ISOLATIONS OF GEODESIC PLANES IN HYPERBOLIC 3-MANIFOLDS

AMIR MOHAMMADI AND HEE OH

Abstract. We present a quantitative isolation property of properly immersed geodesic planes in a geometrically finite hyperbolic 3-manifold. Our estimates are polynomial in the shadow constants, tight areas, and densities of the Bowen-Margulis-Sullivan measures of geodesic planes. The proof is based on the quantitative non-concentration property of the Patterson-Sullivan density and the analysis of a Margulis function measuring the distance between two planes.

Contents

1. Introduction 1
2. Notation and preliminaries 7
3. Shadow constants 8
4. Tight area of a geodesic surface 15
5. Averaging operator 17
6. A linear algebra lemma 21
7. Height function $\omega$ 22
8. Averaging operator for the height function 25
9. Construction of a Margulis function 28
10. Quantitative isolation of a closed orbit 34
11. Appendix: Proof of Theorem 1.1 in the compact case 35

References 37

1. Introduction

Let $\mathbb{H}^3$ denote the hyperbolic 3-space, and let $G = \text{PSL}_2(\mathbb{C}) \simeq \text{Isom}^+(\mathbb{H}^3)$. Any oriented complete hyperbolic 3-manifold can be presented as the quotient $M = \Gamma \backslash \mathbb{H}^3$ where $\Gamma$ is a torsion-free discrete subgroup of $G$. A geodesic plane in $M$ is the image of a totally geodesic immersion $\mathbb{H}^2 \to M$. Let $H := \text{PSL}_2(\mathbb{R})$. Via the identification of $X := \Gamma \backslash G$ with the oriented frame bundle $FM$, a geodesic plane is the image of a unique $\text{PSL}(2, \mathbb{R})$-orbit under the base point projection map

$$\pi : X \simeq FM \to M.$$
Moreover a properly immersed geodesic plane in $M$ corresponds to a closed $H$-orbit in $X$.

In this paper, we obtain a quantitative isolation result of a closed $H$-orbit for \textit{geometrically finite} hyperbolic 3-manifolds. We fix a left invariant Riemannian metric on $G$, which projects to the hyperbolic metric on $\mathbb{H}^3$. This induces the distance $d$ on $X$ so that the canonical projection $G \to X$ is a local isometry. We use this Riemannian structure on $G$ to define the volume of a closed $H$-orbit in $X$. For a closed subset $S \subset \Gamma \backslash G$ and $\varepsilon > 0$, $B(S, \varepsilon)$ denotes the $\varepsilon$-neighborhood of $S$.

\textbf{Finite volume case.} We first state the result for the finite volume case. In this case, Ratner [15] and Shah [18] independently showed that every $H$-orbit is either closed or dense in $X$. Moreover, every closed orbit has finite volume and there are only countably many closed $H$-orbits in $X$. Mozes and Shah [13] proved that an infinite sequence of closed $H$-orbits becomes equidistributed in $X$. Our questions concerns the following quantitative isolation property: for a given closed $H$-orbit $Y$ in $X$,

(1) how close can $Y$ approach another closed $H$-orbit?
(2) given $\varepsilon > 0$, what portion of $Y$ enters into in the $\varepsilon$-neighborhood of another closed $H$-orbit?

When $X$ is not compact, the distance between two closed $H$-orbits may be zero, e.g., if they both have cusps going into the same cuspidal end of $X$. To remedy this issue, we use the thick-thin decomposition of $X$.

For $\varepsilon > 0$, we denote by $X_\varepsilon$ the $\varepsilon$-thick part of $X$, i.e.,

$$X_\varepsilon = \{ x \in X : \text{inj}(x) \geq \varepsilon \}$$

where $\text{inj}(x)$ denotes the injectivity radius of $x$.

It turns out that, in the finite volume case, the volume of a closed orbit is the only complexity which measures the quantitative isolation property. The following theorem was proved by Margulis in an unpublished note:\footnote{The notation $A \ll B$ and $A \gg B$ means $A \leq CB$ and $A \geq CB$ for some absolute constant $C > 0$ depending only on $\Gamma$.}

\textbf{Theorem 1.1 (Margulis).} Let $\Gamma$ be a lattice in $G$. Let $Y \neq Z$ be closed $H$-orbits in $X$.

(1) For all $0 < \varepsilon < 1$,

$$d(Y \cap X_\varepsilon, Z) \gg \varepsilon^9 \text{Vol}(Y)^{-3} \text{Vol}(Z)^{-3}.$$  

(2) For any $0 < \varepsilon < 1$,

$$m_Y(Y \cap B(Z, \varepsilon)) \ll \varepsilon^{1/3} \text{Vol}(Z)$$

where $m_Y$ denotes the probability $H$-invariant measure on $Y$.

In both statements, the implied constant depends only on $\Gamma$. 
Remark 1.2. By recent works ([12], [1]), only when $\Gamma$ is an arithmetic lattice, there may be infinitely many closed $H$-orbits. Theorem 1.1 is also proved in [5, Lemma 10.3], based on the effective ergodic theorem which relies on the arithmeticity of $\Gamma$ via uniform spectral gap on closed $H$-orbits. Margulis’s proof does not rely on the arithmeticity and yields a sharper exponent. It is based on the construction of a certain function which measures the distance from a closed orbit, and also satisfies a coarse version of an inequality reminiscent to the one satisfied by super harmonic functions. A similar function appeared first in the work of Eskin, Mozes and Margulis in the study of quantitative Oppenheim conjecture [7], and later in [6], [4] and [8].

Infinite volume case. We now turn to the case of a geometrically finite hyperbolic 3-manifold $M = \Gamma \backslash \mathbb{H}^3$ of infinite volume. We denote by $\text{core } M$ the convex core of $M$, and fix $\varepsilon_X$ to be $\min_{q \in \text{core } M} \text{inj}(q)$ if $\text{core } M$ is compact, and the maximum $\varepsilon_0$ so that the $\varepsilon_0$-thin part of $\text{core } M$ consists of disjoint horoballs, otherwise. We call the $\varepsilon_X$-thick part of $\text{core } M$ the compact core of $M$. We denote by $\text{RF } M$ the renormalized frame bundle and for $\varepsilon > 0$, consider the following $\varepsilon$-thick part of $\text{RF } M$:

$$\text{RF}_\varepsilon M := \text{RF } M \cap X_\varepsilon.$$  

Let $Y = yH$ be a closed $H$-orbit and let $\Delta_Y := \text{Stab}_H(y)$. The associated hyperbolic surface $S_Y := \Delta_Y \backslash \mathbb{H}^2$ is always geometrically finite [14]. We assume $S_Y$ is non-elementary; otherwise, we cannot expect an isolation phenomenon for $Y$, as there may be a continuous family of parallel closed elementary $H$-orbits.

Let $0 < \delta_Y \leq 1$ denote the critical exponent of $S_Y$. We define

$$\delta_Y^\circ = \begin{cases} 
\delta_Y & \text{Y has no cusp} \\
2\delta_Y - 1 & \text{otherwise};
\end{cases}$$

note that $0 < \delta_Y^\circ \leq 1$, and $\delta_Y^\circ = 1$ if and only if $S_Y$ has finite area.

We now state the main result of this paper, which generalizes Theorem 1.1 to all geometrically finite manifolds:

**Theorem 1.4.** Let $M$ be a geometrically finite hyperbolic 3-manifold. Let $Y$ be a non-elementary closed $H$-orbit in $X$, and denote by $\mu_Y$ the probability Bowen-Margulis-Sullivan measure on $Y$.

1. There exists $N > 0$ such that for all $0 < \varepsilon < \varepsilon_X$ and for any non-elementary closed $H$-orbit $Z \neq Y$,

$$d(Y \cap \text{RF}_\varepsilon M, Z) \gg \left(\frac{\nu_{Y,\varepsilon}}{\max\{1, s_Y\} \cdot \text{area}(Z)}\right)^{N/\delta_Y^\circ},$$

where

- $\nu_{Y,\varepsilon} = \min_{y \in Y \cap \text{RF}_\varepsilon M} \text{mu}(B_Y(y, \varepsilon))$ where $B_Y(y, \varepsilon)$ is the $\varepsilon$-ball around $y$ in the induced metric on $Y$.
- $s_Y$ is the shadow constant of $S_Y$.
• area_t Z is the tight area of S_Z relative to M.

(2) There exists N > 0 such that for all 0 < ε < ε_X,

\[ m_Y(Y \cap B(Z, \varepsilon)) \ll (\max\{1, s_Y\})^N \cdot \varepsilon^{\delta_Y/N} \cdot (\text{area}_t Z)^N \]

In both statements, the exponent N and the implied constants depend only on Γ.

We now introduce geometric invariants of closed orbits: the shadow constant and the tight area of a properly immersed geodesic surface. When X has finite volume, it turns out that the shadow constant \( s_Y \) is 1, and the tight area of Z satisfies \( \text{area}_t Z \approx \text{area} Z \); hence Theorem 1.4 recovers Theorem 1.1. For a general geometrically finite case, these notions are new.

**Definition 1.6 (Shadow constant).** Let \( \Lambda Y \subset \partial \mathbb{H}^2 \) denote the limit set of \( \Delta Y \), \( \{\nu_p : p \in \mathbb{H}^2\} \) denote the Patterson-Sullivan density for \( \Delta Y \), and \( B_p(\xi, \varepsilon) \) denote the \( \varepsilon \)-neighborhood of \( \xi \in \partial \mathbb{H}^2 \) with respect to the Gromov metric at \( p \). We define the shadow constant of \( S_Y \) as follows:

\[
\begin{align*}
\tag{1.7}
s_Y := \sup_{\xi \in \Lambda Y, p \in [\Lambda Y, \xi], 0 \leq \varepsilon \leq 1} \frac{\nu_p(B_p(\xi, \varepsilon))}{\varepsilon^{\delta_Y} \nu_p(B_p(\xi, 1))},
\end{align*}
\]

where \( [\Lambda Y, \xi] \) is the union of all geodesics connecting points in \( \Lambda Y \) to \( \xi \).

Observe that the shrinking exponent of \( \nu_p(B_p(\xi, \varepsilon)) \) is given by \( \delta_Y \) rather than by \( \delta_Y \). We show that \( 0 < s_Y < \infty \) in Corollary 3.14; moreover, note that the above definition is an invariant of \( Y \), i.e., it is independent of the choice of \( y \), see Lemma 3.16.

**Definition 1.8 (Tight area of \( S \)).** For a properly immersed geodesic surface \( S \) of \( M \), we define the tight-area of \( S \) relative to \( M \) as follows:

\[
\text{area}_t(S) := \text{area}(S \cap \mathcal{N}(\text{core } M))
\]

where \( \mathcal{N}(\text{core } M) := \{p \in M : d(p, q) \leq \text{inj}(q) \text{ for some } q \in \text{core } M\} \) is the tight neighborhood of core \( M \).

Note that the tight area essentially amounts to the area of the portion of \( S \) which enters either the cuspidal regions of core \( M \), or the 1-neighborhood of the compact core of \( M \). For \( M \) geometrically finite, the flares of \( S \) can enter the cuspidal regions of core \( M \) essentially only in cusp-like shapes, and this geometric feature implies the finiteness of the tight area of \( S \) relative to \( M \) (Lemma 4.6).

**Remark 1.9.**

(1) If \( Y \) is convex cocompact, then \( v_{Y, \varepsilon} \asymp \varepsilon^{1+2\delta_Y} \). In general,

\[
\varepsilon^{1+2\delta_Y} \ll v_{Y, \varepsilon} \ll \varepsilon^{4\delta_Y} - 1
\]

where the implied constants depend only on \( Y \) (cf. [9]). For instance, if \( Y \) has finite volume, \( v_{Y, \varepsilon} \asymp \varepsilon^3 \text{Vol}(Y)^{-1} \) with the implied constant independent of \( Y \).
(2) Our proof yields a more general version of Theorem 1.4(1): Let \( y \in Y_0 \cap X \) and \( z \in Z \). If \( y \notin B_Z(z, \varepsilon) \), then

\[
d(y, z) \gg \left( \frac{m_Y(B_Y(y, \varepsilon))}{\max\{1, s_Y\} \cdot \text{area}_Z} \right)^{N/\delta_Y^n}
\]

where \( Z \) is allowed to be equal to \( Y \).

A geometrically finite hyperbolic 3-manifold \( M \) is called *acylindrical* if its compact core does not have any essential discs or cylinders which are not boundary parallel. Note that this is a topological condition. When \( M \) is convex cocompact acylindrical, there is a uniform positive lower bound for \( \delta_Y = \delta_X^\varepsilon \) for all non-elementary closed \( H \)-orbits \( Y \) \([10]\); therefore the dependence of \( \delta_Y \) can be removed in Theorem 1.4.

Examples of \( X \) with infinitely many closed \( H \)-orbits are provided by the following theorem which can be deduced from \([10], [11], [3]\):\

**Theorem 1.10.** Let \( M_0 \) be an arithmetic hyperbolic 3-manifold with a properly immersed geodesic plane. Any geometrically finite acylindrical hyperbolic 3-manifold \( M \) which covers \( M_0 \) contains infinitely many non-elementary properly immersed geodesic planes.

It is easy to construct examples of \( M \) satisfying the hypothesis of this theorem. For instance, if \( M_0 \) is an arithmetic hyperbolic 3-manifold with a properly imbedded compact geodesic plane \( P \), \( M_0 \) is covered by a geometrically finite acylindrical manifold \( M \) whose convex core is isometric to \( P \).

We discuss some of the main ingredients of the proof of Theorem 1.4. First consider the case when \( X = \Gamma \backslash G \) is compact. Let \( \varepsilon_0 = \varepsilon_X/2 \). Let \( Y \) and \( Z \) be two closed orbits of \( H \) in \( X \). Fixing \( y \in Y \), the set

\[
I_Z(y) := \{ v \in \mathfrak{sl}_2(\mathbb{R}) - \{0\} : \|v\| \leq \varepsilon_0, \ y \exp(v) \in Z \}
\]

keeps track of all points of \( Z \cap B(y, \varepsilon_0) \) in the transversal direction to \( H \).

The following function \( f_Z : Y \to [2, \infty) \), called the Margulis function, encodes the information on the distance \( d(y, Z) \):

\[
f_Z(y) = \begin{cases} \sum_{v \in I_Z(y)} \|v\|^{-1/3} & \text{if } I_Z(y) \neq \emptyset \\ \varepsilon_0^{-1/3} & \text{otherwise} \end{cases}
\]

Note that even though we are eventually interested in controlling the size of the smallest \( v \in I_Z(y) \), the above definition of \( f_Z \) as a sum is a more stable object to consider. Indeed, for \( h \in H \), the value \( f_Z(yh) \) can be traced back from the value \( f_Z(y) \) and the size of \( h \), whereas the smallest vector in \( I_Z(yh) \) and the smallest vector in \( I_Z(y) \) are related only via an a priori unknown constant depending on \( Z \); namely, we would need to choose \( \varepsilon_0 = \varepsilon_0(Z) \) so that the \( \varepsilon_0 \)-neighborhood of \( y \) contains at most one sheet of \( Z \).

Since \( f_Z(y) \) and \( f_Z(yh) \) are comparable for \( h \in H \) near identity, Theorem 1.1 follows if we can show that the value of \( f_Z \) is controlled by the appropriate
complexity of the orbit \( Z \) on average, i.e.,
\[
(1.11) \quad m_Y(f_Z) \ll \text{Vol}(Z)
\]
where \( m_Y \) is the \( H \)-invariant probability measure on \( Y \).

The estimate (1.11) is proved by establishing the following super-harmonicity type inequality: there exist \( t > 0 \) and \( b > 1 \) such that for all \( y \in Y \),
\[
(1.12) \quad A_t f_Z(y) \leq \frac{1}{2} f_Z(y) + b \text{Vol}(Z)
\]
where \((A_t f_Z)(y) = \int_0^1 f_Z(yu_s a_t)ds, u_s = \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}, a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}\).

The proof of (1.12) is based on the simple linear algebra lemma, which says that in a finite dimensional irreducible representation \( \rho : \text{SL}_2(\mathbb{R}) \rightarrow \text{GL}(V) \) (we apply it to the adjoint representation), the norm \( \|\rho(u_s a_t)(v)\| \) (for \( t \gg 1 \)) of a non-zero vector \( v \in V \) grows on average for \( s \in [0, 1] \). In the Appendix §11 we give the proof of Theorem 1.1 in the compact case. It may be helpful to read the Appendix first.

When \( X \) is of finite volume with cusps, there is no lower bound for the injectivity radius on \( X \), and this forces to incorporate the height of \( y \), which is essentially \( \omega(y) = \text{inj}(y)^{-1} \), in the definition of \( f_Z \), and hence establish (1.12) for a suitable power of \( \omega \) as well.

Essentially, we would like to adapt this argument for a general geometrically finite manifold of infinite volume. As the \( H \)-invariant measure on \( Y \) is infinite in general, the correct measure to control the average behavior of the orbit \( Y \) is the Bowen-Margulis-Sullivan probability measure, which we denote by \( m_Y \). Recall that the complexity of a closed orbit \( Y \) was described only in terms of its volume in the finite volume case. In general, due to the fractal nature of the BMS measure on \( Y \) as well as the subtle fluctuation/dependence property of the Patterson-Sullivan density on the viewpoints, it turns out that there are more geometric invariants of \( Y \) which play roles in the quantitative isolation of \( Y \) relative to \( Z \), such as the shadow constant of \( Y \) and the tight area of \( Z \). We remark that in most of the existing literature, the operator \( A_t \) is defined as the averaging operator over large spheres in \( \mathbb{H}^2 \), but we consider averages over expanding pieces of the horocycles; horocyclic averages are more amenable to the change of variables and iteration arguments for Patterson-Sullivan measures. Finally, let us highlight that when \( \Delta_Y \) is a lattice, \( m_Y \) is \( H \)-invariant and (1.11) follows by simply integrating (1.12) with respect to \( m_Y \). In general, the BMS measure is not \( H \)-invariant, and we appeal to equidistribution results of expanding horocycles, in order to conclude an estimate similar to (1.11) from inequalities like (1.12) which we show hold for all large \( t \). Note, however, that since our function \( f_Z \) is not a bounded function in the presence of cusps, we need to control the growth rate of \( f_Z \) and \( \omega \) to apply equidistribution theorems.

We end this introduction with an outline of the paper. In §2 we fix some notation and conventions to be used throughout the paper. In §3 we discuss
Sullivan’s shadow lemma with precise description of the multiplicative constants and show the finiteness of the shadow constant. In §4, we show the finiteness of the tight area of a properly immersed geodesic surface. In §5, the definition of the averaging operator and a basic property of this operator are discussed. In §6 we prove a lemma from linear algebra; this lemma is a key ingredient to prove a local version of our main inequality. §7 is devoted to the study of the height function in \( X \). In §8, we discuss the averaging operator applied to the height function. The definition of the desired Margulis function is given in §9, where we also prove all the necessary estimates for the proof of Theorem 1.4. In §10, we give a proof of Theorem 1.4. In the Appendix §11, we provide a short proof of Theorem 1.1 in the compact case.

2. Notation and preliminaries

We set \( G = \text{PSL}_2(\mathbb{C}) \simeq \text{Isom}^+(\mathbb{H}^3) \), and \( H = \text{PSL}_2(\mathbb{R}) \). We fix \( \mathbb{H}^2 \subset \mathbb{H}^3 \) so that \( \{ g \in G : g(\mathbb{H}^2) = \mathbb{H}^2 \} = H \). We fix a point \( o^* \in \mathbb{H}^2 \subset \mathbb{H}^3 \) and a unit tangent vector \( v_{o^*} \in T_{o^*}(\mathbb{H}^3) \) once and for all. Denote by \( K_0 \) and \( M_0 \) respectively the stabilizer subgroups of \( o^* \) and \( v_{o^*} \). The isometric action of \( G \) on \( \mathbb{H}^3 \) induces identifications \( G/K_0 = \mathbb{H}^3 \), \( G/M_0 = T^1\mathbb{H}^3 \), and \( G = F\mathbb{H}^3 \) where \( T^1\mathbb{H}^3 \) and \( F\mathbb{H}^3 \) denote, respectively, the unit tangent bundle and the oriented frame bundle over \( \mathbb{H}^3 \).

Let \( A \) denote the following one-parameter subgroup:

\[
A = \left\{ a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} : t \in \mathbb{R} \right\}.
\]

The right translation action of \( A \) on \( G \) induces the geodesic/frame flow on \( T^1\mathbb{H}^3 \) and \( F\mathbb{H}^3 \), respectively. Let \( v_{o^*}^{\pm} \in S^2 \) denote the forward and backward end points of the geodesic given by \( v_{o^*} \). For \( g \in G \), we define

\[
g^{\pm} = g(v_{o^*}^{\pm}) \in S^2.
\]

Let \( \Gamma < G \) be a discrete subgroup. We denote by \( \Lambda \) the limit set of \( \Gamma \). A point \( \xi \in \Lambda \) is called \textit{parabolic} if \( \xi \) is fixed by a parabolic element of \( \Gamma \), and \textit{radial} iff there exists a geodesic ray \( \xi_t \) toward \( \xi \) which accumulates on a compact subset of \( M \).

We assume that \( \Gamma \) is geometrically finite, which is equivalent to the condition that \( \Lambda \) is the union of the radial limit points and parabolic limit points \( \Lambda_{rad} \cup \Lambda_p \).

We set \( M := \Gamma \backslash \mathbb{H}^3 \) to be the associated hyperbolic 3-manifold, and set

\[
X := \Gamma \backslash G.
\]

We denote by \( X_0 = \text{RFM} \) the renormalized frame bundle, i.e.,

\[
\text{RFM} = \{ [g] \in X : g^{\pm} \in \Lambda \}.
\]
Thick-thin decomposition of $X_0$. We fix a Riemannian metric $d$ on $G$ which induces the hyperbolic metric on $\mathbb{H}^3$. By abuse of notation, we use $d$ to denote the distance function on $X$ induced by $d$, as well as on $M$. For a subset $S \subset \mathfrak{g}$ and $\varepsilon > 0$, $B_{\mathfrak{g}}(S, \varepsilon)$ denotes the set $\{x \in \mathfrak{g} : d(x, S) \leq \varepsilon\}$. When $S = \{e\}$, we simply write $B_{\mathfrak{g}}(\varepsilon)$ instead of $B_{\mathfrak{g}}(S, \varepsilon)$. When there is no room for confusion for the ambient space $\mathfrak{g}$, we omit the subscript $\mathfrak{g}$.

For $x \in X$, the injectivity radius of $x$, which we denote by $\text{inj}(x)$, is the supremum $r > 0$ such that the map $g \mapsto xg$ is injective on the ball $B_G(e, r)$.

For $\varepsilon > 0$, we set
$$X_{\varepsilon} := \{x \in RF_M : \text{inj}(x) \geq \varepsilon\}.$$ As $M$ is geometrically finite, either $X_0$ is compact or $X_0 - X_{\varepsilon_0}$ is contained in finitely many disjoint horoballs for all small $\varepsilon > 0$. Set
$$(2.1) \quad \varepsilon_0 = \begin{cases} \min_{q \in RF_M} \text{inj}(q) & \text{if } X_0 \text{ is compact} \\ \max\{\varepsilon_0 : X_0 - X_{\varepsilon_0} \text{ is contained in horoballs}\} & \text{otherwise} \end{cases}$$

Let $\xi_1, \ldots, \xi_\ell$ be representatives of $\Gamma$-orbits in $\Lambda_p$ so that $RF_M - X_{\varepsilon_0}$ is contained in a disjoint union of finitely many horoballs $h_{\xi_1}, \ldots, h_{\xi_\ell}$. We denote by $N$ the expanding horospherical subgroup for $A$:
$$N = \left\{ u_s = \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} : s \in \mathbb{C} \right\}.$$ Each horoball $h_i = h(\xi_i)$ is of the form
$$h_i = \left[ g_i \right] N A_{(-\infty, -T_i]} K_0$$
for some $T_i \geq 1$, where $g_i \in G$ is an element with $g_i^{-1} = \xi_i$ and $A_{[-T, -\infty)} = \{a_t : -\infty < t \leq -T\}$.

**Convention.** We will use the notation $A \asymp B$ when the ratio between the two lies in $[C^{-1}, C]$ for some absolute constant $C \geq 1$. By an absolute constant, we mean $C$ depends at most on $G$ and $\Gamma$ in general. We write $A \ll B^\ast$ (resp. $A \ll \ast B$) to mean that $A \leq C B^L$ (resp. $A \leq C \cdot B$) for some absolute constants $C > 0$ and $L > 0$ — when $\Gamma$ is a lattice, the exponent $L$ depends only on $G$.

### 3. Shadow constants

Let $\mathcal{H}$ denote the collection of all closed $H$-orbits $yH$ in $X$ such that $\text{Stab}_H(y)$ is non-elementary, or equivalently, $\text{Stab}_H(y)$ is Zariski dense in $H$. Fixing $Y \in \mathcal{H}$, we will introduce various measures and constants associated to $Y$ in this section.

We choose, once and for all,
$$(3.1) \quad y_0 \in Y \cap X_{\varepsilon_0} \quad \text{and} \quad \Delta_Y := \text{Stab}_H(y_0).$$

As $\Gamma$ is geometrically finite, the subgroup $\Delta_Y$ is a geometrically finite subgroup of $H$. We denote by $\Lambda_Y \subset \partial \mathbb{H}^2$ the limit set of $\Delta_Y$. We denote by
\[ 0 < \delta_Y \leq 1 \] the critical exponent of \( \Delta_Y \), equivalently the Hausdorff dimension of \( \Lambda_Y \). We remark that \( \delta_Y = 1 \) if and only if \( Y \) has finite Riemannian volume.

Let \( S_Y := \Delta_Y \setminus \mathbb{H}^2 \). The convex core of \( S_Y \), denoted by \( \text{core}(S_Y) \), is given by \( \Delta_Y \setminus \text{hull}(\Lambda_Y) \). The compact core of \( S_Y \) is the minimal connected surface \( C_Y \) such that \( \text{core}(S_Y) - C_Y \) is a union of disjoint cusps. We denote by \( d_Y \) the diameter of the compact core of \( S_Y \).

Consider the proper map
\[ \tilde{f}_Y : \Delta_Y \setminus H = T^1(S_Y) \to Y \subset X \] given by \([h] \to y_0 h\), and let \( f_Y : S_Y \to M \) be the induced proper immersion.

Letting \( F_Y \subset \mathbb{H}^2 \) be the Dirichlet domain for \( \Delta_Y \) containing \( o^* \), we fix \( o_Y \in \text{hull}(\Lambda_Y) \cap F_Y \) so that its image \([o_Y]\) in \( S_Y \) is the base point of \( y_0 \) via \( f_Y \). As \( y_0 \in Y \cap X_{\epsilon_0} \), \([o_Y]\) \( \in C_Y \). Therefore, the choice of \( o_Y \) (equivalently the choice of \( y_0 \)) is unique up to the distance \( d_Y \). If there is no confusion, we drop the index \( Y \) from the notation when referring to \( o_Y \).

Patterson-Sullivan measure. We denote by \( \{ \nu_p = \nu_{Y,p} : p \in \mathbb{H}^2 \} \) the Patterson-Sullivan density for \( \Delta_Y \), normalized so that \( \nu_{o_Y} \) has mass one: for all \( \gamma \in \Delta_Y, p,q \in \mathbb{H}^2, \xi \in \Lambda_Y \),
\[ \frac{d\gamma_*\nu_p}{d\nu_p}(\xi) = e^{-\delta_Y \beta_{\xi}(\gamma^{-1}(p),p)} \quad \text{and} \quad \frac{d\nu_q}{d\nu_p}(\xi) = e^{-\delta_Y \beta_{\xi}(q,p)} \]
where \( \beta_{\xi}(\cdot,\cdot) \) denotes the Busemann function.

Set \( U := \left\{ u_s = \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} : s \in \mathbb{R} \right\} = N \cap H \) which is the expanding horocyclic subgroup of \( H \). Using the parametrization \( s \mapsto u_s \) we may identify \( U \) with \( \mathbb{R} \). Note that for all \( s,t \in \mathbb{R}, \)
\[ a_{-t} u_s a_t = u_{e^t s}. \]

For any \( h \in H \), the restriction of the visual map \( g \mapsto g^+ \) is a diffeomorphism between \( hU \) and \( S^1 - \{h^{-}\} \). Using this diffeomorphism, we can define a measure on \( hU \):
\[ d\mu_{hU}^{PS}(hu_r) = e^{\delta_Y \beta_{(hu_r)^+}(p,hu_r(p))} d\nu_p(hu_r)^+; \]
this is independent of the choice of \( p \in \mathbb{H}^2 \).

We simply denote this by
\[ d\mu_h^{PS}(r) = d\mu_{hU}^{PS}(hu_r). \]

Note that these measures depend on the \( U \) orbits but not on the individual points. By the \( \Delta_Y \)-invariance and the conformal property of the PS-density, we have
\[ d\mu_h^{PS}(O) = d\mu_{\gamma h}^{PS}(O) \]
for any $\gamma \in \Delta_Y$ and for any bounded Borel set $O \subset \mathbb{R}$; therefore $\mu_y^{\text{ps}}(O)$ is well-defined for $y \in \Delta_Y \setminus H \simeq Y$.

For any $y \in \Delta_Y \setminus H$ and any $s \in \mathbb{R}$, we have:

\[(3.3) \quad \mu_y^{\text{ps}}([-e^s, e^s]) = e^{\delta_Y s} \mu_{y a_{-s}}^{\text{ps}}([-1, 1]).\]

**Bowen-Margulis-Sullivan measure** $m_Y$. We denote by $m_Y$ the Bowen-Margulis-Sullivan probability measure on $\Delta_Y \setminus H = T^1(Y)$, which is the unique probability measure of maximal entropy (that is $\delta_Y$) for the geodesic flow. We set

$Y_0 = \text{supp}(m_Y)$, i.e., the non-wandering set of the geodesic flow. We note that $Y_0 = f_Y^{-1}(Y \cap RFM)$. We will also use the same notation $m_Y$ to denote the push-forward of the measure to the $H$-orbit $Y$ in what follows.

**Shadow lemma.** The convex core of $S_Y$ is the union of the compact core $C_0 := C_Y$ and the union of disjoint horoballs, say, $C_1, \ldots, C_m$.

We choose the lifts $\tilde{C}_i \subset F_Y \cap \text{hull}(\Lambda_Y)$, $0 \leq i \leq m$, so that $\Delta_Y \setminus \Delta_Y \tilde{C}_i = \tilde{C}_i$. In particular, $\partial \tilde{C}_0$ intersects $\tilde{C}_i$ as an interval. Let $\xi_i \in \Lambda_Y$ be the base point of the horoball $\tilde{C}_i$, i.e., $\xi_i = \partial \tilde{C}_i \cap \mathbb{S}^1$. Let $F_i$ be a minimal interval for the action of $\text{Stab}_{\Delta_Y}(\xi_i)$ in $\Lambda_Y - \{\xi_i\}$ such that $\partial Y = \partial \tilde{C}_0$ lies on a geodesic connecting a point of $F_i$ to $\xi_i$.

For $p \in \mathbb{H}^2$, let $d_p$ denote the Gromov distance on $\mathbb{S}^1$: for $\xi \neq \eta \in \mathbb{S}^1$,

\[d_p(\xi, \eta) = e^{-(\beta_1(p, \eta) + \beta_0(p, \eta))/2}\]

where $y$ is any point on the geodesic connecting $\xi$ and $\eta$.

We set

\[B_p(\xi, r) = \{\eta \in \mathbb{S}^1 : d_p(\eta, \xi) \leq r\} \]

For $t \geq 0$, we denote by $V(p, \xi, t)$ the set of all $\eta \in \mathbb{S}^1$ such that the distance between $p$ and the orthogonal projection of $\eta$ onto the geodesic $[p, \xi]$ is at least $t$. For some absolute constant $c_0 > 0$, we have

\[V(p, \xi, t + c_0) \subset B_p(\xi, e^{-t}) \subset V(p, \xi, t - c_0) \quad \text{for all } t \geq c_0.\]

The following is a uniform version of Sullivan's shadow lemma [21]. The proof of this proposition can be found in [20, Thm. 3.4]; since the dependence on the multiplicative constant is important to us, we give a sketch of the proof while making the dependence of constants explicit.

**Proposition 3.4.** There exists a constant $c \propto e^{d_Y}$ such that for all $\xi \in \Lambda_Y$, $p \in \tilde{C}_0$ and for all $t > 0$,

\[c^{-1} \cdot \nu_p(F_{\xi_t}) \beta_Y e^{-\delta_Y t + (1 - \delta_Y)d(\xi_t, \Delta_Y(p))} \leq \nu_p(V(p, \xi, t)) \leq c \cdot \nu_p(F_{\xi_t}) e^{-\delta_Y t + (1 - \delta_Y)d(\xi_t, \Delta_Y(p))}\]

where

- $\{\xi_t\}$ is the unit speed geodesic ray $[p, \xi]$ so that $d(p, \xi_t) = t$;
- $F_{\xi_t} = \mathbb{S}^1$ if $\xi_t \in C_0$, and $F_{\xi_t} = F_{\xi_i}$ if $\xi_t \in C_i$ for $1 \leq i \leq m$;
Proof. As we allow the constant $c \asymp e^{d_Y}$ and $d_Y = \text{diam} \hat{C}_0$, it suffices to prove the claim when $p = o \in \hat{C}_0$. We fix $\xi \in \Delta_Y$ and let $\xi_t$ denote the unit speed geodesic from $o$ to $\xi$. By the $d_Y$-conformality of $\{\nu_p : p \in \mathbb{H}^2\}$, we have
\[
\nu_o(V(o, \xi, t)) \asymp e^{-d_Y t} \nu_{\xi_t}(V(o, \xi, t))
\]
with an absolute implied constant. Therefore it suffices to show
\[
\nu_{\xi_t}(V(o, \xi, t)) \asymp e^{(1-\delta_Y) d(\xi_t, \Delta_Y(o))}
\]
while making the dependence of the implied constant explicit.

Claim A. If $\xi_t \in \Delta_Y \hat{C}_0$, then
\[
e^{-\delta_Y d_Y} \inf_{\eta \in \Delta_Y} \nu_o(B(\eta, e^{-d_Y})) \ll \nu_{\xi_t}(V(o, \xi, t)) \ll e^{\delta_Y d_Y}
\]
where the implied constants are absolute.

By the assumption, there exists $\gamma \in \Delta_Y$ such that $d(\xi_t, \gamma o) \leq d_Y$. Hence it follows from the conformality of $\{\nu_p\}$
\[
e^{-\delta_Y d_Y} \nu_{\xi_t}(V(o, \xi, t)) \leq \nu_{\gamma o}(V(o, \xi, t)) = \nu_o(V(\gamma^{-1} o, \gamma^{-1} \xi, t)) \leq e^{\delta_Y d_Y} \nu_{\xi_t}(V(o, \xi, t)).
\]

The first inequality shows the upper bound since $|\nu_o| = 1$. The lower bound follows from the second inequality; indeed
\[
V(\gamma^{-1} o, \gamma^{-1} \xi, t) = V(\gamma^{-1} \xi_t, \gamma^{-1} \xi, 0)
\]
and the latter contains $B_o(\gamma^{-1} \xi, e^{-d_Y})$, since $d(o, \gamma^{-1} \xi_t) \leq d_Y$.

Claim B. Let $\xi$ be a parabolic limit point in $\Lambda_Y$. Assume that for some $i$, $\xi_t \in \hat{C}_i$ for all large $t$.

We claim:
\[
\nu_{\xi_t}(V(o, \xi, t)) \asymp e^{\delta_Y d_Y} \nu_o(F_\xi) \cdot e^{(1-\delta_Y)(d(\xi_t, \Delta_Y(o)) + d_Y)},
\]
and
\[
\nu_{\xi_t}([S^1 - V(o, \xi, t)) \asymp e^{\delta_Y d_Y} \nu_o(F_\xi) \cdot e^{(1-\delta_Y)(d(\xi_t, \Delta_Y(o)) + d_Y)}.
\]

Let $s_i \geq 0$ be such that $\xi_{s_i} \in \partial \hat{C}_i$. Then for all $t \geq s_i$,
\[
|d(\xi_t, \Delta_Y(o)) - (t - s_i)| \leq d_Y.
\]
Hence for (3.6), it suffices to show
\[
(3.8) \quad \nu_{\xi_t}(V(o, \xi, t)) \asymp e^{(1-\delta_Y)(t-s_i)} \nu_o(F_\xi).
\]
Note that
\[
\nu_{\xi_t}(V(o, \xi, t)) \asymp \sum_{p \in \Gamma_\xi, \partial F_\xi \cap V'(o, \xi, t) \neq \emptyset} \nu_{\xi_t}(pF_\xi)
\]
where $\Gamma_\xi = \text{Stab}_{\Delta_Y}(\xi)$. Since $pF_\xi \cap V(o, \xi_t, t) \neq \emptyset$, then $d(o, po) \geq 2t - k$ for some absolute $k$, and hence

$$\nu_\xi(V(o, \xi, t)) \asymp \sum_{d(o, po)\geq 2t} \nu_\xi(pF_\xi).$$

Let $F_\xi^+$ denote the image of $F_\xi$ on the horocycle based at $\xi$ passing through $o$ via the visual map. We use the fact that if $d(o, po) \geq 2t$, then for all $\eta \in F_\xi$,

$$|\beta_\eta(p^{-1}\xi_t, \xi_t) - d(o, po) + 2t| \ll \text{diam}F_\xi^+ \leq d_Y$$

(cf. proof of [20, Lemma 2.9]). Since $\nu_\xi(pF_\xi) = \int_{F_\xi} e^{-\delta_Y \beta_{po}(\xi_t, p\xi_t)} d\nu_\xi(\eta)$, $\nu_\xi(F_\xi) \asymp e^{-\delta_Y t} \nu_\xi(F_\xi)$, we deduce, with multiplicative constant $\asymp e^{\delta_Y d_Y}$,

$$\sum_{d(o, po)\geq 2t} \nu_\xi(pF_\xi) \asymp e^{\delta_Y t} \sum_{d(o, po)\geq 2t} e^{2\delta_Y d(o, po)} \nu_\xi(F_\xi)$$

$$\times \nu_\xi(F_\xi) e^{\delta_Y t} \sum_{d(o, po)\geq 2t} e^{-\delta_Y d(o, po)}$$

$$\times \nu_\xi(F)e^{(1-\delta_Y)t}$$

using $a_n := \#\{p \in \Gamma_\xi : n < d(o, po) \leq n + 1\} \asymp e^{n/2}$ in the last estimate. This proves (3.6). shadow

The estimate (3.8) follows similarly now using

$$\nu_\xi(S^1 - V(o, \xi, t)) \asymp \sum_{d(o, po)\leq 2t} \nu_\xi(pF_\xi).$$

and $\sum_{n=0}^{[2t]} a_n e^{-\delta_Y n} \asymp e^{(1-2\delta_Y)t}$.

The next two remaining cases are deduced from Claims A and B.

**Claim C.** When $\xi$ be a parabolic limit point, (3.6) holds with multiplicative constant $\asymp e^{\delta_Y}$. (see the proof of [20, Prop. 3.6]).

**Claim D.** If $\xi_t \in \Delta_Y C_i$ for some $i$, then (3.6) holds with multiplicative constant $\asymp e^{\delta_Y}$. (see the proof of [20, Lemma 3.8]).

**Proposition 3.9.** Let $Y \in \mathcal{H}$. There exists $R_Y \asymp e^{\delta_Y}$ such that for all $y \in Y_0$, we have

$$R_Y \beta_Y e^{(1-\delta_Y) d(C_Y, \pi(y))} \leq \mu_y^{PS}([-1, 1]) \leq R_Y e^{(1-\delta_Y) d(C_Y, \pi(y))}$$

where $\pi$ denotes the base point projection $Y \rightarrow S_Y = \Delta_Y \setminus \mathbb{H}^2$.

**Proof.** We give an argument which is a slight modification of the proof of [19, Prop. 5.1]. Since the map $y \mapsto \mu_y^{PS}([-1, 1]}$ is continuous on $Y_0$ and $\{[h] \in Y_0 : h^- \text{ is a radial limit point of } \Delta_Y\}$ is dense in $Y_0$, it suffices to prove the claim for $y = [h]$ assuming that $h^-$ is a radial limit point.

Recall that $\mu_y^{PS}([-1, 1]) = e^{\delta_Y t} \mu_{y_{-t}}^{PS}([-e^{-t}, e^{-t}])$ for all $t \in \mathbb{R}$. Let $t \geq 0$ be the minimal number so that $\pi(y_{a_{-t}}) \in C_Y$. Then

$$d(\pi(y), C_Y) \leq d(\pi(y), \pi(y_{a_{-t}})) \leq d_Y + d(\pi(y), C_Y).$$
On the other hand,
\[\mu_{ha_{-t}}[-e^{-t}, e^{-t}] \asymp \nu_{ha_{-t}(o)}(V(ha_{-t}(o), h^+, t))\]
with an absolute implied constant (cf. [20, Lemma 4.4]).

Since \(ya_{-t} \in C_Y, F_{ha_{-t}} = S^1\). So \(\nu_{ha_{-t}}(F_{ha_{-t}}) = |\nu_{ha_{-t}}| \asymp 1\). Therefore, for some implied constant \(\asymp e^{*d_Y}\), we have
\[\beta_Y e^{-\delta_Y t + (1-\delta_Y)d(\pi(y), \pi(ya_{-t}))} \ll \nu_{ha_{-t}(o)}(V(ha_{-t}, h^+, t)) \ll e^{-\delta_Y t + (1-\delta_Y)d(\pi(y), \pi(ya_{-t}))}.\]

This estimate and (3.10), therefore, imply that
\[\beta_Y e^{(1-\delta_Y)d(\pi(y)C_Y)} \ll \mu_y^{ps}([-1,1]) \ll e^{(1-\delta_Y)d(\pi(y)C_Y)}\]
with the implied constant \(\asymp e^{*d_Y}\), proving the claim. \(\Box\)

**The shadow constant \(s_Y\).** We set

\[
(3.11) \quad \delta_Y^\circ = \begin{cases} 
\delta_Y & \text{if } Y \text{ is convex cocompact} \\
2\delta_Y - 1 & \text{otherwise}
\end{cases}
\]

Note that \(0 < \delta_Y^\circ \leq 1\) since \(\delta_Y > 1/2\) when \(Y\) is not convex cocompact.

We use the following variant of Sullivan’s shadow lemma, obtained by Schapira-Maucourant ([21], [19]):

**Corollary 3.12.** We have

\[0 \leq \sup_{y \in Y_0, 0 \leq \varepsilon \leq 1} \frac{\mu_y^{PS}([-\varepsilon, \varepsilon])}{\delta_Y^{\circ} \mu_y^{PS}([-1,1])} \leq R_Y^2 \cdot \beta_Y^{-1} < \infty,\]

where \(R_Y\) is as in Proposition 3.9.

**Proof.** By (3.3), we have \(\mu_y^{ps}([-\varepsilon, \varepsilon]) = \varepsilon^{\delta_Y} \mu_y^{ps}([-1,1])\). Hence the case when \(Y\) is convex cocompact follows by Proposition 3.9.

Now suppose \(Y\) has a cusp. Let \(y \in Y_0\); using the triangle inequality, we get that \(d(ya_{-\log \varepsilon}, C_Y) - d(y, C_Y) \leq -\log \varepsilon\). Therefore, by Proposition 3.9 we have
\[
\frac{\mu_y^{ps}([-1,1])}{\mu_y^{ps}([-1,1])} \leq R_Y^2 \cdot \beta_Y^{-1} \cdot e^{(1-\delta_Y)(d(ya_{-\log \varepsilon}, C_Y) - d(y, C_Y))} \leq R_Y^2 \cdot \beta_Y^{-1} \cdot \varepsilon^{\delta_Y - 1}.
\]

In consequence, we have
\[
\frac{\mu_y^{ps}([-\varepsilon, \varepsilon])}{\mu_y^{ps}([-1,1])} \leq R_Y^2 \cdot \beta_Y^{-1} \cdot \varepsilon^{2\delta_Y - 1}
\]
which establishes the claim. \(\Box\)

Recall that the shadow constant of \(Y\) was defined in (1.6):

\[
(3.13) \quad s_Y = \sup_{\xi \in \Delta_Y, \rho \in [\Delta_Y, \xi], 0 \leq \varepsilon \leq 1} \frac{\nu_p(B_p(\xi, \varepsilon))}{\epsilon^{\delta_Y} \nu_p(B_p(\xi, 1))}.
\]
Corollary 3.14. Set
\[
\begin{equation}
(3.15)
p_Y = \max \left\{ 1, \sup_{y \in Y_0, 0 \leq \varepsilon \leq 1} \frac{\mu^\text{PS}_y([-\varepsilon, \varepsilon])}{\varepsilon \mu^\text{PS}_y([-1, 1])} \right\} \in [1, \infty].
\end{equation}
\]
Then
\[
\max\{1, s_Y\} \times p_Y \ll e^{* \delta_Y} \beta_Y^{-1}.
\]

Proof. If \( y = [h] \), then \( \mu_y([-r, r]) = \int e^{-\delta_Y \beta_{h \cdot u^+}} d\nu_{h(o)}(h \cdot u^+) \). Since
\[
|\beta_{h \cdot u^+}(h, h \cdot u_r)| \leq d(e, u_r),
\]
we have that for all \( 0 < r < 1 \),
\[
\mu_y([-r, r]) \sim \nu_{h(o)}(B_{h(o)}(h^+, r))
\]
where the implied constant is independent of \( r \). Hence \( p_Y \propto \max\{1, s_Y\} \).

The claimed inequality is proved in Corollary 3.12. \( \square \)

The constants \( d_Y \) and \( s_Y \) are defined using the group \( \Delta_Y \), which in turn is defined using our fixed base point \( y_0 \), see (3.1). We now show that these two quantities are indeed independent of the choice of \( y_0 \).

Lemma 3.16. Let \( y_0 \) and \( \Delta_Y \) be as in (3.1). Let \( y = y_0 h^{-1} \in Y \) for some \( h \in H \) and let \( \Delta'_Y = \text{Stab}_H(y) \). Define \( d'_Y \) and \( s'_Y \) as above using \( \Delta'_Y \) instead of \( \Delta_Y \). Then \( d_Y = d'_Y \) and \( s_Y = s'_Y \).

Proof. Note that \( \Delta'_Y = h \Delta_Y h^{-1} \). Therefore, \( h \mathcal{F}_Y \) is a fundamental domain for \( \Delta'_Y \). Moreover, \( h \mathcal{C}_0 \) is compact, \( h \tilde{C}_i \) is a horoball for every \( 1 \leq i \leq m \), and \( h \tilde{C}_i \cap h \tilde{C}_j = \emptyset \) for all \( 1 \leq i \neq j \leq m \). In consequence, since the action of \( h \) on \( \mathbb{H}^2 \) is an isometric action, we get that
\[
d'_Y = \text{diam}(h \mathcal{C}_0) = \text{diam}(\mathcal{C}_0) = d_Y.
\]

We now show that \( s'_Y = s_Y \). Let \( \{\nu_{p} : p \in \mathbb{H}^2\} \) denote the Patterson-Sullivan density for \( \Delta_Y \). Then the family \( \{\nu'_p := h_{*} \nu_{h^{-1} p}\} \) is the Patterson-Sullivan density for \( \Delta'_Y \). To see this, first note that the limit set of \( \Delta'_Y \) equals \( h \Lambda_Y \). Let now \( \xi \in \Lambda_Y \), then
\[
\frac{d\left((h \gamma h)^{-1} p\right)_{*} \nu'_{p}}{d\nu'_{p}}(h \xi) = \frac{d\left((h \gamma) p\right)_{*} \nu_{h^{-1} p}}{d\nu_{h^{-1} p}}(h \xi) \\
= \frac{d\gamma_{*} \nu_{h^{-1} p}}{d\nu_{h^{-1} p}}(\xi) = e^{-\delta_Y \beta_{\xi}(\gamma^{-1}(h^{-1} p), h^{-1} p)} \\
= e^{-\delta_Y \beta_{\xi}(h \gamma^{-1}(p), p)}.
\]
The above, in view of the uniqueness of the Patterson-Sullivan density (up to scalar), thus implies that \( \{\nu'_p := h_{*} \nu_{h^{-1} p}\} \) is the Patterson-Sullivan density for \( \Delta'_Y \).

Now for every \( 0 < \varepsilon \leq 1 \) and every \( \xi \in \Lambda_Y \), we have
\[
\nu'_p(B_{p}(h\xi, \varepsilon)) = h_{*} \nu_{h^{-1} p}(B_{p}(h\xi, \varepsilon)) = \nu_{h^{-1} p}(h^{-1} B_{p}(h\xi, \varepsilon)).
\]
Moreover, by the equivariance properties of the Gromov distance, we have 

\[ h^{-1} B_p(h_\xi, \varepsilon) = B_{h^{-1} p}(\xi, \varepsilon). \]

We conclude that \( s_Y = s'_Y \). \( \square \)

4. Tight area of a geodesic surface

For a closed subset \( Q \subset M \), we define the tight neighborhood of \( Q \) by

\[ N(Q) := \{ p \in M : d(p, q) \leq \text{inj}(q) \text{ for some } q \in Q \}. \]

For a properly immersed geodesic surface \( S \) of \( M \), we define the tight-area of \( S \) relative to \( M \) as follows:

\[ \text{area}_t(S) := \text{area}(S \cap N(\text{core}(M))). \]

Proposition 4.1. For a properly immersed non-elementary geodesic surface \( S \) of \( M \), we have

\[ \text{area}_t(S) < \infty. \]

Proof. Since \( S \) is a geometrically finite surface, \( S \) is the union of its convex core and finitely many connected components, called flares. Since \( \text{core}(S) \) has finite area, it suffices to show that for the image of each flare, say \( F \), of \( S \), the area of \( F \cap N(\text{core}(M)) \) is finite.

Let \( R \) be the maximum of the injectivity radius of points in the compact core of \( M \). Since \( N(\text{core}(M)) \) is contained in the \( R \)-neighborhood of the compact core of \( M \) and finitely many tight neighborhoods of disjoint horoballs, it suffices to show that the intersection of \( F \) with the tight neighborhood of a horoball in \( \text{core}(M) \) has finite area.

We denote by \( M_{\text{cpt}} \subset \text{core}(M) \) the compact core of \( M \). Fix a horoball \( h \) in \( \text{core}(M) \), i.e., a connected component of \( \text{core}(M) - M_{\text{cpt}} \). We may assume that \( h \) is covered by the horoball based at \( \infty \): \( h = \{(x_1, x_2, y) \in \mathbb{H}^3 : y > c\} \) for some \( c > 0 \).

Since \( \infty \) is a bounded parabolic fixed point, there exists \( c_0 > 0 \) such that \( \{(x_1, x_2, y) : |x_1| < b, |x_2| < b, y > c\} \) projects onto \( h \). Since \( \text{inj}(x) \cong y^{-1} \) for \( x = (x_1, x_2, y) \in h \), the tight neighborhood \( N(h) \) is covered by

\[ \tilde{h}_{\text{chimney}} := \{(x_1, x_2, y) : |x_1| < b + c_0, |x_2| < b + c_0, y > c\} \]

for some absolute constant \( c_0 > 0 \).

Let \( p : \mathbb{H}^3 \to M = \Gamma \setminus \mathbb{H}^3 \) be the projection map. We set

\[ h_{\text{chimney}} := p(\tilde{h}_{\text{chimney}}), \]

and call it the chimney in \( h \). We claim that \( F \cap N(h) \) is contained in the union of an \( R_0 \)-neighborhood of \( \text{core}(S) \) and a cusp-like region with finite area which is isometric to \( \{(x, y) \in \mathbb{H}^2 : |x| < b', y > c\} \) and arises as \( F \cap h_{\text{chimney}} \). This establishes the claim, since \( \text{core}(S) \) has finite area.

Fix a geodesic plane \( P \subset \mathbb{H}^3 \) which covers \( S \) and let \( \Delta = \text{Stab}_\Gamma(P) \). In the rest of the proof, we denote by \( \Lambda \subset \partial P \) the limit set of \( \Delta \). Since \( S \) is properly immersed, there exists some \( R_1 > 0 \) so that

\[ P \cap p^{-1}(S \cap B(M_{\text{cpt}}, 1)) \subset B_P(\text{hull}(\Lambda), R_1); \]
recall that $S \cap B(M_{\text{cpt}}, 1) \neq \emptyset$.

We will show that there exists $R_2 > 0$ depending only on $R_1$ so that

\begin{equation}
F \cap N(h) \subset B_S(\text{core}(S), R_2) \cup h_{\text{chimney}}.
\end{equation}

Let $z \in F \cap N(h)$ and let $\tilde{z} \in P$ be any lift of $z$. Fix a complete geodesic \{\xi_t\} $\subset P$ with $\xi_0 = \tilde{z}$. If any of the geodesic rays $\xi_{(-\infty, 0)}$ and $\xi_{(0, \infty)}$ projects into $h$ under $p$, then at least one of $\xi_{\pm \infty}$ belongs to the orbit $\Gamma(\infty)$. As \{\xi_{\pm \infty}\} $\subset \partial P$, it follows that $\gamma_0(\infty) \in \partial P$ for some $\gamma_0 \in \Gamma$; hence $\gamma_0^{-1}P$ is a vertical plane. Since \{\gamma h : \gamma \in \Gamma\} is a disjoint collection of horoballs and $\gamma_0^{-1}P \cap h \neq \emptyset$, we get

\[ z \in p(\gamma_0^{-1}P \cap \tilde{h}_{\text{chimney}}) \subset h_{\text{chimney}}. \]

Now suppose that neither of the geodesic rays $\xi_{(-\infty, 0)}$ and $\xi_{(0, \infty)}$ is completely contained in $h$ under $p$. Since $\partial h \subset M_{\text{cpt}}$, there exists $-\infty < t_1 < 0 < t_2 < \infty$ such that $\xi_{t_1}, \xi_{t_2} \in P \cap p^{-1}(M_{\text{cpt}})$. By the choice of $R_1$, we have $\xi_{t_1}, \xi_{t_2} \in B_P(h(\Lambda), R_1)$. For each $i = 1, 2$, let $\xi_i' \in h(\Lambda)$ be the unique closest points to $\xi_{t_i}$ in $h(\Lambda)$. Let $a$ be the union of three geodesic segments:

\[ a := (\xi_{t_1}, \xi_{t_1}') \cup (\xi_{t_1}', \xi_{t_2}') \cup (\xi_{t_2}', \xi_{t_2}). \]

Then $a$ is a quasi-geodesic with both multiplicative and additive constants depending only on $R_1$; note also that since $h(\Lambda)$ is convex, $(\xi_{t_1}', \xi_{t_2}') \subset h(\Lambda)$ and hence $a \subset B_P(h(\Lambda), R_1)$. Thus, by the Morse lemma from hyperbolic geometry, there exists some $R_1' > 0$, depending only on $R_1$, so that the Hausdorff distance between $a$ and $(\xi_{t_1}, \xi_{t_2})$ is at most $R_1'$. In particular, we get that

\[ \tilde{z} = \xi_0 \in B_P(h(\Lambda), R_1 + R_1'). \]

By setting $R_2 := R_1 + R_1'$, we get $z \in B_S(\text{core}(S), R_2)$, as desired in (4.2). \hfill $\square$

We record the following byproduct of the proof of the above proposition, which is of independent interest.

**Corollary 4.3.** For any properly immersed geodesic surface $S \subset M$, there exists $R \geq 0$ such that

\[ S \cap N(\text{core } M) \subset B_S(\text{core}(S), R) \cup N(h^\dagger_{\text{chimney}}) \]

where $h^\dagger_{\text{chimney}}$ is the union of chimneys over all horoballs in core $M$.

This corollary says that the portion of $S$, especially of the flares of $S$, staying in the "tight" neighborhood of core $M$ can go to infinity only in cusp-like shapes, by visiting the chimneys of horoballs of core $M$. This is not true any more if we replace the tight neighborhood of core $M$ by the 1-neighborhood of core $M$. More precisely if $\Lambda(\Gamma)$ contains a parabolic limit point of rank one which is not stabilized by any element of $\pi_1(S)$, then some infinite area region of $S$ can stay in the 1-neighborhood of core $M$. This
situation may be compared to the presence of divergent geodesics in finite area setting.

**Tight volume of** $Z \in \mathcal{H}$. We consider the following neighborhood of $X_0$:

\[(4.4) \quad X'_0 := \{ x \in X : \exists x_0 \in X_0 \text{ with } d(x, x_0) \leq \text{inj}(x_0) \},\]

and set

\[ M'_0 := \pi(X'_0) \]

where $\pi : X \to M$ is the base-point projection map.

Note that $M'_0 \subset N(\text{core } M)$, in particular, it is contained in the 1-neighborhood of core $M$, and hence has finite volume.

We set

\[(4.5) \quad \tau_Z := \max\{1, \text{area}(S_Z \cap M'_0)\}.\]

Recalling the definition of the tight area of $S_Z$ from (1.8) and the fact that $M'_0 \subset N(\text{core } M)$, note that

\[ \text{area}(S_Z \cap M'_0) \leq \text{area}_t(S_Z). \]

This, together with the fact that $\text{area}_t S_Z \geq (\varepsilon_3)^2$ where $\varepsilon_3$ is the Margulis constant for $\mathbb{H}^3$, implies that $\tau_Z \ll \text{area}_t(S_Z)$.

**Corollary 4.6.** For $Z \in \mathcal{H}$, we have

\[ 1 \leq \tau_Z \ll \text{area}_t(S_Z) < \infty. \]

5. **Averaging operator**

In this section we define an averaging operator and investigate some of its basic properties. Fix a closed orbit $Y \in \mathcal{H}$.

**Definition 5.1.** For each $t \in \mathbb{R}$ and $\rho > 0$, we define the operator $A_{t, \rho} : C(Y_0) \to C(Y_0)$ by

\[(5.2) \quad (A_{t, \rho} \psi)(y) := \frac{1}{\mu_y^\text{PS}([-\rho, \rho])} \int_{-\rho}^{\rho} \psi(yu_t a_t) d\mu_y^\text{PS}(r).\]

For the sake of notational ease, we set $A_t := A_{t, 1}$.

This is well-defined by the continuity of the map $y \mapsto \mu_y^\text{PS}([-\rho, \rho])$ for $\rho > 0$.

We remark that in [7], the averaging operator $A_t$ was defined using the integral over the translates $\text{SO}(2)a_t$, whereas we use the integral over the translates $U_{[-\rho, \rho]}a_t$ of a horocyclic piece. The proof of the following proposition, which is an analogue of [7, §5.3], is the main reason for our digression from their definition, as the handling of the PS-measure on $U$ is much harder than that of the PS-measure on $\text{SO}(2)$ in performing change of variables.

**Proposition 5.3.** For every $\sigma > 1$ there exists some $0 < c_0 = c_0(\sigma) < 1$ so that the following holds. Suppose that $F : Y_0 \to [2, \infty)$ is a continuous function satisfying the following properties:
(a) For all $h \in B_H(2)$ and $y \in Y_0$, we have
\[
\sigma^{-1}F(y) \leq F(yh) \leq \sigma F(y).
\]
(b) There exists $\tau \geq 2$ and $D_0 = D_0(\tau) > 0$ such that, for all $y \in Y_0$ and $1 \leq \rho \leq 2$, we have
\[
A_{\tau, \rho}F(y) < c_0 \cdot F(y) + D_0.
\]
Then we have:
1. For any $0 < c < 1$, there exists $t_0 = t_0(\sigma, c) > 0$, and $D = D(\sigma) > 0$ so that for all $t > t_0$ and all $y \in Y_0$, we have
\[
A_tF(y) < c \cdot F(y) + D.
\]
2. For any $y \in Y_0$, we have
\[
\limsup_{t \to \infty} A_tF(y) \leq 1 + D.
\]

Proof. For any $y \in Y_0$ and for any $\rho > 0$, set
\[
b_y(\rho) := \mu_y^{PS}([-\rho, \rho]).
\]
For simplicity, we set $b_y = b_y(1)$.

In view of Proposition 3.9, we have
\[
\sup_{y \in Y_0} \frac{b_y(2)}{b_y} \leq d \cdot p_y
\]
where $d$ is an absolute constant. Fix $\sigma > 1$, and let $0 < c_0 = c_0(\sigma) < 1$ be so that
\[
\kappa := 2c_0\sigma dp_y < 1. \tag{5.5}
\]

Let $D_0$ be as in the assumption (b) and put
\[
\hat{D} := 2D_0 dp_y. \tag{5.6}
\]

In the sequel we will use the fact that $\sum_{n \geq 1} e^{-n\tau} \leq 1/2$ — recall that $\tau \geq 2$. The main step in the proof is the following estimate.

Claim 5.7. For any $1 \leq \rho \leq 2$ and any $y \in Y_0$, we have
\[
A_{(n+1)\tau, \rho}F(y) \leq \kappa A_{n\tau, \rho + e^{-n\tau}}F(y) + \hat{D}. \tag{5.8}
\]

Let us first assume Claim 5.7 and finish the proof. Iterating (5.8) with $\rho = 1$ and using the fact that $\sum e^{-n\tau} \leq 1/2$, we have that for any $n \geq 2$,
\[
A_{n\tau}F(y) \leq \kappa^{n-1}A_{\tau, 1 + \sum e^{-j\tau}}F(y) + \hat{D}(1 + \kappa + \cdots + \kappa^{n-2})
\leq \kappa^{n-1}(c_0 F(y) + D_0) + \hat{D}(1 + \kappa + \cdots + \kappa^{n-2})
\leq \kappa^n F(y) + \hat{D}
\]
where $\hat{D} = \frac{\hat{D}}{1 - \kappa}$ and for the second inequality we used our assumption (b) with $\rho = 1 + \sum e^{-j\tau}$. 

Repeatedly applying the assumption in part (a), there exists some \( \sigma_1 = \sigma_1(\sigma) > 1 \) so that
\[
\sigma_1^{-1} F(y) \leq F(y + h) \leq \sigma_1 F(y)
\]
for all \( h \in \{ u_r : |r| \leq 2 \} \{ a_t : |t| \leq \tau \} \{ u_r : |r| \leq 2 \} \) and all \( y \in Y_0 \).

Let \( n_c = n_c(\sigma) \geq 1 \) be large enough so that \( \kappa^n \sigma_1 < c \) for all \( n \geq n_c \). Let now \( t \geq n_c \tau \) and write \( t = n \tau + \ell \) where \( 0 \leq \ell < \tau \). Then
\[
A_t F(y) = \frac{1}{b_z} \int_{-1}^1 F(y u_r a_{n \tau + \ell}) d\mu_y^{PS}(r) \leq \sigma_1 A_{n \tau} F(y).
\]

Since \( \sigma_1 \kappa^n < c \), we get
\[
A_t F(y) \leq \sigma_1 \kappa^n F(y) + \sigma_1 \tilde{D} \leq c F(y) + D
\]
where \( D = \sigma_1 \tilde{D} \).

In consequence, assuming Claim 5.7, we obtain part (1) with \( t_c = n_c \tau \) and \( D \) as above.

To prove part (2), apply part (1) with \( c = 1 / F(y) \).

**Proof of Claim 5.7.** We now turn to the proof of (5.8). To ease the notation, we prove this with \( \rho = 1 \); the proof in general is similar.

By our assumption in part (b), we have
\[
A_t F(y) \leq c_0 F(y) + D_0 \leq \left( \frac{c_0 \sigma}{b_y} \int_{-1}^1 F(y u_r) d\mu_y^{PS}(r) \right) + D_0.
\]

Set \( \rho_n := e^{-n \tau} \). Let \( \{ [r_j - \rho_n, r_j + \rho_n] : j \in J \} \) be a covering of \([-1, 1] \cap \text{supp} \( \mu_y^{PS} \)\) with \( r_j \in [-1, 1] \cap \text{supp} \( \mu_y^{PS} \)\) and with multiplicity bounded by 2. For each \( j \in J \), let \( z_j := y u_{r_j} \). Then
\[
\sum_j b_{z_j}(\rho_n) = \sum_j \mu_y^{PS} (z_j u_{[r_j - \rho_n, r_j + \rho_n]}) \leq 2b_y(2).
\]

Moreover, we can compute
\[
A_{(n+1)\tau} F(y) = \frac{1}{b_y} \int_{-1}^1 F(y u_r a_{(n+1)\tau}) d\mu_y^{PS}(r)
\]
\[
\leq \frac{1}{b_y} \sum_j \int_{-\rho_n}^{\rho_n} F(z_j u_r a_{(n+1)\tau}) d\mu_{z_j}(r)
\]
\[
= \frac{1}{b_y} \sum_j \int_{-\rho_n}^{\rho_n} F(z_j a_{n \tau} u_{r a_{\tau}}) d\mu_{z_j}^{PS}(r).
\]

We now make the change of variables \( t = re^{n \tau} \). Therefore, in view of (5.11), we have
\[
A_{(n+1)\tau} F(y) \leq \frac{1}{b_y} \sum_j \frac{b_{z_j}(\rho_n)}{b_{z_j} a_{\tau}} \int_{-1}^1 F(z_j a_{n \tau} u_{t a_{\tau}}) d\mu_{z_j}^{PS}(t).
\]
Applying (5.9) with the base point \( z_j a_{n, \tau} \), we get from the above that

\[
A_{(n+1)\tau} F(y) \leq \frac{1}{b_y} \sum_j b_j(\rho_n)c_0\sigma \int_{-\rho_n}^{\rho_n} F(z_j u_{n, \tau}) d\mu_{z_j}^{PS}(t) + \frac{1}{b_y} \sum_j b_j(\rho_n) D_0.
\]

By (5.10), we have

\[
\frac{1}{b_y} \sum_j b_j(\rho_n) D_0 \leq \hat{D};
\]

see (5.6). Therefore, reversing the change of variable, i.e., now letting \( t = e^{-n\tau}r \), we get from (5.12) the following:

\[
A_{(n+1)\tau} F(y) \leq \frac{1}{b_y} \sum_j c_0\sigma \int_{-\rho_n}^{\rho_n} F(z_j u_{n, \tau}) d\mu_{z_j}^{PS}(t) + \hat{D}
\]

\[
\leq \frac{2c_0\sigma}{b_y} \int_{-\rho_n}^{\rho_n} F(y u_{n, \tau}) d\mu_{y}^{PS}(r) + \hat{D}
\]

\[
= \frac{2c_0\sigma b_y(1 + \rho_n)}{b_y \cdot (1 + \rho_n)} \int_{-\rho_n}^{\rho_n} F(y u_{n, \tau}) d\mu_{y}^{PS}(r) + \hat{D}
\]

\[
\leq \kappa A_{n, \tau, 1 + \rho_n} F(y) + \hat{D}.
\]

as \( \kappa \geq \sup_{y \in Y_0} \frac{2c_0\sigma b_y(2)}{b_y} \).

\[\square\]

**Lemma 5.14.** If \( F : Y_0 \to [2, \infty) \) is a continuous function satisfying (1) and (2) in Proposition 5.3, then \( F \in L^1(Y_0, m_Y) \).

**Proof.** The claim follows from the Birkhoff’s ergodic theorem as we now explicate. For every \( R > 2 \), let \( F_R : Y_0 \to \mathbb{R} \) be given by \( F_R(y) := \min\{F(y), R\} \). As \( F_R \) is bounded, it belongs to \( L^1(Y_0, m_Y) \). Let \( R_0 \) be large enough so that \( \int F_R \, dm_Y > 0 \) for all \( R > R_0 \); fix some \( R > R_0 \).

Therefore, in view of the A-ergodicity of \( m_Y \), for \( m_Y \)-a.e. \( y \in Y_0 \), we have

\[
\frac{1}{T} \int_0^T F_R(ya_t) \, dt \to \int F_R \, dm_Y.
\]

Hence, using Egorov’s theorem, for every \( \varepsilon > 0 \) there exists a subset \( Y_{\varepsilon}' \subset Y_0 \) with \( m_Y(Y_{\varepsilon}') > 1 - \varepsilon^2 \) and some \( T_\varepsilon > 1 \) so that for every \( y \in Y_{\varepsilon}' \) and all \( T > T_\varepsilon \) we have

\[
\frac{1}{T} \int_0^T F_R(ya_t) \, dt \geq \frac{1}{2} \int F_R \, dm_Y.
\]

Now by the maximal ergodic theorem [17, Thm. 17], if \( \varepsilon \) is small enough, there exists a subset \( Y_\varepsilon \subset Y_{\varepsilon}' \) with \( m(Y_\varepsilon) > 1 - \varepsilon \) so that for all \( y \in Y_\varepsilon \) we have

\[
\mu_y^{PS}\{r \in [-1, 1] : y u_r \in Y_\varepsilon'\} > \frac{1}{2} \mu_y^{PS}([-1, 1]).
\]
Altogether, if \( y \in Y_\varepsilon \) and \( T > T_\varepsilon \), we have
\[
(5.15) \quad \frac{1}{\mu_y^\text{PS}([-1, 1])} \int_{-1}^{1} \frac{1}{T} \int_{0}^{T} F_R(yu_t a_t) dt \, d\mu_y^\text{PS}(r) > \frac{1}{4} \int F_R \, dm_Y.
\]

Fix some \( y \in Y_\varepsilon \) and apply Proposition 5.3(1) with this \( y \) and \( r = 1/4 \). Then for all large enough \( t \), we have
\[
\frac{1}{\mu_y^\text{PS}([-1, 1])} \int_{-1}^{1} F_R(yu_t a_t) d\mu_y^\text{PS}(r) \leq \frac{1}{\mu_y^\text{PS}([-1, 1])} \int_{-1}^{1} F(yu_t a_t) d\mu_y^\text{PS}(r) < \frac{1}{4} + D.
\]

Averaging this over \( t \) and using (5.15), we get that
\[
\int F_R \, dm_Y < F(y) + 4D.
\]

Since \( R > R_0 \) is arbitrary, the claim follows. \( \square \)

We will also use the following equidistribution theorem in the sequel.

**Theorem 5.16** (Equidistribution of expanding horocycles [16], [14]). For any bounded \( \psi \in C(Y_0) \), \( y \in Y_0 \), and \( \rho > 0 \), we have
\[
\lim_{t \to +\infty} A_{t, \rho} \psi(y) = m_Y(\psi).
\]

### 6. A Linear Algebra Lemma

In this section, it is more convenient to use the isomorphism between \( G \) and \( \text{SO}(Q) \) for the quadratic form
\[
Q(x_1, x_2, x_3, x_4) = 2x_1x_4 + x_2^2 + x_3^2.
\]

We consider the standard representation of \( G \) on \( \mathbb{R}^4 \). We have \( H = \text{Stab}_G(\varepsilon_3) \simeq \text{SO}(2, 1)^{\circ} \), and
\[
A = \{ a_t = \text{diag}(e^t, 1, 1, e^{-t}) \} < H.
\]

Set \( V := \mathbb{R} e_1 \oplus \mathbb{R} e_2 \oplus \mathbb{R} e_4 \). Then the action of \( H \) on \( V \) is isomorphic to the adjoint representation of \( H \) on its Lie algebra; in particular, it is irreducible.

**Lemma 6.1.** For any \( Y \in \mathcal{H} \), \( 0 < s < \frac{s_0}{2} \), and \( 1 \leq \rho \leq 2 \), we have
\[
(6.2) \quad \sup_{y \in Y_0, \|v\| = 1} \frac{1}{\mu_y^\text{PS}([-\rho, \rho])} \int_{-\rho}^{\rho} \frac{1}{\|vu_t a_t\|^s} d\mu_y^\text{PS}(r) \leq c_1 \|y\| e^{-st/2}
\]
where \( c_1 \) depends only on \( V \).

**Proof.** Since \( \mu_y^\text{PS}[-\rho, \rho] = \rho^{s_0} \mu_y^\text{PS}[-\log, \rho] [-1, 1] \) and \( Y_0 \) is \( A \)-invariant, it suffices to prove the claim for \( \rho = 1 \). We write
\[
V = V^- \oplus V^0 \oplus V^+
\]
where \( V^0 = \{ v \in V : vu_t = v \text{ for all } t \in \mathbb{R} \} \), \( V^+ \) denotes the sum of eigenspaces with eigenvalue \( > 1 \) and \( V^- \) denotes the sum of eigenspaces with eigenvalue \( < 1 \) for \( A = \{ a_t \} \).
Let \( p^+ : V \rightarrow V^+ \) denote the natural projection. For any unit vector \( v \in V \) and any \( \varepsilon > 0 \), define
\[
D^+(v, \varepsilon) = \{ r \in [-1, 1] : |p^+(vu_r)| \leq \varepsilon \}.
\]

We claim that the Lebesgue measure of \( D^+(v, \varepsilon) \) is at most \( c \cdot \varepsilon^{1/2} \) where \( c \) is an absolute constant.

To see this, note that the map \( r \mapsto p^+(vu_r) \) is a polynomial map of degree at most 2. Moreover, since \( \|v\| = 1 \), we have maximum of the absolute value of the coefficients of \( p^+ \) is bounded from below by an absolute constant. In particular, \( \sup_{r \in [-1,1]} \|p^+(vu_r)\| \geq c' \) for some absolute \( c' > 0 \) (depending only on 2). The claim now follows using Lagrange’s interpolation, see [2] for a more general statement.

Together with the fact that \( D^+(v, \varepsilon) \) is a union of at most 2 intervals, this claim implies that \( D^+(v, \varepsilon) \) may be covered by \( \preceq 1 \) many intervals of length \( \varepsilon^{1/2} \). Now by the definition of \( p_Y \) in (3.15), we get
\[
\mu^\text{PS}_y(D^+(v, \varepsilon)) \leq c_0 p_Y \varepsilon^{s_0/2} \mu^\text{PS}_y([-1,1])
\]
where \( c_0 \) is an absolute constant.

Recall that \( s < s_0 := \frac{\delta_Y}{2} \). Note also that \( \|vu_r a_t\| \geq e^{t/2} \|p^+(vu_r)\| \). In consequence, for all \( 0 < \varepsilon \leq 1/2 \) and all \( \|v\| = 1 \), we have
\[
\frac{1}{\mu^\text{PS}_y([-1,1])} \int_{r \in D^+(v, \varepsilon)} \frac{1}{\|vu_r a_t\|^s} d\mu^\text{PS}_y(r) \leq c_0 p_Y \varepsilon^{s_0} \cdot (e^{t/2} \varepsilon/2)^{-s} = c_0 2^s p_Y \varepsilon^{s_0 - s} e^{-st/2}.
\]

Summing up the geometric series we get
\[
\frac{1}{\mu^\text{PS}_y([-1,1])} \int_{r \in D^+(v, 1/2)} \frac{1}{\|vu_r a_t\|^s} d\mu^\text{PS}_y(r) \leq \frac{c_0 p_Y}{2^{-s} - 2^{-s_0}} e^{-st/2}.
\]

On the other hand,
\[
\frac{1}{\mu^\text{PS}_y([-1,1])} \int_{r \in [-1,1] \setminus D^+(v, 1/2)} \frac{1}{\|vu_r a_t\|^s} d\mu^\text{PS}_y(r) \leq 2e^{-st/2}.
\]

Combining (6.4) and (6.5) we get that
\[
\frac{1}{\mu^\text{PS}_y([-1,1])} \int_{-1}^{1} \frac{1}{\|vu_r a_t\|^s} d\mu^\text{PS}_y(r) \leq \left( \frac{c_0 p_Y}{2^{-s} - 2^{-s_0}} + 2 \right) e^{-st/2},
\]
as was claimed in the lemma.

\[\square\]

7. Height function \( \omega \)

In this section we use the thick-thin decomposition of RFM to define the notion of height of \( x \in \text{RFM} \). For this purpose, we continue to use the isomorphism between \( G \) and \( \text{SO}(Q)^\circ \), and consider the standard representation of \( G \) on \( \mathbb{R}^4 \). By the choice of the quadratic form \( Q \), \( Q(e_1) = 0 \) and the stabilizer of \( e_1 \) is equal to \( MN \).
We assume that $\Gamma$ is not convex cocompact, so that $\Lambda_p \neq \emptyset$; otherwise the content of this section is void. We fix $\xi_1, \cdots, \xi_\ell$ a set of $\Gamma$-representatives in $\Lambda_p$, and choose $g_i \in G$ so that $g_i^{-1} = \xi_i$. Then

$$\text{Stab}_G(\xi_i) = g_iAMG_0Ng_i^{-1}.$$ 

For each $1 \leq i \leq \ell$, let $v_i \in \mathbb{R}^4$ be a unit vector on the line $\mathbb{R}e_1g_i^{-1}$. Then

$$g_iM_0Ng_i^{-1} = \text{Stab}_G(v_i)$$

and by Witt’s theorem,

$$\{v \in \mathbb{R}^4 : v \neq 0, Q(v) = 0\} = v_iG \simeq g_iM_0Ng_i^{-1}\backslash G.$$ 

**Lemma 7.2.** For each $1 \leq i \leq \ell$, the orbit $v_i\Gamma$ is a closed (and hence discrete) subset of $\mathbb{R}^4$.

**Proof.** The condition $\xi_i \in \Lambda_p$ implies that $\Gamma \backslash \Gamma g_iM_0N$ is a closed subset of $X$. Equivalently, $\Gamma g_iM_0N$, as well as $\Gamma g_iM_0Ng_i^{-1}$ is closed in $G$. Therefore, its inverse $g_iM_0Ng_i^{-1}$ is a closed subset of $G$. By (7.1), $v_i\Gamma \subset \mathbb{R}^4$ is a closed subset of $v_iG = \{v \in \mathbb{R}^4 : v \neq 0, Q(v) = 0\}$.

It is now enough to show that $v_i\Gamma$ does not accumulate on $0$. Suppose on the contrary that there exists an infinite sequence $v_i\gamma_\ell \to 0$ for some $\gamma_\ell \in \Gamma$. Using the Iwasawa decomposition $G = g_iNAGg_i^{-1}K_0$, we may write $\gamma_\ell = g_i n_\ell a_{t_\ell} g_i^{-1} k_\ell$ with $n_\ell \in N, t_\ell \in \mathbb{R}$ and $k_\ell \in K_0$. Since

$$v_i\gamma_\ell = e^{t_\ell}(e_1g_i^{-1}k_\ell),$$

the assumption that $v_i\gamma_\ell \to 0$ implies that $t_\ell \to -\infty$.

On the other hand, as $\xi_i \in \Lambda_p$, $\text{Stab}_G(\xi_i) = \Gamma \cap g_iAMG_0Ng_i^{-1}$ contains a parabolic element, say, $\gamma'$. Note that $n_0 := g_i^{-1}\gamma' g_i$ is then an element of $N$, as any parabolic element of $AMG_0N$ belongs to $N$ in the group $G \simeq \text{PSL}_2(\mathbb{C})$. Now observe that

$$\gamma_\ell^{-1}\gamma' \gamma_\ell = k_\ell^{-1} g_i a_{-t_\ell}(n_\ell^{-1} g_i^{-1} \gamma' g_i n_\ell) a_{t_\ell} g_i^{-1} k_\ell = k_\ell^{-1} g_i (a_{-t_\ell} n_0 a_{t_\ell}) g_i^{-1} k_\ell$$

as $N$ is abelian.

Since $t_\ell \to -\infty$, the sequence $a_{-t_\ell} n_0 a_{t_\ell}$ converges to $e$. Since $k_\ell^{-1} g_i$ is a bounded sequence, it follows that, up to passing to a subsequence, $\gamma_\ell^{-1}\gamma' \gamma_\ell$ is an infinite sequence converging to $e$, which contradicts the discreteness of $\Gamma$. 

We let $\| \cdot \|$ be the Euclidean norm on $\mathbb{R}^4$. For each $1 \leq i \leq \ell$, define the function $\omega_i : X_0 \to [2, \infty)$ as follows: for $[g] \in X_0$,

$$\omega_i([g]) = \max_{\gamma \in \Gamma} \left\{ \frac{1}{2, \|v_i\gamma g\|} \right\};$$

this is well-defined by Lemma 7.2.

**Definition 7.3** (Height function). Define the height function $\omega : X_0 \to [2, \infty)$ by

$$\omega(x) := \max_{1 \leq i \leq \ell} \omega_i(x).$$
It will be convenient to introduce the following notation:

**Notation 7.4.** Let $Q \subset G$ be a compact subset.

1. Let $d_Q \geq 1$ be the infimum of all $d \geq 1$ such that for all $g \in Q$ and $v \in \mathbb{R}^4$,
   \begin{equation}
   d^{-1}||v|| \leq ||vg|| \leq d||v||.
   \end{equation}
   Note that $d_Q \asymp \max_{g \in Q} ||g||$, up to an absolute multiplicative constant.

2. We also define $c_Q \geq 1$ to be the infimum of all $c \geq 1$ such that for any $x \in X_0$, $g \in Q$ with $xg \in X_0$, and for all $1 \leq i \leq \ell$
   \begin{equation}
   c^{-1}\omega_i(x) \leq \omega_i(xg) \leq c\omega_i(x).
   \end{equation}
   We note that $c_Q \asymp \max_{g \in Q} ||g||$ up to an absolute multiplicative constant.

**Proposition 7.7.** For all $x \in X_0$,
   \[ \omega(x) \asymp \text{inj}(x)^{-1} \]
   with an absolute implied constant. In particular, $\omega : X_0 \to [2, \infty)$ is a proper map satisfying that for all $x \in RFM$,
   \[ \omega(x) \asymp d(x, X_{x_0} \cap RFM) \]
   where the implied constant is independent of $x \in RFM$.

**Proof.** Recall that a horoball in $X_0$ is of the form $X_0 \cap [g]NA_{(-\infty, T]}K_0$ for some $T \gg 1$. In view of the thick-thin decomposition, it suffices to show the claim for $[g] = X_0 \cap [g]NA_{(-\infty, T]}K_0$.

Let $g = \gamma g g u a_{-t} \gamma_k$ where $u a_{-t} \gamma_k \in NA_{(-\infty, T]}K_0$ and $\gamma g \in \Gamma$; so inj([g]) $\asymp e^{-t}$. Note that
   \begin{equation}
   e^{-t} = ||e_1 u a_{-t} \gamma_k|| = ||e_1 g_j^{-1} \gamma g u a_{-t} \gamma_k|| = ||e_1 g_j^{-1} \gamma g^{-1} g ||.
   \end{equation}

In view of the definition of $\omega$ and $\omega_i$, the claim will follow if we show the following: there exists some $c \asymp 1$ so that for all $1 \leq i \leq \ell$ and $\gamma \in \Gamma$ such that $\gamma^{-1} h_i \neq \gamma g h_j$, we have
   \begin{equation}
   ||e_1 g_i^{-1} \gamma g || \geq c
   \end{equation}
   where $h_j = g_j NA_{(-\infty, T]}K_0$.

Write $g_i^{-1} = u_i a_i k_i$ and let $\tau \geq 0$ be so that $|t_i| = e^\tau$ for all $i$. Assume now that there exists some $i$ so that
   \[ ||e_1 g_i^{-1} \gamma g || \leq e^{-T-\tau-1}. \]

Then $\gamma g = g_i u a_{-s} g_i^{-1} k'$ where $u' \in N$, $k' \in K_0$ where $s \geq T + \tau + 1$. In view of the choice of $\tau$, this implies that $g \in \gamma^{-1} h_i$. Since $g \in \gamma g h_j$, we get that $i = j$ and $\gamma^{-1} h_j = \gamma g h_j$.

This implies that (7.9) holds with $c = e^{-T-\tau-1}$ and finishes the proof. \hfill \Box
Using [9, Prop. 5.5], one can show that the restriction of \( \omega^s \) to \( Y \) belongs to \( L^1(Y, m_Y) \) if and only if \( 0 < s < 2\delta_Y - 1 \). To keep this paper as self-contained as possible, however, we will use Lemma 5.14 for the estimates we need in Proposition 8.6.

8. Averaging operator for the height function

In this section, we fix a closed \( H \)-orbit \( Y \in \mathcal{H} \) and fix

\[
0 < s < \frac{\delta_Y^*}{2}.
\]

The main goal of this section is to prove the following property of the averaging operator applied to the height function \( \omega^s \) (restricted to \( Y_0 \)):

**Theorem 8.1.** (1) There exists \( 0 < D_1 \ll p_Y^* \) with the following property: for any \( 0 < c \leq 1/2 \), there exists some \( 0 < t(c, p_Y) \ll |\log c| + \log p_Y \) such that for any \( y \in Y_0 \) and all \( t \geq t(c, p_Y) \),

\[
A_t \omega^s(y) \leq c \cdot \omega^s(y) + D_1. \tag{8.2}
\]

(2) For every \( y \in Y_0 \), we have

\[
\limsup_{t \to \infty} A_t \omega^s(y) \leq 1 + D_1. \tag{8.3}
\]

The proof of this theorem is based on Proposition 5.3; indeed the function \( \omega^s \) satisfies the condition in Proposition 5.3(a) by (7.6). We will now show that it also satisfies the condition in Proposition 5.3(b).

**Lemma 8.4.** For every \( 0 < c \leq 1/2 \), there exist some \( 0 < \tau \ll |\log c| + \log p_Y \) and some \( 1 \leq D_2 \ll p_Y^*/c^* \) so that for all \( y \in Y_0 \) and all \( 1 \leq \rho \leq 2 \),

\[
A_{\tau, \rho} \omega^s(y) \leq c \cdot \omega^s(y) + D_2. \tag{8.5}
\]

**Proof.** Since \( Q(e_1) = 0 \) and \( G = SO(Q)^o \), we have \( Q(e_1 g) = 0 \) for every \( g \in G \); in particular, there exists an absolute constant \( \eta > 0 \) so that for every unit vector \( v \in e_1 G \), we have

\[
\|v - e_3\| \geq \eta. \tag{8.5}
\]

For every \( v \in \mathbb{R}^4 \), let us write \( v = v_0 + v_1 \) where \( v_0 \in \mathbb{R}e_3 \) and \( v_1 \in V = \mathbb{R}e_1 \oplus \mathbb{R}e_2 \oplus \mathbb{R}e_4 \).

Let \( v \in e_1 g \) be as above. Moreover, since \( \mathbb{R}^4 = \mathbb{R}e_3 \oplus V \) as an \( H \)-representation and since \( e_3 \) is \( H \)-invariant, we have \( vh = v_0 + v_1 h \in \mathbb{R}e_3 \oplus V \) for all \( h \in H \). Therefore,

\[
\frac{1}{\mu^\text{PS}([-\rho, \rho])} \int_{-\rho}^{\rho} \frac{1}{|v_{u^r a_t}|^s} d\mu^\text{PS}_y(r) \leq \frac{1}{\mu^\text{PS}([-\rho, \rho])} \int_{-\rho}^{\rho} \frac{1}{|v_{1 u^r a_t}|^s} d\mu^\text{PS}_y(r)
\]

\[
\leq c_1 p_Y e^{-st/2} \|v_1\|^{-s} \text{ by Lemma 6.1}
\]

\[
\leq d_1 p_Y e^{-st/2} \|v\|^{-s} \text{ by (8.5)}
\]

for some absolute constant \( d_1 > 0 \).
Let \( \tau \) be given by the equation:
\[
d_1 p_Y e^{-s \tau/2} = c.
\]
We claim that the lemma holds with this choice of \( \tau \). To see this, we compare \( \omega(yu, a_r) \) and \( \omega(y) \) for \( r \in [-2, 2] \). Setting
\[
Q := \{ u_r : |r| \leq 2 \} \{ a_t : |t| \leq \tau \} \{ u_r : |r| \leq 2 \},
\]
we have \( c_Q \times e^\tau \times p_Y^2 / c^2 \), see (7.6).

Recall the constant \( \varepsilon_0 \) which satisfies
\[
\varepsilon_0^{-1} \gg \sup_{y \in X_{c_0} \cap Y_0} \omega(y).
\]
We consider two cases.

**Case 1:** \( \omega(y) \leq 2c_Q / \varepsilon_0 \). In this case, for \( h \in Q \) with \( yh \in Y_0 \),
\[
\omega(yh) \leq 2c_Q / \varepsilon_0.
\]
Hence, the claim in this case follows if we choose \( D_2 > 2c_Q / \varepsilon_0 \).

**Case 2:** \( \omega(y) > 2c_Q / \varepsilon_0 \). For some \( 1 \leq i \leq \ell \), we have
\[
\omega_i(y) > 2c_Q / \varepsilon_0, \quad \text{and hence } y \in h_i.
\]
By the definition of \( c_Q \), we have
\[
\omega_i(yh) > 2\varepsilon_0, \quad \text{and hence } yh \in h_i
\]
for all \( h \in Q \) with \( yh \in Y_0 \). Choose \( g_0 \in G \) so that \( y = [g_0] \). In view of (7.9) and (7.8), there exists \( \gamma \in \Gamma \) such that simultaneously for all \( h \in Q \) with \( yh \in Y_0 \),
\[
\omega(yh) = \| v_i \gamma g_0 h \|^{-1}.
\]
Therefore, we have
\[
A_{\tau, \rho} \omega^s(y) = \frac{1}{\mu_y^{PS}([-\rho, \rho])} \int_{-\rho}^{\rho} \frac{1}{\| v_i \gamma u_r a_r \|^s} d\mu_y^{PS}(r)
\leq c_1 p_Y e^{-s \tau/2} \| v_i \gamma \|^{-s} = c \cdot \omega^s(y).
\]

The proof is now complete. \( \square \)

**Proof of Theorem 8.1.** The conditions in Proposition 5.3 holds with \( F = \omega^s \); indeed, condition (a) follows from (7.6) and condition (b) holds in light of Lemma 8.4. The theorem thus follows from Proposition 5.3. \( \square \)

**Equidistribution for an unbounded function \( \omega^s \).** Using Theorem 8.1 we will deduce the following proposition from Theorem 5.16 and Lemma 5.14.

**Proposition 8.6.** For any \( y \in Y_0 \), we have
\[
\lim_{t \to \infty} A_t \omega^s(y) = m_Y(\omega^s).
\]
Proof. As was mentioned before, we will prove this proposition using the equidistribution Theorem 5.16. Note, however, that $\omega$ is not a bounded function and we may not apply the aforementioned equidistribution results directly. We use the fact that $s < \frac{2\delta}{Y-1}$ and choose $s_0$ so that $s + s_0 < \frac{2\delta}{Y-1}$ to approximate $\omega^s$ by a bounded function.

Let us now turn to the details. Let $R > 1$ be a parameter. Let $0 \leq \chi_R \leq 1$ be a continuous function which is equal to 1 on $\{x \in RFM : \omega(x) \geq R + 1\}$ and is equal to 0 on $\{x \in RFM : \omega(x) \leq R\}$. Set

$$h_R := \omega^s - \omega^s\chi_R;$$

the function $h_R$ is a continuous function with compact support.

By Theorem 5.16, we have

$$\lim_t \omega^s(y) = m_Y(h_R).$$

Moreover, in view of Theorem 8.1 that $\omega^s$ satisfies (1) and (2) in Proposition 5.3. Hence, using Lemma 5.14, we have

$$|m_Y(\omega^s) - m_Y(h_R)| \rightarrow 0 \text{ as } R \rightarrow \infty.$$

Let $s_1 = s + s_0$; then $s_1 < \frac{2\delta}{Y-1}$. Note that, using $\omega^{-s_0} \leq R^{-s_0}$ on $\{\omega(x) \geq R\}$, we have

$$(\omega^s\chi_R)(y) \leq R^{-s_0} \cdot \omega^{s_1}(y),$$

and hence

$$A_t(\omega^s\chi_R)(y) \leq R^{-s_0} \cdot A_t\omega^{s_1}(y).$$

Now by Theorem 8.1(2), applied with $s_1$, we have

$$\limsup_t A_t\omega^{s_1}(y) \leq 1 + D_1$$

which in view of the above computation implies

$$A_t(\omega^s\chi_R)(y) \leq R^{-s_0}(1 + D_1).$$

Therefore for any $R > 1$, using (8.7),

$$\limsup_t |A_t\omega^{s}(y) - m_Y(\omega^s)|$$

$$\leq \limsup_t |A_t(\omega^s \cdot \chi_R)| + \limsup_t |A_t(h_R) - m_Y(h_R)| + |m_Y(h_R) - m_Y(\omega^s)|$$

$$\leq R^{-s_0}(1 + D_1) + |m_Y(h_R) - m_Y(\omega^s)| \text{ by (8.10).}$$

Taking $R \rightarrow \infty$ and using (8.8), we get

$$\limsup_t |A_t\omega^{s}(y) - m_Y(\omega^s)| = 0.$$

The proof of complete. □
Remark 8.11. Since no horoball can contain a complete geodesic, it follows that $Y_0$ intersects $X_{\epsilon_0}$. This fact, however, is not used in the discussion here. Indeed, similar to the proof of Theorem 1.4(2), we get a non-divergence result from Theorem 8.1 and Proposition 8.6 with explicit constants depending on $\Gamma$ and $p_Y$.

9. Construction of a Margulis function

Throughout this section, we fix $Y, Z \in H$ and

$$0 < s < \frac{\delta_0}{2}.$$ 

The Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ of $G$ decomposes as $\mathfrak{sl}_2(\mathbb{R}) \oplus i\mathfrak{sl}_2(\mathbb{R})$. We write $V = i\mathfrak{sl}_2(\mathbb{R})$ and consider the action of $H$ on $V$ via the adjoint representation; so $vh = Ad(h^{-1})(v)$ for $v \in V$ and $h \in H$. We use the relation $g(exp v)h = gh exp(vh)$ which is valid for all $g \in G, v \in V, h \in H$.

For any $E \geq 1$ and any $y \in Y_0$, we define

$$(9.1) \quad I_Z(y, E) = \{ v \in V : \|v\| \leq \omega(y)^{-E} p_Y^{-E}, y \exp(v) \in Z \}.$$ 

Recall the following notation from (4.4) and (4.5):

$$X_0' = \{ x \in X : \exists x_0 \in X_0 \text{ with } d(x, x_0) \leq \text{inj}(x) \}, \quad \text{and} \quad \tau_Z = \max \{ 1, \text{area}(S_Z \cap \pi(X_0')) \}.$$ 

Lemma 9.2. There exist $E, L, c_0 \geq 1$ so that

$$\sup_{y \in Y_0} \# I_Z(y, E) \leq c_0 p_Y^L \cdot \tau_Z.$$ 

Proof. Fix $0 < D_1 \ll p_Y^*$ as given by Theorem 8.1 for $Y$. Set

$$\mathcal{K} := \{ x \in Y_0 : \omega(x) \leq 1 + D_1 \}.$$ 

Let $E_0 > 1$ be given by the conditions that $\exp : \mathfrak{sl}_2(\mathbb{C}) \to G$ is a local diffeomorphism on the $2^{-E_0}$-neighborhood of 0, and for any $x \in X$

$$\inf \{ \text{inj}(x \exp v) : \|v\| \leq 2^{-E_0+1} \} \geq \frac{1}{2} \text{inj}(x).$$

There exists $c \geq 1$ such that for any $x \in X$ and $v \in V$ with $\|v\| \leq 2^{-E_0}$,

$$d(x, x \exp(v)) \leq c \|v\|^c.$$

Recall from Lemma 7.7, that $\omega^{-1}(x) \asymp \text{inj}(x)$ for all $x \in X_0$. Increasing $E_0$ if necessary, we may assume that for all $x \in X_0$,

$$c \cdot \omega(x)^{-cE_0} < \text{inj}(x).$$

Since $p_Y \geq 1$, it follows that for any $y \in Y_0$ and $v \in I_Z(y, E_0)$,

$$(9.3) \quad d(y, y \exp(v)) \leq c \|v\|^c \leq c \cdot \omega(y)^{-cE_0} p_Y^{-cE_0} \leq c \cdot \omega(y)^{-cE_0} \text{inj}(y).$$

We first claim that there exists some $c_0 \geq 1$ so that for any $y \in \mathcal{K}$,

$$(9.4) \quad \# I_Z(y, E_0) \leq c_0 p_Y^* \cdot \tau_Z.$$
Note that as \( y \in Y_0 \subset X_0 \) and \( v \in I_Z(y, E_0) \), we have from (9.3) that
\[
y \exp v \in Z \cap X'_0.
\]
Since \( \omega \geq 2 \), for all \( v \in I_Z(y, E_0) \) we have \( \text{inj}(y \exp v) \geq \frac{1}{2} \text{inj}(y) \). Let
\[
\rho = \min\{1, \frac{1}{2} \text{inj}(y)\}. \quad \text{Then}
\]
\[
\#I_Z(y, E_0) \cdot \text{Vol}(B_H(y, \rho)) \leq \text{Vol}(\bigcup B_H(y \exp v, \rho) : v \in I_Z(y, E_0)) \leq \tau_Z
\]
since \( \rho \leq \text{inj}(y \exp v) \) and \( y \in X_0 \). Since \( \text{inj}(y)^{-1} \times \omega(y) \leq 1 + D_1 \leq p_Y^* \) by Proposition 7.7 and the dimension of \( Z \) is equal to 3, we get that \( \text{Vol}(B_H(y, \rho))^{-1} \ll p_Y^* \); hence (9.4) follows.

Let now \( y \in Y_0 - K \). By taking \( c = \frac{1}{2\omega(y)} \) in Theorem 8.1, there exists some \( t := t(y) \leq \star \log(\omega(y)p_Y) \) so that
\[
A_t \omega^s(y) < 1 + D_1.
\]
Therefore \( y' := yu_{r_0}a_t \) belongs to \( K \) for some \( r_0 \in \text{supp}(\mu^*_{y}) \cap [-1, 1] \).

Note that
\[
\sup_{|r| \leq 1, s \leq \log(\omega(y)p_Y)^t} \|v(u_r a_s)\| \leq c' \cdot (\omega(y)p_Y)^\star \cdot \|v\|
\]
for some absolute constant \( c' \geq 1 \). Since \( p_Y \geq 1 \) and \( \omega \geq 2 \), for \( E := E_0 + \star \), we have if \( \|v\| \leq \omega(y)^{-E}p_Y^{-E} \), then
\[
\sup_{|r| \leq 1, s \leq \log(\omega(y)p_Y)^t} \|v(u_r a_s)\| \leq c' \omega(y)^{-E+\star}p_Y^{-E+\star} \leq \frac{1}{2} \omega(y)^{-E_0}p_Y^{-E_0} \leq \omega(y')^{-E_0}p_Y^{-E_0}
\]
using \( \omega(y) \geq \omega(y') \).

Since \( y(\exp v)u_{r_0}a_t = y' \exp(v(u_{r_0}a_t)) \in Z \) and \( y' \in K \subset Y_0 \), we thus get that
\[
y(\exp v)u_{r_0}a_t = y' \exp(v(u_{r_0}a_t)) \in Z \cap X'_0.
\]
Altogether, it follows that the map \( v \mapsto v(u_{r_0}a_t) \) gives an injective map from \( I_Z(y, E) \) into \( I_Z(y', E_0) \). Therefore by (9.4),
\[
\#I_Z(y, E) \leq \#I_Z(y', E_0) \leq c_0p_Y^* \cdot \tau_Z
\]
as was claimed. \( \square \)

**Definition 9.5.** Define \( f_s := f_s, y, Z : Y_0 \to (0, \infty] \) by
\[
(9.6) \quad f_s(y) := \begin{cases}
\omega(y)^sE p_Y^sE & \text{if } I_Z(y) = \emptyset \\
\sum_{v \in I_Z(y)} \|v\|^{-s} & \text{otherwise}
\end{cases}
\]
where \( E > 0 \) as in Lemma 9.2 and \( I_Z(y) = I_Z(y, E) \).

Note that
\[
(9.7) \quad f_s(y) \geq \omega(y)^sE p_Y^sE
\]
for all \( y \in Y_0 \) and that if \( Z \neq Y \), we have \( f(y) < \infty \) for all \( y \in Y_0 \).
Lemma 9.8. Let $Q \subset H$ be a compact subset. Then for any $y \in Y_0$ and $h \in Q$ such that $yh \in Y_0$, we have

$$f_{s/E}(yh) \leq \sum_{v \in I_Z(yh)} \|vh\|^{-s/E} + c_0 c_Q d_Q^{s/E} \tau_Z p_Y^{L+s} \omega(y)^s$$

where $c_Q$ and $d_Q$ are as in §7.4, $L$ and $c_0$ are as in Lemma 9.2, and the sum is understood as 0 in the case when $I_Z(y) = \emptyset$.

Proof. Let $y \in Y_0$ and $h \in Q$ with $yh \in Y_0$. If $I_Z(yh) = \emptyset$, then by (7.6), we have

$$f_{s/E}(yh) = \omega(y)^s p_Y^s \leq c_Q \omega(y)^s p_Y^s$$

proving the claim.

Now suppose that $I_Z(yh) \neq \emptyset$. Setting

$$\varepsilon := \omega(y)^s p_Y^{-E}/d_Q,$$

we write

$$f_{s/E}(yh) = \sum_{v \in I_Z(yh)} \|v\|^{-s/E}$$

$$= \sum_{v \in I_Z(yh), \|v\|<\varepsilon} \|v\|^{-s/E} + \sum_{v \in I_Z(yh), \|v\|\geq\varepsilon} \|v\|^{-s/E}.$$  \hfill (9.9)

Since by Lemma 9.2, $\# I_Z(yh) \leq c_0 p_Y^{L+1} \tau_Z$, we have

$$\sum_{v \in I_Z(yh), \|v\|\geq\varepsilon} \|v\|^{-s/E} \leq (c_0 p_Y^{L+1} \tau_Z) \varepsilon^{-s/E}$$

$$\leq (c_0 d_Q^{s/E} \tau_Z p_Y^{L+s}) \omega(y)^s.$$  \hfill (9.10)

If there is no $v \in I_Z(yh)$ with $\|v\| \leq \varepsilon$, then this proves the claim by (9.9). If $v \in I_Z(yh)$ satisfies $\|v\| < \varepsilon$, then

$$\|vh^{-1}\| \leq d_Q \varepsilon = \omega(y)^{-E} p_Y^{-E};$$

in particular, $vh^{-1} \in I_Z(y)$. Therefore, by setting $v' = vh^{-1},$

$$\sum_{v \in I_Z(yh), \|v\|<\varepsilon} \|v\|^{-s/E} \leq \sum_{v' \in I_Z(y)} \|v'h\|^{-s/E}.$$  \hfill (9.12)

Together with (9.10), this finishes the proof.

Definition 9.11 (Definition of $F_{s,\lambda}$). For $\lambda \geq 1$, we define $F_{s,\lambda} = F_{s,\lambda,Y,Z} : Y_0 \to \mathbb{R}$ as follows:

$$F_{s,\lambda}(y) = f_{s/E}(y) + \lambda \omega^s(y).$$  \hfill (9.13)

Theorem 9.13. There exists $0 < \lambda \ll p_Y^{s+} \tau_Z$ so that the following holds: there exists $D \ll p_Y^{s+} \tau_Z$ and for every $0 < c \leq 1/2$ there exists some $t_c \ll |\log c| + \log p_Y$ so that for all $t \geq t_c$, we have

$$A_t F_{s,\lambda}(y) \leq c F_{s,\lambda}(y) + D$$

for all $y \in Y_0$. 

This theorem will follow from Proposition 5.3, once we verify that the conditions in that proposition hold. Let us begin with the following log-continuity statement.

**Lemma 9.14.** There exists $1 \leq \sigma \ll \rho_Y^s \tau_Z^s$ so that

$$\sigma^{-1} F_{s, \lambda}(y) \leq F_{s, \lambda}(yh) \leq \sigma F_{s, \lambda}(y)$$

for all $y \in Y_0$ and all $h \in B_H(1)$ so that $yh \in Y_0$.

**Proof.** Since $B_H(1)^{-1} = B_H(1)$, it suffices to show the inequality $\leq$. By Lemma 9.8, applied with $c := c_B H(1)$ and $d := d_B H(1)$, we have that for all $h \in B_H(1)$ with $yh \in Y_0$,

$$f_{s/E}(yh) \leq \sum_{v \in I_Z(y)} \|v\|^{-s/E} + c_0 c^s d^{s/E} \tau_Z p_{Y_Y}^L \omega(y)^s$$

$\leq d^{s/E} \sum_{v \in I_Z(y)} \|v\|^{-s/E} + c_0 c^s d^{s/E} \tau_Z p_{Y_Y}^L \omega(y)^s$.

Recall that $\lambda \ll \rho_Y^s \tau_Z^s$. If $I_Z(y) = \emptyset$, $f_{s/E}(y) \ll \rho_Y^s \tau_Z^s \omega(y)$; hence

$$F_s(yh) \ll \rho_Y^s \tau_Z^s \omega(y)^s \ll \rho_Y^s \tau_Z^s (f_{s/E}(y) + \lambda \omega(y)) \ll \rho_Y^s \tau_Z^s F_s(y)$$

where all the implied constants are absolute constants.

Suppose now that $I_Z(y) \neq \emptyset$. Then

$$f_{s/E}(yh) \leq d^{s/E} f_{s/E}(y) + c^s d^{s/E} \tau_Z p_{Y_Y}^L \omega(y)^s.$$ 

Hence $F_s(yh) \ll \rho_Y^s \tau_Z^s F_s(y)$; again, the implied constants are absolute and we used the fact that $\lambda \ll \rho_Y^s \tau_Z^s$. $\square$

We also have the following lemma which is an analogue of Lemma 8.4.

**Lemma 9.15.** Let the notation be as in Theorem 9.13. For every $0 < c \leq 1/2$, there exist

- $\lambda \ll \rho_Y^s \tau_Z^s c^{-s}$ and
- $\tau \ll |\log c| + \log \rho_Y$

so that for all $y \in Y_0$ and all $1 \leq \rho \leq 2$, we have

$$A_{\tau, \rho} F_{s, \lambda}(y) \leq c F_{s, \lambda}(y) + \lambda D_1$$

where $D_1 > 0$ is as in Theorem 8.1.

**Proof.** Let $y \in Y_0$. Apply Theorem 8.1 with $c/2$. Thus, there exists $\tau_1 \ll \log |c| + \log \rho_Y$ so that for all $\tau \geq \tau_1$ and $1/2 \leq \rho \leq 2$, we have

$$A_{\tau, \rho} \omega(y) < \frac{c}{2} \omega(y) + D_1.$$ 

By Lemma 6.1, for every non-zero vector $v \in V$, we have

$$\sup_{1 \leq \rho \leq 2} \frac{1}{H_y^{\text{PS}}[\rho - \rho, \rho]} \int_{-\rho}^{\rho} \frac{\rho}{\|v u, a_i\|^{-s/E}} d\mu_y^{\text{PS}}(r) \leq c_1 \rho_Y e^{-st/(2E)} \|v\|^{-s/E}$$

where $c_1 \geq 1$ is an absolute constant.
Let $\tau \geq \tau_1$ be so that
\[ c_1 p_Y e^{-s\tau/(2E)} = c. \]
We will show that the claim holds for this $\tau$ and
\[ \lambda := 2\hat{c} e^{rs(1+1/E)\tau Z p_Y^{L+s}} \]
where $\hat{c} \geq 1$ is an absolute constant which will be determined momentarily.

Note that since $e^{s\tau/(2E)} = c_1 p_Y e^{-s\tau/(2E)} \geq 2$, we have $\lambda \geq 1$.

As we have done before, the argument is based on comparing $f_{s/E}(yu_r a_\tau)$ and $f_{s/E}(y)$ for $r \in (-1, 1)$ such that $yu_r a_\tau \in Y_0$.

Let $Q \subset H$ be the compact subset \{ $u_r : |r| \leq 2$\}$\{ a_t : |t| \leq \tau$\}$\{ u_r : |r| \leq 2 \}$. Then $c_Q \approx e^{\tau}$ and $d_Q \approx e^{\tau}$ up to an absolute multiplicative constant, where $c_Q$ and $d_Q$ are as in §7.4.

Hence, by Lemma 9.8, we have that for any $|r| \leq 2$,
\[ f_{s/E}(yu_r a_\tau) \leq \sum_{v \in I_Z(y)} \| vu_r a_\tau \|^{-s/E} + \hat{c} e^{rs(1+1/E)\tau Z p_Y^{L+s}} \omega(y)^s \]
where $\hat{c}$ is an absolute constant.

Use this $\hat{c}$ to define $\lambda$. We now average this over $[-\rho, \rho]$ with respect to $\mu^{PS}_y$. Then using (9.16) and (9.17), we get that
\[ (9.18) \quad A_{\tau, p} f_{s/E}(y) \leq c \cdot f_{s/E}(y) + \hat{c} e^{rs(1+1/E)\tau Z p_Y^{L+s}} \omega(y)^s \]
recall that $c_1 p_Y e^{-s\tau/(2E)} = c$.

Then by (9.16) and (9.18), we have
\[ A_{\tau, p} F_{s, \lambda}(y) = A_{\tau, p} f_{s/E}(y) + A_{\tau, p} \omega^s(y) \]
\[ \leq c \cdot f_{s/E}(y) + \frac{c_1}{2} \omega^s(y)^s + \frac{c_1}{2} \omega^s(y)^s + \lambda D_1 \]
\[ = c \cdot F_{s, \lambda}(y) + \lambda D_1. \]

Since $\lambda \ll p_Y^{L+s+E} \tau Z e^{-E}$, this proves the claim. \hfill \Box

Proof of Theorem 9.13. In view of Lemma 9.14 and Lemma 9.15, the theorem follows from Proposition 5.3(1). \hfill \Box

Corollary 9.19. Let the notation be as in Theorem 9.13. For any $y \in Y_0$ we have
\[ \limsup_{t} A_t F_{s, \lambda}(y) \leq 1 + D. \]

Proof. In view of Lemma 9.14 and Lemma 9.15, the theorem follows from Proposition 5.3(2). \hfill \Box

Theorem 9.20. For any $y \in Y_0$, we have
\[ \lim_{t \to \infty} A_t F_{s/2, \lambda}(y) = m_Y(F_{s/2, \lambda}). \]
In particular, $m_Y(F_{s/2, \lambda}) \leq 1 + D$. 

Proof. The proof is similar to the proof of Proposition 8.6 and is based on Theorem 5.16, Theorem 9.13, and Lemma 5.14.

Recall that $Z \neq Y$. Hence, $F_{\bullet, \lambda} : Y_0 \to [2, \infty)$ is a continuous function for $\bullet = s, s/2$. For every $R > 0$, let

$$F_R = F_{s/2, \lambda} - F_{s/2, \lambda} \cdot \chi_R$$

where $\chi_R$ is a continuous function which equals 1 on $\{y \in Y_0 : F_{s/2, \lambda} \geq R + 1\}$ and equals 0 on $\{y \in Y_0 : F_{s/2, \lambda} \leq R\}$. Then by Theorem 5.16, for any $y \in Y_0$, we have

$$\lim_{t \to \infty} A_tF_R(y) = m_Y(F_R).$$

Moreover, in view of Theorem 9.13 and Corollary 9.19, $F_{s/2, \lambda}$ satisfies (1) and (2) in Proposition 5.3. Hence, using Lemma 5.14, we have

$$|m_Y(F_{s/2, \lambda}) - m_Y(F_R)| \to 0 \quad \text{as} \quad R \to \infty.$$ 

We now compute that

$$\limsup_t |A_tF_{s/2, \lambda}(y) - m_Y(F_{s/2, \lambda})| \leq \limsup_t |A_t(F_{s/2, \lambda} \cdot \chi_R)| + \limsup_t |A_t(F_R) - m_Y(F_R)| + |m_Y(F_R) - m_Y(F_{s/2, \lambda})|.$$

The second and third term on the right side of the above tend to 0 by (9.21) and (9.22). We thus estimate $\limsup_t |A_t(F_{s/2, \lambda} \cdot \chi_R)|$.

First note that

$$F_{s/2, \lambda} \cdot \chi_R \leq R^{-1}(F_{s/2, \lambda} \cdot \chi_R)^2.$$ 

Now by the Cauchy-Schwartz inequality and the definition of $F_{s/2, \lambda}$, we have

$$(F_{s/2, \lambda} \cdot \chi_R(y))^2 \ll p^{s, \tau_Z}(y)^2 \begin{cases} \omega(y)^{s^E} + \lambda^2 \omega^s(y) & \text{if } I_Z(y) = \emptyset \\ \sum_{v \in I_Z(y)} \|v\|^{-s} + \lambda^2 \omega^s(y) & \text{otherwise} \end{cases}$$

where we used Lemma 9.2 to control $\#I_Z(y)$.

Since $\lambda \ll p^{s, \tau_Z}$, we get that $(F_{s/2, \lambda} \cdot \chi_R(y))^2 \ll p^{s, \tau_Z} F_{s, \lambda}$. This and (9.23) now imply that

$$F_{s/2, \lambda} \cdot \chi_R \ll p^{s, \tau_Z} R^{-1} F_{s, \lambda}.$$ 

Applying the averaging operator $A_t$ to the above estimate, we get that

$$A_t(F_{s/2, \lambda} \cdot \chi_R) \ll p^{s, \tau_Z} A_t(R^{-1} F_{s, \lambda}).$$ 

Since $A_t F_{s, \lambda} < D + 1$ (see Corollary 9.19), it follows that

$$\limsup_t |A_t(F_{s/2, \lambda} \cdot \chi_R)| \ll R^{-1},$$

where the multiplicative constant is independent of $R$. The first claim now follows by letting $R \to 0$.

The second claim follows from the first claim and Corollary 9.19. □
10. Quantitative isolation of a closed orbit

In this section, we prove Theorem 1.4 in two installations.

**Theorem 10.1** (Isolation in distance). There exist \( N > 0 \) depending on \( \Gamma \) and an absolute constant \( c > 0 \) such that the following holds: Let \( Y \neq Z \) belong to \( \mathcal{H} \). For any \( 0 < \varepsilon < \varepsilon_X \) and \( y \in Y_0 \cap X_{\varepsilon} \), we have

\[
d(y, Z) \geq c \cdot \rho_Y^{-N/\delta_Y} \tau_Z^{-N/\delta_Y} m_Y(B(y, \varepsilon))^{N/\delta_Y}.
\]

**Proof.** Fix \( F := F_{s,\lambda,Y,Z} \) as in Theorem 9.13 for a fixed \( s = \frac{1}{4} \delta_Y \). By Lemma 9.14, there exists \( \sigma \ll \rho_Y \tau_Z \) so that

\[
\sigma^{-1} F(y) \leq F(yh) \leq \sigma F(y)
\]

for all \( y \in Y_0 \) and all \( h \in B_H(\varepsilon_X) \) so that \( yh \in Y_0 \).

By Theorem 9.20 and Corollary 9.19, we have

\[
m_Y(F) \leq 1 + D.
\]

Let \( 0 < \varepsilon \leq \varepsilon_X \). For all \( y \in Y_0 \cap X_{\varepsilon} \), we have

\[
m_Y(F) \leq \frac{\int_{x \in B_H(\varepsilon)} F(x) dm_Y(x)}{m_Y(B(y, \varepsilon))} \leq \frac{\sigma m_Y(F)}{m_Y(B(y, \varepsilon))} \leq \frac{\sigma D + 1}{m_Y(B(y, \varepsilon))}.
\]

In view of the definition of \( F \), we get that

\[
(10.2) \quad f_{s/E,Y,Z}(y) \leq \frac{(D + 1)\sigma}{m_Y(B(y, \varepsilon))} - \lambda \omega^s(y) \leq \frac{(D + 1)\sigma}{m_Y(B(y, \varepsilon))}.
\]

It follows from the definition of \( f_{s,Y,Z} \) that

\[
f_{s/E,Y,Z}(y) \geq d(y, Z)^{-s/E}.
\]

Since \( s = \delta_Y^2/4 \),

\[
d(y, Z) \geq \left( \frac{m_Y(B(y, \varepsilon))}{(D + 1)\sigma} \right)^{4E/\delta_Y^2}.
\]

Now the claim follows since \( D \) and \( \sigma \) are \( \ll \rho_Y^t \tau_Z^t \). \( \square \)

**Theorem 10.3** (Isolation in measure). There exist \( N > 0 \) depending on \( \Gamma \) and an absolute constant \( c > 0 \) so that the following holds. Let \( 0 < \varepsilon \leq 1/2 \).

For all \( Y, Z \in \mathcal{H} \) with \( Z \neq Y \), we have

\[
m_Y\{y \in Y_0 : d(y, Z) \leq \varepsilon \} \leq c \cdot \rho_Y^N \tau_Z^N \varepsilon^{\delta_Y^2/N}.
\]

**Proof.** Fix \( s := \frac{1}{4} \delta_Y \) and let \( F := F_{s,\lambda,Y,Z} \) as was done in the proof of Theorem 10.1. For any \( \eta > 0 \), define

\[
\Omega_\eta := \{y \in Y_0 : F(y) > \eta^{-1}\};
\]

note that if \( y \in Y_0 \) satisfies either \( d(y, Z) \leq \eta^{E/s} \) or \( \omega(y) > \eta^{-1/s} \), then \( y \in \Omega_\eta \).
By the definition of $\Omega_\eta$, we have

$$\frac{m_Y(\Omega_\eta)}{\eta} \leq \int_{\Omega_\eta} Fdm_Y \leq m_Y(F).$$

By Theorem 9.20, $m_Y(F) \leq D + 1$. Hence we get that

(10.4) $$m_Y\{y \in Y_0 : d(y, Z) \leq \eta^{E/s}\} \leq m_Y(\Omega_\eta) \leq (D + 1)\eta.$$

Recall that $\varepsilon^{s/E} = \delta^{\delta/(4E)}$ and $D \ll p_Y^*\tau_Z^*$. The claim thus follows from (10.4) applied with $\varepsilon = \eta^{E/s}$. □

**Proof of Theorem 1.4.** By Corollaries 3.14 and 4.6, we have $s_Y \asymp p_Y$ and $\tau_Z \ll area_t(S_Z)$, respectively. Theorem 1.4 thus follows from Theorems 10.1 and 10.3 □

11. **Appendix: Proof of Theorem 1.1 in the compact case**

In this section we present the proof of Theorem 1.1 when $X$ is compact. As was mentioned in the introduction, this case is due to G. Margulis. His original proof used the operator $A_\sigma\psi(y) = \int_H \psi(yh)d\sigma(h)$ where $\sigma$ is any measure whose support generates a Zariski dense subgroup of $H$, but we use the horocyclic average as in (11.4).

Let $Y \neq Z$ be two closed orbits of $H$ in $X = \Gamma\backslash G$. Fix small $0 < \varepsilon_0 < 1/2$ so that $0 < 2\varepsilon_0 < \min_{x \in X} inj(x)$.

Define $f_Z : Y \to [2, \infty)$ as

$$f_Z(y) = \begin{cases} \sum_{v \in I_Z(y)} \|v\|^{-1/3} & \text{if } I_Z(y) \neq \emptyset \\ \varepsilon_0^{-1/3} & \text{otherwise} \end{cases}$$

where

$$I_Z(y) = \{v \in i\mathfrak{sl}_2(\mathbb{R}) - \{0\} : \|v\| \leq \varepsilon_0 \& y \exp(v) \in Z\}.$$

**Remark 11.1.** With the proof given in Lemma 6.1, the exponent $1/3$ can be replaced with any positive number $s$ strictly less than $1/2$, which is the reciprocal of the highest weight of the adjoint representation of $H$ — this is needed to ensure the validity of the convergence of the geometric series in (6.4) in Lemma 6.1. One can give a slightly different argument which allows for any exponent $0 < s < 1$, see [7, Lemma 5.1].

We use the following special case of Lemma 6.1: For any $v \in i\mathfrak{sl}_2(\mathbb{R})$ with $\|v\| = 1$, we have

(11.2) $$\int_0^1 ds \frac{ds}{\|vu_s a_i\|^{1/3}} \leq Ce^{-t/3}$$

where $vh = Ad(h)(v)$ for $h \in H$. 
Remark 11.3. It is worth noting that Lemma 6.1 considers symmetric interval $[-1, 1]$ — this is necessary in the infinite volume setting; indeed the half interval $[0, 1]$ may even be a null set for $\mu_y^{PS}$ for some $y$, see (3.2) for the notation.

For $t > 0$ and $y \in Y$, define

\begin{equation}
A_t f_Z(y) = \int_0^1 f_Z(y u_s a_t) ds.
\end{equation}

Proposition 11.5. There exist absolute constants $\tau > 0$ and $b > 1$ such that for all $y \in Y$,

\begin{equation}
A_\tau f_Z(y) \leq \frac{1}{2} f_Z(y) + b \text{Vol}(Z).
\end{equation}

Proof. Let $C$ be as in (11.2), and let $\tau$ be given by the equation $Ce^{-\tau/3} = 1/2$. We compare $f_Z(y u_r a_t)$ and $f_Z(y)$ for $r \in [0, 1]$. Let $C_1$ be large enough so that $\|vh\| \leq C_1 \|w\|$ for all $v \in i\mathbb{I}_2(\mathbb{R})$ and all $h \in \{u_s : s \in [0, 1]\} \{a_t : |t| \leq \tau\} \{u_s : s \in [-1, 1]\}$.

Let $v \in I_Z(y u_t a_\tau)$ be so that $\|v\| < \varepsilon_0/C_1$. Then $\|va_\tau u_s\| \leq \varepsilon_0$; in particular, $va_\tau u_s \in I_Y(y)$.

In the following, if $I_Z(\cdot) = \emptyset$, the sum is interpreted as to equal to $\varepsilon_0^{-1}$. In view of the above observation and the definition of $f$, we have

\begin{equation}
f_Z(y u_t a_\tau) = \sum_{v \in I_Z(y u_t a_\tau)} \|v\|^{-1/3}
= \sum_{\|v\| < \varepsilon_0/C_1} \|v\|^{-1/3} + \sum_{\|v\| \geq \varepsilon_0/C_1} \|v\|^{-1/3}
\leq \sum_{v \in I_Z(y)} \|vu_t a_\tau\|^{-1/3} + \sum_{\|v\| \geq \varepsilon_0/C_1} \|v\|^{-1/3}
\end{equation}

Moreover, note that $\#I_Z(y u_t a_\tau) \asymp \text{Vol}(Z)$. Hence,

\begin{equation}
\sum_{\|v\| \geq \varepsilon_0/C_1} \|v\|^{-1/3} \ll \text{Vol}(Z) \varepsilon_0^{-1/3}.
\end{equation}

We now average (11.7) over $[0, 1]$. Then using (11.8) and (11.2) we get

\[ A_\tau f_Z(y) \leq \frac{1}{2} f_Z(y) + O(\text{Vol}(Z)); \]

as was claimed in (11.6). \qed

Let $m_Y$ be the $H$-invariant probability measure on $Y$:

Corollary 11.9.

\begin{equation}
m_Y(f_Z) \leq 2b \text{Vol}(Z).
\end{equation}

Proof. Since $m_Y$ is an $H$-invariant probability measure, $m_Y(A_\tau f_Z) = m_Y(f_Z)$. Hence the claim follows by integrating (11.6) with respect to $m_Y$. \qed
Proof of Theorem 1.1. There exists $\sigma = \sigma(\varepsilon_0) > 0$ such that for any $h \in B_H(\varepsilon_0)$ and $y \in Y$, $f_Z(y) \leq \sigma f_Z(yh)$. Hence

$$f_Z(y) \leq \sigma \cdot \frac{\int_{B_H(\varepsilon_0)} f_Z(yh) dm_Y(yh)}{m_Y(B(y, \varepsilon_0))} \leq \sigma \cdot \frac{m_Y(f_Z)}{m_Y(B(y, \varepsilon_0))} \leq 2b \varepsilon_0^{-3} \text{Vol}(Y) \text{Vol}(Z).$$

Since $\text{dist}(y, Z)^{-1/3} \leq f_Z(y)$, this shows Theorem 1.1(1):

$$d(y, Z) \gg \text{Vol}(Z)^{-3} \text{Vol}(Y)^{-3}.$$  

Theorem 1.1(2) follows as

$$m_Y\{y \in Y : d(y, Z) \leq \varepsilon\} \leq m_Y\{y \in Y : f_Z(y) \geq \varepsilon^{-1/3}\} \leq \varepsilon^{1/3} m_Y(f_Z) \leq 2b \text{Vol}(Z) \varepsilon^{1/3}.$$

REFERENCES


Mathematics Department, UC San Diego, 9500 Gilman Dr, La Jolla, CA 92093

\textit{E-mail address}: ammohammadi@ucsd.edu

Mathematics department, Yale University, New Haven, CT 06520 and Korea Institute for Advanced Study, Seoul, Korea

\textit{E-mail address}: hee.oh@yale.edu