QUANTITATIVE RECURRENCE AND LARGE DEVIATIONS
FOR TEICHMULLER GEODESIC FLOW

JAYADEV S. ATHREYA
UNIVERSITY OF CHICAGO

Abstract. We prove quantitative recurrence and large deviations results for
the Teichmüller geodesic flow on connected components of strata of the moduli
space $Q_g$ of holomorphic unit-area quadratic differentials on a compact genus
$g \geq 2$ surface.

Keywords: Moduli spaces, geodesic flow, large deviations.

Acknowledgements: Research partially supported by NSF grant DMS-0244542.

1. Introduction

Let $\Sigma_g$ be a compact surface of genus $g \geq 2$. Let $Q_g$ be the moduli space of
unit-area holomorphic quadratic differentials on $\Sigma_g$. That is, a point $q \in Q_g$ is an
equivalence class of pairs $(M, \omega)$, where $M$ is a genus $g$ Riemann surface, and $\omega$
is a holomorphic quadratic differential on $M$, i.e., a tensor with the form $f(z)dz^2$
in local coordinates, such that $\int_M |\omega| = 1$. Two pairs $(M_1, \omega_1)$ and $(M_2, \omega_2)$ are
equivalent if there is a biholomorphism $f : M_1 \to M_2$ such that $f_*\omega_1 = \omega_2$.

Given a pair $q \in Q_g$, one obtains (via integration of the form) an atlas of charts
to $\mathbb{C} \cong \mathbb{R}^2$, with transition maps of the form $z \mapsto \pm z + c$. Similarly, given such
an atlas of charts, one obtains a holomorphic quadratic differential by pulling back
the form $dz^2$ on $\mathbb{C}$.

These charts allow us to define a $SL(2, \mathbb{R})$ action on $Q_g$ (and $\tilde{Q}_g$) given by linear
post-composition with charts.

$Q_g$ is naturally the unit cotangent bundle to $M_g$, the moduli space of Riemann
surfaces. The fiber over each point $M \in M_g$ is the vector space of holomorphic
quadratic differentials on $M$.

The space $Q_g$ is naturally stratified by integer partitions $\beta$ of $4g - 4$. Strata are
not always connected, however, they have at most finitely many components \[19\],
and are invariant under $SL(2, \mathbb{R})$. For the rest of our paper we work with one of
these connected components, call it $Q$.

Without loss of generality, we will study strata of squares of abelian differentials
(holomorphic 1-forms). Otherwise, we pass to a double cover. Since our results
hold for all $g \geq 2$, we simply consider the (higher-genus) stratum defined in this
way. Each stratum $Q$, while non-compact, is endowed with a continuous, ergodic,$SL(2, \mathbb{R})$-invariant probability measure $\mu_Q$.

More details about the basics of moduli spaces and quadratic differentials can
be found in, e.g., \[10\] \[19\] \[23\].

We are particularly interested in the action of the standard subgroups $K$ (maxi-
mal compact) and $A$ (diagonal matrices) of $SL(2, \mathbb{R})$. These are both one-parameter

\[2000 \text{ Mathematics Subject Classification.} \quad \text{32G15, 37A10.} \]
subgroups, and can be described as follows:

\[ K = \left\{ r_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} : 0 \leq \theta < 2\pi \right\} \]

\[ A = \left\{ g_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} : t \in \mathbb{R} \right\}. \]

We will often be interested also in the action of the semigroup \( A^+ = \{ g_t : t \geq 0 \} \).

The action of \( K \), known as the circle flow, preserves the underlying holomorphic structure, so it acts as identity when projected to \( \mathcal{M}_g \). The action of \( A \) is known as Teichmüller geodesic flow, since the projection of an \( A \)-orbit yields a geodesic in the Teichmüller metric on \( \mathcal{M}_g \) (and in fact, all Teichmüller geodesics arise this way).

Masur [17] and Veech [25] independently showed the Teichmüller geodesic flow is ergodic with respect to \( \mu_Q \) and even mixing. Veech [26] showed that it was ‘measurably Anosov’, and more recently Forni [11] has obtained explicit formulas regarding the hyperbolic behavior. His results, which imply that as long as trajectories remain in a compact set their hyperbolicity can be controlled, provide much of the motivation for our research.

Our main results are concerned with the following scenario. Fix \( q \in \mathbb{Q} \), and consider the ‘circle’ \( Kq = \{ r_\theta q : 0 \leq \theta < 2\pi \} \). This set is endowed with a natural probability measure \( \nu \), coming from the Haar measure on \( K \sim S^1 \). What is the recurrence behavior of a trajectory \( \{ g_t r_\theta q \} \) when \( \theta \) is chosen at random according to \( \nu \)?

This type of situation was first considered by Kerckhoff-Masur-Smillie [13], in order to understand dynamics of billiards in rational angle Euclidean polygons. They developed a dictionary between the dynamics of the straight-line flow on the singular Euclidean surface determined by a quadratic differential \( q \) and the recurrence behavior of the geodesic trajectory \( g_t q \) in the stratum \( \mathcal{Q} \). The \( K \)-action here corresponds to changing the direction of the straight line flow without changing the underlying surface. Thus, making a statement about a fixed \( q \) and almost all directions allowed one to say something about straight line flow in almost all directions for a given flat surface, in analogy with Weyl’s equidistribution theorem on the torus.

To apply these results to billiards, one follows the unfolding procedure of Zemljakov and Katok [27], which translate questions about dynamics of billiards in rational angle polygons to that of flows on an associated singular Euclidean surface. However, the set of surfaces arising from billiards is of measure 0 in every stratum, so making statements about almost every point \( q \) does not suffice. To make statements about billiards, we need statements that hold for all \( q \in \mathbb{Q} \).

The main result in [13] is that for all \( q \in \mathbb{Q} \), and almost every \( \theta \in [0, 2\pi) \), the geodesic trajectory \( \{ g_t r_\theta q \}_{t \geq 0} \) is recurrent in \( \mathcal{Q} \). As a corollary, one obtains that the directional flow for Euclidean polygonal billiards is uniquely ergodic for almost every direction. Further results in this direction include estimates on the Hausdorff dimension of divergent and bounded trajectories ([3, 15]), and further relations between dynamics of the straight line flow and recurrence of the associated trajectory ([4, 5]).

Our main results concern finer recurrence behavior of geodesic trajectories. In particular, for any fixed \( q \in \mathbb{Q} \), we estimate the measure of the set of angles such
that the associated $A^+$ orbit $\{g_t r \theta q\}_{t \geq 0}$ ‘behaves poorly’ for a length of time $T$. For us, poor behavior means that the trajectory is spending a lot of time in a neighborhood of the cusp of $Q$. Masur [18] proved a statistical result, known as a logarithm law in this situation, in analogy with earlier results on symmetric spaces due to Sullivan [22] and Kleinbock-Margulis [14].

We will construct a proper (that is, unbounded off compact sets) continuous function $V : Q \to \mathbb{R}^+$, and consider the recurrence behavior of $\{g_t r \theta q\}_{t \geq 0}$ to the family of compact sets $C_l := \{q : V(q) \leq l\}$, which form an exhaustion of $Q$ as $l$ varies. Our main results can be summarized as follows:

**Theorem 1.1.** Fix notation as above. Then

1. For all $l$ sufficiently large and all $q \notin C_1$, there are positive constants $c_1 = c_1(l, q), c_2(l)$, with
   \[ \nu\{\theta : g_t r \theta q \notin C_1, 0 \leq t \leq T\} \leq c_1 e^{-c_2 T} \]
   for all $T$ sufficiently large. That is, the probability that a random geodesic trajectory has not visited $C_1$ by time $T$ decays exponentially in $T$.

2. For all $l, S, T$ sufficiently large and all $q \in Q$, there are positive constants $c_3 = c_3(l, q), c_4 = c_4(l)$, with
   \[ \nu\{\theta : g_t r \theta q \notin C_1, S \leq t \leq S + T\} \leq c_3 e^{-c_4 T}. \]
   That is, the probability that a random geodesic trajectory does not enter $C_1$ in the interval $[S, S + T]$ decays exponentially in $T$.

3. Let $q \in Q$. For any $0 < \lambda < 1$, there is a $l \geq 0$, and $0 < \gamma < 1$, such that for all $T$ sufficiently large (depending on all the above constants)
   \[ \nu\{\theta : \frac{1}{T}|\{0 \leq t \leq T : g_t r \theta q \notin C_1\}| > \lambda\} \leq \gamma^T. \]

Result (3) above may be thought of as a large deviations result for the Teichmuller flow. While ergodicity guarantees that $\frac{1}{T}|\{0 \leq t \leq T : g_t q \in C_1\}| \to \mu_Q(C_1)$ for $\mu_Q$-almost every $q \in Q$, our result gives explicit information about the likelihood of bad trajectories. Notice, however, this is not a traditional large deviations result, which estimates the probability of a deviation of any $\epsilon > 0$ from the ergodic average. Other interesting results concerning deviations are due to Bufetov [2], who proved a central limit theorem for this flow.

In [11], Forni related this type of fine recurrence behavior for the geodesic trajectory $\{g_t q\}_{t \geq 0}$ to deviation of ergodic averages for the straight line flow on the flat surface associated to $q$. He proved that for almost every $q \in Q$, the associated flow has the same deviation behavior. However, his result gives no information on billiards, as the set of quadratic differentials arising from billiards are a set of measure zero. Since our results provide fine recurrence information about the geodesic trajectory $\{g_t r \theta q\}_{t \geq 0}$ for all $q \in Q$, and almost all $\theta$, we conjecture the following:

**Conjecture.** For all rational-angle Euclidean polygons, the deviation of ergodic averages for the billiard flow is the same for almost all directions, and depends only on the $\text{SL}(2, \mathbb{R})$ orbit of the associated quadratic differential.

It was also shown in [11] that as long as a geodesic trajectory stays within a compact set, the rate of expansion/contraction in the tangent space along the trajectory is bounded away from 1. Thus our results can be used to obtain explicit
estimates on the hyperbolicity of the flow along specific trajectories. Avila-Gouezel-Yoccoz [1] used exponential return estimates to a different family of compact sets to prove the exponential rate of mixing for the Teichmuller geodesic flow, which was the original motivation for this research.

Other quantitative recurrence results for dynamics on Teichmuller spaces were obtained by Minsky-Weiss [20] for the case of Teichmuller horocycle flow.

We also have a collection of results for a certain class of random walks, defined as follows: Fix $\tau > 0$. Given that we are at a point $q \in \mathbb{Q}$, the next point in our trajectory will be chosen at random according to Haar measure on $S^1$ from the ‘circle’ of radius $\tau$, $\{g_r \cdot q : 0 \leq \theta < 2\pi\}$. Note that $\mu_Q$ is a stationary measure for this walk, since it is $SL(2, \mathbb{R})$-invariant. Let $\{X_n\}_{n=0}^\infty$ denote the random walk generated this way.

**Remark:** By hyperbolic geometry, one can see that trajectories of this walk closely approximate geodesics. Thus, understanding recurrence properties of the walk gives one insight into properties of the flow.

We define $P_q(E) := \text{Prob}(E|X_0 = q)$ for any event $E$ defined on the trajectory $\{X_n\}_{n=0}^\infty$ starting at $X_0 = q$. Also, for any measurable $A \subset \mathbb{Q}$, we define $P_n(q,A) := P_q(X_n \in A)$. That is, $P_n(q,\cdot)$ is the probability distribution of $X_n$ given $X_0 = q$.

**Theorem 1.2.** For all $\tau$ sufficiently large, we have:

1. There is an $l(\tau)$ such that for all $l > l(\tau)$, and all $q \notin C_l$, there are constants $c_5 = c_5(q,l,\tau), c_6 = c_6(l,\tau)$, such that
   $$P_q(X_i \notin C_l : 1 \leq i \leq n) \leq c_5 e^{-c_6 n}.$$

2. For all compact $C \subset \mathbb{Q}$, $\epsilon > 0$, there is a $C_l \supset C$ such that $\forall q \in C$, and for all $m \geq 0$,
   $$P_q(X_m \in C_l) > 1 - \epsilon.$$

3. For all $\epsilon > 0$, there is a $l = l(\epsilon) > 0$ such that $\forall q \in Q$, there is an $M(q)$ such that for all $m > M(q)$,
   $$P_q(X_m \in C_l) > 1 - \epsilon.$$

The rest of the paper is organized as follows: In the next section, we give more detailed statements of our main results, and construct the function $V$. We also give a version of our results for general $SL(2, \mathbb{R})$-actions. In section 3 we prove our results for random walks. In section 4 we collect some technical lemmas about change of polar coordinates on hyperbolic space. In section 5 we construct the required lemmas from the theory of large deviations. In section 6 we prove our main theorems for the flow.

## 2. Statement of Results

The precise statements for the flow regard a family of proper functions $V_\delta$, $0 < \delta < 1$, and the compact sets $C_{\delta,l} := \{q : V_\delta(q) \leq l\}$. We have:

### 2.1. Flow Results

**Theorem 2.1.** For every $1 > \delta > 0$ there is a proper (i.e., unbounded off compact sets), smooth, $K$-invariant function $V_\delta : Q \to \mathbb{R}^+$ and positive constants $t_0 =$...
Let $q \in Q$. For every $\epsilon > 0$, $1 > \delta > 0$, there are positive constants $S_0, T_1, l_1$ depending on $\delta$, $1 > \delta'' = \delta''(l) > \delta$, and $\alpha = \alpha(q)$ such that
\[
\nu\{ \theta : g_r \varrho \theta q \notin C_{\delta,l}, 0 \leq t \leq S + T \} \leq \alpha e^{-(1-\delta'')(S + T)},
\]
for all $S > S_0$, $T > T_1$, and $l > l_1$, with
\[
\alpha \leq 8(1 + \epsilon) \sup_{q \in (0, 2\pi)} \frac{ab V'(g_r \varrho \theta q)}{l},
\]
where $b$ depends only on the curvature of $\mathbb{H}^2$.

Here, $S_0$ depends only on the choice of curvature for the hyperbolic plane $\mathbb{H}^2$, and $T_1$ is the maximum of $t_0$ from Theorem 2.1 and a $T_2$ depending only on curvature.

Theorem 2.3. Let $q \in Q$. For any $0 < \lambda < 1$, and any $0 < \delta < 1$ there is a $l \geq 0$, and $\gamma < 1$, such that
\[
\nu\{ \frac{1}{T} \{ 0 \leq t \leq T : g_r \varrho \theta q \notin C_{\delta,l} \} > \lambda \} \leq \gamma^T,
\]
for all $T$ sufficiently large.

This theorem uses a technical tool from the theory of large deviations, Proposition 7.1, which makes it difficult to track the dependence of $\gamma$ on $l$ and $\lambda$. Clearly, if an $l$ works for a fixed $\lambda_0$ and $\delta$, it works for all $\lambda > \lambda_0$, and any larger $l$ (with the same $\delta$) will work for $\lambda_0$. Similarly, the same $l$ will work for the same $\lambda_0$ and any smaller $\delta$.

In a personal communication, Forni posed the following question: for every $q$ in $Q$, and almost all $\theta$, does there exist a $\zeta > 0$ such that the geodesic trajectory $\{ g_r \varrho \theta q \}_{t \geq 0}$ spends at least a proportion $\zeta$ of its time in a fixed compact set? The following corollary answers in the affirmative.

Corollary 2.4. Let $q \in Q, \zeta > 0$. Fix $\epsilon > 0$. Let $l$ be such that Theorem 2.3 is satisfied with $\lambda = 1 - \zeta, \epsilon = \epsilon$. Then for $\nu$-almost every $\theta$,
\[
\limsup_{T \to \infty} \frac{1}{T} \{ 0 \leq t \leq T : g_r \varrho \theta q \notin C_{\delta,l} \} \leq \lambda.
\]

The proof of this result follows from an application of the Borel-Cantelli lemma and Theorem 2.3.

2.2. Random walks on $Q$. Let $\{ X_n \}_{n=0}^{\infty}$ be our trajectory. That is, $X_{n+1} = g_r \varrho \theta_n X_n$, where $\{ \theta_n \}_{n=0}^{\infty}$ are independent and identically distributed (i.i.d.) according to the uniform distribution on $[0, 2\pi)$ (equivalently, $X_{n+1} = g_r k_n X_n$, where $\{ k_n \}_{n=0}^{\infty}$ are i.i.d. according to Haar measure on $K$). This implies $\{ X_n \}$ is a Markov chain, with transition probability function $P(x, A) = \nu\{ \theta : g_r \varrho \theta q \in A \}$.
Theorem 2.5. Fix $0 < \delta < 1$. Let $\{C_{\delta,l}\}$ be as in Theorem 2.4. Then there is a $\tau_0 = \tau_0(\delta) > 0$ such that for all $\tau > \tau_0$, there is a $\tilde{l}_0 = \tilde{l}_0(\delta,\tau)$ such that for all $l > \tilde{l}_0$, there is a $\gamma = \gamma(l,\delta) < 1$ such that for all $q \notin C_{\delta,l}$, and for all $n \geq 1$,

$$P_q(X_1 \notin C_{\delta,l} : 0 \leq j \leq n) \leq \frac{V_{\delta}(q)}{l} \gamma^n.$$ 

This is completely analogous to Theorem 2.4, and indeed the proof of this result will be essential to the proof of Theorem 2.6.

For the rest of the section, fix $1 > \delta > 0$, and $\tau > \tau_0(\delta)$.

Theorem 2.6. Let $\epsilon > 0$. Then:

1. For all compact $C \subset Q$, there is a $C_{\delta,l} \supset C$ such that $\forall q \in C$, and for all $m \geq 0$, 

   $$P_q(X_m \in C_{\delta,l}) > 1 - \epsilon.$$

2. There is a $l > 0$ such that $\forall q \in Q$, there is an $M(q)$ such that for all $m > M(q)$,

   $$P_q(X_m \in C_{\delta,l}) > 1 - \epsilon.$$

Finally, we have:

Corollary 2.7. Fix $q \in Q$, $0 < \lambda < 1$, and $\epsilon > 0$. Then there is a $l = l(q,\lambda,\delta,\epsilon) > 0$ so that, for all $n$ sufficiently large,

$$P_q \left( \frac{1}{n} \left| \{1 \leq i \leq n : X_i \in C_{\delta,l}\} \right| > \lambda \right) > 1 - \epsilon$$

This is analogous to Corollary 2.4

Remark: $l$ can be chosen uniformly as $q$ varies over a compact set.

2.3. $SL(2,\mathbb{R})$-actions. Let $SL(2,\mathbb{R})$ act continuously on a topological space $X$. Suppose there is a family of $K$-invariant functions $V_\delta : X \to \mathbb{R}$, $0 < \delta < 1$ satisfying the following properties:

- For all $x \in X$, consider the function $V_{\delta,x} : SL(2,\mathbb{R}) \to \mathbb{R}^+$ defined by $V_{\delta,x} = V_\delta(gx)$. Note that by $K$-invariance, we can view $V_{\delta,x}$ as a function on $\mathbb{H}^2 = SO(2) \backslash SL(2,\mathbb{R})$. We require the following: For all $\sigma > 1$, there exists a $\kappa > 0$ such that for any $p \in \mathbb{H}^2$ with $d(p,i) < \kappa$, and any $h \in SL(2,\mathbb{R})$

   $$\sigma^{-1}V_{\delta,x}(h) \leq V_{\delta,x}(ph) \leq \sigma V_{\delta,x}(h).$$

(2.8)

We say such a function is logsmooth.

- For all $1 > \delta > 0$, there is a constant $\tilde{c} = \tilde{c}(\delta)$ such that for all sufficiently large $\tau$, there is a $\tilde{b} = \tilde{b}(\tau,\delta)$ such that for all $x \in X$,

   $$a(x) := \int_0^{2\pi} V_\delta(g_r \tau_0 x) d\nu(\theta) \leq \tilde{c}e^{-(1-\delta)\tau}V_\delta(x) + \tilde{b}.$$  

(2.9)

All our results on recurrence to the sublevel sets $C_{\delta,l}$ from subsections 2.1 and 2.2 while constructed for the space $Q$ hold for any space $X$ with the above properties.

The key lemma needed to prove our results will be the construction of a family of functions $\{V_\delta\}$ on $Q$ satisfying the above requirements.

Lemma 2.10. There is a family of smooth, proper functions $V_\delta : Q \to \mathbb{R}^+$, $0 < \delta < 1$ satisfying equations (2.8) and (2.9).
We require the following technical lemma from [8] (page 465, Lemma 7.5).

**Lemma.** For all \( \delta > 0 \) there are logsmooth functions \( V_0, \ldots, V_n : Q \to \mathbb{R}^+ \) such that \( V_0 \) is proper, and for every \( \tau > 0 \), there are constants \( w = w(\tau, \delta), \tilde{b}' = \tilde{b}'(\tau, \delta) \), and \( \tilde{c}' = c'(\delta) \), independent of \( \tau \) such that for all \( 0 \leq i \leq n \), and \( \forall q \in Q \),

\[
(A_\tau V_i)(q) \leq \tilde{c}' e^{-(1-\delta)\tau} V_i(q) + w \sum_{j=i+1}^{n} V_j(q) + \tilde{b}'.
\]

**Remark:** In fact \( V_0(q) = \max \left( 1, \frac{1}{l(q)^{1-\delta}} \right) \), where \( l(q) \) denotes the length of the shortest saddle connection on \( q \). Recall that a saddle connection is a geodesic (in the metric determined by \( q \)) connecting two zeroes of \( q \).

**Proof of Lemma 2.10**

Fix \( \delta, \tau > 0 \). Let \( \tilde{c} = 2\tilde{c}' \lambda_0 = \frac{\tilde{c}}{\tilde{c}} \), \( \lambda_i = \left( \frac{\tilde{c}}{\tilde{c}} + 1 \right)^{i-1} \), for \( 0 \leq i \leq n \). Note that

\[
\sum_{i=0}^{j-1} \lambda_i \leq \frac{\tilde{c}}{2w} \lambda_j.
\]

Set \( \tilde{b} = \tilde{b}' \sum_{i=0}^{n} \lambda_i \). Let \( V_0(q) = \sum_{i=0}^{n} \lambda_i V_i(q) \). Then

\[
(A_\tau V_0)(q) = \sum_{i=0}^{n} \lambda_i (A_\tau V_i)(q) \leq \sum_{i=0}^{n} \lambda_i \left( \tilde{c}' e^{-(1-\delta)\tau} V_i(q) + w \sum_{j=i+1}^{n} V_j(q) + \tilde{b}' \right)
\]

\[
= \sum_{i=0}^{n} \lambda_i \tilde{c}' e^{-(1-\delta)\tau} V_i(q) + w \sum_{j=1}^{n} \left( \sum_{i=0}^{j-1} \lambda_i \right) V_j(q) + \tilde{b}
\]

\[
= \tilde{c}' e^{-(1-\delta)\tau} \lambda_0 V_0(q) + \sum_{i=1}^{n} V_i(q) \left( \lambda_i \tilde{c}' e^{-(1-\delta)\tau} + w \sum_{i=0}^{j-1} \lambda_i \right) + \tilde{b}
\]

\[
= \tilde{c}' e^{-(1-\delta)\tau} \lambda_0 V_0(q) + \sum_{j=1}^{n} 2\lambda_j \tilde{c}' e^{-(1-\delta)\tau} V_j(q) + \tilde{b}
\]

\[
\leq \tilde{c}' e^{-(1-\delta)\tau} V_0(q) + \tilde{b}.
\]

Thus we have constructed a family of functions \( V_0 \) satisfying equations (2.8), and \( (A_\tau V_i) \). That \( V_0 \) satisfies (2.8) follows from the logsmoothness of the \( V_i \)'s noted in [8] (where they are called \( \alpha_i \)'s) on page 471, at the beginning of the proof of Proposition 7.2. Finally, \( V_0 \) is proper since \( V_0 \) is proper.

\[ \square \]

3. **General Markov Chain results**

Our goal in this section is to recall some results from the theory of Markov Chains which will allow us to prove Theorems 2.8, 2.9. We also give some applications to random walks on homogeneous spaces, as considered in [7]. Many of these results can be found, in greater generality, in [21].

Before stating our main results, we recall some basic notation: If \( \{X_n\} \) is a Markov chain on \((S, \mathcal{S})\), \( S \) the state space, and \( \mathcal{S} \) the \( \sigma \)-algebra, then, for any event \( E \) (an event \( E \) is a set in the product \( \sigma \)-algebra) and any starting point
Proposition 3.1. Let $V$ be a proper function for all $x \in \tau \{\text{algebra. Let we write} \}$ given that our starting point was $x$. Similarly, given a measurable subset $C \subset S$, we write $P^n(x, C) = P_x(X_n \in C)$, and we define the first hitting time of $C$ by

$$\tau_C := \inf \{n \geq 1 : X_n \in C\}.$$ 

The main result of this section is:

**Proposition 3.1.** Let $S$ be a non-compact topological space, and $S$ its Borel $\sigma$-algebra. Let $\{X_n\}_{n=0}^\infty$ be a Markov chain on $(S, S)$. Suppose there exists a smooth, proper function $V : S \to \mathbb{R}^+$ and constants $0 < c < 1$ and $b > 0$ such that

$$(PV)(x) := E(V(X_1)|X_0 = x) \leq cV(x) + b,$$

for all $x \in S$. Then

1. For any $l > 0$ and $x \notin C_l := \{y \in S : V(y) \leq l\}$,

$$p_n(x) := P_x(\tau_{C_l} > n) \leq \frac{V(x)}{l} \left(c + \frac{b}{l}\right)^n,$$

for all $n \geq 0$.

2. For all $\epsilon > 0$, and for all compact $C \subset S$, there is an $l$ such that $\forall x \in C, m \geq 0$,

$$P^m(x, C_l) > 1 - \epsilon.$$ 

3. For all $\epsilon > 0$ there is an $l > 0$ such that $\forall x \in S$, there is an $M(x) > 0$ such that for $m > M(x)$,

$$P^m(x, C_l) > 1 - \epsilon.$$ 

4. For all $x \in S$, $0 < \lambda < 1$, and $\epsilon > 0$, there is a $l = l(q, \lambda, \epsilon) > 0$ so that, for all $n$ sufficiently large,

$$P_x \left(\frac{1}{n} \{1 \leq i \leq n : X_i \in C_l\} > \lambda\right) > 1 - \epsilon.$$ 

If we fix $\lambda$ and $\epsilon$, then $l$ can be chosen uniformly as $x$ varies over a compact set.

**Proof of Theorems 2.5-2.6 and Corollary 2.7:** Combine Proposition 3.1 with Lemma 2.10. For Theorem 2.5, fix $\delta > 0$ and let $\tau_0$ be such that $c = e^{-1(1 - \delta)\tau_0} < 1$, we obtain our result with $\gamma = \left(c + \frac{b}{l}\right)$ for $l > \frac{1}{\gamma}$. \qed

**Proof of Proposition 3.1**

- **Proof of (1):** Let $B_n := \{\tau_{C_l} > n\}$. Then $p_n(x) = P_x(B_n)$. For $n \geq 0$,

$$l p_n \leq E_x(V(X_n) : B_n) =: D_n$$

since on $B_n$, $X_n \notin C_l$, i.e. $V(X_n) \geq l$.

Now, $B_n \subset B_{n-1}$, so

$$D_n \leq E_x(V(X_n) : B_{n-1}) = E_x((PV)(X_{n-1}) : B_{n-1})$$

where we are using the Markov property in the 2nd line.

Now we can apply our condition

$$(PV)(X_{n-1}) \leq cV(X_{n-1}) + b.$$
This yields

\[ D_n \leq cD_{n-1} + bp_{n-1}, \]

and using the observation that \( p_{n-1} \leq \frac{D_{n-1}}{l} \), we obtain the recurrence relation

\[ D_n \leq \left( c + \frac{b}{l} \right) D_{n-1}. \]

Iterating this, we obtain

\[ D_n \leq D_0 \left( c + \frac{b}{l} \right)^n. \]

Since \( D_0 = V(X_0) = V(x) \), and \( p_n \leq \frac{D_n}{l} \), we obtain our result. Note that the result is only meaningful if \( (c + \frac{b}{l}) < 1 \), which is equivalent to setting \( l > \frac{b}{1-c} \).

- **Proof of (2):** We have \((PV)(x) \leq cV(X) + b\) for \( c < 1 \). Iterating this, we get that \((P^mV)(x) \leq e^mV(x) + b'\), where \( b' \) does not depend on \( m \) or \( x \). Set \( l = \sup_{g \in \Gamma} \frac{V(g)}{b'} \). Then we have that

\[ lp^m(x, C') \leq E_x(V(X_m)) \leq e^mV(x) + b', \]

so we get that \( P^m(x, C') < \epsilon \) as desired. □

- **Proof of (3):** For the third property, we select \( l = 2b'/\epsilon \), where \( b' \) is as above. For \( m \) sufficiently large, \( e^mV(x) \leq b' \), so by the argument in part (2), we can get our conclusion. □

- **Proof of (4):** Fix \( x \in S \), \( 0 < \lambda < 1 \), and \( \epsilon > 0 \). By part (3), there is an \( l > 0 \) so that for \( n \) sufficiently large, \( P(X_n \in C_I) > 1 - \epsilon' \), where \( 0 < \epsilon' < \frac{2}{3} \epsilon (1 - \lambda) \). Set \( S_n = \frac{1}{n} \sum_{i=1}^{n} \chi_{C_I}(X_i) \), where \( \chi_{C_I} \) is the indicator function of \( C_I \). Then, for any \( \lambda < 1 \),

\[
E(S_n) \leq \lambda P(S_n \leq \lambda) + P(S_n > \lambda) = \lambda + (1 - \lambda)P(S_n > \lambda).
\]

Thus, we have

\[ P(S_n > \lambda) \geq \frac{E(S_n) - \lambda}{1 - \lambda}. \]

Now, for \( n \) sufficiently large \( E(S_n) \geq 1 - \frac{3}{4} \epsilon' \), thus,

\[ P(S_n > \lambda) \geq \frac{1 - \frac{3}{4} \epsilon' - \lambda}{1 - \lambda} > 1 - \epsilon. \]

The fact that \( l \) can be chosen uniformly as \( x \) varies over a compact set follows from part (2). □

These types of questions were considered for random walks on homogeneous spaces in [7] by Eskin and Margulis. They constructed a function \( V \) on their state space satisfying the conditions of Proposition [3] and used this to draw conclusions (2) and (3). Conclusions (1) and (4) appear to be new results for these walks.

Precisely, we have the following:
Theorem. Let $G$ be a semisimple Lie group, and $\Gamma$ a non-uniform lattice. Let $\mu$ be a probability measure on $G$ satisfying the conditions of Theorem 2.1 in \cite{7}. Consider the Markov chain $\{X_n\}_{n=0}^\infty$ defined on $G/\Gamma$ by the measure $\mu$:

$$X_{n+1} = g_n X_n,$$

with $\{g_n\}_{n=0}^\infty$ an i.i.d. (with distribution $\mu$) sequence of elements of $G$. Then there is a function $V : G/\Gamma \to \mathbb{R}^+$ satisfying the conditions of Proposition \ref{prop:1}. Thus, conclusions (1)-(4) of the Proposition are satisfied.

4. Polar coordinates and shadowing

We require two lemmas about change of polar coordinates in the hyperbolic plane $\mathbb{H}^2 = SO(2)\setminus SL(2,\mathbb{R})$. We fix two positive numbers $t_1, t_2$, and basepoints $i$ and $z_0 = i.g_{t_1} r_0$ (these will correspond to our basepoint $q$ and an arbitrary $q_0$ in its $SL(2,\mathbb{R})$-orbit, projected to $SO(2)\setminus Q$). We let $d(.,.)$ denote distance in the hyperbolic plane.

Consider the circle of radius $t_2$ around $z_0$, defined by $\{z_\phi = i.g_{t_2} r_0 g_{t_1} r_0 : 0 < \phi < 2\pi\}$. We say that $t_2, \phi$ are the polar coordinates of $z_\phi$ based at $z_0$. For each $\phi$, we define $D = D_{t_1, t_2}(\phi)$ and $\Psi = \Psi_{t_1, t_2}(\phi)$ by $z_\phi = i.g_{D(\phi)} r_\theta \Psi(\phi)$, i.e. $D(\phi), \theta + \Psi(\phi)$ are the polar coordinates of $z_\phi$ based at $i$. Note that $D, \Psi$ are independent of $\theta$.

Geometrically, $D(\phi)$ is the distance and $\Psi(\phi)$ is the angle (measured clockwise from the the geodesic connecting $i$ to $z_0$) of the geodesic segment connecting $i$ to $z_\phi$. Hyperbolic trigonometry (the laws of sines and cosines, applied to the triangle formed by the points $i$, $z_0$, and $z_\phi$) yield:

\begin{align}
\cosh D(\phi) &= \cosh t_1 \cosh t_2 + \sinh t_1 \sinh t_2 \cos \phi, \\
\sin \Psi(\phi) &= \frac{\sinh t_2}{\sinh D(\phi)} \sin \phi.
\end{align}

If $t_2 > t_1$, $i$ lies inside the circle of radius $\tau$ around $z_0$, and thus, the map $\Psi$ is both one-to-one and onto. If $t_1 > t_2$, the point $i$ is outside the circle, and $\Psi$ is neither one-to-one or onto. In this case, the image is an interval, with boundary points such that the geodesic determined by those angles intersects the circle of radius $t$ tangentially. In the interior of the interval, each point $\psi$ has two preimages, call them $\phi_1, \phi_2$ one such that $D(\phi_1) \approx t_1 - t_2$, and one such that $D(\phi_2) \approx t_1 + t_2$. For our applications, we will only be concerned with $\phi_2$.

The key technical lemma is as follows:

**Lemma 4.3.** Let $A \subset [0, 2\pi)$ be a measurable set. Then, for every $\epsilon > 0$, there are $\tau_1, \tau_2$, such that for all $t_1 > \tau_1, t_2 > \tau_2$ the neighborhood $U = \Psi_{t_1, t_2}(\pm \pi/2, \pi/2)$ of $0$ satisfies

$$\frac{\nu(U \cap \Psi(A))}{\nu(U)} \leq 4(1 + \epsilon)\nu(A).$$

**Proof:** For this estimate, we need to control the behavior of the derivative $\Psi'$. More precisely, we need to control ratios $\Psi'(\phi_1)/\Psi'(\phi_2)$, with $\phi_1, \phi_2 \in \Psi^{-1} U$, so we can compare $\nu$ and $\Psi_* \nu$, where $\Psi_* \nu(E) = \nu(\Psi^{-1} E)$.

We have the following claim:
Claim. Let $\eta > 0$. For $t_1, t_2$ sufficiently large,
\[
\frac{e^{-t_1}}{2}(1 - \eta) \leq |\Psi'(\phi)| \leq e^{-t_1}(1 + \eta),
\]
for $\phi \in [-\pi/2, \pi/2]$.

Proof of Claim:
Implicit differentiation of equations (4.1) and (4.2) yield:
\[
D'(\phi) \sinh D(\phi) = -\sinh t_1 \sinh t_2 \sin \phi
\]
and
\[
\Psi'(\phi) \cos \Psi(\phi) = \sinh t_2 \frac{\cos \sinh D(\phi) + \sin^2 \phi \coth D(\phi) \sinh t_1 \sinh t_2}{\sinh^2 D(\phi)}.
\]

Let $\kappa > 0$. Let $t_1, t_2$ be large enough so that for all $\phi \in [-\pi/2, \pi/2]$,
(1) $D(\phi) > t_1 + t_2 - \kappa$,
(2) $\cosh(D(\phi)) > 1 - \kappa$
(3) $\cos(\Psi(\phi)) > 1 - \kappa$
(4) $|\sinh t_1 \sinh t_2 - \sinh D(\phi)| \leq \kappa$
(5) $1 - \kappa < \frac{2 \sinh t_1}{e^{t_1/2}}, \frac{2 \cosh t_1}{e^{t_1/2}} < 1 + \kappa$ for $i = 1, 2$.

That we can achieve the above inequalities follows from hyperbolic geometry and the basic properties of sinh and cosh.

Let $\eta > 0$. Using the above inequalities and some basic algebra, we can select $\tau_1, \tau_2$ such that for all $t_1 > \tau_1, t_2 > \tau_2$, we have
\[
(1 - \eta) \leq \frac{\Psi'(\phi)}{e^{-t_1}(2 \cos \phi + \sin^2 \phi)} \leq (1 + \eta).
\]

Remark: The expression $e^{-t_1}(2 \cos \phi + \sin^2 \phi)$ is obtained by replacing the quantities in equation (4.4) with their approximations (1)-(5).

Now, let $f(\phi) = 2 \cos \phi + \sin^2 \phi$. We have $1 \leq f(\phi) \leq 2$ for $\phi \in [-\pi/2, \pi/2]$. Thus,
\[
\frac{e^{-t_1}}{2}(1 - \eta) \leq \Psi'(\phi) \leq (1 + \eta)e^{-t_1},
\]
completing the proof of the claim.

To complete the proof of the lemma, let $\epsilon > 0$ and $\eta$ be such that $\frac{1 + \eta}{1 - \eta} \leq 1 + \epsilon$. By the claim, we know that the proportion of measure of any set in $[-\pi/2, \pi/2]$ cannot be expanded by more than $2\frac{1 + \eta}{1 - \eta} < 2(1 + \epsilon)$ under $\Psi$, since that is the maximum of $\frac{\Psi'(\phi_1)}{\Psi'(\phi_2)}$ for $\phi_1, \phi_2 \in [-\pi/2, \pi/2]$. Now, since $\nu([-\pi/2, \pi/2]) = 1/2$, we have
\[
\frac{\nu(A \cap [-\pi/2, \pi/2])}{\nu([-\pi/2, \pi/2])} \leq 2\nu(A),
\]
which yields
\[
\frac{\nu(\Psi(A) \cap U)}{\nu(U)} \leq 2(1 + \epsilon) \cdot 2\nu(A) = 4(1 + \epsilon)\nu(A).
\]
Our second main lemma is as follows: let $\tau > 0$. For any $t > 0$, define
\[ I(t) := \{0 \leq \theta < 2\pi : d(i.g_t, i.g_\theta) \leq 3\tau \}. \]

Fix $\kappa > 0$. For any $\theta \in I(t)$, define
\[ U_\theta := \{0 \leq \phi < 2\pi : D_{t,\tau}(\phi) > 1 - \kappa, (\Psi_{t,\tau}(\phi) + \theta) \in I(t + \tau) \}. \]

Let $L_0 = \nu(U_\theta)$. The following lemma is proved in [3]. For notational convenience, we write drop the subscripts for $T$ and $S$.

**Lemma 4.6.** There is a constant $c'' > 0$ such that for all $\kappa > 0$, there is a $\tau > 0$ such that for all $\theta \in I(t)$, the map $\Psi|_{L_\theta}$ is a diffeomorphism onto its image, and, making the substitution $\psi = \Psi(\phi)$, we have
\[ c'' \leq \nu(L_\theta) \leq \int_{L_\theta} d\phi = \int_{\Psi(L_\theta)} \frac{d\phi}{d\psi} |d\psi|. \]

**Proof:** [3], page 467, Lemma 7.6.

**Remarks:**
- $d\theta$ denotes $d\nu(\theta)$, and since we have normalized $\nu$ to be a probability measure, we do not need to divide by $2\pi$.
- In [3], the sets $I(t)$ are defined by $I(t) := [-\rho e^{-t}, \rho e^{-t}]$ for some positive constant $\rho$, and they require $\rho$ to be large enough so that the diameter of the set $J(t) = \{i.g_\theta : \theta \in I(t)\}$ is at least $2\tau$. By hyperbolic geometry, our sets $I(t)$ are of this form, and they obviously satisfy the required condition.

We have the following Corollary, also from [3]:

**Corollary 4.7.** Let $f : SL(2, \mathbb{R}) \rightarrow \mathbb{R}$ be a logsmooth $SO(2)$-invariant function. Fix $\sigma > 1$. Let $\kappa > 0$ be as in equation 2.8. Fix $\tau$ so that Lemma 4.6 holds. Then there is a $c' > 0$, independent of $\tau$, such that
\[ \int_{I(t+\tau)} f(g_{t+\tau}r_\theta)d\theta \leq c'\sigma \int_{I(t)} (A_\tau f)(g_tr_\theta)d\theta. \]

**Proof:** [3], page 468, Lemma 7.7.

5. Large Deviations

In this section, we prove the key technical lemma for our main large deviations result Theorem 2.8. We assume some familiarity with the theory of conditional expectation. Excellent references include [24, 5].

**Proposition 5.1.** Let $\{\tau_i\}_{i=0}^\infty$ be a sequence of positive real-valued random variables on a probability space $(\Omega, \mathcal{F}, P)$. Let $\{\mathcal{F}_i\}$ be a filtration of $\mathcal{F}$ such that for all $i$, $\tau_i \in \mathcal{F}_i$, i.e., $\tau_i$ is $\mathcal{F}_i$-measurable. Suppose there exist positive random variables $\eta, \xi$ with $E\eta < E\xi$, $E\xi > 0$, and a real number $\theta_0 > 0$ such that, for all $0 \leq \theta < \theta_0$:
1. $E(e^{\theta\tau_1}|\mathcal{F}_{2i-1}) \leq E(e^{\theta\eta})$
2. $E(e^{-\theta(t_1+\tau_2)}|\mathcal{F}_{2i-1}) \leq E(e^{-\theta\xi})$

Let $T_n = \sum_{i=0}^n \tau_i$. Let
\[ X(t) = \begin{cases} 1 & T_{2i-1} \leq t < T_{2i} \\ 0 & \text{otherwise} \end{cases} \]
Then, \( \forall \lambda > E\eta / E\xi \), there is a \( 0 < \gamma < 1 \) such that for all \( T \) sufficiently large
\[
P \left( \frac{1}{T} \int_0^T X(t)dt > \lambda \right) \leq \gamma^T.
\]

For our application, \( \tau_{2i} \) will be the time a random geodesic spends outside the compact set, and \( \tau_{2i-1} \) the time inside on the \( i \)th ‘sojourn’. \( \eta \) is a stochastic upper bound for \( \tau_{2i} \), and \( \xi \) as a stochastic lower bound for the length of a ‘cycle’. Thus condition 1 should be thought of as a stochastic upper bound for the time spent outside and condition 2 a stochastic lower bound for the total time of a cycle. In our application, we will give a stronger, deterministic lower bound: in fact, we will construct our variables so that \( \tau_{2i-1} > C \), for some fixed \( C \). This clearly implies condition 2, simply by taking \( \xi = C \). For condition 1, we will use Theorem 2.1 to show that \( \tau_{2i} \) cannot grow too large.

**Proof:**

Define \( f(\theta) := E(e^{\theta \eta}) \) and \( g(\theta) := E(e^{-\theta \xi}) \).

Let \( N(T) = \sup \nu \{ k : \tau_{2k} \leq T \} \). Then,
\[
P \left( \int_0^T X(t)dt > \lambda T \right) \leq P \left( \int_0^T X(t)dt > \lambda T \text{ and } N(T) \leq cT \right) + P \left( N(T) > cT \right)
\]
(5.2)
\[
= P \left( \sum_{i=0}^{\lfloor cT/2 \rfloor} \tau_{2i} > \lambda T \right) + P \left( N(T) > cT \right).
\]

We estimate each of these terms in turn. Let \( n := \lfloor cT/2 \rfloor \). Then, since
\[
P \left( \sum_{i=0}^{\lfloor cT/2 \rfloor} \tau_{2i} > \lambda'n \right) \geq P \left( \sum_{i=0}^{\lfloor cT/2 \rfloor} \tau_{2i} > \lambda T \right) \geq P \left( \sum_{i=0}^{\lfloor cT/2 \rfloor} \tau_{2i} > \lambda'(n + 1) \right)
\]
to estimate the first term it suffices to estimate \( P \left( \sum_{i=0}^{n} \tau_{2i} > \lambda'n \right) \), where \( \lambda' = 2\lambda/c \). Now, for any \( \theta_0 > \theta \geq 0 \)
\[
P \left( \sum_{i=0}^{n} \tau_{2i} > \lambda'n \right) = P \left( e^{\theta \sum_{i=0}^{n} \tau_{2i}} > e^{\theta \lambda'n} \right)
\]
\[
\leq e^{-\theta \lambda'n} E \left( e^{\theta \sum_{i=0}^{n} \tau_{2i}} \right)
\]
\[
\leq e^{-\theta \lambda'n} f(\theta)^n.
\]
(5.3)

The last inequality follows from the first condition in our theorem, and the fact that each \( \tau_i \) is \( \mathcal{F}_i \)-measurable. Since equation holds for any \( \theta_0 > \theta \geq 0 \), we have
\[
P \left( \sum_{i=0}^{n} \tau_{2i} > \lambda'n \right) \leq \inf_{\theta_0, \theta \geq 0} \left( f(\theta)e^{-\theta \lambda'} \right)^n.
\]

Let \( \lambda' > E\eta \). Then, letting
\[
F(\theta) = f(\theta)e^{-\theta \lambda'} = E \left( e^{\theta(\eta - \lambda')} \right),
\]
we get \( F(0) = 1 \),
\[
F'(\theta) = E \left( e^{-\theta(\eta - \lambda')}(\eta - \lambda') \right).
\]
This implies that
\[ F'(0^+) = E(\eta - \lambda') < 0. \]
Thus, there is a \( 0 \leq \theta_1 \leq \theta_0 \) such that \( F(\theta_1) := \gamma' < 1 \). Plugging this into equation (5.3) yields the estimate for our first term.

To estimate the second term, let \( \xi_i = \tau_{2i-1} + \tau_{2i} \). Fix \( c > 1/E\xi \). By a similar argument to that above, we obtain
\[ P(N(T) > cT) \leq \inf_{0 \leq \theta \leq \theta_0} e^{\theta T} g(\theta)^c T. \]
Let
\[ G(\theta) = g(\theta)e^{\theta/c} = E\left(e^{-\theta(\xi - 1/c)}\right). \]
Once again, as above, we obtain
\[ G'(0^+) = -E(\xi - 1/c) \leq 0. \]
Thus, there exists \( \theta_2 \) with \( G(\theta_2) = \gamma'' < 1 \), and so we have our desired estimate. \( \square \)

**Corollary 5.4.** With notation as above,
\[ \limsup_{T \to \infty} \frac{1}{T} \int_0^T X(t)dt \leq \lambda \]
with probability 1 for all \( \lambda > E\eta/E\xi \).

In order to prove this corollary, we need the following technical lemma:

**Lemma 5.5.** Let \( 0 < \gamma < 1 \). Let \( U : \mathbb{R}^+ \to \mathbb{R}^+ \) be such that for all sequences \( \{a_n\}_{n=0}^{\infty} \) with \( \sum_{n=0}^{\infty} \gamma a_n \) convergent,
\[ \limsup_{n \to \infty} U(a_n) \leq c, \]
for some \( c > 0 \). Then
\[ \limsup_{T \to \infty} U(T) \leq c. \]

**Proof:** We proceed by contradiction. Suppose \( \limsup_{T \to \infty} U(T) > c \). Then, there is a sequence of time \( t_n \), \( t_n \to \infty \), such that \( U(t_n) > c \). Take a subsequence \( t_{nk} \), where \( n_k \) is such that \( t_n > k \) for all \( n \geq n_k \). Such a subsequence exists since \( t_n \) diverges. Now, letting \( a_k = t_{nk} \), note that \( a_k > k \), so \( \sum_{k=0}^{\infty} \gamma^{a_k} \) is convergent. So \( \limsup_{k \to \infty} U(a_k) \leq c \). But by definition, \( U(a_k) > c \) for all \( k \). This is a contradiction. \( \square \)

We now proceed with the proof of Corollary. Let \( U(T) = \frac{1}{T} \int_0^T X(t)dt \). Let \( \gamma \) be as in the conclusion of Proposition. Then, for any sequence \( a_n \) we have
\[ \sum_{n=0}^{\infty} P(U(a_n) > \lambda) \leq \sum_{n=0}^{\infty} \gamma a_n. \]
Thus, if \( \sum_{n=0}^{\infty} \gamma a_n \) converges, by the Borel-Cantelli lemma, \( \limsup_{n \to \infty} U(a_n) \leq \lambda \), with probability one. Applying Lemma we have our result. \( \square \)
6. Proofs of main results

We fix the following notation for the rest of this section: fix a \(\delta > 0\), and fix \(q \in \mathbb{Q}\). For \(h \in SL(2, \mathbb{R})\), we define \(V_{\delta,q}(h) = V_{\delta}(hq)\). \(V_{\delta,q}\) is \(K\)-invariant, and thus can be viewed as a function on \(\mathbb{H}^2\). We define the sets \(C_{\delta,l}(q) = \{ z \in \mathbb{H}^2 : V_{q}(z) \leq l \}\). In the rest of this section we work in this \(\mathbb{H}^2\), identifying \(g_r\theta q \in Q\) with \(i.g_r\theta \in \mathbb{H}^2\).

For notational convenience, we drop the subscripts \(\delta\) and \(q\), and write \(V\) and \(C_l\) for \(V_{\delta,q}, C_{\delta,l}(q)\). Furthermore, all distances are measured in \(\mathbb{H}^2\).

Proof of Theorem 2.1: Let \(q \notin C_l\). We want to estimate the measure of the sets \(B'(T,l,q) = \{ \theta : i.g_r\theta \notin C_l, 0 \leq t \leq T \}\).

For technical reasons, we will instead study the sets \(B(T) := B(T,l,q) = \{ \theta : \exists \phi \in B'(T,l,q) \text{ such that } d(i.g_r\theta, i.g_r\phi) \leq 3\tau \}\), where we will specify \(\tau\) shortly. Note that, by definition, and logsmoothness of \(V\), there is some \(a_\tau \geq 1\) such that \(\phi \in B(T,l,q)\) implies that \(V(i.g_r\phi) > l/a_\tau =: l''\) for all \(0 \leq t \leq T\).

Let \(B_{n\tau} = B(n\tau, l,q)\), and \(p_{n\tau} = \nu(B_{n\tau})\). We have

\[
l'' p_{n\tau} \leq \int_{B_{n\tau}} V(i.g_{n\tau}r_\theta)d\theta =: D_{n\tau}.
\]

Our main lemma is as follows

Lemma 6.1. For all \(\delta > 0\), and all \(\tau\) sufficiently large there are constants \(c = c(\delta), b = b(\tau, \delta)\), so that

\[
D_{n\tau} \leq ce^{-(1-\delta)\tau} + b
\]

Proof: Note that by definition \(B_{n\tau}\) is a union of arcs of the form \(I(n\tau)\), and as such, both Lemma 4.6 and Corollary 4.7 apply (we are also using the fact that \(V\) is logsmooth). Fix \(\sigma > 1\), and let \(\tau\) be such that we can apply corollary 4.7. Setting \(c = \sigma c'\hat{c}\) and \(b = \sigma c'b\), we obtain our result. \(\square\)

Let \(\delta' = \delta + \sup_{\tau_0 \leq \tau \leq \tau_0} a(\tau),\)

and set \(l' = l/a\).

Let

\[
\delta' = \delta + \sup_{\tau_0 \leq \tau \leq \tau_0} \frac{1}{\tau} \ln \left( c + \frac{b}{l'} e^{(1-\delta)\tau} \right).
\]
It is easy to check that $\delta' < 1$, and that it is decreasing as a function of $l$. There is some $\tau_0 \leq \tau \leq 2\tau_0$ and $n \in \mathbb{N}$ such that $t = n\tau$. We have

$$p_{n\tau} \leq \frac{V_\delta(q)}{l^\nu} \left(\frac{ce^{-l\delta'}}{b}ight)^n.$$ 

Rewriting this, we obtain

$$p_t \leq \frac{V_\delta(q)}{l^\nu} e^{-l\delta'}.$$ 

\[ \square \]

**Proof of Theorem 2.2:** Let $q \in Q$. Consider the circle of radius $S + T$, $\{i, g_{S^2 + T}r_\theta : 0 \leq \theta < 2\pi\}$. Given $T > 0$, we want to show that the set $B = B_{S,T}(q, l) = \{\theta : i, g_{r_\theta} \notin C_l, S \leq t \leq S + T\}$ has exponentially small measure in $T$ for sufficiently large $l$, $S$, and $T$.

Given $\theta_0 \in B$, let $z_0 = i, g_{r_\theta_0}$, and consider the circle $\{i, g_{r_\theta} g_{\theta_0} : 0 \leq \phi < 2\pi\}$ of radius $T$ around it. By Theorem 2.1, we know that for most (the complement is exponentially small in $T$) directions $\phi$ on this circle, $i, g_{r_\theta} g_{\theta_0} \in C_l$ for some $t < T$.

Our idea is as follows: there is a small neighborhood $U$ of $\theta_0$ such that each geodesic trajectory $\{i, g_{r_\theta} : l_0 \leq \theta \leq l_0 + T\}$, $\theta \in U$ is closely shadowed by a piecewise geodesic of the form $\gamma = \gamma_{\theta_0, \phi}, \theta = \Psi_{S,T}(\phi)$, where

$$\gamma(t) = \begin{cases} i, g_{r_\theta_0} & 0 \leq t \leq S \\ i, g_{t^2 - \theta_0} g_{r_\theta_0} & S < t < S + T \end{cases}$$

By closely shadowed, we mean that $d(\gamma(t), i, g_{r_\theta})$ is small for all $t$. Thus, if $V(\gamma(t)) \leq l$ for some $S \leq S + T$, we have that $V(i, g_{r_\theta}) \leq \tilde{l}$, for some $\tilde{l} > l$.

Now, for all but a small set of $\phi$, $\gamma(t) \in C_l$ for some $S \leq t \leq S + T$. Thus, in a small neighborhood of $\theta_0$, we have a (large-proportioned) collection of angles which are not in $\tilde{B} = B_{S,T}(q, \tilde{l})$.

To make this rigorous, fix $\epsilon, \delta > 0$, and let $A := A(\theta_0) = B_{0,T}(z_0, l) = \{\phi : V(i, g_{r_\theta} g_{\theta_0}) > l, 0 \leq t \leq T\}$, $d$ be the maximum thickness of a hyperbolic triangle. By logsmoothness of $V$, there is a $b \geq 1$ such that $d(z_1, z_2) \leq d$ implies

$$V(z_1) \leq V(z_2).$$

Let $l_1 = bl_0$, where $l_0$ is as in the conclusion of Theorem 2.1. For any $l$, let $\tilde{l} = l/b$.

Setting $\delta'' = \delta' \tilde{l}$, and $\nu' = \tilde{l}^{-1} = a^{-1}l$, where $a = a(\delta)$ is as in Theorem 2.1, yields

$$\nu(A) \leq \frac{V(z_0)}{l'} e^{-l\delta'}. $$

Let $T, S$ be large enough so that we can apply Lemma 2.1 with $\epsilon$, and $t_2 = T$, $t_1 = S$. This gives a neighborhood $U$ of $\theta_0$ in which the proportion of angles

$$\frac{\nu(\Psi(A) \cap U)}{\nu(U)} \leq 4(1 + \epsilon) \sup_{\theta \in [0, 2\pi]} \frac{V(i, g_{r_\theta})}{l'} e^{-l\delta'}. $$

Now, by the thinness of triangles in hyperbolic geometry, given $\theta \in U$, there is a $\phi$ with $\Psi(\phi) = \theta$ such that

$$d(i, g_{r_\theta}, \gamma_{\theta_0, \phi}(t)) \leq d$$

for $t \in [0, S + T]$.

For all $\phi \notin A$, we have

$$\gamma_{\theta_0, \phi}(t_0) \in C_l,$$
for some $S + T > t_0 > S$. Thus,
\[ g_{t_0} \rho q \in C_l, \]
so $\theta \notin B$.

Thus, given any $\theta_0 \in B$, we have produced a neighborhood $U$ s.t.
\[ \frac{\nu(BC \cap U)}{\nu(U)} > 1 - 4(1 + \epsilon) \sup_{\theta \in [0, \pi]} \frac{V(\i.g \rho \theta)}{l'} e^{-(1 - \delta')t}. \]

To complete the proof, we need the following standard lemma (see, for example, [9]):

**Lemma 6.2.** Let $B \subset [0, 2\pi)$ be a measurable set such that for all $b \in B$, there is a $\delta_b > 0$, so that $U_b = [b - \delta_b, b + \delta_b] \subset [0, 2\pi)$ satisfies
\[ \frac{\nu(U_b \cap B)}{\nu(U_b)} < \epsilon. \]

Then
\[ \nu(B) \leq 2\epsilon. \]

Applying the lemma to our set $B$, we obtain
\[ \nu(B) \leq 8(1 + \epsilon) \sup_{\theta \in [0, 2\pi]} \frac{V(\i.g \rho \theta)}{l'} e^{-(1 - \delta')t}. \]

\[ \square \]

**Proof of Theorem 2.3.**

Our strategy is as follows: Given a direction $\theta$, consider the successive departures and returns of the geodesic trajectory $\{i.g \rho \theta\}_{t \geq 0}$ to the compact set $C_l$. Theorem 2.1 implies that the probability any departure is long is small, and thus, we can try and apply Proposition 5.1, to the ‘random variables’ defined by the length of sojourns inside and outside the compact set.

We proceed as follows: Let $d$ be as in the proof of Theorem 2.2. Fix $\delta > 0$, and let $l_0$ be as in Theorem 2.1. Let $l > l_0$ be such that $d(C_l, C_{l_0}) > 2d$, and define $C = C(l) = d(C_l, C_{l_0}) - 2d$. Define $l_0(\theta) = 0$, and set
\[ t_{2n}(\theta) = \inf\{t > t_{2n-1} : \exists \phi \text{ such that } d(i.g \rho \phi, i.g \rho) < d, i.g \rho \phi \in C_{l_0} \} \]
and
\[ t_{2n+1}(\theta) = \inf\{t > t_{2n} : \exists \phi \text{ such that } d(i.g \rho \phi, i.g \rho) < d, i.g \rho \phi \notin C_1 \}, \]
for $n \geq 0$. Define $\tau_i(\theta) = t_i - t_{i-1}$. Now fix $C' > 0$, and define auxiliary functions $\tau'_i$ by
\[ \tau'_{2i} := \begin{cases} 0 & \tau_{2i} \leq C' \\ \tau_{2i} & \tau_{2i} > C' \end{cases} \]
and
\[ \tau'_{2i-1} := \begin{cases} \tau_{2i-1} + \tau_{2i} & \tau_{2i} \leq C' \\ \tau_{2i-1} & \tau_{2i} > C' \end{cases} \]
Define $t'_i := \sum_{j=1}^i \tau'_j$. Note that $\tau_{2i} + \tau_{2i-1} = \tau'_{2i} + \tau'_{2i-1}$, so $t'_{2i} = t_{2i}$.

The $t'_i$’s should be thought of as the entry and departure times of the trajectory $\{i.g \rho \theta\}_{t \geq 0}$ from our compact sets. For technical reasons, they are defined as the
first time when a nearby trajectory leaves a larger compact set \((C_1)\), or re-enters a smaller one \((C_0)\). The \(\tau_{2i}\)'s measure time spent after departing the larger set before returning to the smaller, and the \(\tau_{2i-1}\)'s measure the time spent after returning to the smaller before departing the larger. The auxiliary \(\tau'_i\)'s are defined to exclude short (\(\leq C'\)) sojourns. Thus, if \(t'_{2i} < t < t'_{2i+1}\) we are within distance \(C'\) of the larger compact set, and thus still within a compact set. Thus, for all such \(t\), \(V(i.g_0 r_0) \leq \hat{l} = \hat{l}(C')\), i.e.,

\[
X(t) \geq \chi_{C'}(i.g_0 r_0).
\]

Thus, it suffices to show that for all \(\lambda > 0\) there is a \(C' > 0\) so that \(\nu \left( \frac{t}{\lambda} \int_0^t X(s)ds > \lambda \right)\) decays exponentially in \(T\).

We will apply Proposition 5.1 to \(\{\tau'_i\}_{i \geq 0}\), with \(\Omega = S^1, \mathcal{F}\) the standard \(\sigma\)-algebra, and \(P = \nu\) the Haar measure. Let \(\mathcal{F}_n = \sigma(t'_1, \ldots, t'_n)\) be the \(\sigma\)-algebra generated by \(t'_1, \ldots, t'_n\). Clearly \(\tau'_n\) is \(\mathcal{F}_n\)-measurable. Note that since \(d(C'_1, C_0) = 2d + C\),

\[
\tau'_{2i+1} \geq \tau_{2i+1} = t_{2i+1} - t_{2i} > C
\]

(since \(i.g_{t_{2i}} r_0\) is within distance \(d\) of \(C_0\) and \(i.g_{t_{2i+1}} r_0\) is within distance \(d\) of \(C'_1\)). Thus, condition 2 of the proposition is satisfied, with \(\xi = C\).

It remains to check condition 1. We will show

\[
(6.3) \quad \nu(\tau'_{2i} > t) \mathbb{1}_{\mathcal{F}_{2i-1}} \leq a_1 e^{-a_2 t}
\]

for some \(a_1, a_2 > 0\) and all \(t > C'\).

By definition, \(t'_{i-1}(x)\) is an interval, and thus \(\mathcal{F}_i\) is generated by these intervals. Fix \(\theta\). Let

\[
I_n(\theta) = \{ \phi : t'_{i}(\phi) = t'_i(\theta) \text{ for all } 0 \leq i \leq n \}.
\]

It suffices to show that there are \(a_1, a_2 > 0\) such that, for all \(t > C'\),

\[
(6.4) \quad \nu(\tau_{2n} > t) \mathbb{1}_{\mathcal{F}_{2n-1}}(\theta) \leq a_1 e^{-a_2 t}.
\]

Fix \(\phi \in I_{2n-1}(\theta)\) with \(\tau_{2i}(\phi) > t\). Let \(z_\phi = i.g_{t_{2i-1}} r_\phi\). Note that this point is within distance \(d\) of the boundary of \(C'_0\), thus it is both outside \(C_0\) and still contained within a compact set. Consider the circle of radius \(t\) around \(z_\phi\) and the associated map \(\Psi = \Psi_{t,2i-1}\) to the circle at \(i\).

Let \(A = A(\phi) = \{ \theta : i.g_\theta r_\phi g_{t_{2i-1}} r_\phi \notin C_0 : 0 \leq s \leq t \}\). Since \(z_\phi\) is contained in a compact set, we can pick \(c_1 = c_1(l), c_2 = c_2(l)\) independent of \(\phi\) such that

\[
\nu(A) \leq c_1 e^{-c_2 t},
\]

for all \(t\) sufficiently large. By picking \(C'\) large enough, we can get this to hold for all \(t > C'\). For the rest of this section, let \(t > C'\).

Applying Lemma 1.3 with \(\epsilon = 1\), we obtain a neighborhood \(U\) of \(\phi\), with

\[
\frac{\nu(\Psi_A \cap U)}{\nu(U)} \leq 16c_1 e^{-c_2 t}.
\]

Note that \(U \subset \chi_{2n-1}(\theta)\), since \(\forall 0 \leq t \leq t_{2i-1}, \theta' \in U\),

\[
d(i.g_\theta r_\phi, i.g_{\theta'} r_\phi) < d.
\]

Finally, observe that \(\tau_{2i}(\theta') < t\) for all \(\theta' \in U - \Psi(A(\phi))\), since \(\theta' \notin \Psi(A)\) implies that there is some \(0 \leq s \leq t\) such that \(V(i.g_s r_{\theta'} g_{t_{2i-1}} r_\phi) \leq l_0\), with \(\Psi(\psi) = \theta'\). Thus, since \(d(i.g_\theta r_\phi g_{t_{2i-1}} r_\phi, i.g_{\theta'+t_{2i-1}} r_{\theta'}) \leq d\), we have \(V(i.g_{\theta'+t_{2i-1}} r_{\theta'}) < l\), and thus \(\tau_{2i}(\theta') < t\). Once again applying Lemma 1.3, we obtain equation (6.3), with \(a_1 = 32c_1\) and \(a_2 = c_2\).
Let $C'$ be large enough so that $a_1 e^{-a_2 C'} < 1$. Let $\eta$ be a non-negative function on $S^1$ such that

$$\nu\{ \theta : \eta(\theta) = 0 \} = a_1 e^{-a_2 C'}$$

and

$$\nu\{ \theta : \eta(\theta) > t \} = a_1 e^{-a_2 t}$$

for all $t > C'$. Condition 1 is then clearly satisfied, since $\nu(\tau_{2i} = 0) \geq a_1 e^{-a_2 C'}$, and $\nu(\tau_{2i} > t) \leq a_1 e^{-a_2 t}$ for $t > C'$. Now, note that $E(\eta) = \frac{a_1}{a_2} e^{-a_2 C'}$, and $E(\xi) = C$. Thus, by enlarging $C'$, we can make $E(\eta)/E(\xi)$ arbitrarily small. Precisely for any $\lambda > 0$, take $C'$ so that

$$E(\eta)/E(\xi) \leq \lambda.$$ 

Then, setting $\tilde{l} = \tilde{l}(C')$ and applying Proposition 5.1, we obtain that there is a $\gamma < 1$ so that

$$\nu\{ \theta : \frac{1}{T}|\{0 \leq t \leq T : g_t r_{0q} \notin C_1\}| > \lambda\} \leq \nu\left( \frac{1}{T} \int_0^T X(t) dt > \lambda \right) \leq \gamma T,$$

for all $T$ sufficiently large. \hfill \Box

**Proof of Corollary 2.4:** Applying Corollary 5.4 to $X(t)$, we obtain our result. \hfill \Box

**Acknowledgements:** I would like to thank my advisor, Professor Alex Eskin, for his guidance throughout this project. I would also like to thank Professors Howard Masur, Giovanni Forni, Steven Lalley and Krishna Athreya for valuable discussions, and my colleague Matthew Day for help with hyperbolic trigonometry. Thanks are also due to the anonymous referee, whose detailed comments and suggestions greatly improved this paper.

**References**

[1] A. Avila, S. Gouezel, and J. -C. Yoccoz, *Exponential mixing for the Teichmuller flow*, preprint, 2005. arxiv.org/math.DS/0511614

[2] A. Bufetov, *Decay of Correlations for the Rauzy-Veech-Zorich Induction Map on the Space of Interval Exchange Transformations and the Central Limit Theorem for the Teichmueller Flow on the Moduli Space of Abelian Differentials*, preprint, 2005. arxiv.org/math.DS/0506222

[3] Y. Cheung, *Slowly divergent geodesics in moduli space*, Conform. Geom. Dyn., vol. 8, 167-189, 2004.

[4] Y. Cheung and H. Masur, *A divergent Teichmuller geodesic with uniquely ergodic vertical foliation*, preprint, 2005, arxiv.org/math.DS/0501296

[5] Y. Cheung and H. Masur, *Minimal nonergodic directions on genus 2 translation surfaces*, to appear in Ergodic Theory and Dynamical Systems, 2005. arxiv.org/math.DS/0501286

[6] R. Durrett, *Probability: Theory and Examples, Second Edition*, Duxbury Press, 1996.

[7] A. Eskin and G. Margulis, *Recurrence properties of random walks on finite volume homogeneous manifolds*, in *Random walks and geometry*, 431–444, Walter de Gruyter GmbH and Co. KG, Berlin, 2004.

[8] A. Eskin and H. Masur, *Asymptotic Formulas on Flat Surfaces*, Ergodic Theory and Dynamical Systems, v.21, 443-478, 2001.

[9] L. C. Evans and R. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, 1991.

[10] B. Farb, D. Margalit, *A primer on mapping class groups*, in preparation.

[11] G. Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Annals of Math., v.155, no. 1, 1–103, 2002.

[12] S. Katok, *Fuchsian Groups*, Chicago Lectures in Mathematics, University of Chicago Press, 1992.
20 J. S. ATHREYA

[13] S. Kerckhoff, H. Masur, and J. Smillie, *Ergodicity of billiard flows and quadratic differentials*, Annals of Mathematics, v. 124, no. 2, 293-311, 1986.

[14] D. Y. Kleinbock and G. A. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math., v. 138, 451-494, 1999.

[15] D. Kleinbock and B. Weiss, *Bounded geodesics in moduli space*, Int. Math. Res. Not., no. 30, 1551-1560, 2004.

[16] M. Kontsevich and A. Zorich, *Connected components of the moduli spaces of Abelian differentials with prescribed singularities*, Invent. Math. v. 153, no. 3, 631-678, 2003.

[17] H. Masur, *Interval exchange transformations and measured foliations*, Annals of Mathematics, v. 115, 169-200, 1982.

[18] H. Masur, *Logarithm law for geodesics in moduli space*, Contemporary Mathematics, v. 150, 229-245, 1993.

[19] H. Masur and S. Tabachnikov, *Rational Billiards and Flat Structures*, Handbook of Dynamical Systems, v. 1A, 1015-1089, 2002.

[20] Y. Minsky, and B. Weiss, *Non-divergence of horocyclic flows on moduli space*, J. Reine Angew. Math. v. 552, 131-177, 2002.

[21] S.P. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, 1993

[22] D. Sullivan, *Disjoint Spheres, Approximation by Imaginary Quadratic Numbers, and the Logarithm Law for Geodesics*, Acta Math., v. 149, 215-238, 1982.

[23] K. Strebel, *Quadratic Differentials*, Springer-Verlag, 1984.

[24] S. R. S. Varadhan, *Large Deviations and Applications*, CBMS-NSF Regional Conference Series in Applied Mathematics v. 46, SIAM, 1984.

[25] W. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Annals of Mathematics, v. 115, 201-242, 1982.

[26] W. Veech, *Teichmüller Geodesic Flow*, Annals of Mathematics, v. 124, 441-530, 1986.

[27] A. Zemljakov and A. Katok, *Topological transitivity of billiards in polygons*, Mat. Zametki, v. 19, no. 2, 291-300, 1975.

E-mail address: jayadev@math.uchicago.edu

5734 S. University Avenue, Chicago, IL 60637