LOGARITHM LAWS FOR UNIPOTENT FLOWS, II

JAYADEV S. ATHREYA AND GREGORY A. MARGULIS
(Communicated by Anatole Katok)

Abstract. We prove analogs of the logarithm laws of Sullivan and Kleinbock-Margulis in the context of unipotent flows. In particular, we prove results for horospherical actions on homogeneous spaces $G/\Gamma$.

1. INTRODUCTION

Homogeneous dynamics is the study of the action of subgroups $H \subset G$ acting on homogeneous spaces $X = G/\Gamma$, where $G$ is a semisimple Lie group and $\Gamma$ is a lattice. A particularly important set of examples is the action of unipotent subgroups $H$.

In the setting where $\Gamma$ is a non-uniform lattice, the space $X$ is finite-volume but non-compact, the recurrence properties of unipotent actions have been well-studied. In contrast, the study of excursions to the non-compact part of the space is a relatively recent phenomenon, beginning with our paper [5], which studied one-parameter unipotent actions on the space of lattices. Subsequently, there have been several studies of the subject, including [3, 4, 6, 13].

In the present work, we consider actions of certain horospherical subgroups and one-parameter actions of unipotent flows. Our main results are logarithm laws governing the rate of excursions of large pieces of orbits to the non-compact part of the space.

1.1. Preliminaries and notation. Let $G$ be a connected semisimple Lie group without compact factors and $\Gamma \subset G$ be an irreducible non-uniform, torsion-free, lattice. Let $\mu$ denote the probability measure on $G/\Gamma$ arising from Haar measure on $G$. Let $A_0$ denote a maximal $\mathbb{R}$-diagonalizable subgroup and $\{a_t\}$ a one-parameter subgroup of $A_0^+$, the positive Weyl chamber in $A_0$. Let $\|\cdot\|$ denote a Weyl-group invariant norm on $A_0$. Using the Cartan decomposition $G = KA_0^+K$ we can extend this norm to all of $G$ by

$$\|g\| = \|a^+(g)\|,$$

where

$$g = k_1(g)a^+(g)k_2(g).$$
Following [1], we define a *norm-like pseudometric* on $G$ associated to this norm by
\[ d(g, h) = \| gh^{-1} \|. \]
Note that this is by construction bi-$K$-invariant and right $G$-invariant, so it descends to a metric on $X$, which we also denote $d$.

### 1.2. Horospherical excursions.

We define the *expanding* horospherical subgroup
\[ H := \{ h \in G : a_{-t} ha_t \xrightarrow{t \to +\infty} 1 \} \]
associated to $\{a_t\}$. Let $B \subset H$ be a non-empty, bounded, open subset. For $t > 0$, form the expanding family of subsets of $H$
\[ B_t := a_{\log t} B a_{-\log t}. \]
Let $d_X$ denote the induced distance function on $X = G/\Gamma$. We will drop the subscripts when it is clear on which space we are measuring distances. The main excursion statistic we are concerned with is the asymptotic behavior of the quantities
\[ \beta_t(x) := \sup_{b \in B_t} d_X(bx, x). \]
Dani [8] showed that $Hx$ is dense for all $x$ such that $\{a_{-t}x\}_{t \geq 0}$ is non-divergent. We have
\[ \limsup_{t \to \infty} \beta_t(x) = \infty. \]
We have the following theorem:

**Theorem 1.1.** Suppose the distance $d$ arises from a $Q$-reducing norm with $\|a_t\|_Q = t$. Then
\[ \limsup_{t \to \infty} \frac{\beta_t(x)}{\log t} = 1 \]
for $\mu$-almost every $x$.

**Remark 1.** We define the notion of $Q$-reducing norm and $\|\cdot\|_Q$ in §2.1.

### 1.3. One-parameter subgroups.

We have a similar result for actions of one-parameter subgroups:

**Theorem 1.2.** Let $\{u_t\}_{t \in \mathbb{R}} \subset G$ denote a one-parameter unipotent subgroup. Fix a reference point $y \in X$. Then there is a $c > 0$ (independent of $y$) such that for $\mu$-a.e. $x \in X$,
\[ \limsup_{t \to \infty} \frac{d_X(u_t x, y)}{\log t} = c. \]

We prove this theorem in §5.
1.4. **Diagonalizable actions.** For diagonalizable flows, the statistical properties of cusp excursions have been studied using dynamical techniques beginning with Sullivan [19] (in the context of finite volume hyperbolic manifolds) and later, in the more general context of the actions of diagonalizable subgroups on non-compact finite-volume homogeneous spaces, by Kleinbock-Margulis [14]. The term **logarithm law** was introduced by Sullivan, and more general Borel-Cantelli type results were obtained by Kleinbock-Margulis.

1.5. **Organization.** The paper is organized as follows. In §2, we state the most general version of our main result. In §3, we collect technical results on tori and divergent trajectories required for our proofs. In §4, we use these technical results to prove our main theorem on horospherical actions, as well as related corollaries on hyperbolic surfaces. In §5, we prove our results on one-parameter flows.

### 2. Norms and Excursions

In this section, we state our main result, Theorem 2.1, from which Theorem 1.1 follows, and carefully explain the class of norms for which our results apply. We also state results on the size of the exceptional sets in the context of hyperbolic surfaces in §2.3.

2.1. **Reducing norms.** We define the class of $Q$-reducing norms. If the $R$-rank of $G$ is equal to 1, then there is (up to scaling and conjugation) a unique 1-parameter $R$-diagonalizable subgroup $A_0$. We allow any norm (for example, the norm arising from a $K$-invariant Riemannian metric) on the Lie algebra $a_0$ of $A_0$ normalized so that

$$a_t = e^{t z}, \quad z \in a, \|z\| = 1.$$  

If the $R$-rank of $G$ is at least 2, we assume, by the Arithmeticity Theorem [16, Chapter IX], that $G = G(R)^{\circ}$ and that $\Gamma$ is commensurable to $G(Z)$, where $G$ is a semisimple algebraic $Q$-group. We assume that our maximal $R$-diagonalizable subgroup $A_0$ contains a maximal $Q$-diagonalizable subgroup $A$. That is, let $A = S(R)^{\circ}$ be a maximal $Q$-diagonalizable subgroup, where $S$ is a maximal $Q$-split torus, and let $a$ denote its Lie algebra. Let $T = T(R)^{\circ}$ be the maximal $R$-diagonalizable subgroup containing $A$, where $T$ is a maximal $Q$-split torus containing $S$, and let $a_0$ denote its Lie algebra. Let $M = M(R)^{0}$ denote the $Q$-anisotropic part of $A_0$. We can write $a_t = d_t m_t$, with $d_t \in A^{+}, m_t \in M^{+}$. This decomposition is orthogonal at the level of Killing forms, see Appendix A. Any Weyl-group invariant norm on $A_0$ can be expressed as

$$\|a\| = \sup_{(\rho,V) \in I} c_{\rho} \log \|\rho(a^{-1})\|_{op},$$  

where $I$ is a countable collection of $R$-representations $(\rho,V)$, $c_{\rho}$ are positive scalars, and the operator norm is taken with respect to a natural $K$-invariant inner product on $V$. This can be seen by approximating the unit norm ball for $\|\cdot\|$ by supporting hyperplanes through points which are scalar multiples of elements of the weight lattice. We define the associated projection norm $\|\cdot\|_Q$
by the norm of the orthogonal projection to the $Q$-split torus. We say a norm is $Q$-reducing (or just reducing) if this projection is norm-decreasing.

2.1.1. Examples of reducing norms. We describe three important classes of $Q$-reducing norms.

2.1.1.1. R-rank 1. By construction, all norms on rank 1 groups are reducing.

2.1.1.2. Q-norms. We say a norm $\| \cdot \|$ is a $Q$-norm if it can be written as

$$\| a \| = \sup_{(\rho, V) \in J} c_{\rho} \log \| \rho(a^{-1}) \|_{op},$$

where $J$ is a countable collection of $Q$-representations. By construction, these norms are reducing. Note that our reduction $\| \cdot \|$ is by construction a $Q$-norm.

2.1.1.3. Killing form. By the orthogonality of the decomposition into $Q$-split and anisotropic parts with respect to the Killing form, the pseudometric associated to the norm on $A$ given by the Killing form is $Q$-reducing.

2.1.1.4. Space of lattices. In [5] and [3], we studied cusp excursion for flows on the space of unimodular lattices $X_n = SL(n, \mathbb{R})/SL(n, \mathbb{Z})$ as measured by the function $\alpha_1(\Lambda) = \max_{0 \neq \nu \in \Lambda} 1/\| \nu \|$, where $\| \cdot \|$ is a fixed norm on $\mathbb{R}^n$, and $\Lambda = g\mathbb{Z}^n$ is viewed as a point in $X_n$ via the identification $g\mathbb{Z}^n \rightarrow gSL(n, \mathbb{Z})$. Our current results generalize this setting, as $\log \alpha_1(\Lambda)$ is comparable to $d(\Lambda, \Lambda_0)$, where $d(\cdot, \cdot)$ is the distance arising on $X_n$ from the operator norm $\| \cdot \|_{op}$ on the Weyl chamber $A^+ \sim \mathbb{R}^{n-1}_+$ coming from the standard action of $SL(n, \mathbb{R})$ on $\mathbb{R}^n$, and the operator norm is measured using the norm in the definition of $\alpha_1$.

2.2. Excursion results. We keep notation as in §1. Given $x \in G/\Gamma$, let

$$\omega(x) := \omega^-(x, a_t, d, \Gamma) := \limsup_{t \to +\infty} \frac{d_X(a_{-t}x, x)}{t}$$

denote the linear divergence rate of the backward geodesic trajectory of $x$. Our main theorem is

**Theorem 2.1.**

$$\limsup_{t \to -\infty} \frac{\beta_t(x)}{\log t} \leq 1 + \omega(x).$$

If $\{a_{-t}x\}_{t \geq 0}$ is non-divergent, then

$$\limsup_{t \to -\infty} \frac{\beta_t(x)}{\log t} \geq 1.$$

We will prove this theorem in §4. Combining this result with the Kleinbock-Margulis logarithm law [14, Theorem 1.7], which implies that $\omega^-(x) = 0$ for $\mu$-almost every $x \in G$, we obtain Theorem 1.1.
2.2.1. **Strategy of proof.** To understand the excursion of the piece of horospherical orbit

\[ B_t x = a_{\log t} B a_{-\log t} x, \]

pull back by \( a_{-\log t} \) to obtain the piece of \( H \)-orbit of \( a_{-\log t} x \)

\[ B a_{-\log t} x. \]

Since \( H \) is horospherical, it encodes all possible forward (that is, \( t > 0 \)) \( a_{\log t} \)-behaviors. In particular, it will contain points which diverge at a maximal (linear) rate under \( a_{\log t} \), which we have normalized to be 1. See §3 for the proof and the fact that it is achieved on a dense set of points. This gives (2.2). If the point \( x \) itself has deep (linear) cusp excursions under \( a_{-\log t} \), then we could go even further into the cusp, so our upper bound (2.1) depends on the base point \( x \). We can summarize our discussion as follows: the excursion rate of a horospherical orbit of \( x \) depends on the maximal rate on \( X \) of forward divergence under \( a_{t} \) and the backward divergence rate under \( a_{t} \) of the basepoint \( x \). Similar arguments were used in [6].

2.3. **Hyperbolic surfaces.** Specializing to \( G = SL(2, \mathbb{R}) \) with \( H = \{ h_s \}_{s \in \mathbb{R}} \), where

\[ h_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \]

we have that

\[ a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}. \]

We take \( B = \{ h_s \}_{s \in (0,1)} \), and so \( B_t = \{ h_s \}_{s \in (0,t)} \) \( (a_t h_s a_{-t} = h_{se^t}) \), and obtain a sharp result for the horocycle flow on the unit tangent bundle of a general non-compact finite volume hyperbolic surface. Let \( \Gamma \subset SL(2, \mathbb{R}) \) be a non-uniform lattice. Let \( d \) denote the distance on the hyperbolic surface \( S = \mathbb{H}^2 / \Gamma \) (\( \mathbb{H}^2 \) denotes the upper-half plane with constant curvature \( -1 \)), and let \( p : M \rightarrow S \) be the natural projection from \( M = SL(2, \mathbb{R}) / \Gamma \).

**Corollary 2.2.** [6, Corollary 12] Let \( H = \{ h_s \}_{s \in \mathbb{R}} \). Fix \( y \in S \). Then for all \( x \in S \), almost all \( \tilde{x} \in p^{-1}(x) \),

\[ \limsup_{s \to \infty} \frac{d(p(h_s \tilde{x}), y)}{\log s} = 1. \]  

Moreover, for all \( \tilde{x} \in M \) such that \( H \tilde{x} \) is not closed,

\[ \limsup_{s \to \infty} \frac{d(p(h_s \tilde{x}), y)}{\log s} \geq 1. \]

The following proposition shows that while (2.5) holds for almost every point, the inequality in (2.6) is strict for a (topologically) large set of points:

**Proposition 2.3.** Let \( \Gamma \subset SL(2, \mathbb{R}) \) be a non-uniform lattice and \( H = \{ h_s \}_{s \in \mathbb{R}} \) be as in (2.3). For any \( y \in \mathbb{H}^2 / \Gamma \), the set

\[ E = \left\{ x \in SL(2, \mathbb{R}) / \Gamma : \limsup_{t \to \infty} \frac{d(p(h_t x), y)}{\log s} = 2 \right\}. \]
contains a dense set of second Baire category.

Note that in the metric on $H^2$, $d(p(h_s), i) = 2 \log s$, where by abuse of notation, $p : SL(2, \mathbb{R}) \rightarrow H^2 = SO(2) \setminus SL(2, \mathbb{R})$ is the projection $p(g) = SO(2)g$,
so 2 is the maximum value this limsup can attain. In fact, for any sequence $r_n \rightarrow \infty$ in $SL(2, \mathbb{R})$,
$$d(p(r_n), i) \approx 2 \log |r_n|$$
($\approx$ means the ratio goes to 1), where $|g|$ is the supremum of the matrix entries of $g$. By (2.1), the set $B$ must consist of trajectories which diverge at rate 1 under $a_{-t}$, that is, they must satisfy
$$\limsup_{t \rightarrow \infty} d(a_{-t} \tilde{x}, y) = 1.$$
show the result for norms arising from such a representation \((\rho, V)\). Passing to a finite index subgroup if necessary, we have that \(\rho(\Gamma) \subset GL(n, \mathbb{Z})\), where we identify \(V\) with \(\mathbb{R}^n\) by choosing an orthonormal (with respect to a \(\rho(K)\)-invariant inner product) basis of eigenvectors \(v_1, \ldots, v_n\) for elements of \(a \in A^+\). We order the eigenvectors so that \(v_1\) is the lowest weight,

\[
\rho(a^{-1})v_1 = \chi(a^{-1})v_1,
\]

where \(\chi\) is a \(G\)-character of \(A^+\). That is, \(v_1\) is the vector achieving

\[
\|\rho(a^{-1})\|_\text{op} = \|\rho(a^{-1})v_1\|_V.
\]

Here \(\|\cdot\|_V\) denotes the norm on \(V\) arising from the inner product.

We will estimate \(d(a_t x, x_0)\), where \(x_0\) is the identity coset and \(x = g \Gamma\) with \(g \in G(\mathbb{Q})\). That is, we need to estimate

\[
\min_{\gamma \in \Gamma} \|\rho((a_t g \gamma)^{-1})\|_\text{op}.
\]

Writing \(a_t = d_t m_t\), with \(d_t \in A^+\), we see that since \(\Gamma\) preserves the lattice \(\Lambda\) generated by \(v_1, \ldots, v_n\), and \(g \in G(\mathbb{Q})\) has \(\rho(g) \in GL_n(\mathbb{Q})\), we have that for any \(\gamma \in \Gamma\), there is a \(v \in \Lambda\) so that \(g \gamma v\) has a non-zero (and thus bounded below in terms of the denominators of \(\rho(g)\)) projection to \(\mathbb{R}v_1 \subset V\). This is getting contracted by \(d_t\) at a maximal rate (which we have normalized to be \(e^{-t}\)). This rate of contraction is a \(Q\)-character, and since \(Q\)-characters are trivial on \(M\) by definition, we have that \(a_t = d_t m_t\) also contracts at rate \(e^{-t}\), yielding our result.

3.2. \(\mathbb{R}\)-rank 1. If the \(\mathbb{R}\)-rank of \(G\) is 1, we apply standard reduction theory [10] and the density of orbits of parabolic subgroups ([18, Lemma 8.5]) to obtain a dense set of points diverging under \(a_t\) at rate 1. See also [8, 20] for more details on divergent trajectories.

4. Horospherical actions

The proof of Theorem 2.1 splits naturally into an upper and lower bound:

4.1. Lower bound.

**Lemma 4.1.** For all \(x \in G/\Gamma\) with \(\{a_{-t} x\}_{t \geq 0}\) non-divergent,

\[
(4.1) \quad \limsup_{t \to \infty} \frac{\beta_t(x)}{\log t} \geq 1.
\]

*Proof:* Given the piece of orbit \(B_{e^T} x\), we want to show that it has moved depth \(t\) into the cusp. Write \(B_{e^T} x = a_t B a_{-t} x\). If \(\{a_{-t} x\}\) is non-divergent, we can take some \(T\) so that \(a_{-T} x\) is in a compact set. Using the fact that forward divergent \(a_t\)-trajectories are dense, we can find a divergent trajectory (moving at rate 1) in a ‘thickening’ of the orbit \(B a_{-t} x\) in the directions transverse to \(H\). Since \(a_t\) does not expand the directions transverse to \(H\), the divergent trajectory (which will be approximately depth \(T\) into the cusp after applying \(a_T\)) will be near \(B_{e^T} x\), so there is some \(\hat{h} \in B_{e^T}\) with \(hx\) almost depth \(T\) into the cusp, as desired. To make this argument precise, we need to use the following lemma.
**Lemma 4.2.** Let $C \subset G/\Gamma$ be compact with non-empty interior, and $\epsilon, \phi > 0$. Then there is a $T_{c, \epsilon, \phi}$ such that

$$\{ x : d(a_t x, x) > (1 - \phi)t - T_{c, \epsilon, \phi} \text{ for all } t > 0 \}$$

is $\epsilon$-dense in $C$.

*Proof of Lemma 4.2.* Note that by Proposition 3.1,

$$\{ x \in G/\Gamma : \exists T(x) \text{ such that } d(a_t x, x) > (1 - \phi)t - T(x) \text{ for all } t > 0 \}$$

is dense in $G/\Gamma$.

Now let $\epsilon > 0$, $C \subset G/\Gamma$ compact. Let $\{B(x, \epsilon)\}_{x \in C}$ be the cover of $C$ by open $\epsilon$-metric balls. Since $C$ is compact, we can take a finite subcover $\{D_1, D_2, \ldots, D_n\}$, where each $D_i = B(x_i, \epsilon)$. For $1 \leq i \leq n$, there is an $x_i \in D_i$, $T(x_i) > 0$, such that

$$d(a_t x_i, y) > (1 - \phi)t - T(x_i).$$

Let $T_{c, \epsilon, \phi} = \max_{1 \leq i \leq n} T(x_i)$. Now for all $x \in C$ there is an $x_i$ such that $d(x_i, x) < \epsilon$, and for all $1 \leq i \leq n$, $d(a_t x_i, y) > (1 - \phi)t - T_{c, \epsilon, \phi}$, so we have our result. \qed

*Proof of Lemma 4.1.* Let $H^{-0}$ be the subgroup associated to the neutral/stable directions for $a_t$ ($t > 0$). Let $x \in G/\Gamma$ be such that $a_{-t}x$ is non-divergent. Thus, there is a compact $C'' \subset G/\Gamma$ with non-empty interior and $t_n \to \infty$ so that $a_{-t_n} x \in C''$ for all $n$. For $G'$ a subgroup of $G$, $g_0 \in G'$, $r > 0$, we define

$$B_{G'}(g_0, r) := \{ g \in G' : d_{G'}(g_0, g) < r \}.$$

Let $\epsilon_1$ be such that for all $\epsilon < \epsilon_1$, there are $\epsilon^+, \epsilon^-$,

$$B_G(\epsilon) = B_{H^{-0}}(\epsilon^-) B_H(\epsilon^+).$$

Let $C' = \overline{BC''}$. Let $b_0 \in B$, $\epsilon_0 > 0$ such that $B_{H^{-0}}(b_0, \epsilon_0) = B_H(\epsilon_0) b_0 \subset B$. There is $0 < \epsilon < \epsilon_1$ and an $\epsilon'$ so that (perhaps shrinking $\epsilon_0$) we can write

$$B_{G'}(\epsilon) = B_{H^{-0}}(\epsilon') B_{H}(\epsilon_0).$$

Let $C = \overline{B_G(\epsilon) C'}$. Now $b_0 x_n \in C'$, so $B_G(\epsilon) b_0 x_n \in C$. Shrinking $\epsilon$ if necessary, we have

$$B_G(\epsilon)b_0 x_n = B(b_0 x_n, \epsilon).$$

Fix $\phi > 0$, and let $T = T_{c, \epsilon, \phi}$. There is an $x_n' \in B(b_0 x_n, \epsilon)$ so that

$$d(a_{t_n} x_n', x_n') > (1 - \phi)t_n - T.$$

We can write $x_n' = h^{-1} b_0 x_n$ for $h^{-1} \in B_{H^{-0}}(\epsilon')$. Now we have

$$a_{t_n} x_n' = h^{-1} b_n x,$$

where

$$h_n = a_{t_n} h^{-1} a_{-t_n} \in B_{H^{-0}}(\epsilon')$$

and

$$b_n = a_{t_n} b_0 a_{-t_n} \in B_{\epsilon_0}.$$

Thus, we have

$$d(b_n x, y) \geq d(a_{t_n} x_n', x_n') - \epsilon \geq (1 - \phi)t_n - T - \epsilon,$$
so
\[ \lim_{n \to \infty} \frac{d(b_n x, x)}{t_n} \geq (1 - \phi) \]
(note that since \( x'_n \) varies in a compact set, it does not matter in the limit whether we measure distance from \( x \) or \( x'_n \)). Thus,
\[ \lim_{n \to \infty} \frac{\beta_{e^{t_n}}(x)}{t_n} \geq 1 - \phi, \]
which, since \( \phi > 0 \) was arbitrary, yields Lemma 4.1.

4.2. **Upper bound.**

**Lemma 4.3.** For all \( x \in G/\Gamma \),
\[ (4.2) \quad \limsup_{t \to \infty} \sup_{h \in B} \frac{d(h x, y)}{\log t} \leq 1 + \omega(x). \]

**Proof.** Let \( \epsilon > 0 \). By the definition of \( \omega := \omega(x) \), and the boundedness of \( B \) for all \( t \) sufficiently large, for all \( b \in B \),
\[ d(b a_{-\log t} x, x) < (\omega + \epsilon) \log t. \]

By definition
\[ d(a_{\log t} b a_{-\log t}, b a_{-\log t}) \leq \log t. \]

Combining these two inequalities, and using the triangle inequality, we have, for all \( b \in B \) and \( t \) sufficiently large,
\[ d(a_{\log t} b a_{-\log t} x, x) < (\omega + 1 + \epsilon) \log t. \]

Since \( \epsilon \) was arbitrary, we have our result.

**Proof of Theorem 2.1.** Combine Lemmas 4.3 and 4.1.

4.3. **Hyperbolic geometry.** In this subsection we prove Corollary 2.2 and Proposition 2.3.

**Proof of Corollary 2.2.** Apply Theorem 2.1 to \( G = SL(2, \mathbb{R}) \), with \( H, B, \) and \( \{a_i\} \) as in §2.3. Note that the Riemannian metric on \( H^2 = K \backslash G \) is coarsely isometric to the normlike metric induced by the norm it induces on \( A \).

**Proof of Proposition 2.3.** We need the following lemma, which exploits properties of divergent geodesic trajectories:

**Lemma 4.4.** Let \( F \subset SL(2, \mathbb{R})/\Gamma \) be a non-empty open set, and let \( y \in H^2 / \Gamma \). There is a \( C = C(F) \) such that for all \( T > 0 \) there is \( z \in F, t > T \) such that
\[ (4.3) \quad d(p(h_t z), y) > 2 \log t - C. \]

**Proof.** By the density of divergent geodesic trajectories there is a \( c > 0 \) and a \( z \in F \) such that \( d(p(g_s z), y) > s - c \) for all \( s > 0 \), where here and for the rest of this section we use the notation
\[ g_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}. \]
Fix lifts $y_0 \in \mathbb{H}^2$ of $y$ and $z_0 \in SL(2, \mathbb{R})$ of $z$ to fundamental domains for $\Gamma$ respectively. More precisely, suppose without loss of generality $z_0$ is $i$ with the upward pointing tangent vector, i.e., $z_0 = e \in SL(2, \mathbb{R})$. Then
\[ p(g_s z_0) = p(g_s) = SO(2)g_s, \]
and if $v_s = r_{\theta_s} \in SO(2)$ is a unit tangent vector (based at $i = p(z_0)$) determining a horocycle connecting $e^{i} = p(g_s)$ and $i$, $v_s$ can be chosen to approach the upward pointing tangent vector as $s \to \infty$, that is, $\theta_s \to 0$. In addition, if $t = t_s$ is the time it takes for the horocycle to reach $e^{i}$, we have
\[ SO(2)g_s = SO(2)h_t r_{\theta_s}, \]
that is, there is a $\theta'_s$ such that
\[ h_t = r_{\theta'_s} g_s r_{\theta_s}, \]
(this is simply the Cartan (or KAK) decomposition). Direct calculation shows
\[ s \sim 2 \log t_s, \]
that is, the ratio tends to 1 as $t_s \to \infty$. Thus, for $s \gg 0$, $r_{\theta_s} z \in A$, and
\[ d(p(h_t r_{\theta_s} z), y) = d(p(g_s z), y) > s - c > 2 \log t_s - C, \]
for some possibly larger $C$.

To complete the proof of the proposition, define $f_T : SL(2, \mathbb{R})/\Gamma \to [0, 2)$ by
\[ f_T(x) = \sup_{2 \leq t \leq T} \frac{d(p(h_t x), y)}{\log t}. \]
The function $f_T(x)$ is increasing in $T$, and bounded, so
\[ f_\infty(x) = \lim_{T \to \infty} f_T(x) \]
is well defined. The $f_T$’s are continuous in $x$ for $T < \infty$, but $f_\infty$ is not. We have
\[ E = \{ x : f_\infty(x) = 2 \} = \bigcap_{k=1}^{\infty} \bigcup_{n=0}^{\infty} \left\{ x : f_n(x) > 2 - \frac{1}{k} \right\}. \]
For each $k$,
\[ \bigcup_{n=0}^{\infty} \left\{ x : f_n(x) > 2 - \frac{1}{k} \right\} \]
is dense by Lemma 4.4, and open by the continuity of $f_n$. Thus $E$ is a countable intersection of open dense sets, as desired.
5. Borel-Cantelli Lemmas

We prove Theorem 1.2, using a generalization of the Borel-Cantelli lemma. The first example of a logarithm law can be derived from the classical Borel-Cantelli lemma (see, for example, [9, Chapter 14]) as follows. Fix \( \lambda > 0 \). Let \( \{Y_n\}_{n=0}^{\infty} \) be independent identically distributed (i.i.d.) exponential random variables with parameter \( \lambda \). That is, for any \( t > 0 \),

\[
P(Y_n > t) = e^{-\lambda t}.
\]

Let \( \{r_n\}_{n=0}^{\infty} \) be a sequence of positive real numbers. Applying the convergence case of the Borel-Cantelli lemma to the sequence of random variables

\[
X_n := \begin{cases} 1 & Y_n > r_n \\ 0 & \text{otherwise,} \end{cases}
\]

implies \( Y_n > r_n \) infinitely often if and only if \( \sum_{n=0}^{\infty} e^{-\lambda r_n} = \infty \). As a corollary, one obtains that almost surely

\[
\limsup_{n \to \infty} \frac{Y_n}{\log n} = \frac{1}{\lambda}.
\]

To prove Theorem 1.2, we use the following (relatively standard) generalization of the Borel-Cantelli lemma to weakly dependent sequences.

**Proposition 5.1.** Let \((S, \Omega, P)\) be a probability space. Let \( X_n : S \to \{0, 1\} \) be a sequence of \( 0-1 \) random variables on \( S \), with \( P(X_n = 1) = p_n \). Also define \( p_{i,j} := P(X_i X_j = 1) \). Suppose

1. \( \sum_{n=1}^{\infty} p_n = \infty \).
2. There is a function \( \psi(m) \) such that for all \( m > 0 \),

\[
\sup_n |p_{n,n+m} - p_n p_{n+m}| \leq \psi(m).
\]
3. \( \lim_{n \to \infty} \sum_{m=1}^{n} \frac{\psi(m)(n-m)}{(\sum_{i=1}^{n} p_i)^2} = 0. \)

Then

\[
P\left( \sum_{n=0}^{\infty} X_n = \infty \right) = 1.
\]

**Proof.** Given measurable \( X : S \to \mathbb{R} \), we write

\[
E(X) := \int_S X dP \quad \text{and} \quad V(X) := E(X^2) - E(X)^2
\]

for the expectation and the variance of \( X \) respectively. Let \( J_n = \sum_{i=1}^{n} X_i \) and

\[
Y_n = \frac{J_n}{\sum_{i=1}^{n} p_i} = \frac{J_n}{E(J_n)}.
\]

We will show that \( Y_n \) converges in probability to 1, that is, for any \( \epsilon > 0 \),

\[
P(|Y_n - 1| > \epsilon) \to 0.
\]
This implies the existence of a subsequence \( n_k \) such that \( Y_{n_k} \to 1 \) with probability 1, and thus, that

\[ J_{n_k} \to \infty. \]

Since \( E(Y_n) = 1 \), it suffices to show that \( V(Y_n) \to 0 \), that is, that \( Y_n \to 1 \) in \( L^2 \). Now,

\[ V(Y_n) = \frac{V(J_n)}{E(J_n)^2}. \]

We have

\[ V(J_n) = V\left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} V(X_i) + \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j), \]

where \( \text{Cov}(X_i, X_j) := |p_{i,j} - p_i p_j| \) is the covariance of \( X_i \) and \( X_j \). By property (2),

\[ \text{Cov}(X_i, X_j) \leq \psi(|j-i|), \]

so

\[ V(J_n) \leq \sum_{i=1}^{n} p_i + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} \psi(j-i) = \sum_{i=1}^{n} p_i + 2 \sum_{i=1}^{n} \sum_{m=1}^{n-i} \psi(m) = \sum_{i=1}^{n} p_i + 2 \sum_{m=1}^{n} \psi(m)(n-m). \]  

(5.1)

Dividing by \( \left( \sum_{i=1}^{n} p_i \right)^2 \), the two right-hand terms go to zero (by properties (1) and (3) respectively).

We apply this result to group actions on homogeneous spaces. Given \( y \in G/\Gamma \), we define a sequence of functions \( Y_n : G/\Gamma \to \mathbb{R}^+ \) by \( Y_n(x) = d(u_n x, y) \). Given a sequence of numbers \( \{r_n\}_{n \in \mathbb{N}} \), we set

\[ X_n(x) := \begin{cases} 1 & Y_n(x) > r_n \\ 0 & \text{otherwise.} \end{cases} \]

In order to apply Proposition 5.1 to our context, we must estimate two quantities, the measure decay rate of the cusp neighborhoods \( \mu(x : d(x, y) > t) \) and the covariances for the random variables \( X_n \).

The first estimate follows from Kleinbock-Margulis [14, Proposition 5.1], which yields the existence of a \( k > 0 \) so that, in our notation,

\[ C_1 e^{-kr_n} \leq p_n = \mu(x : X_n(x) = 1) \leq C_2 e^{-kr_n}. \]

In order to estimate the covariances, we control the matrix coefficients of the sequence \( \{u_k\} \) under the regular representation of \( G \) on \( G/\Gamma \) via the following:

**Proposition 5.2.** [14, Lemma 4.2 and Corollary 3.5] There are constants \( C > 0 \), \( 0 < \beta \) such that for all \( n, m \in \mathbb{N} \),

\[ |p_{n,n+m} - p_n p_{n+m}| \leq C p_n p_{n+m} m^{-\beta}, \]

where \( p_{i,j} = \mu(x : X_i(x) X_j(x) = 1) \).
Remark 2. If we were able to obtain \( \beta \geq 2 \), we would in fact be able to prove \( \alpha = 1 \) in the statement of Theorem 1.2 following Proposition 4.1 in [14]. However, the proof in [14] does not yield \( \beta \geq 2 \).

We will not prove Proposition 5.2 in this paper, instead referring the interested reader to the appropriate sections of [14].

Proof of Theorem 1.2. Let \( r_n > \frac{1}{k} \log n \). Then \( p_n \) is summable, so for almost all \( x \), \( X_n = 1 \) only finitely often, yielding our upper bound. For our lower bound we apply Proposition 5.1 to our sequence \( X_n \), with \( \psi(m) = m^{-\beta} \). It is a simple calculation that for any \( \gamma < \beta/2 \), setting \( r_n = \frac{\gamma}{k} \log n \) will yield

\[
\lim_{n \to \infty} \frac{\sum_{m=1}^{n} \psi(m)(n-m)}{(\sum_{i=1}^{n} p_i)^2} = 0.
\]

Using Proposition 5.1, we have, for \( \mu \)-a.e. \( x \),

\[
\frac{\beta}{2k} \leq \limsup_{n \to \infty} \frac{d(u_n x, y)}{\log n} \leq \frac{1}{k}.
\]

By continuity, if \( |t - n| < 1 \), there is a universal \( C \) so that

\[|d(u_n x, y) - d(u_t x, y)| < C,\]

so

\[
\limsup_{n \to \infty} \frac{d(u_n x, y)}{\log n} = \limsup_{t \to \infty} \frac{d(u_t x, y)}{\log t},
\]

and so we have, as desired,

\[
\frac{\beta}{2k} \leq \limsup_{t \to \infty} \frac{d(u_t x, y)}{\log t} \leq \frac{1}{k}.
\]

Finally, note that

\[
\limsup_{t \to \infty} \frac{d(u_t x, y)}{\log t}
\]

is a measurable \( u_t \)-invariant function on \( G/T \). Thus, if the \( u_t \)-action is ergodic, it must be constant almost everywhere. If \( u_t \) is not ergodic, it must act trivially in some factor of \( G \) by the Moore ergodicity theorem [17], and thus we can reduce to the ergodic case.

Our proof shows that \( c = \frac{\alpha}{k} \), where \( 0 < \alpha < 1 \). It would be very interesting to find examples of unipotent subgroups where \( \alpha < 1 \), though we suspect that such subgroups do not exist.

Appendix A. Decomposition of tori

Lemma A.1. Let \( G \) be a semisimple algebraic \( \mathbb{Q} \)-group. Let \( S \) denote a maximal \( \mathbb{Q} \)-split torus, and let \( T \) be the maximal \( \mathbb{R} \)-split torus containing \( S \). Up to isogeny, \( T = S \times M \), where \( M \) is a \( \mathbb{Q} \)-anisotropic torus. Moreover, at the level of Lie algebras, this decomposition is orthogonal with respect to the Killing form.
The decomposition is a standard fact about tori, see, for example, [12, Section 34.3, page 219] or [7, Proposition 8.15, page 219]. However, the fact that the decomposition is orthogonal with respect to the Killing form seems to not be written up in the literature.

Proof. Let $\mathbb{G}_m$ denote the multiplicative group. Let
\begin{align}
X_* (\mathbb{T}) &= \text{hom}(\mathbb{G}_m, \mathbb{T}) \\
X^* (\mathbb{T}) &= \text{hom}(\mathbb{T}, \mathbb{G}_m)
\end{align}
be the sets of one-parameter subgroups and characters of $\mathbb{T}$ respectively. Note that since $\mathbb{T}$ is $\mathbb{R}$-split, these homomorphisms will be defined over $\mathbb{R}$.

There is a natural pairing between $X^* (\mathbb{T})$ and $X_* (\mathbb{T})$ given as follows: let $x \in X_* (\mathbb{T})$, $y \in X^* (\mathbb{T})$, then $y \cdot x \in \text{hom}_\mathbb{Q} (\mathbb{G}_m, \mathbb{G}_m)$, so $y \cdot x (\alpha) = \alpha^n$ for some $n \in \mathbb{Z}$. The map $(x, y) \mapsto n$ yields a pairing
\[ X_* \times X^* \to \mathbb{Z}. \]

We denote the image of $(x, y)$ under this map by $\langle x, y \rangle$. We can write $\mathbb{T}$ as a product of one-parameter subgroups. That is, there is a basis $x_1, \ldots, x_m$ of $X_* (\mathbb{T})$ so that
\[ \mathbb{T} = \prod_{i=1}^m \text{Im}(x_i). \]

For $1 \leq i \leq m$, let $y_i \in X^* (\mathbb{T})$ be given by $y_i |_{\text{Im}(x_i)} = e$, where $e$ is the identity element, and trivial on $\text{Im}(x_j)$ for $j \neq i$. We thus obtain a dual basis $y_1, \ldots, y_m$ of $X^* (\mathbb{T})$ with $\langle x_i, y_j \rangle = \phi_{ij}$.

Let $K$ be the splitting field of $\mathbb{T}$. $\text{Gal}(K/\mathbb{Q})$ acts on $X_* (\mathbb{T})$ and $X^* (\mathbb{T})$ via the action on the coefficients of the regular functions used to define the homomorphisms. The actions are dual with respect to the pairing.

We have that the $\mathbb{Q}$-split part of $\mathbb{T}$ corresponds to the fixed points of this action on $X_*$. Similarly, the $\mathbb{Q}$-anisotropic part is the kernel of the fixed points of the action on $X^*$ (that is, the kernel of the $\mathbb{Q}$-characters).

In particular, the action of $\text{Gal}(K/\mathbb{Q})$ on the $\mathbb{Q}$-anisotropic part of $X_* (\mathbb{T})$ is fixed-point free. We now recall how to define the Killing form on $X_* (\mathbb{T})$. Let $x, y \in X_* (\mathbb{T})$. We have the associated differentials
\[ dx, dy : \text{Lie}(\mathbb{G}_m) \to \text{Lie}(\mathbb{T}). \]

We choose a distinguished element $I \in \text{Lie}(\mathbb{G}_m)$, and define
\begin{align}
B(x, y) := &-\text{tr} (\text{ad}(dx(1)), \text{ad}(dy(1))) \end{align}

$B(x, y)$ is $\text{Gal}(K/\mathbb{Q})$ invariant. Moreover, if $x$ is in the $\mathbb{Q}$-split part and $y$ in the anisotropic part, we have $B(x, y) = 0$.

The first assertion follows from the definition of the action, and the second can be seen as follows: since
\[ \sum \sigma y = 0, \]
we have $B(x, \sum \sigma y) = 0$. Here, the sums are taken over all non-trivial $\sigma \in \text{Gal}(K/Q)$. Since $\text{Gal}(K/Q)$ acts trivially on the split part, we have

$$B(x, \sum \sigma y) = \sum B(x, \sigma y) = \sum B(\sigma x, \sigma y).$$

Since $B$ is invariant $B(\sigma x, \sigma y) = B(x, y)$, so $B(x, y) = 0$, as desired. \hfill \Box

Acknowledgments. We thank Herbert Abels, Yitwah Cheung, Manfred Einsiedler, Alex Eskin, Roger Howe, Dmitry Kleinbock, Enrico Leuzinger, Yair Minsky, Nimish Shah, Barak Weiss and Dave Witte Morris for useful discussions, and the anonymous referee for their helpful comments. J.S.A. thanks Yale University for its hospitality in the 2012-13 academic year and the Mathematical Sciences Research Institute, where he was a member of the program on Geometric and Arithmetic Aspects of Homogeneous Dynamics in Spring 2015.

References

[1] H. Abels and G. Margulis, Coarsely geodesic metrics on reductive groups, in Modern Dynamical Systems and Applications, Cambridge Univ. Press, Cambridge, 2004, 163–183.
[2] H. Abels and G. Margulis, preprint.
[3] J. S. Athreya, Cusp excursions on parameter spaces, J. London Math. Soc., 87 (2013), 741–765.
[4] J. S. Athreya and Y. Cheung, A Poincaré section for horocycle flow on the space of lattices, Int. Math. Res. Notices, 2014, no. 10, 2643–2690.
[5] J. S. Athreya and G. Margulis, Logarithm laws for unipotent flows, I, Journal of Modern Dynamics, 3 (2009), 359–378.
[6] J. S. Athreya and F. Paulin, Logarithm laws for strong unstable foliations in negative curvature and non-Archimedean Diophantine approximation, Groups, Geometry, and Dynamics, 8 (2014), 285–309.
[7] A. Borel, Linear Algebraic Groups, 2nd enlarged edition, Springer Verlag, New York, 1991.
[8] S. G. Dani, Divergent trajectories of flows on homogeneous spaces and Diophantine approximation, J. Reine Angew. Math., 359 (1985), 55–89.
[9] W. Feller, An Introduction to Probability Theory and Its Applications, 1, Wiley, 1957.
[10] H. Garland and M. S. Raghunathan, Fundamental domains for lattices in (R-)rank 1 semi-simple Lie groups, Annals of Math. (2), 92 (1970), 279–326.
[11] J. Humphreys, Introduction to Lie Algebras and Representation Theory, Springer-Verlag, New York, 1972.
[12] J. Humphreys, Linear Algebraic Groups, 2nd printing, Springer-Verlag, New York-Heidelberg, 1975.
[13] D. Kelmer and A. Mohammadi, Logarithm laws for one parameter unipotent flows, Geom. Funct. Anal., 22 (2012), 756–784.
[14] D. Y. Kleinbock and G. A. Margulis, Logarithm laws for flows on homogeneous spaces, Invent. Math., 138 (1999), 451–494.
[15] E. Leuzinger, Geodesic rays in locally symmetric spaces, Differential Geometry and its Applications, 6 (1996), 55–65.
[16] G. A. Margulis, Discrete Subgroups of Semisimple Lie Groups, Springer-Verlag, Berlin-New York, 1991.
[17] C. C. Moore, Ergodicity of flows on homogeneous spaces, Amer. J. Math., 88 (1966), 154–178.
[18] G. D. Mostow, Strong Rigidity of Locally Symmetric Spaces, Annals of Math. Studies, Princeton Univ. Press, 1973.
[19] D. Sullivan, Disjoint spheres, approximation by quadratic numbers and the logarithm law for geodesics, Acta Mathematica, 149 (1982), 215–237.
[20] B. Weiss, Divergent trajectories on noncompact parameter spaces, *Geom. and Funct. Anal.*, **14** (2004), 94–149.

**Jayadev S. Athreya** <jathreya@uw.edu>: Department of Mathematics, University of Washington, Seattle, WA 98195, USA

**Gregory A. Margulis** <margulis@math.yale.edu>: Department of Mathematics, Yale University, New Haven, CT 06520, USA