

$$\sum_{k=1}^{n+1} 2^{kr} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) < \sum_{k=1}^{n_0-1} 2^{2kr} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}})$$

$$+ \sum_{k=n_0}^{n+1} 2^{2kr} \frac{\pi \epsilon c(\mu)}{2^{(k-1)r}} 2^r (2^r - 1)$$

$$< \sum_{k=1}^{n_0-1} 2^{2kr} \mu(\mathcal{N}_{\epsilon/2^{(k-1)r}} \setminus \mathcal{N}_{\epsilon/2^{kr}}) + 2^{4r} R \pi \epsilon c(\mu).$$

As R becomes large, the term containing R dominates, and letting $r \to 0$ we get the upper bound

$$\frac{2\sqrt{\epsilon c(\mu)}}{\sqrt{\pi}} \geq \lim_{R \to \infty} \frac{\|\Psi_R\|_{L^2}}{\sqrt{R}},$$

finishing the proof for the L^2-norm.   □

The next lemma will need the regularity in (4.4) of μ and the following theorem of Eskin and Masur: For any stratum $\mathcal{Q}(\alpha)$ and any $0 < \delta < 1$ there exist constants $m_2(\alpha, \delta) > 0$ such that the number $s(q, \kappa)$ of saddle connections shorter than $\kappa < 1$ is bounded above by

$$s(q, \kappa) \leq m_2 \left(\frac{\kappa}{\ell_q(\beta)} \right)^{1+\delta} \tag{5.4}$$

where β is the shortest saddle connection for q. It should be noted that while Eskin and Masur state the theorem for strata of holomorphic 1-forms, it is also true for strata of quadratic differentials by passing to the canonical double cover.

Fix the constant δ in the Eskin–Masur theorem and choose $a > 1$ such that $a < 2/(1 + \delta)$. Let $\mathcal{N}' \subset \mathcal{N}_{\epsilon/R} \setminus \mathcal{N}_{\epsilon/R^a}$ be the subset of quadratic differentials such that apart from the shortest saddle connection, all other short saddle connections satisfy $\ell_q^2(\beta_i) \geq \epsilon/R$. We define $\Psi' : \mathcal{N}' \to \mathbb{R}_{\geq 0}$ by

$$\Psi'(q) = \frac{2}{\pi}(R_2 + \cdots + R_j).$$

We define a slightly more complicated truncation

$$\overline{\Psi}_R = \overline{\Psi}(\chi_1 - \chi_R) + \Psi'.$$

In this particular case of Theorem 1.9, the extra term Ψ' allows us to keep track of excursions that are concurrent with the largest excursion if it exceeds $T(\log T)^c$.

Lemma 5.5. *There exists a constant $B > 0$ such that for R large enough,*

$$\|\overline{\Psi}_R\|_{L^1} \leq \|\Psi_R\|_{L^1} + B.$$

The L^2-norms satisfy

$$\|\overline{\Psi}_R\|_{L^2} \leq \|\Psi_R\|_{L^2} + o(\sqrt{R}).$$

Proof. Consider the subset $\mathcal{N}_{k,j}$ of $\mathcal{N}_{\epsilon/2^{k-1}} \setminus \mathcal{N}_{\epsilon/2^k}$ of those q such that the length of the second shortest saddle connection β_2 (not $\mathcal{N}$-parallel to β_1) satisfies $\epsilon/2^{j-1} \geq \ell_q^2(\beta_2) > \epsilon/2^j$ where $j \leq k$. By regularity (4.4),

$$\mu(\mathcal{N}_{k,j}) \leq \frac{m_1 \epsilon^2}{2^{k-1} 2^{j-1}}.$$

For a quadratic differential q in $\mathcal{N}_{k,j}$, using the bound (5.4) on the number of short saddle connections we get the pointwise bound

$$\overline{\Psi}_R(q) - \Psi_R(q) < \sum_{i=1}^{j} 2^i m_2 \left(\frac{2^k}{2^{i-1}} \right)^{(1+\delta)/2}$$

$$< 2m_2 2^{k(1+\delta)/2} \sum_{i=1}^{j} 2^{(i-1)(1-\delta)/2} < m_3 2^{k(1+\delta)/2} 2^{j(1-\delta)/2}$$

for some constant $m_3 > 0$. Thus

$$\int_{\mathcal{N}_{k,j}} (\overline{\Psi}_R - \Psi_R) \, d\mu < m_3 2^{k(1+\delta)/2} 2^{j(1-\delta)/2} \frac{m_1 \epsilon^2}{2^{k-1} 2^{j-1}} = \frac{4m_1 m_3 \epsilon^2}{2^{k(1-\delta)/2} 2^{j(1+\delta)/2}}.$$

Summing over j from 1 to k (when $2^{k-1} > R$ it suffices to sum till the smallest k' such that $2^{k'} > R$, in which case the sum would be even smaller), we get

$$\int_{\mathcal{N}_{\epsilon/2^{k-1}} \setminus \mathcal{N}_{\epsilon/2^k}} (\overline{\Psi}_R - \Psi_R) \, d\mu < \frac{m_4}{2^{k(1-\delta)/2}}$$

for some constant $m_4 > 0$. Let n_a be the smallest integer such that $2^{n_a} \geq R^a$. The bound for the integral established above implies that

$$\|\overline{\Psi}_R\|_{L^1} - \|\Psi_R\|_{L^1} < \sum_{k=1}^{n_a} \frac{m_4}{2^{k(1-\delta)/2}}.$$

The sum on the right hand side is bounded from above independently of n_a, which proves the lemma for L^1-norms.

The same pointwise bound above implies

$$\int_{\mathcal{N}_{k,j}} \overline{\Psi}_R^2 \, d\mu$$

$$< \int_{\mathcal{N}_{k,j}} \Psi_R^2 \, d\mu + 2m_3 2^{k(1+\delta)/2} 2^{j(1-\delta)/2} \int_{\mathcal{N}_{k,j}} \Psi_R \, d\mu + m_3^2 2^{k(1+\delta)} 2^{j(1-\delta)} \mu(\mathcal{N}_{k,j})$$

$$< \int_{\mathcal{N}_{k,j}} \Psi_R^2 \, d\mu + 2m_3 2^{k(1+\delta)/2} 2^{j(1-\delta)/2} \frac{2^k m_1 \epsilon^2}{2^{k-1} 2^{j-1}} + m_3^2 2^{k(1+\delta)} 2^{j(1-\delta)} \frac{m_1 \epsilon^2}{2^{k-1} 2^{j-1}}$$

$$= \int_{\mathcal{N}_{k,j}} \Psi_R^2 \, d\mu + \frac{1}{2^{j(1+\delta)/2}} 8m_1 m_3 \epsilon^2 2^{k(1+\delta)/2} + \frac{1}{2^{j\delta}} 4m_1 m_3^2 \epsilon^2 2^{k\delta}.$$

Summing over j from 1 to k we get

$$\int_{\mathcal{N}_{\epsilon/2^{k-1}} \setminus \mathcal{N}_{\epsilon/2^k}} \overline{\Psi}_R^2 \, d\mu < \int_{\mathcal{N}_{\epsilon/2^{k-1}} \setminus \mathcal{N}_{\epsilon/2^k}} \Psi_R^2 \, d\mu + m_5 2^{k(1+\delta)/2} + m_6 2^{k\delta}$$

for some constants $m_5, m_6 > 0$. Summing over k from 1 to n_a we get

$$\|\overline{\Psi}_R\|_{L^2}^2 < \|\Psi_R\|_{L^2}^2 + \frac{m_7}{2} 2^{n_a(1+\delta)/2} + \frac{m_8}{2} 2^{n_a \delta} < \|\Psi_R\|_{L^2}^2 + m_7 R^{a(1+\delta)/2} + m_8 R^{a\delta}$$

for some constants $m_7, m_8 > 0$. Recall that we chose $a > 1$ to satisfy $a(1 + \delta) < 2$, which implies $a\delta < a(1 + \delta)/2 < 1$. Thus, the corresponding terms on the right hand side are $o(R)$, from which the lemma follows for L^2-norms. □

We will justify the choice of the cutoff R^a used for truncation by the following sharper lemma. The lemma is a special case of a continuous time version of the analog of the Borel–Bernstein theorem [17, Theorem 30], [1].

Lemma 5.6. *For any $1/2 < c < 1$ and μ-almost every $q \in \mathcal{N}$ there is T_0 depending on q such that for all $T > T_0$, all excursions $E(\gamma, H)$ till time T satisfy*

$$E(\gamma, H) < T(\log T)^{2c}.$$

Proof. We recall some notation. For $R \geq 1$, we denote by ϕ_R the characteristic function of $\mathcal{N}_{2\epsilon/R} \setminus \mathcal{N}_{\epsilon/R}$. Choose A close to 1 so that estimate (5.3) is satisfied with $r = 1$, i.e. $A > 1$ is chosen close enough to 1 such that

$$\frac{1}{2} < \frac{2 - A^2}{A} < \frac{2A^2 - 1}{A} < 2.$$

The given choice of A implies that for all R large enough

$$\|\phi_R\|_{L^1} < \frac{2\pi \epsilon c(\mu)}{R} \quad \text{and hence} \quad \|\phi_R\|_{L^2}^2 < \frac{2\pi \epsilon c(\mu)}{R}.$$

Define $n : \mathbb{R} \to \mathbb{N}$ so that for T satisfying $2^k \leq T < 2^{k+1}$, $n(T)$ is the largest integer such that $2^n \leq k^{2c} 2^k$, i.e. $n = \lfloor k + 2c \log k \rfloor$. Choose a constant $c_1 > 1/2$ for Theorem 2.4 to satisfy $c_1 < c$. Fix $m > 1$. The upper bound in Theorem 2.4 implies that for μ-almost every q,

$$\int_0^T \phi_{2^n}(v(\gamma_t)) \, dt \leq mT \|\phi_{2^n}\|_{L^1} + T^{1/2} (\log T)^{c_1} (\|\phi_{2^n}\|_{L^2}^2 - \|\phi_{2^n}\|_{L^1}^2)^{1/2}$$

for all T large enough depending on q and m, and with $n(T)$ defined as above. By passing to a larger T if necessary, we can ensure that the upper bounds above are also satisfied by $\|\phi_{2^n}\|_{L^1}$ and $\|\phi_{2^n}\|_{L^2}^2$. This gives the upper bound

$$\int_0^T \phi_{2^n}(v(\gamma_t)) \, dt \leq mT \frac{2\pi \epsilon c(\mu)}{2^n} + T^{1/2} (\log T)^{c_1} \left(\frac{2\pi \epsilon c(\mu)}{2^n} \right)^{1/2}$$

$$< \frac{B_1}{(\log T)^{2c}} + \frac{B_2}{(\log T)^{c - c_1}}$$

for some constants B_1, $B_2 > 0$. Notice that if T is large enough, the right hand side is less than $\log 2$. But if an excursion till time satisfies $E(\gamma, H) \geq T(\log T)^{2c}$ then γ must spend time at least $\log 2$ in $\mathcal{N}_{\epsilon/2^{n-1}} \setminus \mathcal{N}_{\epsilon/2^n}$, which proves the lemma. $\square$

Proof of Theorem 1.9 when V is all saddle connection holonomies. Fix c satisfying $1/2 < c < 1$. Observe that Proposition 3.2 holds in this setting. That is, along μ-typical geodesics there is at most a single excursion larger than $T(\log T)^c$ till time T for T large enough. In fact, as noted in Remark 3.4, the proof of Proposition 3.3 does not need the precise asymptotic for $\mu(\mathcal{N}_{\epsilon/R})$ as $R \to \infty$, just that $\mu(\mathcal{N}_{\epsilon/R})$ decays as $1/R$ up to some multiplicative constant.

If $2^k \leq T < 2^{k+1}$, then let $n = n(T) = \lfloor k + c \log_2 k \rfloor$. Replicating the exact argument in the proof of Theorem 1.2, we use Lemma 5.1 to conclude that for μ-almost every $q \in \mathcal{N}$,

$$\lim_{T \to \infty} \frac{1}{T \log T} \int_0^T \Psi_{2^n}(v(\gamma_t)) \, dt = 2\epsilon c(\mu)$$

where γ is the Teichmüller geodesic ray with $v(\gamma_0) = q$. Lemma 5.5 implies that the above limit holds when Ψ_{2^n} is replaced by $\overline{\Psi}_{2^n}$. By Lemma 5.6, for T large enough the largest excursion till time T is smaller than $T(\log T)^{2c} < T^a$. Hence, up to an additive error whose dependence on T will be described below,

$$E(\gamma, T) - \max_{k \leq N} E(\gamma, H_k) \asymp \int_0^T \overline{\Psi}_{2^n}(v(\gamma_t)) \, dt.$$

The additive error is bounded above by the sum of the additive errors arising from individual excursions. First, we analyze the partial excursions (if they exist).

The partial excursions correspond to configurations of saddle connections that are short in $q_T = v(\gamma_T)$. If the q_T-length squared of the small saddle connection in such a configuration is ϵ/R then by (3.7), the additive error for the associated partial excursion is at most R. Let β_1 be the shortest saddle connection in q_T. Let β_2 be the second shortest non-parallel saddle connection. Lemma 5.6 implies that if T is large enough then $\ell_{q_T}^2(\beta_1) \geq \epsilon/(T(\log T)^{2c})$. Proposition 3.2 implies that $\ell_{q_T}^2(\beta_2) \geq \epsilon/(T(\log T)^c)$. Let n be the smallest positive integer such that $2^n \geq T(\log T)^{2c}$. Let m be the smallest positive integer such that $2^m \geq T(\log T)^c$. The Eskin–Masur bound (5.4) implies that the total additive error is bounded above by

$$\sum_{k=0}^{m-1} 2^{k+1} m_2 \left(\frac{2^{n/2}}{2^{k/2}} \right)^{1+\delta} = \sum_{k=0}^{m-1} 2m_2 2^{k(1-\delta)/2} 2^{n(1+\delta)/2} \leq B_3 2^{m(1-\delta)/2} 2^{n(1+\delta)/2}$$

$$\leq 4B_3 T^{(1-\delta)/2} (\log T)^{c(1-\delta)/2} T^{(1+\delta)/2} (\log T)^{2c(1+\delta)/2}$$

$$= 4B_3 T (\log T)^{c(3+\delta)/2}$$

for some constant $B_3 > 0$. For this bound to be of a lower order, c must satisfy $c < 2/(3 + \delta)$. Note that $2/(3 + \delta) > 1/2$. So the condition can be satisfied.

By (3.6), the additive error from an individual complete excursion is uniform. This means that the additive error from all complete excursions is at most linear in the number N of horoballs that γ intersects till time T. We claim that N grows linearly in T. Let

$\mathcal{S}_T$ be the saddle connections for q whose length squared gets shorter than ϵ in time less than T along the Teichmüller geodesic ray determined by q. Then the necessary conditions on the (x, y)-coordinates of the holonomy vectors of the saddle connections in $\mathcal{S}_T$ are $|xy| \leq \epsilon^2/2$, $y \leq e^T$ and $x < \epsilon$. By [1, Theorem 1.6] the number of such vectors for μ-almost every q is linear in T, thus proving the claim. $\qquad\qquad\square$

5.7. The general case

Let V be an $\mathrm{SL}(2, \mathbb{R})$-invariant locus. Suppose that for q there are j short saddle connections no two of which are $\mathcal{N}$-parallel with holonomy in $V(q)$. Index the saddle connections $\beta_1, \ldots, \beta_j$ in the order of increasing q-lengths and let $\ell_q^2(\beta_1) = \epsilon/R_1, \ldots, \ell_q^2(\beta_j) = \epsilon/R_j$. We define the functions

$$\Psi^V(q) = \frac{2}{\pi} R_1 \quad \text{and} \quad \overline{\Psi}^V(q) = \frac{2}{\pi}(R_1 + \cdots + R_j).$$

We can define truncations of these functions in an analogous way using the characteristic function of $\mathcal{N}(V)_{\epsilon/R}$. However, the shortest saddle connection β for q may not have holonomy in $V(q)$ and be shorter than β_1 as above. Hence, some care is required in defining the truncations. The crucial point is that in light of Lemma 5.6, we can impose a lower bound on the length of β in defining the truncations. As in the previous case, we use a sloppier lower bound than what Lemma 5.6 implies. This simplifies the expressions in the estimates.

Let $a > 1$ be such that $a < 2/(1 + \delta)$. In particular, $a\delta < 1$. Let $\mathcal{N}(R, a) \subset \mathcal{N}(V)_\epsilon \setminus \mathcal{N}(V)_{\epsilon/R}$ be the subset of those q such that the shortest saddle connection β satisfies $\ell_q^2(\beta) \geq \epsilon/R^a$. Let $\chi_{R,a}$ denote the characteristic function of $\mathcal{N}(R, a)$.

Let $\mathcal{N}'(V) \subset \mathcal{N}(V)_{\epsilon/R} \setminus \mathcal{N}(V)_{\epsilon/R^a}$ be the subset of quadratic differentials such that $\ell_q^2(\beta) \geq \epsilon/R^a$ and apart from β_1 all other short saddle connections with holonomy in $V(q)$ satisfy $\ell_q^2(\beta_i) \geq \epsilon/R$. Let $(\Psi^V)' : \mathcal{N}'(V) \to \mathbb{R}$ be defined as

$$(\Psi^V)'(q) = \frac{2}{\pi}(R_2 + \cdots + R_j).$$

We define $\Psi_R^V = \Psi^V \chi_{R,a}$ and $\overline{\Psi}_R^V = \overline{\Psi}^V \chi_{R,a} + (\Psi^V)'$. Again, the extra term $(\Psi^V)'$, analogous to Ψ' earlier, allows us to track excursions for saddle connections in V that happen during the largest excursion with holonomy in V, should it exceed $T(\log T)^c$.

For $2^k \leq R$ we have the estimate

$$\mu\big(\mathcal{N}(V)_{\epsilon/2^{k-1}} \setminus \mathcal{N}(V)_{\epsilon/2^k}\big) - \mu\big(\mathcal{N}(R, a) \cap \mathcal{N}(V)_{\epsilon/2^{k-1}} \setminus \mathcal{N}(V)_{\epsilon/2^k}\big) \leq \frac{m_1 \epsilon^2}{2^{k-1} R^a}.$$

This means here each term in the summations for lower and upper bound for L^1-norm in Lemma 5.2 changes by at most $2^k m_1 \epsilon^2/(2^{k-1} R^a) = 2 m_1 \epsilon^2/R^a$. Hence the summations change by at most $2nm_1\epsilon^2/R^a < m_6(\log R)/R$ for some constant $m_6 > 0$. This implies

$$\lim_{R \to \infty} \frac{\|\Psi_R^V\|_{L^1}}{\log R} = 2\epsilon c(V, \mu).$$

Similarly each term in the summations for lower and upper bound for L^2-norms changes by at most $2^{2k} m_1 \epsilon^2 / (2^{k-1} R^a) = m_1 \epsilon^2 2^{k+1} / R^a$, and hence the summations change by at most $4 m_1 \epsilon^2 2^n / R^a < m_7 / R^{a-1}$ for some constant $m_7 > 0$. This implies

$$\lim_{R \to \infty} \frac{\|\Psi_R^V\|_{L^2}}{\sqrt{R}} = 2 \frac{\sqrt{\epsilon c(V, \mu)}}{\sqrt{\pi}}.$$

Lemma 5.8. *There exists a constant B_V such that for R large enough,*

$$\|\overline{\Psi}_R^V\|_{L^1} \leq \|\Psi_R^V\|_{L^1} + B_V.$$

The L^2-norms satisfy

$$\|\overline{\Psi}_R^V\|_{L^2} \leq \|\Psi_R^V\|_{L^2} + o(\sqrt{R}).$$

Proof. Consider $\mathcal{N}(R, a) \cap \mathcal{N}(V)_{\epsilon/2^{k-1}} \setminus \mathcal{N}(V)_{\epsilon/2^k}$ and let $\mathcal{N}_{k,j}^V$ be its subset consisting of those q for which (among the collection of non-parallel saddle connections with holonomies in $V(q)$) the second shortest saddle connection β_2 satisfies $\epsilon/2^{j-1} \geq \ell_q^2(\beta_2) > \epsilon/2^j$ where $j \leq k$.

We further partition $\mathcal{N}_{k,j}^V$ into two sets, $\mathcal{N}_{k,j}^V(1) \cup \mathcal{N}_{k,j}^V(2)$, depending on whether the shortest saddle connection for q has holonomy in $V(q)$ or not, i.e. $\mathcal{N}_{k,j}^V(1)$ is the subset of q for which β_1 is the shortest saddle connection, and $\mathcal{N}_{k,j}^V(2)$ is when it is not. On $\mathcal{N}_{k,j}^V(1)$ the integral

$$\int_{\mathcal{N}_{k,j}^V(1)} (\overline{\Psi}_R^V - \Psi_R^V) \, d\mu$$

is bounded from above just as in Lemma 5.5.

Let n_a be the smallest integer such that $2^{n_a} \geq R^a$. For $q \in \mathcal{N}_{k,j}^V(2)$ suppose that the shortest saddle connection β satisfies $\epsilon/2^{p-1} > \ell_q^2(\beta) \geq \epsilon/2^p$ where $2^k \leq 2^p \leq 2^{n_a}$. The measure of the subset of such q is bounded above by

$$\frac{m_1 \epsilon^2}{2^{p-1} 2^{k-1}}.$$

The number of short saddle connections whose q-length squared is at least $\epsilon/2^{i-1}$ is bounded above by

$$m_2 \left(\frac{2^p}{2^{i-1}} \right)^{(1+\delta)/2}.$$

This gives the pointwise bound

$$\overline{\Psi}_R^V(q) - \Psi_R^V(q) < \sum_{i=1}^{j} 2^i m_2 \left(\frac{2^p}{2^{i-1}} \right)^{(1+\delta)/2} = 2 m_2 2^{p(1+\delta)/2} \sum_{i=1}^{j} 2^{(i-1)(1-\delta)/2}$$

$$< m_9 2^{p(1+\delta)/2} 2^{j(1-\delta)/2}$$

for some constant $m_9 > 0$, implying that

$$\int_{\mathcal{N}_{k,j}^V(2)} (\overline{\Psi}_R^V - \Psi_R^V)\, d\mu < \sum_{p=k}^{n_a} m_9 2^{p(1+\delta)/2} 2^{j(1-\delta)/2} \frac{m_1 \epsilon^2}{2^{p-1} 2^{k-1}} < \frac{m_{10}}{2^k}$$

for some constant $m_{10} > 0$. Thus adding up the upper bounds for the integrals on $\mathcal{N}_{k,j}^V(1)$ and $\mathcal{N}_{k,j}^V(2)$ we get

$$\int_{\mathcal{N}_{k,j}^V} (\overline{\Psi}_R^V - \Psi_R^V)\, d\mu < \frac{4m_1 m_3 \epsilon^2}{2^{k(1-\delta)/2} 2^{j(1+\delta)/2}} + \frac{m_{10}}{2^k}$$

Summing over j from 1 to k we get

$$\int_{\mathcal{N}(R,a) \cap \mathcal{N}(V)_{\epsilon/2^{k-1}} \setminus \mathcal{N}(V)_{\epsilon/2^k}} (\overline{\Psi}_R^V - \Psi_R^V)\, d\mu < \frac{m_{11}}{2^{k(1-\delta)/2}} + \frac{m_{10} k}{2^k}$$

for some constant $m_{11} > 0$. Summing over k from 1 to n_a observe that the sum of the right hand side is bounded independently of n, which proves the lemma for L^1-norms.

The pointwise bound also implies

$$\int_{\mathcal{N}_{k,j}^V(2)} [(\overline{\Psi}_R^V)^2 - (\Psi_R^V)^2]\, d\mu < \sum_{p=k}^{n_a} 2m_9 2^{p(1+\delta)/2} 2^{j(1-\delta)/2} \frac{2^k m_1 \epsilon^2}{2^{p-1} 2^{k-1}}$$

$$+ \sum_{p=k}^{n_a} m_9^2 2^{p(1+\delta)} 2^{j(1-\delta)} \frac{m_1 \epsilon^2}{2^{p-1} 2^{k-1}}$$

$$< \frac{m_{12}}{2^k} + \frac{m_{13} 2^{n_a \delta}}{2^{k\delta}}$$

for some constants $m_{12}, m_{13} > 0$. The corresponding upper bound for $\mathcal{N}_{k,j}^V(1)$ is identical to that in Lemma 5.5 and is of the form

$$\int_{\mathcal{N}_{k,j}^V(1)} [(\overline{\Psi}_R^V)^2 - (\Psi_R^V)^2]\, d\mu < \frac{m_{14} 2^{k(1+\delta)/2}}{2^{j(1+\delta)/2}} + \frac{m_{15} 2^{k\delta}}{2^{j\delta}}$$

for some constants $m_{14}, m_{15} > 0$. Adding up the bounds for $\mathcal{N}_{k,j}^V(1)$ and $\mathcal{N}_{k,j}^V(2)$ and summing over j from 1 to k we get

$$\int_{\mathcal{N}(R,a) \cap \mathcal{N}(V)_{\epsilon/2^{k-1}} \setminus \mathcal{N}(V)_{\epsilon/2^k}} [(\overline{\Psi}_R^V)^2 - (\Psi_R^V)^2]\, d\mu$$

$$< \frac{m_{12} k}{2^k} + \frac{m_{13} k 2^{n_a \delta}}{2^{k\delta}} + m_{14} 2^{k(1+\delta)/2} + m_{15} 2^{k\delta},$$

and when $2^{k-1} > R$,

$$\int_{\mathcal{N}'(V) \cap \mathcal{N}(V)_{\epsilon/2^{k-1}} \setminus \mathcal{N}(V)_{\epsilon/2^k}} [(\overline{\Psi}_R^V)^2 - (\Psi_R^V)^2] \, d\mu$$

$$< \frac{m_{12}k}{2^k} + \frac{m_{13}k2^{n_a\delta}}{2^{k\delta}} + m_{14}2^{k(1+\delta)/2} + m_{15}2^{k\delta}.$$

Summing over k from 1 to n_a we get

$$\|\overline{\Psi}_R^V\|_{L^2}^2 - \|\Psi_R^V\|_{L^2}^2 < \frac{m_{16}}{2}2^{n_a\delta} + \frac{m_{17}}{2}2^{n_a(1+\delta)/2} + m_{18} < m_{16}R^{a\delta} + m_{17}R^{a(1+\delta)/2} + m_{18}$$

for some constants $m_{15}, m_{16}, m_{17}, m_{18} > 0$. The condition on a implies that the right hand side is $o(R)$. Thus the lemma follows for L^2-norms. □

Proof of Theorem 1.9. Fix c satisfying $1/2 < c < 1$. As in the earlier case when V is all saddle holonomies, Proposition 3.2 holds for $\mathcal{N}(V)$ for the same reason. It asserts that for μ-almost every q, the Teichmüller geodesic ray corresponding to q has at most a single excursion (in horoballs given by saddle connections with holonomies in V) till T that is larger than $T(\log T)^c$ for all T large enough depending on q. Moreover, by Lemma 5.6 the largest excursion till time T cannot exceed $T(\log T)^{2c}$.

The latter fact implies that up to additive error our truncation $\overline{\Psi}_{2^n}^V$ satisfies

$$\int_0^T \overline{\Psi}_{2^n}^V(v(\gamma_t)) \, dt \asymp E(\gamma, T) - \max_{k \leq N_V} E(\gamma, H_k)$$

where the additive error is of order lower than $T \log T$ for the same reason as earlier. Theorem 2.4 and Lemma 5.8 conclude the proof of Theorem 1.9 in the general case, the precise argument a replica of earlier proofs. □

References

[1] Athreya, J., Parrish, A., Tseng, J.: Ergodic theory and diophantine approximation for linear forms and translation surfaces. Nonlinearity **29**, 2173–2190 (2016) Zbl 1365.37036 MR 3538408

[2] Avila, A., Gouëzel, S.: Small eigenvalues of the Laplacian for algebraic measures in moduli space, and mixing properties of the Teichmüller flow. Ann. of Math. (2) **178**, 385–442 (2013) Zbl 1287.58016 MR 3071503

[3] Avila, A., Gouëzel, S., Yoccoz, J.-C.: Exponential mixing for the Teichmüller flow. Publ. Math. Inst. Hautes Études Sci. **104**, 143–211 (2006) Zbl 1263.37051 MR 2264836

[4] Avila, A., Matheus, C., Yoccoz, J.-P.: SL(2, ℝ)-invariant probability measures on the moduli spaces of translation surfaces are regular. Geom. Funct. Anal. **23**, 1705–1729 (2013) Zbl 1316.37016 MR 3132901

[5] Avila, A., Resende, M.-J.: Exponential mixing for the Teichmüller flow in the space of quadratic differentials. Comment. Math. Helv. **87**, 589–638 (2012)   Zbl 1267.37033 MR 2980521

[6] Avila, A., Viana, M.: Simplicity of Lyapunov spectra: proof of the Zorich–Kontsevich conjecture. Acta Math. **198**, 1–56 (2007) Zbl 1143.37001 MR 2316268

[7] Boissy, C., Lanneau, E.: Dynamics and geometry of Rauzy–Veech induction for quadratic differentials. Ergodic Theory Dynam. Systems **29**, 767–816 (2009)   Zbl 1195.37030 MR 2505317

[8] Brin, M., Stuck, G.: Introduction to Dynamical Systems. Cambridge Univ. Press (2002) Zbl 1314.37002 MR 1963683

[9] Diamond, H., Vaaler, J.: Estimates for partial sums of continued fraction partial quotients. Pacific J. Math. **122**, 73–82 (1986) Zbl 0589.10056 MR 0825224

[10] Eskin, A., Masur, H.: Asymptotic formulas on flat surfaces. Ergodic Theory Dynam. Systems **21**, 443–478 (2001) Zbl 1096.37501 MR 1827113

[11] Eskin, A., Masur, H., Zorich, A.: Moduli spaces of Abelian differentials: The principal boundary, counting problems, and the Siegel–Veech constants. Publ. Math. Inst. Hautes Études Sci. **97**, 61–179 (2003) Zbl 1037.32013 MR 2010740

[12] Eskin, A., Mirzakhani, M.: Invariant and stationary measures for the SL(2, ℝ) action on moduli space. arXiv:1302.3320v4 (2016)

[13] Eskin, A, Mirzakhani, M., Mohammadi, A.: Isolation, equidistribution and orbit closures for the SL(2, ℝ) action on moduli space. Ann. of Math. (2) **182**, 673–721 (2015) Zbl 1357.37040 MR 3418528

[14] Filip, S.: Splitting mixed Hodge structures over affine invariant manifolds. Ann. of Math. (2) **183**, 681–713 (2016) Zbl 1342.14015 MR 3450485

[15] Gadre, V., Maher, J., Tiozzo, G.: Word length statistics for Teichmüller geodesics and singularity of harmonic measure. Comment. Math. Helv. **92**, 1–36 (2017)   Zbl 1362.32009 MR 3615033

[16] Hubbard, J., Masur, H.: Quadratic differentials and foliations. Acta Math. **142**, 221–274 (1979) Zbl 0415.30038 MR 0523212

[17] Khinchin, A.: Continued Fractions. Univ. of Chicago Press, Chicago, IL (1964) Zbl 0117.28601 MR 0161833

[18] Kontsevich, M., Zorich, A.: Connected components of the moduli space of abelian differentials with prescribed singularities. Invent. Math. **153**, 631–678 (2003)   Zbl 1087.32010 MR 2000471

[19] Lanneau, E.: Connected components of the strata of the moduli spaces of quadratic differentials. Ann. Sci. École Norm. Sup. (4) **41**, 1–56 (2008) Zbl 1161.30033 MR 2423309

[20] Masur, H.: Uniquely ergodic quadratic differentials. Comment. Math. Helv. **55**, 255–266 (1980) Zbl 0436.30034 MR 0576605

[21] Masur, H.: Interval exchange transformations and measured foliations. Ann. of Math. (2) **115**, 169–200 (1982) Zbl 0497.28012 MR 0644018

[22] Masur, H.: Logarithm law for geodesics in moduli space. In: Mapping Class Groups and Moduli Spaces of Riemann Surfaces (Göttingen, 1991/ Seattle, WA, 1991), Contemp. Math. 150, Amer. Math. Soc., Providence, RI, 229–245 (1993) MR 1234267

[23] Masur, H., Smillie, J.: Hausdorff dimension of sets of nonergodic foliations. Ann. of Math. **134**, 455–543 (1991) Zbl 0774.58024 MR 1135877

[24] Masur, H., Zorich, A.: Multiple saddle connections on flat surfaces and the principal boundary of the moduli space of quadratic differentials. Geom. Funct. Anal. **18**, 919–987 (2008) Zbl 1169.30017 MR 2439000

[25] Matheus, C.: Some quantitative versions of Ratner's mixing theorems. Bull. Braz. Math. Soc. (N.S.) **44**, 469–488 (2013) Zbl 1305.37017 MR 3124746

[26] Moore, C.: Exponential decay of correlation coefficients for geodesic flows. In: Group Representations, Ergodic Theory, Operator Algebras and Mathematical Physics (Berkeley, CA, 1984), Math. Sci. Res. Inst. Publ. 6, Springer, New York, 163–181 (1987) Zbl 0625.58023 MR 0880376

[27] Ratner, M.: The rate of mixing for geodesic and horocycle flows. Ergodic Theory Dynam. Systems **7**, 267–288 (1987) Zbl 0623.22008 MR 0896798

[28] Sullivan, D.: Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics. Acta Math. **149**, 215–237 (1982) Zbl 0517.58028 MR 0688349

[29] Veech, W.: The Teichmüller geodesic flow. Ann. of Math. (2) **124**, 441–530 (1986) Zbl 0658.32016 MR 0866707

[30] Vorobets, Y.: Periodic geodesics on generic translation surfaces. In: Algebraic and Topological Dynamics, Contemp. Math. 385, Amer. Math. Soc., Providence, RI, 205–258 (2005) Zbl 1130.37015 MR 2180238

[31] Wright, A.: Cylinder deformations in orbit closures of translation surfaces. Geom. Topol. **19**, 413–438 (2015) Zbl 1318.32021 MR 3318755