CLASSIFICATION OF JOININGS FOR KLEINIAN GROUPS

AMIR MOHAMMADI AND HEE OH

ABSTRACT. We classify all locally finite joinings of a horospherical subgroup action on \( \Gamma \backslash G \) when \( G = \text{PSL}_2(\mathbb{R}) \) or \( \text{PSL}_2(\mathbb{C}) \) and \( \Gamma \) is a geometrically finite Zariski dense subgroup of \( G \). This generalizes Ratner’s 1983 joining theorem for the case when \( \Gamma \) is a lattice in \( G \).

CONTENTS

1. Introduction .......................... 1
2. Basic properties of PS-measure ....... 7
3. Uniformity in the equidistribution of PS-measure 11
4. Equidistribution of non-closed \( U \)-orbits 15
5. Window lemma for horospherical averages 22
6. Rigidity of \( AMU \)-equivariant maps 28
7. Joining classification .................. 34
References ............................... 52

1. INTRODUCTION

Furstenberg introduced in his 1967 influential paper [10] the notion of a joining which has become an indispensable tool in ergodic theory. Ratner obtained in 1983 the joining classification for horocycle flows on a finite volume quotient of \( G = \text{PSL}_2(\mathbb{R}) \) [31]; this precedes her measure classification theorem for unipotent flows on a homogeneous space \( \Gamma \backslash G \) where \( G \) is any connected Lie group and \( \Gamma \) is a lattice in \( G \).

Classification problem of locally finite invariant measures for general unipotent flows on an infinite volume homogeneous space \( \Gamma \backslash G \) is quite mysterious and even a conjectural picture is not clear at present. However if \( G \) is a simple group of rank one, there are classification results for locally finite measures on \( \Gamma \backslash G \) invariant under a horospherical subgroup of \( G \), when \( \Gamma \) is

---

2010 Mathematics Subject Classification. Primary 11N45, 37F35, 22E40; Secondary 37A17, 20F67.

Key words and phrases. Geometrically finite hyperbolic groups, Joining, Burger-Roblin measure, Bowen-Margulis-Sullivan measure.

Mohammadi was supported in part by NSF Grant #1200388.
Oh was supported in part by NSF Grant #1068094.
either geometrically finite ([4, 33, 43]) or when \( \Gamma \) is a special kind of geometrically infinite subgroup ([2, 18, 19, 35, 36]). In this article, our goal is to extend Ratner’s joining theorem to an infinite volume homogeneous space \( \Gamma \backslash G \) where \( G = \text{PSL}_2(F) \) for \( F = \mathbb{R}, \mathbb{C} \) and \( \Gamma \) is a geometrically finite and Zariski dense subgroup of \( G \). As far as we know, this is the first measure classification result in homogeneous spaces of infinite volume for unipotent subgroups which are not horospherical subgroups.

Letting \( n = 2, 3 \), respectively, for \( F = \mathbb{R}, \mathbb{C} \), the group \( G \) is the group of orientation preserving isometries of the real hyperbolic space \( \mathbb{H}^n \) and a discrete subgroup \( \Gamma < G \) is geometrically finite if it admits a finite-sided fundamental domain in \( \mathbb{H}^n \). For \( n = 2 \), geometrically finite groups are just finitely generated discrete subgroups of \( \text{PSL}_2(\mathbb{R}) \). Geometrically finite groups are a natural generalization of lattices; for instance, the fundamental group of any finite volume hyperbolic manifold with totally geodesic boundary is geometrically finite. Any group generated by inversions with respect to a finite number of disjoint vertical hemispheres in \( \mathbb{H}^n \) is geometrically finite. The symmetry group of an Apollonian circle packing is another example of a geometrically finite subgroup of \( \text{PSL}_2(\mathbb{C}) \). Thurston showed that any finitely generated subgroup of a geometrically finite subgroup of \( \text{PSL}_2(\mathbb{C}) \) which is not a lattice is also geometrically finite (cf. [24]). Hence we have quite a big class of such subgroups.

We recall the notion of a joining. Traditionally the joining has been considered only for a finite measure. However in an infinite volume homogeneous space, we cannot restrict ourselves to finite joinings since most of interesting invariant measures under a subgroup action are infinite measures.

Throughout the introduction, we let \( U \) be a horospherical subgroup of \( G \), i.e., for a one-parameter diagonalizable subgroup \( A = \{a_s\} \) of \( G \),

\[
U = \{g \in G : a_s g a_{-s} \to e \text{ as } s \to +\infty\}.
\]

For \( i = 1, 2 \), let \( \Gamma_i < G \) be a discrete subgroup and \( \mu_i \) a locally finite \( U \)-invariant Borel measure on \( \Gamma_i \backslash G \). For a subset \( S \) of \( G \), we set \( \Delta(S) := \{(s, s) : s \in S\} \) for the diagonal embedding of \( S \) into \( G \times G \). Set \( X := \Gamma_1 \backslash G \times \Gamma_2 \backslash G \) and denote by \( \pi_i \) the canonical projection of \( X \) to \( \Gamma_i \backslash G \) for \( i = 1, 2 \).

**Definition 1.1.** A locally finite \( \Delta(U) \)-invariant measure \( \mu \) on \( X \) is called a \( U \)-joining for the pair \( (\mu_1, \mu_2) \) if the push-forward \( (\pi_i)_* \mu \) is proportional to \( \mu_i \) for each \( i = 1, 2 \).

In our joining classification theorem, we will take \( \mu_i \) to be the Burger-Roblin measure \( m^\text{BR}_{\Gamma_i} \) on \( \Gamma_i \backslash G \). The reason for this is that for \( \Gamma \) geometrically finite and Zariski dense, the Burger-Roblin measure \( m^\text{BR}_{\Gamma} \) is the unique locally finite \( U \)-invariant ergodic measure in \( \Gamma \backslash G \) which is non-trivial \(^1\), by

\(^1\)A \( U \)-invariant ergodic measure is considered trivial if its support is contained in a closed \( U \)-orbit.
the work of Burger [4], Roblin [33] and Winter [43]. Therefore the Burger-Roblin measure for $\Gamma \backslash G$, which we will call the BR-measure for simplicity, plays the role of the Haar measure in Ratner’s joining theorem for $\Gamma$ a lattice.

Unlike the Haar measure, the BR measure depends heavily on $\Gamma$. We denote by $\delta$ the critical exponent of $\Gamma$, i.e., the abscissa of convergence of the Poincare series $P_\Gamma(s) = \sum_{\gamma \in \Gamma} e^{-sd(o,\gamma(o))}$ for $o \in \mathbb{H}^n$, and by $\Lambda(\Gamma)$ the limit set, which is the set of the accumulation points of an orbit of $\Gamma$ in the geometric boundary $\partial(\mathbb{H}^n)$. We give a definition of the BR measure using the Iwasawa decomposition $G = KAU$ (see section 2 for more details). Let $\nu_o$ be the Patterson-Sullivan measure on $\partial(\mathbb{H}^n)$ viewed from $o \in \mathbb{H}^n$ where $K := \text{Stab}_G(o)$, which is unique up to a scalar multiple. Via the transitive action of $K$ on $\partial(\mathbb{H}^n)$, we may lift $\nu_o$ to $K$, which we denote by $\nu_o$ by abuse of notation. Then we first define a locally finite measure $\tilde{m}_\Gamma^{BR}$ on $G$ as follows:

$$\tilde{m}_\Gamma^{BR}(\psi) := \int_{KAU} \psi(ka_u) e^{-\delta s} d\nu_o(k) ds du$$

where $du$ and $ds$ are the Lebesgue measures on $U$ and $\mathbb{R}$ respectively. The measure $\tilde{m}_\Gamma^{BR}$ is clearly right $U$-invariant. It is also left $\Gamma$-invariant, as follows from the conformal properties of $\nu_o$. We denote by $m_\Gamma^{BR}$ the measure on $\Gamma \backslash G$ induced from $\tilde{m}_\Gamma^{BR}$. We note that $m_\Gamma^{BR}$ is an infinite measure if $\Gamma$ is of infinite co-volume in $G$.

In what follows, we assume that at least one of $\Gamma_1$ and $\Gamma_2$ has infinite co-volume in $G$; otherwise, the joining classification was already proved by Ratner. Under this assumption, the product measure $m_\Gamma^{BR_1} \times m_\Gamma^{BR_2}$ is never a $U$-joining (with respect to the pair $(m_\Gamma^{BR_1}, m_\Gamma^{BR_2})$), since its projection to $\Gamma \backslash G$ is an infinite multiple of $m_\Gamma^{BR_i}$ for at least one of $i = 1, 2$. On the other hand, a finite cover self-joining provides an example of $U$-joining.

**Definition 1.2** (Finite cover self-joining). Suppose that for some $g_0 \in G$, $\Gamma_1$ and $g_0^{-1} \Gamma_2 g_0$ are commensurable with each other; this in particular implies that the orbit $[(e,g_0)] \Delta(G)$ is closed in $X$. Now using the isomorphism

$$(\Gamma_1 \cap g_0^{-1} \Gamma_2 g_0) \backslash G \to [(e,g_0)] \Delta(G)$$

given by the map $[g] \mapsto [(g,g_0 g)]$, the push-forward of the BR-measure $m_{\Gamma_1 \cap g_0^{-1} \Gamma_2 g_0}^{BR}$ to $X$ gives a $U$-joining, which we call a finite cover self-joining. If $\mu$ is a $U$-joining, then any translation of $\mu$ by $(e,u_0)$ is also a $U$-joining, whenever $u_0$ belongs to the centralizer of $U$, which is $U$ itself. Such a translation of a finite cover self-joining will also be called a finite cover self-joining.

**Theorem 1.3** (Joining Classification). Let $\Gamma_1$ and $\Gamma_2$ be geometrically finite and Zariski dense subgroups of $G$. Suppose that either $\Gamma_1$ or $\Gamma_2$ is not a lattice in $G$. Then any locally finite $U$-ergodic joining on $X$ is a finite cover self-joining.
If $\mu$ is a $U$-joining and $\mu = \int \mu_x$ is the $U$-ergodic decomposition, then almost every $\mu_x$ must be a $U$-joining, since each $m_{\Gamma U}^{\text{BR}}$ is $U$-ergodic. Therefore the above classification of $U$-ergodic joinings gives a complete description for all possible $U$-joinings.

**Corollary 1.4.** Let $\Gamma_1, \Gamma_2$ be as in Theorem 1.3. Then $X$ admits a $U$-joining measure if and only if $\Gamma_1$ and $\Gamma_2$ are commensurable with each other, up to a conjugation.

See Remark 7.19 for a slightly more general statement where $\Gamma_2$ is not assumed to be geometrically finite.

In the course of our classification theorem, we obtain equidistribution of a non-closed $U$-orbit $xU$ in $\Gamma \backslash G$. When $\Gamma$ is a lattice, it is well-known that such $xU$ is equidistributed with respect to the Haar measure ([7], [32]). For $\Gamma$ geometrically finite, the equidistribution is described by the BR-measure and we need to use the normalization given by the Patterson-Sullivan measure $\mu_x^{\text{PS}}$ on $xU$ (see section 2.2 for a precise definition), which controls the return time of $xU$ to a fixed compact subset. We call a norm $\| \cdot \|$ on $\mathbb{F}$ algebraic if the 1-sphere $\{ \mathbf{t} \in \mathbb{F} : \| \mathbf{t} \| = 1 \}$ is contained in a finitely many union of algebraic varieties.

**Theorem 1.5.** Let $\Gamma$ be a geometrically finite and Zariski dense subgroup of $G$. Suppose that $xU$ is not closed in $\Gamma \backslash G$. Then for any $\psi \in C_c(X)$, we have
\[
\lim_{T \to \infty} \frac{1}{\mu_x^{\text{PS}}(B_U(T))} \int_{B_U(T)} \psi(xu)du = m^{\text{BR}}(\psi)
\]
where $B_U(T) = \{ u \in U : \| u \| < T \}$ is the norm ball in $U \simeq \mathbb{F}$ with respect to an algebraic norm.

Indeed, we prove this theorem in a greater generality where $\Gamma$ is a geometrically finite and Zariski dense subgroup of $G = \text{SO}(n, 1)^\circ$ for any $n \geq 2$ (see Theorem 4.6). Note that there are effective versions of this theorem in the case when $\Gamma$ is a congruence subgroup of $\text{SL}_2(\mathbb{Z})$ ([40], [37]). When an effective mixing of frame flow is available (e.g., see [22] when $\delta > (n-1)/2$ for $n = 2, 3$ and when $\delta > n-2$ for $n \geq 4$), we believe that Theorem 1.5 can also effectivized; of course the effectiveness would depend on certain Diophantine type condition on the backward end point $x^-$ of the geodesic determined by $x$.

Theorem 1.5 was proved for $\Gamma = \text{SO}(2,1)^\circ = \text{PSL}_2(\mathbb{R})$ in ([39], [21]). One of the technical difficulties in extending Theorem 1.5 to a higher dimensional case (even to the case $n = 3$) is the lack of a uniform control of the PS-measure of a small neighborhood of the boundary of $B_U(T)$. For $n = 2$, such a neighborhood has a fixed size independent of $T$, but grows with $T \to \infty$ for $n \geq 3$. In the case when $\Gamma$ is a lattice and the PS-measure on $xU$ is simply the Lebesgue measure, it still has size strictly smaller order than the size of $B_U(T)$. For $\Gamma$ geometrically finite with cusps, it is not always the case and this causes highly non-trivial technical difficulties. We mention that
Theorem 1.5 applied to the Apollonian group can be used to describe the distribution of the accumulation of large circles in an unbounded Apollonian circle packing, whereas the papers [16] and [26] considered the distribution of small circles; we hope to address this application in a separate paper. We refer to Theorem 5.4 for a window version of Theorem 1.5 which is one of the main ingredients in our proof of Theorem 1.3.

Similarly to the finite joining case, any measurable factor map for a $U$-action gives rise to a $U$-joining of the product space (see e.g. [9] for a discussion). This was used by Ratner [31] and Witte [44], and in a different context by Furman [9], in combination with joining classification, to obtain a classification of measurable factors. The same argument applies here and we obtain the following theorem as a corollary of Theorem 1.3.

**Theorem 1.6 (U-factor classification).** Let $\Gamma$ be a geometrically finite and Zariski dense subgroup of $G$. Let $(Y, \nu)$ be a measure space with a locally finite $U$-invariant measure $\nu$. Suppose $\pi : (\Gamma \backslash G, m^\text{BR}_\Gamma) \to (Y, \nu)$ is a measurable $U$-equivariant factor map, that is, $\pi_* m^\text{BR}_\Gamma = \nu$.

Then there exists a subgroup $\Gamma_0$ of $G$ which contains $\Gamma$ and $[\Gamma_0 : \Gamma] < \infty$ and the space $(Y, \nu)$ is isomorphic to $(\Gamma_0 \backslash G, m^\text{BR}_{\Gamma_0})$; moreover, under this isomorphism, the map $\pi$ is conjugated to the natural projection $\Gamma_0 \backslash G \to \Gamma_0 \backslash G$.

**On the proof of Theorem 1.3:** Our proof is modeled on Ratner’s proof [31]. However, the fact that we are dealing with an infinite measure introduces several serious technical difficulties which are dealt with in this paper. Here we briefly describe some of the main steps and difficulties.

The main ingredient in the proof is “polynomial behavior” of unipotent orbits. Roughly speaking: this property, which has been utilized by many others in the past forty years or so, guarantees “slow” divergence of orbits of two “nearby” points under unipotent subgroups, see § 7.1 for more precise discussion of this.

Let $B_{\Delta(U)}(T)$ denote the diagonal embedding $\{(u_t, u_t) : u_t \in B_U(T)\}$ of the ball of radius $T$ in $U$. The above mentioned slow divergence implies that the set of times $t$ with $(u_t, u_t) \in B_{\Delta(U)}(T)$ when the orbits of two nearby points are in the intermediate range (they are roughly distance one apart) has Lebesgue measure comparable to the measure of $B_{\Delta(U)}(T)$. In order to utilize this nice property we also need to know this set is “dynamically non-trivial”, e.g. the two orbits at these times spend some time in a fixed compact set.

In the case of a probability measure, this can be guaranteed using the Birkhoff ergodic theorem. For an infinite measure, however, a priori there may be a sequence of times $T_i \to \infty$ so that the times when $xB_{\Delta(U)}(T_i)$

\[2\] which can be thought of as an algebraic property of unipotent orbits.

\[3\] the available ergodic theorem here is the Hopf ratio ergodic theorem which states that for $\phi, \psi \in L^1(\mu)$ the ratio $\frac{\int \phi dt}{\int \psi dt}$ converges to $\frac{\mu(\phi)}{\mu(\psi)}$ almost all $\{x : \int \psi dt \to \infty\}$.
returns to a compact subset are “concentrated” around the center. This seems to be one of the main issues in the way of obtaining measure rigidity type results for general locally finite (but not finite) measures (see also [22]).

In the case at hand, i.e. joining measure, using the projection to one of the factors, this question translates to a similar question for the action of $U$ and the BR-measure. We then prove a non-concentration property for the action of $U$ (a window version of Theorem 1.5), using Theorem 1.5 and the shadow lemma (see Theorems 4.6 and 5.1, as well as Lemma 5.2).

Using this window type equidistribution result (Corollary 7.6), the construction of a polynomial like map in § 7.1 and the fact that an infinite joining measure cannot be invariant under a subgroup of the form $\{e\} \times V$ with $V < U$ (Lemma 7.13), we draw two important corollaries: 1. The fibers of the projection onto each factor are finite almost surely 2. The measure $\mu$ is quasi invariant under the diagonal embedding of $AM$ (here $M$ is the centralizer of $A$ in $K$), which is a subgroup of $N_{G\times G}(\Delta(U))$. In the course of this argument one needs to control averages of the form $xu_\alpha(t)$ where $\alpha$ is a map with controlled, but not necessarily constant, derivatives. Such consideration are also needed in the finite measure case, however, in the case of infinite measure the analysis is more involved. Indeed we only show some “positive proportion type” statement rather than equidistribution of such orbits, see Lemma 7.7.

One other technical difficulty is that unlike the probability Haar measure case which is ergodic for any non-compact subgroup of $G$ by Moore’s ergodicity theorem, the BR measure is not in general ergodic for an action of a non-compact subgroup and this presents another technical problems we need to address, see §7.4.

In the finite measure case, it is possible to finish using entropy, based on the quasi-invariance by the action of $A$. Such an argument using entropy is not understood in the infinite measure case. Ratner’s proof in [31] avoids a “direct” use of entropy and this is the argument which can be applied here.

Let us recap: as a consequence of the above discussion and after a possible conjugation, we are dealing with a measurable factor (set valued) map from $\Upsilon : \Gamma_1 \backslash G \to \Gamma_2 \backslash G$ which commutes with the action of $U$ and $AM$. We need to show that this set-valued map also commutes with the action of $U^{-}$, the opposite horospherical subgroup. Such a result was proved by Ratner in the lattice case [29] and was generalized by Flaminio and Spatzier [8] to the case of convex cocompact groups for the Bowen-Margulis-Sullivan measure. We use a very close relationship between the BR and BMS measures and show that essentially the same proof as in [8] works here, again, modulo the extra technical difficulties caused by the presence of cusps. This proof

---

4Another substantial (and missing) input in proving measure rigidity type results for Radon measure

5One may argue that a shadow of argument using entropy is still present in Ratner’s argument and thus in our argument.
also relies heavily on the polynomial behavior of unipotent flows and non-concentration of the PS measure. Roughly speaking, one constructs two nearby points \( x \) and \( y = xu_\Gamma \) so that the \( U \)-orbits of \( \Upsilon(x) \) and \( \Upsilon(y)u_\Gamma \) “do not diverge on average”. The fact that the divergence of these two orbits is governed by a polynomial map then implies that the two orbits “do not diverge”. One then concludes that the map commutes with the action of \( U^- \) and completes the proof.

**Notation** In the whole paper, we use the following standard notation. We write \( f(t) \sim g(t) \) as \( t \to \infty \) to mean \( \lim_{t \to \infty} f(t)/g(t) = 1 \). We use the Landau notations \( f(t) = O(g(t)) \) and \( f \ll g \) synonymously to mean that there exists an implied constant \( C > 1 \) such that \( f(t) \leq C \cdot g(t) \) for all \( t > 1 \). We write \( f(t) \asymp g(t) \) if \( f(t) = O(g(t)) \) and \( g(t) = O(f(t)) \). For a space \( X \), \( C(X) \) (resp. \( C^\infty(X) \)) denotes the set of all continuous (reps. smooth) functions on \( X \) and \( C_c(X) \) denotes the set of functions in \( C(X) \) which are compactly supported. For a subset \( \Omega \) of \( X \), \( C(\Omega) \) (resp. \( C^\infty(\Omega) \)) denotes the set of functions in \( C(X) \) (resp. \( C^\infty(\Omega) \)) whose supports are contained in \( \Omega \). For a subset \( B \) in \( X \), \( \partial(B) \) denotes the topological boundary of \( B \) with the exception that \( \partial(\mathbb{H}^n) \) means the geometric boundary of the hyperbolic \( n \)-space \( \mathbb{H}^n \). For a function \( f \) on \( \Gamma \setminus G \) and \( g \in G \), the notation \( g.f \) means the function on \( \Gamma \setminus G \) defined by \( g.f(x) = f(xg) \).

Given a subset \( S \subset G \), we let \( \Delta(S) = \{(s, s) : s \in S\} \). Also given a Lie group \( H \) and a subset \( S \subset H \) we denote by \( \langle S \rangle \) the minimal connected subgroup of \( H \) containing \( S \), and by \( N_H(S) \) the normalizer of \( S \) in \( H \).

**Acknowledgment** We would like to thank M. Einsiedler, Y. Minsky and A. Reid for fruitful conversations.

### 2. Basic properties of PS-measure

Let \( n \geq 2 \) and \( G = \text{SO}(n, 1)^o = \text{Isom}^+(\mathbb{H}^n) \), i.e., the group of orientation preserving isometries of the hyperbolic space \( (\mathbb{H}^n, d_{\text{hyp}}) \). Let \( d_G \), or simply \( d \), denote the left \( G \)-invariant and bi \( K \)-invariant metric on \( G \) which induces the hyperbolic metric \( d_{\text{hyp}} \) on \( G/K = \mathbb{H}^n \). For a subset \( S \subset G \), \( S_\varepsilon \) denotes the \( \varepsilon \)-ball of \( e \) in the metric \( d_G \): \( S_\varepsilon = \{ s \in S : d_G(s, e) < \varepsilon \} \).

We denote by \( \partial(\mathbb{H}^n) \) the geometric boundary of \( \mathbb{H}^n \), i.e., the equivalence classes of geodesic rays. Let \( \Gamma < G \) be a torsion-free discrete subgroup. Let \( \Lambda(\Gamma) \subset \partial(\mathbb{H}^n) \) denote the limit set of \( \Gamma \), and \( \delta \) the critical exponent of \( \Gamma \). The convex core of \( \Gamma \) is the quotient by \( \Gamma \) of the smallest convex subset in \( \mathbb{H}^n \) which contains all geodesics connecting points between \( \Lambda(\Gamma) \). The group \( \Gamma \) is called **geometrically finite** if a unit neighborhood of the convex core of \( \Gamma \) has finite volume. Throughout the paper, we assume that \( \Gamma \) is geometrically finite and Zariski dense.

#### 2.1. Conformal densities

A family of finite measures \( \{\mu_x : x \in \mathbb{H}^n\} \) on \( \partial(\mathbb{H}^n) \) is called a \( \Gamma \)-invariant conformal density of dimension \( \delta_\mu > 0 \) if for
any \( x, y \in \mathbb{H}^n \), \( \xi \in \partial(\mathbb{H}^n) \) and \( \gamma \in \Gamma \),
\[
\gamma_s \mu_x = \mu_{\gamma x} \quad \text{and} \quad \frac{d\mu_y}{d\mu_x}(\xi) = e^{-\delta_x \beta_x(y,x)},
\]
where \( \gamma_s \mu_x(F) = \mu_x(\gamma^{-1}(F)) \) for any Borel subset \( F \) of \( \partial(\mathbb{H}^n) \). Here \( \beta_x(y, x) \) denotes the Busemann function: \( \beta_x(y, x) = \lim_{t \to \infty} d(\xi_t, y) - d(\xi_t, x) \) where \( \xi_t \) is a geodesic ray tending to \( \xi \) as \( t \to \infty \).

We denote by \( \{ \nu_x \} \) the Patterson-Sullivan density, i.e., a \( \Gamma \)-invariant conformal density of dimension \( \delta \), which exists uniquely up to a scalar multiple.

For each \( x \in \mathbb{H}^n \), we set \( m_x \) to be the unique probability measure on \( \partial(\mathbb{H}^n) \) which is invariant under the compact subgroup \( \text{Stab}_G(x) \). Then \( \{ m_x : x \in \mathbb{H}^n \} \) forms a \( G \)-invariant conformal density of dimension \( (n-1) \), which will be referred to as the Lebesgue density.

Denote by \( \{ G^s : s \in \mathbb{R} \} \) the geodesic flow on the unit tangent bundle of \( \mathbb{H}^n \) which, as usual, will be denoted by \( T^1(\mathbb{H}^n) \). For \( u \in T^1(\mathbb{H}^n) \), we denote by \( u^\pm \in \partial(\mathbb{H}^n) \) the forward and the backward endpoints of the geodesic determined by \( u \), i.e., \( u^\pm = \lim_{s \to \pm \infty} G^s(u) \).

Fix \( o \in \mathbb{H}^n \) and \( X_0 \in T_o(\mathbb{H}^n) \). Let \( K := \text{Stab}_G(o) \) and \( M := \text{Stab}_G(X_0) \) so that \( \mathbb{H}^n \) and \( T^1(\mathbb{H}^n) \) can be identified with \( G/K \) and \( G/M \) respectively.

Let \( A = \{ a_s : s \in \mathbb{R} \} \) be the one-parameter subgroup of diagonalizable elements in the centralizer of \( M \) in \( G \) such that \( G^s(X_0) = [M]a_s = [a_sM] \).

For \( g \in G \), define
\[
g^\pm := (gM)^\pm \in \partial(\mathbb{H}^n).
\]
The map \( \text{Viz} : G \to \partial(\mathbb{H}^n) \) given by \( g \mapsto g^+ \) (resp. \( g \mapsto g^- \)) will be called the visual map (resp. the backward visual map).

### 2.2. PS measure on \( U \)-orbits.

Let \( U \) denote the expanding horospherical subgroup, i.e.,
\[
U = \{ g \in G : a_sga_{-s} \to e \text{ as } s \to +\infty \}.
\]
The group \( U \) is a connected abelian group, isomorphic to \( \mathbb{R}^{n-1} \); we use the parametrization \( t \mapsto u_t \) so that for any \( s \in \mathbb{R} \), \( a_sua_s = u_{sa} \).

For any \( g \in G \), the restriction \( \text{Viz}|_{gU} \) is a a homeomorphism between \( gU \) and \( \partial(\mathbb{H}^n) - \{ g^- \} \). Using this diffeomorphism, we can define measures on \( gU \) which are equivalent to conformal densities on the boundary:
\[
\begin{align*}
\mu_{gU}(\text{Leb}) &= e^{(n-1)\beta_{(gut)}(o)} dm_o((gut)^+); \\
\mu_{gU}(\text{PS}) &= e^{\delta_{(gut)}(o)}dm_o((gut)^+).
\end{align*}
\]
The conformal properties of \( \{ m_x \} \) and \( \{ \nu_x \} \) imply that these definitions are independent of the choice of \( o \in \mathbb{H}^n \). The former measure is independent of the orbit \( gU \):
\[
\mu_{gU}(\text{Leb}) = \mu_U(\text{Leb}) = dt
\]
and is simply the Lebesgue measure on $U = \mathbb{R}^{n-1}$, up to a scalar multiple. We call the measure $d\mu_{g}^{\text{PS}}$ the Patterson-Sullivan measure (or simply PS-measure) on $gU$. For simplicity, we write

$$d\mu_{g}^{\text{Leb}}(t) = d\mu_{gU}(gu_{t}) \quad \text{and} \quad d\mu_{g}^{\text{PS}}(t) = d\mu_{gU}^{\text{PS}}(gu_{t})$$

although these measures depend on the orbits but not on the individual points.

These expressions are also useful as we will sometimes consider $\mu_{g}^{\text{PS}}$ as a measure on $U$ in an obvious way, for instance, in the following lemma. Let $\mathcal{M}(U)$ be the set of all regular Borel measures on $U$ endowed with the vague topology: $\mu_{n} \to \mu$ if and only if $\int f d\mu_{n} \to \int f d\mu$ for all continuous functions $f$ on $U$ vanishing at infinity.

The following is proved in [8] for $\Gamma$ convex co-compact but the proof works for $\Gamma$ geometrically finite.

**Lemma 2.1.** [8, Lem. 4.1, Cor. 4.2].

- For $g \in G$, the measure $\mu_{g}^{\text{PS}}$ assigns 0 measure to any proper subvariety of $U$.
- The map $g \mapsto \mu_{g}^{\text{PS}}$ is a continuous map from $\{g \in G : g^{+} \in \Lambda(\Gamma)\}$ to $\mathcal{M}(U)$.

For $x = [g] \in \Gamma \backslash G$, we will define a measure $\bar{\mu}_{x}^{\text{PS}}$ on $(U \cap g^{-1}G) \backslash U$ as follows: for $f \in C_{c}(U)$, let $\bar{f} \in C_{c}((U \cap g^{-1}G) \backslash U)$ be given by $\bar{f}([u]) = \sum_{\gamma_{0} \in U \cap g^{-1}G} f(\gamma_{0}u)$. Then the map $f \mapsto \bar{f}$ is a surjective map from $C_{c}(U)$ to $C_{c}((U \cap g^{-1}G) \backslash U)$, and $\int_{[u] \in U \cap g^{-1}G} \bar{f}([u]) \bar{\mu}_{x}^{\text{PS}}([u]) := \int_{U} f(u) d\mu_{g}^{\text{PS}}(t)$ is well-defined, independent of the choice of $g$, by the $\Gamma$-invariance of the PS density. This defines a locally finite measure on $(U \cap g^{-1}G) \backslash U$ by [28, Ch.1].

Now we denote by $\mu_{x}^{\text{PS}}$ the push-forward of the measure $\bar{\mu}_{x}^{\text{PS}}$ from $(U \cap g^{-1}G) \backslash U$ to $xU$ under the injective map $[u] \mapsto xu$ (cf. [25]). The map $\mu_{x}^{\text{Leb}}$ is defined similarly. We caution that $\mu_{x}^{\text{PS}}$ is not a locally finite measure on $\Gamma \backslash G$ unless $xU$ is a closed subset of $\Gamma \backslash G$. For a compact subset $gB \subset gU$ which injects to $\Gamma \backslash G$, we may consider $d\text{t}$ and $d\mu_{x}^{\text{PS}}(t)$ on $xB$ by the push-forwards of the corresponding measures $d\mu_{g}^{\text{Leb}}(t)$ and $d\mu_{g}^{\text{PS}}(t)$ respectively via the injection $gu_{t} \mapsto xu_{t}$.

As $\mu_{x}^{\text{PS}} = \mu_{xU}^{\text{PS}}$, it follows that for any $\psi \in C_{c}(\Gamma \backslash G)$, a Borel subset $B \subset U$ and any $t \in U$, we have

$$\int_{t+r \in B} \psi(xu_{t+r}) d\mu_{x}^{\text{PS}}(t + r) = \int_{r \in B} \psi(xu_{t+r}) d\mu_{xU}^{\text{PS}}(r). \quad (2.1)$$

For $T > 0$, we set

$$B_{U}(T) := \{u_{t} \in U : t \in [-T/2, T/2]^{n-1}\}.$$
For any $s \in \mathbb{R}$, we have:
\[
\mu_g^{\text{PS}}(B_U(e^s)) = e^{\delta s} \mu_{g\cdot x}^{\text{PS}}(B_U(1));
\]
\[
\mu_g^{\text{Leb}}(B_U(e^s)) = e^{(n-1)s} \mu_{g\cdot x}^{\text{Leb}}(B_U(1)).
\]

The following is an immediate corollary of Lemma 2.1.

**Corollary 2.2.** For a compact subset $\Omega \subset X$ and any $r_0 > 0$,
\[
0 < \inf_{x \in \Omega, x^+ \in \Lambda(\Gamma)} \mu_x^{\text{PS}}(xB_U(r_0)) \leq \sup_{x \in \Omega, x^+ \in \Lambda(\Gamma)} \mu_x^{\text{PS}}(xB_U(r_0)) < \infty.
\]

A limit point $\xi \in \Lambda(\Gamma)$ is called radial if some (and hence every) geodesic ray toward $\xi$ has accumulation points in a compact subset of $\Gamma \delta G$, and parabolic if it is fixed by a parabolic element of $\Gamma$ (recall that an element $g \in G$ is parabolic if the set of fixed points of $g$ in $\partial(\mathbb{H}^n)$ is a singleton.) A parabolic limit point $\xi \in \Lambda(\Gamma)$ is called bounded if $\Gamma \xi$ acts co-compactly on $\Lambda(\Gamma) - \{\xi\}$. We denote by $\Lambda_r(\Gamma)$ and $\Lambda_{bp}(\Gamma)$ the set of all radial limit points and the set of all bounded parabolic limit points respectively. As $\Gamma$ is geometrically finite, we have (see [3])
\[
\Lambda(\Gamma) = \Lambda_r(\Gamma) \cup \Lambda_{bp}(\Gamma).
\]

For $x = \Gamma \setminus g$, we write $x^- \in \Lambda_r(\Gamma)$ if $g^- \in \Lambda_r(\Gamma)$; this is well-defined independent of the choice of $g$. If $x^- \in \Lambda_r(\Gamma)$, then the map $u \mapsto xu$ is injective on $U$.

**Lemma 2.3.** For $x \in \Gamma \setminus G$, we have $x^- \in \Lambda_r(\Gamma)$ if and only if $|\mu_x^{\text{PS}}| = \infty$.

**Proof.** If $x^- \notin \Lambda(\Gamma)$, then $\Lambda(\Gamma)$ is a compact subset of $\partial(\mathbb{H}^n) - \{x^-\}$, and hence $|\mu_x^{\text{PS}}| < \infty$. If $x \in \Lambda_{bp}(\Gamma)$, then $xU$ is a closed subset of $\Gamma \setminus G$ and $\mu_x^{\text{PS}}$ is supported on the quotient of $\{xU \cap \partial(\mathbb{H}^n) : (xu)^+ \in \Lambda(\Gamma) - \{x^-\}\}$ by $\Gamma_{x^-}$, which is compact by the definition of a bounded parabolic fixed point. Hence $|\mu_x^{\text{PS}}| < \infty$. If $x^- \in \Lambda_r(\Gamma)$, then $xa_{s_i}$ lies in a compact subset for an infinite sequence $s_i \rightarrow \infty$. If we set $\beta := \inf_i \mu_{xa_{s_i}}(B_U(1))$ which is positive by Corollary 2.2, then
\[
\mu_x^{\text{PS}}(B_U(e^{s_i})) = e^{s_i \delta} \mu_{xa_{s_i}}(B_U(1)) \geq e^{s_i \delta} \cdot \beta.
\]
Hence $|\mu_x^{\text{PS}}| = \infty$.  \hfill \Box

### 2.3. BMS and BR measures.

Fixing $o \in \mathbb{H}^n$, the map
\[
u \mapsto (u^+, u^-), s = \beta_u - (o, \pi(u))
\]
is a homeomorphism between $T^1(\mathbb{H}^n)$ with
\[
(\partial(\mathbb{H}^n) \times \partial(\mathbb{H}^n) - \{(\xi, \xi) : \xi \in \partial(\mathbb{H}^n)\}) \times \mathbb{R}.
\]

Using this homeomorphism, we define measures $\tilde{\nu}_n \subset \mathcal{B}_G$ and $\tilde{\nu}_{\mathbb{H}} = \tilde{\nu}_n^{\mathbb{H}}$ on $T^1(\mathbb{H}^n)$ as follows:
\[
d\tilde{\nu}_n(\nu) = e^{\delta \nu_+ + (o, \pi(u))} e^{\delta \nu_- - (o, \pi(u))} dv_o(u^+)dv_o(u^-)ds; \text{ and}
\]
\[ dm^{\text{BR}}(u) = e^{(n-1)\beta_{u^+}(\pi(u))} e^{\delta_{u^-}(\pi(u))} \, dm_{\text{o}}(u^+) \, dm_{\text{o}}(u^-) \, ds. \]

The conformal properties of \( \{\nu_x\} \) and \( \{m_x\} \) imply that these definitions are independent of the choice of \( o \in \mathbb{H}^n \). Using the identification \( T^1(\mathbb{H}^n) \) with \( G/M \), we lift the above measures to \( G \) so that they are all invariant under \( M \) from the right. By abuse of notation, we use the same notation \( \tilde{m}^{\text{BR}} \) and \( \tilde{m}^{\text{BMS}} \). These measures are all left \( \Gamma \)-invariant, and hence induce locally finite Borel measures on \( \Gamma \setminus G \), which we denote by \( m^{\text{BMS}} \) (the Bowen-Margulis-Sullivan measure, or simply the BMS-measure) and \( m^{\text{BR}} \) (the Burger-Roblin measure or simply the BR-measure) respectively. Note that the supports of \( m^{\text{BMS}} \) and of \( m^{\text{BR}} \) are given by \( \{x \in \Gamma \setminus G: x^+ \in \Lambda(\Gamma)\} \) and \( \{x \in \Gamma \setminus G: x^- \in \Lambda(\Gamma)\} \) respectively. Sullivan showed that \( m^{\text{BMS}} \) is a finite measure [42].

We also consider the contracting horospherical subgroup
\[ U^- = \{ g \in G : a_s g a_s^{-1} \to e \text{ as } s \to \infty \} \]
and the parabolic subgroup
\[ P = M A U^- \]

We note that \( P \) is precisely the stabilizer of \( X_0^+ \) in \( G \). Given \( g_0 \in G \) define the measure \( \nu \) on \( g_0 P \) as follows
\[ d\nu(g_0 \cdot p) = e^{-\delta_{(g_0 \cdot p)^-}(\pi(p))} \, dm_{\text{o}}(g_0 \cdot p^-) \, ds, \quad \text{for } s = \beta_{(g_0 \cdot p)^-}(\pi(p)) \]
where abusing the notation \( \nu_{\text{o}} \) denotes the \( M \) invariant lift of the \( \nu_{\text{o}} \) to \( K \).

This way we get a product structure of the BMS measure which is an important ingredient in our approach: for any \( g_0 \in G \),
\[ \tilde{m}^{\text{BMS}}(\psi) = \int_{g_0 P} \int_{g_0 P U} \psi(g_0 \cdot p_t) \, \mu_{g_0 P}^{\text{PS}}(t) \, d\nu(g_0 \cdot p) \quad (2.2) \]

3. Uniformity in the equidistribution of PS-measure

For section 3 – 5, we let \( G = SO(n,1)^0 \) for \( n \geq 2 \) and \( \Gamma \) be a torsion-free, Zariski dense, geometrically finite subgroup of \( G \). We set \( X = \Gamma \setminus G \) and denote by \( \pi \) the canonical projection from \( G \) to \( X \). For simplicity, we will assume that \( |m^{\text{BMS}}| = 1 \); this can be achieved by replacing \( \{\nu_{\text{x}}\} \) by a suitable scalar multiple if necessary. The right translation action of \( A \) on \( X \) is mixing with respect to \( m^{\text{BMS}} \), as obtained by Flaminio-Spatzier [8] and Winter [43]:

Theorem 3.1. For any \( \psi_1, \psi_2 \in L^2(m^{\text{BMS}}) \), we have
\[ \lim_{s \to \infty} \int_{\Gamma \setminus G} \psi_1(ga_s) \psi_2(g) \, dm^{\text{BMS}}(g) = m^{\text{BMS}}(\psi_1)m^{\text{BMS}}(\psi_2). \]

We note that for each fixed \( \psi_1 \in L^2(m^{\text{BMS}}) \), the above convergence is uniform on any compact family of functions \( \psi_2 \) in \( L^2(m^{\text{BMS}}) \). More precisely,
we will use the following theorem. For \( \varphi \in C^\infty(\Gamma \backslash G) \), we will consider the following Sobolev norm

\[
S_{\infty, 1}(\varphi) = \|\varphi\|_{\infty} + \sum \|Z(\varphi)\|_{\infty}
\]

where the sum is taken over all \( Z \) in a fixed basis of the Lie algebra of \( G \).

**Theorem 3.2.** Fix \( \psi \in L^2(m_{BMS}) \) and a compact subset \( \Omega \) of \( X \). Let \( \mathcal{F} \subset C^\infty(\Omega) \) be a collection of functions satisfying

\[
\sup_{\varphi \in \mathcal{F}} S_{\infty, 1}(\varphi) < \infty.
\]

Then for any \( \varepsilon > 0 \), there exists \( S > 1 \) such that for all \( s \geq S \) and any \( \varphi \in \mathcal{F} \),

\[
\left| \int_X \psi(ga_s) \varphi(g) dm_{BMS}(g) - m_{BMS}(\psi)m_{BMS}(\varphi) \right| \leq \varepsilon.
\]

**Proof.** By the Ascoli-Arzela theorem, \( \mathcal{F} \) is precompact in \( C(\Omega) \) and hence in \( L^2(m_{BMS}) \) as well. Therefore, if the claim does not hold, we will obtain a counter-example to Theorem 3.1. \( \square \)

Theorem 3.1 can be used to prove the following: for any \( \psi \in C_c(X) \), \( x \in X \), and any bounded Borel subset \( B \subset U \) with \( \mu_{PS}(\partial(B)) = 0 \),

\[
\lim_{s \to \infty} \int_B \psi(xu_t a_s) d\mu_x^{PS}(t) = \mu_x^{PS}(B)m_{BMS}(\psi); \quad \text{and} \quad \text{(3.1)}
\]

\[
\lim_{s \to \infty} e^{(n-1-\delta)s} \int_B \psi(xu_t a_s) dt = \mu_x^{PS}(B)m^{BR}(\psi) \quad \text{(3.2)}
\]

(see [33] and [25] for \( M \)-invariant \( \psi \)'s and [23] for general \( \psi \)'s).

In this paper, we will need uniform versions of these two equidistribution statements; more precisely, the convergence in both statements are uniform on compact subsets. Since the uniformity will be crucial for our purpose, and the uniformity is not as straightforward as in the case when \( \Gamma \) is a lattice, we will give a proof. We will be using the following definitions.

**Definition 3.3.** Let \( \Omega \subset X \) be a compact subset.

- The injectivity radius of \( \Omega \) is defined to be the supremum of \( \varepsilon > 0 \) such that the map \( x \mapsto xg \) is injective for all \( x \in \Omega \) and \( g \in G_{\varepsilon} \).
- For each \( x \in X \) and small \( r_0 > 0 \), we denote by \( \mathcal{F}_x(r_0) \) the collection of \( f \in C_c(xB_U(r_0)) \) such that \( 0 \leq f \leq 1 \) and \( f|_{xB_U(r_0/4)} = 1 \).
- For a positive number \( r_0 > 0 \) smaller than the injectivity radius of \( \Omega \), we set \( \mathcal{F}(\Omega, r_0) := \cup_{x \in \Omega} \mathcal{F}_x(r_0) \).

**Lemma 3.4.** For any compact subset \( \Omega \subset G \), there exists \( R_0 > 0 \) such that

\[
(gB_U(−(R_0))^- \cap \Lambda(\Gamma)) \neq \emptyset \quad \text{for any } g \in \Omega.
\]

**Proof.** The claim follows since the map \( g \mapsto d(g^- \Lambda(\Gamma)) \) is continuous where \( d \) is a spherical distance in the boundary \( \partial(H^n) \). \( \square \)
Theorem 3.5. Fix $\psi \in C_c(X)$, a compact subset $\Omega \subset X$ and a positive number $r_0$ which is smaller than the injectivity radius of $\Omega$. For any $\varepsilon > 0$, there exists $S > 1$ such that for all $f_x \in F_x(r_0)$ with $x \in \Omega$ with $x^+ \in \Lambda(\Gamma)$ and $s > S$, we have:

\begin{equation}
\int_{B_U(r_0)} \psi(xu_t a_s) f_x(xu_t) d\mu_x^P(t) - \mu_x^P(f_x) m^{BMS}(\psi) \leq \varepsilon \cdot \mu_x^P(f_x); \tag{1}
\end{equation}

\begin{equation}
e^{(n-1-\delta)s} \int_{B_U(r_0)} \psi(xu_t a_s) f_x(xu_t) dt - \mu_x^P(f_x) m^{BR}(\psi) \leq \varepsilon \cdot \mu_x^P(f_x). \tag{2}
\end{equation}

Proof. Let $\Omega$ be a compact subset of $G$. For $g \in \Omega$ with $g^+ \in \Lambda(\Gamma)$, let $f_g \in C(gB_U(r_0))$ be a function such that $0 \leq f_g \leq 1$ and $f_g|_{B_U(r_0/4)} = 1$. We denote by $F_g$ the collection of all such functions. Fix a non-negative function $\psi \in C_c(G)$, whose support injects to $\Gamma \setminus G$.

Fixing $\varepsilon > 0$, (1) follows if we show that there exists $S > 1$ such that for all $g \in \Omega$ with $g^+ \in \Lambda(\Gamma)$, $f_g \in F_g$ and $s > S$,

\begin{equation}
\sum_{\gamma \in \Gamma} \int \psi(\gamma gu_t a_s) f_g(gu_t) d\mu_g^P(t) - \mu_g^P(f_g) \tilde{m}^{BMS}(\psi) < \varepsilon \cdot \mu_g^P(f_g). \tag{3.3}
\end{equation}

For each $\eta > 0$, we define

$$
\psi^+_\eta(h) := \sup_{w \in G_\eta} \psi(hw) \text{ and } \psi^-_\eta(h) := \inf_{w \in G_\eta} \psi(hw).
$$

By the continuity of $\psi$, there exists $0 < \eta < \varepsilon$ such that $\tilde{m}^{BMS}(\psi^+_\eta - \psi^-_\eta) \leq \varepsilon$.

Recall $P = MAU^-$ and $P_\eta$ denotes the $\eta$-neighborhood of $e$ in $P$. The basic idea is to try to thicken $xB_U(r_0)$ in the transversal direction of $P$, and then to apply the mixing theorem 3.1. However the transverse measure of $xP_\eta$ may be trivial for all $\eta$ smaller than the injectivity radius of $\Omega$, in which case the thickening approach does not work. So we flow $x$ in the backward direction until we find $xa_{-s}$ for which the local transverse measure is non-trivial uniformly over all $x \in \Omega$. Let’s now begin to explain this process carefully. For any $p \in P_\eta$ and $t \in B_U(1)$, we have

$$
p^{-1}u_t \in u_{\rho_p(t)}PD,
$$

where $\rho_p : B_U(1) \to B_U(1 + O(\eta))$ is a diffeomorphism onto this image and $D = D_\eta$ is a constant depending only on $\eta$.

Let $R_0 > 1$ be as in Lemma 3.4 with respect to $\Omega$. Set

$$
s_0 := \log(R_0\eta^{-1}), g_0 := g_{a_{s_0}}, \text{ and } f_0(g_0u_t) := f_g(g_0u_ta_{s_0}).
$$

Then

$$
\nu_{g_0\rho_0}(g_0P_\eta) > 0 \text{ for all } g \in \Omega
$$
since \((g_0 P_\eta)^-\) contains \((g B_U^- (R_\eta))^-\). Let \(\varphi_{g_0}\) be a function supported in \(g_0 B_U (1 + O(\eta)) P_\eta\) and given by
\[
\varphi_{g_0}(g_0 P_\eta(a_s) t) = \frac{f_0}(g_0 P_\eta a_s) \nu(g_0 P_\eta) e^{\frac{\delta}{\epsilon} (u_0 P_\eta)}.
\]

If \(s > \log(D_\eta^{-1}) + \log(2 R_\eta^{-1})\), then
\[
e^{\delta s_0} \int_{B_U (r_0)} \psi(\gamma g^u t a_s) f_g (g^u t) d\mu^P (t)
\]
\[
= \int_{B_U (1)} \psi(\gamma g^u t a_{s-s_0}) f_0 (g^u t) d\mu^P (t) \leq \\ \frac{1}{\nu(g_0 P_\eta)} \int_{P_0} \int_{t \in B_0} \psi^\frac{1}{2} (\gamma g^u t a_{s-s_0}) f_0 (g^u t) \frac{d\mu^P (t)}{d\mu^P (P_0)} d\mu^P (P_0) d\nu (g_0 P_\eta)
\]
\[
= \int_G \psi^\frac{1}{2} (\gamma g^u t a_{s-s_0}) \varphi_{g_0}(h) d m^BMS (h).
\]

We remark that \((g_0 P_\eta)^- (g_0 P_\eta(t))^- \) and \(d\mu^P (t) = e^{\frac{\delta}{\epsilon} (u_0 P_\eta)}\).

Hence if we set
\[
\Psi^+_\eta (\Gamma h) := \sum_{\gamma \in \Gamma} \psi^+_\eta (\gamma h) \quad \text{and} \quad \Phi^+_g (\Gamma h) = \sum_{\gamma \in \Gamma} e^{-\delta s_0} \varphi_{g_0}(\gamma h),
\]
then
\[
\sum_{\gamma \in \Gamma} \int_B \psi(\gamma g^u t a_s) f_g (g^u t) d\mu^P (t) \leq \langle a_{s-s_0}, \Psi^+_\eta, \Phi^+_g \rangle_{BMS}.
\]

Note that
\[
m^BMS (\Phi^+_g) = e^{-\delta s_0} \mu^P (f_0) = \mu^P (f_g).
\]

We claim that there exists \(S > \log(D_\eta^{-1}) + \log(2 R_\eta^{-1})\) such that for any \(s > S\), and any \(g \in \Omega\) (and \(g_0 = ga_{s_0}\)),
\[
\langle a_{s-s_0}, \Psi^+_\eta, \Phi^+_g \rangle_{BMS} \leq (1 + \epsilon) \cdot m^BMS (\Psi^+_\eta) m^BMS (\Phi^+_g).
\]

The collection \(\{\Phi^+_g : g \in \Omega\}\) may not be a relatively compact subset of \(L^2 (m^BMS)\), but we can find \(\Phi^+_g \in C_c (X) : g \in \Omega\) which are all supported in a fixed compact subset of \(X\), \(\Phi^+_g \leq \Phi^+_g, S_1, \infty (\Phi^+_g)\) is uniformly bounded and \(\Phi^+_g - \Phi^+_g \|L^2 (m^BMS) \leq \epsilon\). Note that by (2.2),
\[
\inf_{g \in \Omega, g^+ \in \Lambda (\Gamma)} m^BMS (\Phi^+_g) \gg \inf_{g \in \Omega, g^+ \in \Lambda (\Gamma)} \mu^P (f_g) > 0.
\]

So the collection \(\{\Phi^+_g : g \in \Omega, g^+ \in \Lambda (\Gamma)\}\) satisfies the conditions for \(\mathcal{F}\) in Theorem 3.2. Therefore there exists \(S > \log(D_\eta^{-1}) + \log(2 R_\eta^{-1})\) such that for all \(s > S\) and for all \(g \in \Omega\) with \(g^+ \in \Lambda (\Gamma)\),
\[
\langle a_{s-s_0}, \Psi^+_\eta, \Phi^+_g \rangle_{BMS} - m^BMS (\Psi^+_\eta) m^BMS (\Phi^+_g) \leq \epsilon.
\]
It follows that
\[ \left| \sum_{\gamma \in \Gamma} \int_{gB} \psi(\gamma g u_t a_s) d\mu_{g}^{PS}(t) - \tilde{m}_{BMS}^{\text{BMS}}(\psi) \mu_{g}^{PS}(B) \right| \ll \varepsilon. \]

This finishes the proof of (3.3) and hence (1). Now the uniformity statement regarding (2) follows from the uniformity of (1); this follows directly from the argument in [25] using the comparison of the transversal intersections. □

4. EQUIDISTRIBUTION OF NON-CLOSED \( U \)-ORBITS

We keep notations \( G, \Gamma, B_U(T), \) etc. from 3. We find it convenient to call a point \( x \in X \) with \( x^+ \in \Lambda(\Gamma) \) a PS-point. Recall that \( \Lambda_r(\Gamma) \) denotes the set of radial limit points of \( \Gamma \). One way of characterizing the set \( \Lambda_r(\Gamma) \) in terms of \( U \)-orbits is that if \( x^- \in \Lambda_r(\Gamma) \), then \( u_t \to xu_t \) is an injective map from \( U \) to \( X \), \( xUM \) is not closed in \( X \) and \( \mu_{x}^{PS} \) is an infinite measure (Lemma 2.3).

The following theorem of Schapira shows that for \( x^- \in \Lambda_r(\Gamma) \), most PS-points of \( xB_U(T) \) come back to a compact subset in a quantitative way.

**Theorem 4.1.** [38] For any \( \varepsilon > 0 \) and any compact subset \( \Omega \subset X \), there exists a compact subset \( Q = Q(\varepsilon, \Omega) \subset X \) such that for any \( x \in \Omega \) with \( x^- \in \Lambda_r(\Gamma) \), there exists some \( T_x > 0 \) such that for all \( T \geq T_x \),
\[
\mu_{x}^{PS}(u_t \in B_U(T) : xu_t \in Q) \geq (1 - \varepsilon) \mu_{x}^{PS}(B_U(T)).
\]

We will also make a repeated use of the following basic fact:

**Lemma 4.2.** [38, Lem 4.5] For a fixed \( \kappa > 1 \), there exists \( \beta > 1 \) such that for any compact \( \Omega \subset X \), there exists \( T_0(\Omega) > 1 \) such that for any \( x \in \Omega \) with \( x^- \in \Lambda_r(\Gamma) \) and for any \( T > T_0(\Omega) \), we have
\[
\mu_{x}^{PS}(B_U(\kappa T)) \leq \beta \cdot \mu_{x}^{PS}(B_U(T)).
\]

4.1. Relative PS-size of a neighborhood of \( \partial(B_U(T)) \). In order to study the distribution of \( xB_U(T) \) either in the PS measure or in the Lebesgue measure, it is of crucial importance to understand the size of a neighborhood of \( \partial(xB_U(T)) \) compared to the size of \( xB_U(T) \) in the PS-measure.

We do not know in general whether the following is true: for \( x \in X \) with \( x^- \in \Lambda(\Gamma) \), there exists \( \rho > 0 \) such that
\[
\lim_{T \to \infty} \frac{\mu_{x}^{PS}(B_U(T + \rho) - B_U(T - \rho))}{\mu_{x}^{PS}(B_U(T))} = 0. \tag{4.1}
\]

Noting that
\[
\frac{\mu_{x}^{PS}(B_U(T + \rho) - B_U(T - \rho))}{\mu_{x}^{PS}(B_U(T))} = \frac{\mu_{xa_{-\log T}}^{PS}(B_U(1 + \rho T^{-1}) - B_U(1 - \rho T^{-1}))}{\mu_{xa_{-\log T}}^{PS}(B_U(1))},
\]
this question is directly related to the uniformity of the size of a neighborhood of a 1-sphere based at \( xa_{-\log T} \) where \( x \) is in a fixed compact subset.
When $\Gamma$ is convex-cocompact, and $x^+ \in \Lambda(\Gamma)$ so that $x \in \text{supp}(m_{\text{BMS}})$, then (4.1) now follows since $\text{supp}(m_{\text{BMS}})$ is compact and any 1-sphere has zero PS-measure, and hence $\mu_{x^+}^{\text{PS}}(B_U(1 + \rho T^{-1}) - B_U(1 - \rho T^{-1}))$ is uniformly small for all large $T$. When there is a cusp in $X$ and $xa_{-\log T}$ visits a cusp, the PS-measure of a ball around $xa_{-\log T}$ depends on which cusp it lies in (see Theorem 5.1 for a precise statement), and it is not clear whether 4.1 holds or not.

We will prove the following weaker result which will be sufficient for our purpose. Let us begin with the following lemma which a consequence of the fact that $\Gamma$ is Zariski dense, together with a compactness argument.

For a coordinate hyperplane $L$ set $L^{(1)} := L \cap B_U(1)$. Define $\Delta_{\rho}(L^{(1)})$ to be the $\rho$-thickening of $L^{(1)}$ in the orthogonal direction.

**Lemma 4.3.** Let $Q \subset X$ be a compact subset. For any $\eta > 0$ there exists some $\rho > 0$ so that

$$\mu_y^{\text{PS}}(\Delta_{\rho}(L^{(1)})) < \eta \cdot \mu_y^{\text{PS}}(yB_U(1)).$$

for all $y \in Q \cap \text{supp}(m_{\text{BMS}})$.

**Proof.** We prove this by contradiction. Suppose the above fails; then there exist $\eta > 0$, a sequence $T_k \to \infty$, $y_k \in Q \cap \text{supp}(\mu_x^{\text{PS}})$ and coordinate hyperplanes $L_k$ so that

$$\mu_{y_k}^{\text{PS}}(\Delta_{1/k}(L_k^{(1)})) \geq \eta \cdot \mu_{y_k}^{\text{PS}}(y_k B_U(1)).$$

Passing to a subsequence we may assume $y_k \to y \in Q \cap \text{supp}(m_{\text{BMS}})$ and $L_k = L$ for all $k$.

Now for every $\alpha > 0$ we have $\Delta_{1/k}(L^{(1)}) \subset \Delta_{\alpha}(L^{(1)})$ for all large enough $k$. Since $\mu_{y_k} \to \mu_y$ by Lemma 2.1, it follows that $\mu_y^{\text{PS}}(\Delta_{\alpha}(L^{(1)})) \geq \eta/2$ for all $\alpha > 0$. This implies $\mu_y^{\text{PS}}(L) > 0$ which contradicts the Zariski density of $\Gamma$. This proves the lemma.

**Lemma 4.4.** Let $\Omega \subset X$ be a compact subset. For any $\epsilon > 0$ there exist some $\rho_0 = \rho_0(\epsilon) > 0$ and $R_0 > 1$ (independent of $\epsilon$) so that

$$\mu_x^{\text{PS}}(B_U(T + \rho_0) - B_U(T - \rho_0)) < \epsilon \cdot \mu_x^{\text{PS}}(B_U(T)).$$

for all $T \geq R_0$ and $x \in \Omega$.

**Proof.** Let $c_1 > 1$ and $T_1 > 1$ be so that for all $x \in \Omega$ and $T \geq T_1$,

$$\mu_x^{\text{PS}}(B_U(T + 2)) \leq \mu_x^{\text{PS}}(B_U(2T)) \leq c_1 \mu_x^{\text{PS}}(B_U(T))$$

as given by Lemma 4.2. Apply Theorem 4.1 with $\Omega$ and $\epsilon/(2c_1)$ and let $Q$ and $T_x \geq 1$ be given by that theorem for each $x \in \Omega$.

Since $\cup_{x \in \Omega, 1 \leq T \leq T_x} xa_{-\log T}$ is compact, there exists $R_0 > 1$ such that

$$\eta_0 := \inf_{x \in \Omega, 1 \leq T \leq T_x} \mu_{xa_{-\log T}}^{\text{PS}}(B_U(R_0)) > 0.$$
Since \( \Gamma \) is Zariski dense, \( \mu_y^{PS}(\partial(B_U(R_0))) = 0 \) for any \( y \in X \). Therefore there exists \( 0 < \rho_1 < 1 \) such that

\[
\sup_{x \in \Omega, 1 \leq T \leq T_x} \mu_x^{PS}(B_U(R_0(1 + \rho_1)) - B_U(R_0(1 - \rho_1))) \leq \varepsilon \cdot \eta_0.
\]

Hence for all \( x \in \Omega \) and for all \( 1 \leq T \leq T_x \),

\[
\begin{align*}
\mu_x^{PS}(B_U(TR_0(1 + \rho_1)) - B_U(TR_0(1 - \rho_1))) \\
= T^\delta \mu_x^{PS}(B_U(R_0(1 + \rho_1)) - B_U(R_0(1 - \rho_1))) \\
\leq T^\delta \cdot \varepsilon \cdot \eta_0 \\
\leq T^\delta \cdot \varepsilon \cdot \mu_x^{PS}(B_U(R_0)) \\
\leq \varepsilon \cdot \mu_x^{PS}(B_U(TR_0));
\end{align*}
\]

in other words, for all \( R_0 \leq T \leq R_0T_x \),

\[
\mu_x^{PS}(B_U(T + \rho_1) - B_U(T - \rho_1)) \leq \mu_x^{PS}(B_U(T + T_1) - B_U(T - T_1)) \leq \varepsilon \cdot \mu_x^{PS}(B_U(T)).
\]

In view of this observation, we assume \( T > R_0T_x \) for the rest of the proof. By our choice of \( Q \) and \( T_x \), we have

\[
\mu_x^{PS}\{u_t \in B_U(T + 2) : xu_t \notin Q\} \leq \frac{\varepsilon}{2^{n_1}} \mu_x^{PS}(B_U(T + 2)) \leq \frac{\varepsilon}{2} \mu_x^{PS}(B_U(T)).
\]

For each \( y \in \partial(xB_U(T)) \cap Q \cap \text{supp}(\mu_x^{PS}) \) and any \( \rho > 0 \), put

\[
\Delta_{\rho,Q,T}(y) = x(B_U(T + \rho) - B_U(T - \rho)) \cap yB_U(1).
\]

Note that for each \( y \in \partial(xB_U(T)) \) there exist some coordinate hyperplane \( L_1(y), \ldots, L_{\ell}(y) \), for some \( 1 \leq \ell \leq n - 1 \), so that

\[
\Delta_{\rho,Q,T}(y) \subset \bigcup_{j} y\Delta_{\rho}(L_j(y))^{(1)}.
\]

Since \( \cup\{yB_U(1) : y \in \partial(xB_U(T)) \cap Q \cap \text{supp}(\mu_x^{PS})\} \) is a covering for \( x(B_U(T + \rho_2) - B_U(T - \rho_2)) \cap \text{supp}(\mu_x^{PS}) \), we can find a finite sub-cover \( \cup_{y \in I(T)} yB_U(1) \) with multiplicity at most \( \kappa \), where \( \kappa \) depends only on the dimension of \( U \); this follows from the Besovitch covering theorem.
Let $\rho_2 > 0$ be the constant given by applying Lemma 4.3 with $\eta = \varepsilon/(2c_1\kappa(n - 1))$. Then using that lemma and (4.2) we have

\[
\mu_x^{PS}(B_U(T + \rho_2) - B_U(T - \rho_2)) \\
\leq \mu_x^{PS}(\{u \in B_U(T + 2) : xu \notin Q\}) + \mu_x^{PS}(\cup_{y \in I(T)} \Delta_{\rho_2, Q, T}(y)) \\
\leq \frac{\varepsilon}{2} \mu_x^{PS}(B_U(T)) + \sum_{y \in I(T)} \mu_y^{PS}(\Delta_{\rho_2, Q, T}(y)) \\
\leq \frac{\varepsilon}{2} \mu_x^{PS}(B_U(T)) + (n - 1) \sum_{y \in I(T)} \mu_y^{PS}(\Delta_{\rho_2}(L^{(1)})) \\
\leq \frac{\varepsilon}{2} \mu_x^{PS}(B_U(T)) + \frac{\varepsilon}{2c_1\kappa} \mu_x^{PS}(\cup_{y \in I(T)} yB_U(1)) \\
\leq \frac{\varepsilon}{2} \mu_x^{PS}(B_U(T)) + \frac{\varepsilon}{2c_1\kappa} \mu_x^{PS}(B_U(T + 2)) \\
\leq \varepsilon \mu_x^{PS}(B_U(T)).
\]

The lemma now follows if we let $\rho_0 = \min\{\rho_1, \rho_2\}$.

As can be easily seen in the above proof, we can take $R_0$ to be any positive number if we are only concerned with the set $\{x \in \Omega : x^+ \in \Lambda(\Gamma)\}$.

4.2. **Equidistribution of a $U$-orbit in PS-measure.** We now prove an equidistribution result:

**Theorem 4.5.** Let $x^- \in \Lambda_\nu(\Gamma)$. Then for any $\psi \in C_c(X)$ we have

\[
\lim_{T \to \infty} \frac{1}{\mu_x^{PS}(B_U(T))} \int_{B_U(T)} \psi(xu_t) d\mu_x^{PS}(t) = m^{BMS}(\psi).
\]

The main ingredients of the following proof are Theorem 3.5, Theorem 4.1 and Lemma 4.4. In view of Theorem 4.1, we only need to focus on the part of $xB_U(T)$ which comes back to a fixed compact subset $Q$, as this part occupies most of the PS-measure. We will use a partition of unity argument for a cover of $xB_U(T) \cap Q$ by small balls centered at PS-points. Each function in the partition of unity will be controlled by Theorem 3.5, here the uniformity in loc. cit is of crucial importance. In this process, we have an error occurring in a small neighborhood of the boundary of $xB_U(T)$ and Lemma 4.4 says this error can be controlled.

More precisely, we proceed as follows.

**Proof.** By the assumption that $x^- \in \Lambda_\nu(\Gamma)$, there exists a compact subset $\Omega \subset X$ and a sequence $s_i \to \infty$ such that $xa_{-s_i} \in \Omega$. Let $Q = Q(\varepsilon, \Omega) \subset X$ be a compact subset given by Theorem 4.1. Let $\rho_0 > 0$ and $R_0 > 1$ be the constants given by Lemma 4.4 applied with $\Omega$ and $\varepsilon > 0$. Let $S_0 > 1$ be the constant provided by Theorem 3.5 applied with $\Omega$, $r_0 = \rho_0$ and $\varepsilon > 0$.

Now choose $s_0 > S_0$ so that $x_0 := xa_{-s_0} \in \Omega$. By Theorem 4.1, there exists $T_0 = T_0(x_0) > R_0$ so that for all $T \geq T_0$ we have

\[
\mu_x^{PS}\{u_t \in B_U(T) : x_0u_t \notin Q\} \leq \varepsilon \cdot \mu_x^{PS}(B_U(T)).
\]

(4.3)
For each $T \geq T_0$, consider a covering of the set $x_0 B_U(T) \cap \text{supp}(\mu_x^{PS}) \cap Q$ by $y B_U(\rho_0)$'s with $y \in x_0 B_U(T) \cap \text{supp}(\mu_x^{PS}) \cap Q$. Take a partition of unity $\{f_y : y \in \mathcal{I}(T)\}$ subordinate to this covering, in particular, each $f_y$ is a continuous function supported in $y B_U(\rho_0)$ such that $0 \leq f_y \leq 1$, $f_y | y B_U(\rho_0/4) = 1$, $\sum_{y \in \mathcal{I}(T)} f_y = 1$ on $x_0 B_U(T - \rho)$, and $\sum_{y \in \mathcal{I}(T)} f_y = 0$ outside $x_0 B_U(T + \rho)$.

Fix $\psi \in C_c(X)$. Without loss of generality, we may assume that $\psi$ is non-negative. For $T \geq T_0$, by applying Theorem 3.5 and Lemma 4.4, we have

$$
\int_{x_0 u_t \in Q \cap x_0 B_U(T)} \psi(x_0 u_t a_{s_0}) d\mu_x^{PS}(t) = 
\sum_{y \in \mathcal{I}(T)} \int_{Q \cap x_0 B_U(T)} \psi(x_0 u_t a_{s_0}) f_y(x_0 u_t) d\mu_x^{PS}(t) 
+ O(\mu_x^{PS}(B_U(T + \rho_0) - B_U(T - \rho_0)))
= \sum_{y \in \mathcal{I}(T)} \mu_x^{PS}(f_y) m^{BMS}(\psi)(1 \pm O(\varepsilon)) + O(\varepsilon \cdot \mu_x^{PS}(B_U(T)))
= \mu_x^{PS}(B_U(T)) \cdot m^{BMS}(\psi) + O(\varepsilon \cdot \mu_x^{PS}(B_U(T)))
$$

where the implied constants depend only on $\psi$.

Hence if $T \geq \varepsilon s_0 T_0$, then

$$
\int_{B_U(T)} \psi(x u_t) d\mu_x^{PS}(t) 
= e^{\delta s_0} \int_{B_U(T^{-s_0})} \psi(x_0 u_t a_{s_0}) d\mu_x^{PS}(t) 
= e^{\delta s_0} \left( \int_{x_0 B_U(T^{-s_0}) \cap Q} \psi(x_0 u_t a_{s_0}) d\mu_x^{PS}(t) + O(\varepsilon \cdot \mu_x^{PS}(B_U(T))) \right)
= e^{\delta s_0} (\mu_x^{PS}(B_U(T^{-s_0})) \cdot m^{BMS}(\psi) + O(\varepsilon \cdot \mu_x^{PS}(B_U(T^{-s_0}))))
= \mu_x^{PS}(B_U(T)) \cdot m^{BMS}(\psi) + O(\varepsilon \cdot \mu_x^{PS}(B_U(T)))
$$

where the implied constants depend only on $\psi$.

This finishes the proof as $\varepsilon > 0$ is arbitrary.

\[\square\]

4.3. **Equidistribution of a $U$-orbit in Lebesgue measure.** We will use a similar idea as in the proof of Theorem 4.5 to show the equidistribution of $x B_U(T)$ in Lebesgue measure. The main difference is that we now need to control the escape of the orbit (measured in Lebesgue measure as opposed to the PS measure as in Theorem 4.5) to flares as well as to cusps. For this, we utilize the idea of comparing the “transversal intersections” of the PS-measure of $x U$ and the Lebesgue measure of $x U$. 
Theorem 4.6. Let $x^+ \in \Lambda_r(\Gamma)$. Then for any $\psi \in C_c(X)$, we have

$$\lim_{T \to \infty} \frac{1}{\mu_x^{PS}(B_U(T))} \int_{B_U(T)} \psi(xu_t)dt = m^{BR}(\psi).$$

Proof. Fix $x$ and $\psi$ as above. Without loss of generality, we may assume that $\psi$ is non-negative. Let $P$ denote the parabolic subgroup $MAU^-$. We call $W = z_{P_{\varepsilon_0}}$ an admissible box if $W$ is the injective image of $P_{\varepsilon_0}U_{\varepsilon_0}$ in $\Gamma \backslash G$ under the map $g \mapsto zg$ and $\mu^{PS}_{z_p}(zpU_{\varepsilon_0}) \neq 0$ for all $p \in P_{\varepsilon_0}$.

Since $\psi$ is compactly supported, using a partition of unity argument, we may assume without loss of generality that $\psi$ is supported in an admissible box $z_{P_{\varepsilon_0}}U_{\varepsilon_0}$ for some $z \in \Gamma \backslash G$ and $0 < \varepsilon_1 \leq \varepsilon_0$.

Since $x^+ \in \Lambda_r(\Gamma)$, there exists a compact subset $\Omega \subset \Gamma \backslash G$ and a sequence $s_i \to \infty$ such that $xa_{-s_i} \in \Omega$. Now fix $\varepsilon > 0$ smaller than $\varepsilon_1$ and $\varepsilon_0$. Let $Q = Q(\varepsilon, \Omega) \subset X$ be a compact subset given by Theorem 4.1. Apply Lemma 4.4 applied with $\Omega$ and $\varepsilon > 0$ and let $\rho_0 > 0$ be so that

$$\mu_x^{PS}(x(B_U(T + 4\rho_0) - B_U(T - 4\rho_0))) < \varepsilon$$

for all $T \geq R_0$ and all $x \in \Omega$.

Let $S_0 > 1$ be the constant provided by Theorem 3.5(2) applied with $\Omega$, $\rho_0 = r_0$, and $\varepsilon > 0$. Now choose $s_0 > S_0$ so that $x_0 := xa_{-s_0} \in \Omega$. By Theorem 4.1, there exists $T_0 = T_0(x_0) > R_0$ so that for all $T \geq T_0$, we have

$$\mu_x^{PS}\{u_t \in B_U(T) : x_0u_t \notin Q\} \leq \varepsilon \cdot \mu_x^{PS}(B_U(T)). \quad (4.4)$$

For each $y \in x_0U$, let $f_y \in C(yB_U(\rho_0))$ be a function such that $0 \leq f_y \leq 1$ and $f_y = 1$ on $yB_U(\rho_0/4)$.

**Claim A:** There exists $c > 1$ such that for any $y \in x_0U$,

$$e^{(n-1-\delta)s_0} \int_{x_0U} \psi(yu_t a_{s_0})f_y(yu_t)dt \ll_y \mu_x^{PS}(f_{y,a_{-s_0}x_0}^+)$$

where $f_{y,a}(yu) := \sup_{u \in U} f_y(yuu')$ for $\eta > 0$. To prove this, define

$$P_\eta(s_0) := \{p \in P_{\varepsilon_1} : yB_U(\rho_0)a_{s_0} \cap zpU_{\varepsilon_0} \neq \emptyset\}.$$

For small $\eta > 0$, we define: for $w \in \text{supp}(\psi)G_{\varepsilon_0}$,

$$\psi_\eta^+(w) := \sup_{g \in G_\eta} \psi(wg), \quad \Psi_\eta^+(wp) = \int_{wpU} \psi_\eta^+(wpu_t)dt;$$

and for $zpu \in z_{P_{\varepsilon_0}}U_{\varepsilon_0}$,

$$\bar{\Psi}_\eta^+(zpu) = \frac{1}{\mu^{PS}_{z_p}(zpU_{\varepsilon_0})} \Psi_\eta^+(zp).$$
Then we have, for some fixed constant $c > 0$,

\[
e^{(n-1)\delta_0} \int_{yB_U(\rho_0)} \psi(yu_t a_{s_0}) f_y(yu_t) dt \\
\ll e^{-\delta_0} \sum_{p \in P_x(s_0)} f_{c^{-s_0} \varepsilon_1}(z p a_{s_0}) \psi_{c^{-s_0} \varepsilon_1}(z p) \quad \text{by } [23, \text{Lemma 6.2}]
\]

\[
\ll \int_U \tilde{\Psi}_{c^{-s_0} \varepsilon_1}^+(yu_t a_{s_0}) f_{c^{-s_0} \varepsilon_1}^+(yu_t) d\mu_{y}^{PS}(t) \quad \text{by } [23, \text{Lemma 6.5}]
\]

\[
\ll \mu_{y}^{PS}(f_{y,c^{-s_0} \varepsilon_1})
\]

where the implied constant depends only on $\Psi$. This implies the claim.

Without loss of generality, from now we assume that $s_0$ is big enough so that $c e^{-s_0} \varepsilon_1 \leq \min\{\rho_0, \varepsilon\}$.

For each $T \geq T_0$, consider a covering of the set $x_0 B_U(T) \cap \text{supp}(\mu_{x_0}^{PS}) \cap Q$ by sets of the form $yB_U(\rho_0)$ with $y \in x_0 B_U(T) \cap \text{supp}(\mu_{x_0}^{PS}) \cap Q$. Take a partition of unity $\{f_y : y \in \mathcal{I}(T)\}$ subordinate to this covering and set $Q_0(T)$ to be the closure of the union $\cup_{y \in \mathcal{I}(T)} yB_U(\rho_0)$.

Using the claim (A) applied to a partition of unity subordinate to the covering by $\rho_0$-balls of the set $(x_0 B_U(T_1 + \rho_0) - x_0 B_U(T_1 - \rho_0))$, and the fact that $c e^{-s_0} \varepsilon_1 \leq \rho_0$, we get that

\[
e^{(n-1)\delta_0} \int_{(x_0 B_U(T_1 + \rho_0) - x_0 B_U(T_1 - \rho_0))} \psi(x_0 u_t a_{s_0}) dt \\
\ll \psi \int_{(x_0 B_U(T_1 + 2\rho_0) - B_U(T_1 - 2\rho_0))} \ll \varepsilon \cdot \mu_{x_0}^{PS}(B_U(T_1)).
\]

Therefore if $T_1 := T e^{-s_0} > T_0$,

\[
e^{(n-1)\delta_0} \int_{x_0 B_U(T_1) \cap Q_0(T_1)} \psi(x_0 u_t a_{s_0}) dt \\
= e^{(n-1)\delta_0} \sum_{y \in \mathcal{I}(T_1)} \int_{Q \cap x_0 B_U(T_1)} \psi(x_0 u_t a_{s_0}) f_y(x_0 u_t) dt \\
+ O \left( e^{(n-1)\delta_0} \int_{(x_0 B_U(T_1 + \rho_0) - x_0 B_U(T_1 - \rho_0))} \psi(x_0 u_t a_{s_0}) dt \right) \\
= \sum_{y \in \mathcal{I}(T_1)} \mu_{x_0}^{PS}(f_y) m^{BR}(\psi) (1 \pm O(\varepsilon)) + O(\varepsilon \cdot \mu_{x_0}^{PS}(B_U(T_1)))
\]

\[
= \mu_{x_0}^{PS}(B_U(T_1)) \cdot m^{BR}(\psi) + O(\varepsilon \cdot \mu_{x_0}^{PS}(B_U(T_1)))
\]

where the implied constants depend only on $\psi$.

By considering a partition of unity $\{f_y : y \in \mathcal{J}(T)\}$ subordinate to the cover $x_0 B_U(T_1) \cap (Q^c \cup Q_0(T_1) \cap)$ by $\rho_0$-balls and applying Claim A, Theorem
4.1, and Lemma 4.4, we have
\[ e^{(n-1-\delta)s_0} \int_{x_0B_U(T_1) \cap (Q^c \cup Q_0(T_1)^c)} \psi(x_0u_t a_{s_0}) \, dt \]
\[ \leq e^{(n-1-\delta)s_0} \sum_{y \in J(T)} \int_{x_0B_U(T_1) \cap (Q^c \cup Q_0(T_1)^c)} \psi(yu_t a_{s_0}) f_y(yu_t) \, dt \]
\[ \ll \sum_{y \in J(T)} \mu^\text{PS}_y(f_y) \]
\[ \ll \mu^\text{PS}_x(x_0B_U(T_1 + 1) \cap Q^c) + \mu^\text{PS}_x(x_0(B_U(T_1 + \rho_0) - B_U(T_1 - \rho_0)) \cap Q) \]
\[ \ll \epsilon \cdot \mu^\text{PS}_x(x_0B_U(T_1)). \]
Observing that
\[ \frac{1}{\mu^\text{PS}_x(B_U(T))} \int_{B_U(T)} \psi(xu_t) \, dt = \frac{e^{(n-1-\delta)s_0}}{\mu^\text{PS}_x(B_U(T_1))} \int_{B_U(T)} \psi(x_0u_t a_{s_0}) \, dt \quad (4.5) \]
we have shown that for all \( T > T_0 e^{s_0} \),
\[ \frac{1}{\mu^\text{PS}_x(B_U(T))} \int_{B_U(T)} \psi(xu_t) \, dt = m^\text{BR}(\psi) + O(\epsilon) \]
which finishes the proof. ∎

Both theorems 4.5 and 4.6 are proved in [21] for the case \( n = 2 \). Theorem 4.5 is also proved in [21] for the unit tangent bundle of a convex cocompact hyperbolic \( n \)-manifold. However as clear from the above proofs, the proof in the convex cocompact case is considerably simpler since the support of \( m^\text{BMS} \) is compact.

Remark 4.7. Although Theorems 4.5 and 4.6 are stated for the norm balls \( B_U(T) \) with respect to the maximum norm on \( U \simeq \mathbb{R}^n \), the only property of the max norm we have used is that \( \{ t \in \mathbb{R}^n : \| t \|_{\text{max}} = 1 \} \) is contained in a finitely many union of algebraic sub-varieties in the proof of Lemma 4.3. In fact, our proofs work for any norm \( \| \cdot \| \) on \( \mathbb{R}^n \) as long as the 1-sphere \( \{ t \in \mathbb{R}^n : \| t \| = 1 \} \) is contained in a finitely many union of algebraic sub-varieties.

Remark 4.8. Theorem 4.6 cannot be made uniform on compact subsets; for instance, if \( x^- \) is very close to a parabolic limit point, the convergence is expected to be slower. However, Egorov’s theorem implies that for a given compact subset \( Q \subset X \) and any \( \varepsilon > 0 \), there exists a compact subset \( Q' \) with \( m^\text{BR}(Q - Q') \leq \varepsilon \) on which the convergence in Theorem 4.6 is uniform. We will use this observation later.

5. Window lemma for horospherical averages

Again, we keep the notations set up in section 3. In this section we first prove that for \( x^- \in \Lambda_r(T) \), the PS-measure of \( xB_U(T) \) is not concentrated
near the center $x$, in the sense that for any $\eta > 0$, there exists $r > 0$ such that for all large $T \gg 1$,
\[
\mu_x^{PS}(xB_U(T) - xB_U(rT)) \geq (1 - \eta)\mu_x^{PS}(xB_U(T)). \tag{5.1}
\]

This is of course immediate in the case when $\Gamma$ is a lattice, in which case $\mu_x^{PS}(B_U(T)) = \mu_x^{\text{Loh}}(B_U(T)) = c \cdot T^{\nu - 1}$ for some fixed $c > 0$. The inequality (5.1) also follows rather easily when $\Gamma$ is convex cocompact since $\mu_x^{PS}(xB_U(T)) \asymp T^{d}$ for all $x \in \text{supp}(m^{\text{BMS}})$. For a general geometrically finite group, our argument is based on Sullivan’s shadow lemma. We remark that (5.1) is not a straightforward consequence of Shadow lemma and finding $r$ in (5.1) is rather tricky as we need to consider several different possibilities for the locations of $x_{a - \log T}$ and $x_{a - \log(rT)}$ in the convex core of $\Gamma$ (see the proof of Lemma 5.2).

5.1. **Shadow lemma.** For $\xi \in \Lambda_p(\Gamma)$, the horoball $\mathcal{H}(\xi, R)$ based at $\xi$ and of depth $R$ is defined to be $\{x \in \mathbb{H}^n : \beta_\xi(o, x) > R\}$ for some $R > 0$.

We note that if we choose $g \in G$ so that $g^- = \xi$, then
\[
\mathcal{H}(\xi, R) = \bigcup_{s \geq R} gUa_s(o).
\]

The rank of $\mathcal{H}(\xi, R)$ is defined to be the rank of a finitely generated abelian subgroup $\Gamma_{\xi} := \text{Stab}_G(\xi)$ and is known to be strictly smaller than $2\delta$.

Denote by $C(\Gamma)^\dagger \subset \Gamma/\mathbb{H}^n$ the convex hull of $\Gamma$. The ”thick-thin” decomposition for $C(\Gamma)^\dagger$ says that there exists a $\Gamma$-invariant disjoint system of horoballs $\{\Gamma H_i^\dagger : 1 \leq i \leq \ell\}$ in $\mathbb{H}^n$ such that whenever $\gamma H_i^\dagger \neq H_j^\dagger$, $d(\gamma H_i^\dagger, H_j^\dagger) \geq \eta$ for some $\eta > 0$ independent of $\gamma$ and $i$, and that the complement of $\bigcup_{i=1}^{\ell} \Gamma H_i^\dagger$ in the convex core $C(\Gamma)^\dagger$ is bounded (see [3]). Via the projection $\pi : G \to \mathbb{H}^n$ given by $g \mapsto g(o)$, we lift this thick-thin decomposition of $C(\Gamma)^\dagger$ to its pre-image $C(\Gamma) \subset X = \Gamma \backslash G$:
\[
C(\Gamma) = C_0 \cup \left( \bigcup_{i=1}^{\ell} \Gamma \backslash H_i \right) \tag{5.2}
\]
where $C_0 \subset X$ is an open bounded subset and $\Gamma H_i = \pi^{-1}(\Gamma H_i^\dagger)$.

The following is a variation of Sullivan’s shadow lemma, obtained by Schapira-Maucourant [21]:

**Theorem 5.1.** Let $\Omega \subset X$ be a compact subset. There exists $\lambda = \lambda(\Omega) > 1$ such that for all $x \in \Omega \cap \text{supp}(m^{\text{BMS}})$ with $x^- \in \Lambda_p(\Gamma)$ and $T > 1$,
\[
\lambda^{-1} T^d e^{k(x, T) - \delta d(C_0, x_{a - \log T})} \leq \mu_x^{PS}(B_U(T)) \leq \lambda T^d e^{k(x, T) - \delta d(C, x_{a - \log T})},
\]
where $k(x, T)$ is the rank of $H_i$ if $x_{a - \log T} \in H_i$ for some $i \geq 1$, and equals 0 if $x_{a - \log T} \in C_0$. 
5.2. Non-concentration property of PS measures.

Lemma 5.2 (Window lemma for PS-measure). Let $\Omega \subset X$ be a compact subset. For any $0 < \eta < 1$, there exists $0 < r = r(\eta, \Omega) < 1$ and $T_0 > 1$ such that for all $x \in \Omega$ with $x^- \in \Lambda_r(\Gamma)$ and $T > T_0$, we have

$$\mu^\text{PS}_x(B_U(rT)) \leq \eta \cdot \mu^\text{PS}_x(B_U(T)). \tag{5.3}$$

Proof. Let $C_0$ be given as (5.2). We first claim that it suffices to prove the following: for any $0 < \eta < 1$, there exists $0 < r < 1$ such that for all $y \in C_0 \cap \text{supp}(m^{\text{BMS}})$ with $y^- \in \Lambda_r(\Gamma)$, and $T > 1$, we have

$$\mu^\text{PS}_y(B_U(rT)) \leq \eta \cdot \mu^\text{PS}_y(B_U(T)). \tag{5.4}$$

Without loss of generality, we may assume $\Omega$ contains $C_0$. By Lemma 3.4, there exists $R_0 > 1$ such that for all $x \in \Omega$,

$$xB_U(R_0) \cap (C_0 \cap \text{supp}(m^{\text{BMS}})) \neq \emptyset.$$

Let $T_0 > R_0r^{-1}$, so that we have $rT + R_0 < (2r)T$ for all $T > T_0$.

If $y = xu_t \in xB_U(R_0) \cap C_0 \cap \text{supp}(m^{\text{BMS}})$, then $x, y \in \Omega$. Hence by Lemma 4.2 applied to $\Omega$ and (5.4), there exists $c_0 > 1$ such that

$$c_0 \mu^\text{PS}_y(B_U(rT)) \leq \eta c_0 \mu^\text{PS}_y(B_U(T)) \leq \eta \mu^\text{PS}_x(B_U(T + R_0)).$$

This proves the claim. Therefore, we need to verify (5.4) only for $x \in C_0 \cap \text{supp}(m^{\text{BMS}})$ with $x^- \in \Lambda_r(\Gamma)$. In particular, $xa_{-\log T} \in C(\Gamma)$.

Let $p_0 := \max_i \text{rank}(H_i)$ and $\lambda = \lambda(\Omega)$ be as given in Theorem 5.1. As remarked before, $2\delta > p_0$.

Set

$$r(\eta) := (\eta \lambda^{-2})^{1+1/(2\delta-p_0)}.$$

Since $\eta \lambda^{-2} < 1$, we have $r(\eta) < \min\{(\eta \lambda^{-2})^{1/(2\delta-p_0)}, \eta \lambda^{-2}\}$.

In view of Theorem 5.1, it suffices to show that $r := r(\eta)$ satisfies the following:

$$\eta \lambda^{-1}T^\delta e(k(x,T)-\delta)d(C_0,xa_{-\log T}) \geq r^\delta \lambda^\delta e(k(x,rT)-\delta)d(C_0,xa_{-\log rT}).$$

or equivalently

$$\eta \lambda^{-2}e(k(x,T)-\delta)d(C_0,xa_{-\log T})e(-k(x,rT)+\delta)d(C_0,xa_{-\log rT}) \geq r^\delta. \tag{5.5}$$

From the triangle inequality, we have

$$d(C_0,xa_{-\log T}) - |\log r| \leq d(C_0,xa_{-\log rT}) \leq d(C_0,xa_{-\log T}) + |\log r|. \tag{5.6}$$

We prove this by considering two cases:

**Case 1:** $k(x,T) \geq k(x,rT)$.
Then
\[
(k(x, T) - \delta)d(C_0, xa_{-\log rT}) - (k(x, rT) - \delta)d(C_0, xa_{-\log rT}) \\
\geq (k(x, rT) - \delta)(d(C_0, xa_{-\log rT}) - d(C_0, xa_{-\log rT})) \\
\geq -|k(x, rT) - \delta| \cdot |\log r|.
\]
Hence the lefthand side of (5.5) is bigger than or equal to \(\eta \lambda^{-2r|k(x,rT) - \delta|}\). Considering two cases \(k(x, rT) \leq \delta\) and \(k(x, rT) > \delta\) separately, it is easy to check that our \(r = r(\eta)\) satisfies \(\eta \lambda^{-2r|k(x,rT) - \delta|} \geq r^\delta\), proving (5.5).

**Case 2:** \(k(x, T) < k(x, rT)\).
We first consider the case when \(k(x, T) = 0\), so that \(d(C_0, xa_{-\log rT}) = 0\) and \(0 < d(C_0, xa_{-\log rT}) \leq |\log r|\) by (5.6). Then the left-hand side of (5.5) becomes
\[
\eta \lambda^{-2}e^{-(k(x,rT) + \delta)}d(C_0, xa_{-\log rT}) \geq \eta \lambda^{-2r|k(x,rT) - \delta|} \geq r^\delta
\]
as before, proving the inequality (5.5).

We now assume that \(k(x, T) \geq 1\). Then \(k(x, rT) \geq 2\), and hence \(\delta > 1\). In this case, \(xa_{-\log T}\) and \(xa_{-\log rT}\) are in two distinct horoballs, and hence there exists \(r \leq \rho \leq 1\) such that \(xa_{-\log rT} \in C_0\). We take a maximum such \(\rho\). Then
\[
d(C_0, xa_{-\log rT}) = d(xa_{-\log (\rho T)}, xa_{-\log rT}) \leq |\log \rho|;
\]
d\((C_0, xa_{-\log rT}) = d(xa_{-\log (\rho T)}, xa_{-\log rT}) \leq \log(\rho r^{-1}).
\]
It follows
\[
e^{\rho T(x, -\log T)} d(C_0, xa_{-\log T}) \geq e^{(1-\delta)}d(C_0, xa_{-\log T}) \geq \rho^{\delta-1}
\]
and
\[
e^{\rho T(x, rT) - \delta}d(C_0, xa_{-\log rT}) \leq \max\{1, (\rho/r)^{\rho T(x, rT) - \delta}\}.
\]
Therefore (5.5) is reduced to the inequality
\[
\eta \lambda^{-2} \rho^{\delta - 1} \geq \rho^\delta \max\{1, (\rho/r)^{\rho T(x, rT) - \delta}\}.
\eqno(5.7)
\]
If \(\max\{1, (\rho/r)^{\rho T(x, rT) - \delta}\} = 1\), since \(\rho > r\), this follows from \(\eta \lambda^{-2r^{\delta - 1}} \geq r^\delta\), which holds, by the definition of \(r = r(\eta)\).

It remains to prove that when \(\rho/r)^{\rho T(x, rT) - \delta} \geq 1\),
\[
\eta \lambda^{-2} \rho^{2\delta - k(x, T)} \geq r^{2\delta - k(x, rT)}.
\eqno(5.8)
\]
By our definition, we have \(r(\eta) \leq \eta^{-2} \lambda^{-1/(2\delta - p_0)}\). Therefore we have
\[
\rho^{2\delta - 1 - k(x, rT)} / r^{2\delta - k(x, rT)} \geq \max\{1, (\rho/r)^{2\delta - k(x, rT)}\}.
\]
The conclusion now follows by taking two cases: \(\rho \leq r(\lambda^2 \eta^{-1})^{1/(2\delta - k(x, rT))}\) and alternatively \(r(\lambda^2 \eta^{-1})^{1/(2\delta - k(x, rT))} \leq \rho \leq 1\). This completes the proof. □
5.3. Equidistribution for windows. We now draw the following corollaries of Lemma 5.2, and Theorems 4.5 and 4.6.

For $\psi \in C_c(\Gamma \setminus G)$ and $T > 0$, we define the notation

$$P_T \psi(x) = \int_{B_U(T)} \psi(xu_t) d\mu_{x}^{PS}(t);$$

$$L_T \psi(x) = \int_{B_U(T)} \psi(xu_t) dt.$$

**Theorem 5.3** (Window lemma for horospherical average). Fix a compact subset $\Omega \subset X$. For any $\eta > 0$, there exists $0 < r = r(\eta, \Omega) < 1$ such that the following holds:

1. For $x \in \Omega$ with $x^- \in \Lambda_r(\Gamma)$ and for any non-negative $\psi \in C_c(X)$ with $m^{BR}(\psi) > 0$, there exists $T_0 = T_0(x, \psi)$ such that

$$L_{rT} \psi(x) \leq \eta \cdot L_T \psi(x) \quad \text{for all } T > T_0.$$

2. For $x \in \Omega$ with $x^- \in \Lambda_r(\Gamma)$ and for any non-negative $\psi \in C_c(X)$ with $m^{BMS}(\psi) > 0$, there exists $T_0 = T_0(x, \psi)$ such that

$$P_{rT} \psi(x) \leq \eta \cdot P_T \psi(x) \quad \text{for all } T > T_0.$$

**Proof.** Let $r = r(\eta/4, \Omega)$ be as in Lemma 5.2. Let $x \in \Omega$ with $x^- \in \Lambda_r(\Gamma)$. By Theorem 4.6, there exists $T_0 = T_0(x, \psi)$ so that for all $T > T_0$,

$$L_{rT} \psi(x) \leq 2m^{BR}(\psi) \mu_{x}^{PS}(B_U(rT))$$

$$L_T \psi(x) \geq \frac{1}{2} m^{BR}(\psi) \mu_{x}^{PS}(B_U(T)).$$

Hence

$$L_{rT} \psi(x) \leq 4 \frac{\mu_{x}^{PS}(B_U(rT))}{\mu_{x}^{PS}(B_U(T))} L_T \psi(x) \leq \eta \cdot L_T \psi(x)$$

by the choice of $r$. This proves (1). (2) is proved similarly using Theorem 4.5 in place of 4.6. \hfill $\square$

**Theorem 5.4** (Equidistribution for window averages). For any compact subset $\Omega \subset X$, the following hold: for any $x \in \Omega$ with $x^- \in \Lambda_r(\Gamma)$ and any $\varphi \in C_c(X)$, we have

1. $$\lim_{T \to \infty} \frac{\int_{B_U(T)-B_U(rT)} \varphi(xu_t) dt}{\mu_{x}^{PS}(B_U(T) - B_U(rT))} = m^{BR}(\varphi); \quad (5.9)$$

2. $$\lim_{T \to \infty} \frac{\int_{B_U(T)-B_U(rT)} \varphi(xu_t) d\mu_{x}^{PS}(t)}{\mu_{x}^{PS}(B_U(T) - B_U(rT))} = m^{BMS}(\varphi). \quad (5.10)$$

where $r = r(\Omega, 1/2)$ be as in Lemma 5.2 for $\eta = 1/2$. 


Proof. By Theorem 4.6, we have
\[ L_T \varphi(x) = m^{BR}(\varphi) \cdot \mu_x^{PS}(B_U(T)) + a_T \quad \text{with} \quad a_T = o(\mu_x^{PS}(B_U(T))); \]
\[ L_{rT} \varphi(x) = m^{BR}(\varphi) \cdot \mu_x^{PS}(B_U(rT)) + b_T \quad \text{with} \quad b_T = o(\mu_x^{PS}(B_U(rT))). \]
Since \( \mu_x^{PS}(B_U(T) - B_U(rT)) \geq \frac{1}{2} \mu_x^{PS}(B_U(T)) \), it follows that \( |a_T| + |b_T| = o(\mu_x^{PS}(B_U(T) - B_U(rT))) \). Hence (1) follows. Similarly (2) can be seen using Theorem 4.5.
\( \square \)

Remark 5.5. In view of remark (4.8), for any compact subset \( Q \) of \( X \) and \( \varepsilon > 0 \), there exists a compact subset \( Q' \) with \( m^{BR}(Q - Q') \leq \varepsilon \) on which the convergences in Theorem 5.4 hold uniformly.

5.4. Remark on measure classification. Burger [4] classified all \( U \)-invariant measures on \( \Gamma \setminus \text{PSL}_2(\mathbb{R}) \) when \( \Gamma \) is convex cocompact with \( \delta > 1/2 \). Roblin [33] generalized Burger’s work in greater generality, and classified all \( UM \)-invariant ergodic measures on \( \Gamma \setminus G \) when \( \Gamma \) is a geometrically finite subgroup of a simple Lie group of rank one. Extending this work, Winter [43] obtained a classification of all \( U \)-invariant ergodic measures on \( \Gamma \setminus G \) when \( \Gamma \) is also assumed to be Zariski dense. In the case of \( G = \text{SO}(n, 1) \) and \( \Gamma \) geometrically finite, we can also deduce this classification result from Theorem 3.5, using the Hopf ratio theorem.

First, recall the Hopf ratio theorem proved by Hochman formulated in a setting we are concerned with:

**Theorem 5.6.** [11] Let \( H \) be a connected Lie group and \( \Gamma \) a discrete subgroup. Let \( \mathbb{R}^k = N \subset H \) be a connected abelian subgroup. Let \( \mu \) be a locally finite \( N \)-invariant ergodic measure on \( \Gamma \setminus G \). Let \( \psi_1, \psi_2 \in C_c(\Gamma \setminus H) \). Suppose \( \psi_2 \geq 0 \). Then for \( \mu \)-almost all \( x \) such that \( \int_{B_N(x)} \psi_2(xu)du = \infty \),
\[ \frac{\int_{B_N(T)} \psi_1(xu)du}{\int_{B_N(T)} \psi_2(xu)du} \to \frac{\mu(\psi_1)}{\mu(\psi_2)}. \]

**Theorem 5.7.** The only ergodic \( U \)-invariant measure on \( \Gamma \setminus G \) which is not supported on a closed orbit of \( MU \) is the \( BR \) measure.

Proof. Let \( \mu \) be such a measure. Let \( \psi \in C_c(X) \) be a non-negative function so that \( \mu(\psi) > 0 \) and \( m^{BR}(\psi) > 0 \). Then since the support of \( \mu \) is not contained in any closed \( MU \)-orbit, there exists \( x \in X \) with \( x^- \in \Lambda_r(\Gamma) \) and the Hopf ratio ergodic theorem holds: for all \( \varphi \in C_c(X) \) we have
\[ \lim_{T \to \infty} \frac{L_T \varphi(x)}{L_T \psi(x)} = \frac{\mu(\varphi)}{\mu(\psi)}. \] (5.11)

Therefore
\[ \frac{\mu(\varphi)}{\mu(\psi)} = \frac{m^{BR}(\varphi)}{m^{BR}(\psi)} \]

It follows that \( \mu \) and \( m^{BR} \) are proportional to each other. \( \square \)
We mention that when $G$ is a general simple group of rank one, we expect an analogue of Theorem 4.6 holds. However in these cases, the horospherical subgroup is not abelian any more and the Hopf ratio theorem is not available for a general non-abelian nilpotent group action (cf. [12]). However a weaker type of the Hopf ratio theorem is still available (see [12, Theorem 1.4]) and together with this, it is plausible that an analogue of Theorem 4.6 would yield an alternative proof for the above mentioned measure classification theorem.

6. Rigidity of $AMU$-equivariant maps

For the rest of the paper, we let $F = \mathbb{R}$ or $\mathbb{C}$ and $G = \text{PSL}_2(F)$. Let

$$U := \left\{ u_t = \begin{pmatrix} 1 & 0 \\ t & 0 \end{pmatrix} : t \in F \right\}, \quad U^- := \left\{ u_{r}^- = \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} : r \in F \right\},$$

and

$$A = \left\{ a_s = \begin{pmatrix} e^{s/2} & 0 \\ 0 & e^{-s/2} \end{pmatrix} : s \in \mathbb{R} \right\}.$$

Let

$$M = \begin{cases} \{e\} & \text{for } G = \text{PSL}_2(\mathbb{R}) \\ \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} : \theta \in \mathbb{R} \right\} & \text{for } G = \text{PSL}_2(\mathbb{C}). \end{cases} (6.1)$$

Since $M$ is considered as a subgroup of $\text{PSL}_2(\mathbb{C})$, two elements which differ by $-1$ are identified.

Let $\Gamma_1$ and $\Gamma_2$ be geometrically finite, and Zariski dense subgroups of $G$ and set $X_i := \Gamma_i \backslash G$ for $i = 1, 2$. We denote by $m_{\Gamma_i}^{\text{BR}}$ and $m_{\Gamma_i}^{\text{BMS}}$ the BR-measure and the BMS-measure on $X_i$, respectively, associated to $\Gamma_i$ for each $i = 1, 2$. We assume that $|m_{\Gamma_1}^{\text{BMS}}| = |m_{\Gamma_2}^{\text{BMS}}| = 1$. When there is no room for confusion, we will omit the subscript $\Gamma_i$ from the notation of these measures.

Suppose

$$v_1, \ldots, v_l : X_1 \to X_2$$

are Borel measurable maps and consider a set-valued map:

$$\Upsilon(x) = \{v_1(x), \ldots, v_l(x)\}. \quad (6.2)$$

The main aim of this section is to prove the following theorem that if $\Upsilon$ is $AMU$-equivariant on a BMS-conull set, it is also $U^-$-equivariant.

**Theorem 6.1.** Suppose that there exists a Borel subset $X' \subseteq \text{supp}(m_{\Gamma_1}^{\text{BMS}})$ with $m_{\Gamma_1}^{\text{BMS}}(X') = 1$ such that for all $x \in X'$ and every $amu_t \in AMU$, we have

$$\Upsilon(xamu_t) = \Upsilon(x)amu_t.$$

Then there exists a BMS-conull subset $X'' \subseteq X'$ such that for all $x \in X''$ and for every $u_{r}^- \in U^-$ with $xu_{r}^- \in X''$, we have

$$\Upsilon(xu_{r}^-) = \Upsilon(x)u_{r}^-.$$
This is proved in [8] for the case $\Gamma$ is convex cocompact and $\ell = 1$; the proof is based on Ratner’s proof of the rigidity of $U$-factors [29] in the lattice case. Here we use similar strategy and generalize this to the case of a geometrically finite group allowing also $\ell \geq 1$. The presence of cusps requires extra care in this extension.

Let us recall that following terminology from [15]. Let $C, \alpha > 0$ and we denote by $| \cdot |$ the absolute value of $F$. A function $f : F^n \to F$ is said to be $(C,\alpha)$-good on a ball $B$ if the following holds: for any ball $V \subset B$ and any $\varepsilon > 0$ we have

$$\ell\{x \in V : f(x) < \varepsilon \} \leq C \left( \frac{\varepsilon}{\sup_{f < \varepsilon} |f|} \right)^{\alpha} \ell(V)$$

(6.3)

where $\ell$ denotes the Lebesgue measure on $F^n$. It follows from Lagrange’s interpolation and induction that if $f$ is a polynomial in $n$ variables and of degree bounded by $d$, then $f$ is $(C,\alpha)$-good on $F^n$ where $C$ and $\alpha$ depend only on $n$ and $d$.

The $(C,\alpha)$-good property for fractal measures was studied in [14]. We need the following lemma (a version of this is [8, Lemma 5.1] for $\Gamma$ convex cocompact); our proof is soft and uses compactness arguments. This can be thought of as a weak form of the $(C,\alpha)$-good property of polynomials. Recall a point $x \in X_1$ is called a BMS-point (resp. a BR-point) if it lies in the support of $m^{BMS}$ (resp. $m^{BR}$).

Lemma 6.2. Let $\mathcal{P}_d$ be the set of polynomial maps $\Theta : U \to F$ with degree at most $d$. For any compact subset $\mathcal{K} \subseteq X_1$, there exists some $C_1 > 0$ depending on $d$ and $\mathcal{K}$ with the following properties:

1. Let $x \in \mathcal{K}$ be a BMS point and let $s \in \mathbb{R}$ be so that $xa_{-s} \in \mathcal{K}$. Then for any $\varepsilon > 0$ and any $\Theta \in \mathcal{P}_d$, we have

$$\frac{1}{\mu^{PS}_x(B_U(e^s))} \int_{B_U(e^s)} |\Theta(t)| d\mu^{PS}_x(t) \geq C_1 \cdot \sup_{t \in B_U(x^s)} |\Theta(t)|.$$  (7.1)

2. Suppose $\mathcal{K}$ is so that $m^{BR}(\partial \mathcal{K}) = 0$. Let $x \in \mathcal{K}$ be a BR point and $s \in \mathbb{R}$ be so that $xa_{-s} \in \mathcal{K}$. Fix $R \geq 1$ so that $\inf_{x \in \mathcal{K}, x^s \in \Lambda(\Gamma)} \mu^{PS}_x(B_U(R)) > 0$. Then for any $\varepsilon > 0$ and any $\Theta \in \mathcal{P}_d$, we have

$$\frac{\int_{B_U(2Re^s)} \chi(xut)|\Theta(t)| dt}{\int_{B_U(2Re^s)} \chi(xut) dt} \geq C_1 \cdot \sup_{t \in B_U(2Re^s)} |\Theta(t)|.$$  (7.2)

Proof. Write $\mathcal{K}^{BMS} = \mathcal{K} \cap \text{supp}(m^{BMS})$. Note that the above statement (1) is invariant under scaling the polynomial $\Theta$. Further, for any $x \in \mathcal{K}^{BMS}$, any $s \in \mathbb{R}$ and all $\Theta \in \mathcal{P}_d$, we have

$$\frac{1}{\mu^{PS}_x(B_U(e^s))} \int_{B_U(e^s)} |\Theta(t)| d\mu^{PS}_x(t) = \frac{1}{\mu^{PS}_x(B_U(e^s))} \int_{B_U(1)} \Theta(e^s t) d\mu_x(e^s t) = \frac{1}{\mu^{PS}_{xa_{-s}}(B_U(1))} \int_{B_U(1)} \tilde{\Theta}(t) d\mu_{xa_{-s}}(t)$$
where $\tilde{\Theta}(t) := \Theta(e^t)$. Suppose now the statement (1) fails. Then we have

- a sequence $x_i \in \mathcal{K}^{BMS}$, a sequence $s_i \to \infty$ such that $y_i := x_i x_{-s_i} \in \mathcal{K}$,
- a sequence $\Theta_i \in \mathcal{P}_d$ with $\sup_{B_U(1)} |\Theta_i| = 1$

so that $\frac{1}{\mu_y^{PS}(B_U(1))} \int_{B_U(1)} |\Theta_i(u_t)|d\mu^{PS}_y(t) \to 0$ as $i \to \infty$.

Passing to a subsequence we may assume that $y_i \to y \in \mathcal{K}^{BMS}$ and $\Theta_i \to \Theta \in \mathcal{P}_d$ with $\sup_{B_U(1)} |\Theta| = 1$. Since the map $x \mapsto \mu_x^{PS}$ is continuous on $\mathcal{K}^{BMS}$ and $0 < \inf_{x \in \mathcal{K}^{BMS}} \mu_x^{PS}(B_U(1)) \leq \sup_{x \in \mathcal{K}^{BMS}} \mu_x^{PS}(B_U(1)) < \infty$, it follows that

$$\int_{B_U(1)} |\Theta(t)|d\mu^{PS}_y(t) = 0.$$

This implies that $\mu_y^{PS}(B_U(1) \cap \{ t : \Theta(t) \neq 0 \}) = 0$ which contradicts the fact that $y \in \text{supp}(m^{BMS})$ in view of Zariski density of $\Gamma$, proving (1).

We now turn to the proof of (2). Recall the notation $\tilde{\Theta}(t) = \Theta(e^t)$, now making a change of variables, we can write

$$\int_{B_U(2R^e)} \chi_{\mathcal{K}}(xu_t)|\Theta(t)|dt = \int_{B_U(2R)} \chi_{\mathcal{K}}(xa_{-s}u_ta_s)|\tilde{\Theta}(t)|dt.$$

Put $x_s = xa_{-s}$. The above using the uniformity statement proved in Theorem 3.5 and the choice of $R$ we get that for all large enough $s$,

$$\frac{\int_{B_U(2R^e)} \chi_{\mathcal{K}}(xu_t)|\Theta(t)|dt}{\int_{B_U(2R^e)} \chi_{\mathcal{K}}(xu_t)dt} \geq C' \frac{1}{\mu_x^{PS}(B_U(2R))} \int_{B_U(2R^e)} |\tilde{\Theta}(t)|d\mu^{PS}_x(t) = C' \frac{1}{\mu_x^{PS}(B_U(2R^e))} \int_{B_U(2R^e)} |\Theta(t)|d\mu^{PS}_x(t) \to 0$$

where $C'$ depends only on $\mathcal{K}$. The claim in (2) now follows from (1). \qed

We also recall the following mean ergodic theorem.

**Theorem 6.3.** [34, Thm. 17] For any Borel set $\mathcal{K}$ of $X_1$ and any $\eta > 0$, there exists some $T_0 > 1$ so that for all $T \geq T_0$, the set

$$\{ x \in X : \frac{1}{\mu_x^{PS}(B_U(T))} \int_{B_U(T)} \chi_{\mathcal{K}}(xu_t)d\mu^{PS}_x(t) \geq (1 - \eta)m^{BMS}(\mathcal{K}) \}$$

has BMS measure at least $1 - \eta$.

**Proof of Theorem 6.1.** Fix a small $\eta > 0$. Then there exists a compact subset $\mathcal{K}_0 \subset X'$ with $m^{BMS}(\mathcal{K}_0) > 1 - \eta$ so that $u_i$ is continuous on $\mathcal{K}_0$ for all $1 \leq i \leq l$ (Lusin’s theorem). For each large $T$, let $\mathcal{K}_T$ be the set of points $x \in X_1$ with

$$\frac{1}{\mu_x^{PS}(B_U(T))} \int_{B_U(T)} \chi_{\mathcal{K}_0}(xu_t)d\mu^{PS}_x(t) \geq (1 - \eta)m^{BMS}(\mathcal{K}_0) \geq 1 - 3\eta. \quad (6.4)$$
By Theorem 6.3, \( m^{BMS}(\mathcal{K}_T) \to 1 \) as \( T \to \infty \). Fix \( T_0 > 1 \) so that \( m^{BMS}(\mathcal{K}_{T_0}) \geq 1 - \eta \) and a compact subset \( \mathcal{K}'_{T_0} \) of \( \mathcal{K}_{T_0} \) with \( m^{BMS}(\mathcal{K}'_{T_0}) \geq 1 - 2\eta \). Set \( \mathcal{K} := \mathcal{K}_0 \cap \mathcal{K}'_{T_0} \). Then

\[ m^{BMS}(\mathcal{K}) \geq 1 - 4\eta. \]

By the ergodicity of \( A \)-flow, which follows from Theorem 3.1, and the Birkhoff ergodic theorem, there exists a conull subset \( X'' \) of \( X' \) such that for all \( x \in X''A \),

\[
\lim_{S \to \infty} \frac{1}{S} \int_0^S \chi_{\mathcal{K}}(xa_s)ds = m^{BMS}(\mathcal{K}) \geq 1 - 2\eta \tag{6.5}
\]

Let \( x \in X'' \) and \( u^-_r \in U^- \) such that \( xu^-_r \in X'' \). We will show \( \Upsilon(xu^-_r) = \Upsilon(x)u^-_r \).

For any \( \epsilon > 0 \) let \( \epsilon' \leq \epsilon/2 \) be so that \( d(v_i(x), v_i(x')) \leq \epsilon/2 \) for all \( 1 \leq i \leq l \) and all \( x, x' \in \mathcal{K} \) with \( d(x, x') \leq \epsilon' \).

By (6.5), we can choose arbitrarily large \( s_0 > 1 \) such that \( xa_{s_0}, xu^-_r a_{s_0} \in \mathcal{K} \), and \( e^{-s_0}r \) is of size at most \( \epsilon' \). Setting \( x_0 := xa_{s_0} \) and \( r_0 := e^{-s_0}r \), it suffices to show that \( \Upsilon(x_0 u^-_r) = \Upsilon(x_0)u^-_r \), thanks to the \( A \)-equivariance. Therefore, we may assume without loss of generality that \( x \in X''A \cap \mathcal{K}, \ xu^-_r \in X''A \cap \mathcal{K}, \ |r| \leq \epsilon' \).

By (6.5), we have a sequence \( s_m \to \infty \) so that \( xu^-_ra_{s_m}, xa_{s_m} \in \mathcal{K} \) for all \( m \). Therefore it suffices to show that for all sufficiently large \( s \gg 1 \) with \( xa_s, xu^-_ra_s \in \mathcal{K} \), we have

\[
\Upsilon(xu^-_ra_s) \subseteq \Upsilon(x) \cdot \{g \in G : d(e, g) \leq e^{-s}\}. \tag{6.6}
\]

So we now suppose that \( xu^-_ra_s, xa_s \in \mathcal{K} \) for some large \( s > 1 \) and for some \( |r| \leq \epsilon' \).

Since \( \Upsilon(xu^-_ra_s) = \Upsilon(x)u^-_ra_s \) is equivalent to \( \Upsilon(xu^-_ra_s) = \Upsilon(x)u^-_ra_s \), we will compare

- \( \Upsilon(xu^-_ra_s) = \Upsilon(x)u^-_ra_s \); and equivalently
- \( \Upsilon(xu^-_ra_s) \in \Upsilon(xa_s) \cap (a_{-s}u^-_ra_s) \); or equivalently
- \( \Upsilon(xu^-_ra_s)(a_{-s}u^-_ra_s)^{-1} \in \Upsilon(xa_s) \); or equivalently.

\( \Upsilon(xu^-_ra_s)u^-_{e^{-r}r} \) with \( \Upsilon(xa_s) \),

since \( (a_{-s}u^-_ra_s)^{-1} = u^-_{e^{-r}r} \) and \( \Upsilon(ya_s) = \Upsilon(y)a_s \) for any \( y \in \mathcal{K} \).

In order to compare \( \Upsilon(xu^-_ra_s)u^-_{e^{-r}r} \) and \( \Upsilon(xa_s) \), we will study the divergence of these points along \( u_t \) flow, and show that for all \( t \in B_U(e^s) \), \( \Upsilon(xu^-_ra_s)u^-_{e^{-r}r} u_t \) and \( \Upsilon(xa_s) u_t \) stay within \( \epsilon \)-distance from each other. This will have an implication on the shape of the element \( g_s \in G \) which measures their difference, i. e. if \( g_s \) is defined by \( \Upsilon(xu^-_ra_s)u^-_{e^{-r}r} = \Upsilon(xa_s)g_s \), then \( d(e, a_s g_s a_{-s}) = O(e^{-r}) \). Since \( \Upsilon \) is a set map, we need to keep track of each \( v_i \) separately.

We now give a rigorous argument for (6.6), assuming that \( xu^-_ra_s, xa_s \in \mathcal{K} \) for some large \( s > 1 \) and for some \( |r| \leq \epsilon' \).
We compute that for all $t$ we have
\[ u_{e^{-s}r}u_t = u_t \begin{pmatrix} 1 + e^{-s}rt & e^{-s}r \\ -e^{-s}rt & 1 - e^{-s}rt \end{pmatrix}. \]

Therefore, for all $|t| \leq e^8$ we have
\[ xu_{e^{-s}r}a_su_t = xa_su_{e^{-s}r}u_t = u_{a_su_{e^{-s}r}(t)}g_r, \]
where $\beta_r(t) = t - \frac{e^{-s}r^2t}{1 + e^{-s}rt}$ and $g_r := g_{r,s} \in AMU^-$ with $d(e, g_r) \leq \varepsilon'$.

Using the $U$-equivariance, we have, for each $1 \leq i \leq t$ and $t \in F$ we have
\[ v_1(xu_{e^{-s}r}a_s)u_{e^{-s}r}u_t = v_1(xu_{e^{-s}r}a_s)u_{a_su_{e^{-s}r}(t)}g_r \]
by (6.7)
\[ = v_{j(i,t)}(xa_su_{e^{-s}r}a_su_{a_su_{e^{-s}r}(t)})g_r \]
for some $1 \leq j(i,t) \leq t$.

This calculation together with (6.7) and the continuity of $v_{j(i,t)}$ on $\mathcal{K}$ imply that for all $|t| \leq e^8$ such that $xa_su_{e^{-s}r}a_su_{a_su_{e^{-s}r}(t)} \in \mathcal{K}$, we have
\[ d(v_1(xu_{e^{-s}r}a_s)u_{e^{-s}r}u_t, v_{j(i,t)}(xa_su_t)) \leq \varepsilon, \]
as well as
\[ d(v_1(xu_{e^{-s}r}a_s)u_{e^{-s}r}u_t, \Upsilon(xa_su_t)) \leq \varepsilon \]
by the $U$-equivariance of $\Upsilon$.

Let
\[ K_2(s, r) = \{ t \in B_U(e^8) : xa_su_{e^{-s}r}a_su_{a_su_{e^{-s}r}(t)} \in \mathcal{K} \}. \]

We claim that
\[ \mu_{xa_s}^{PS}(K_2(s, r)) \geq (1 - c\eta)\mu_{xa_s}^{PS}(B_U(e^8)) \]
for some $c$ depending only on $\mathcal{K}$. Since $xa_s, xu_{e^{-s}r}a_s \in \mathcal{K}$, we get from (6.4) that
\[ \mu_{xa_s}^{PS}(\{ t \in B_U(e^8) : xa_su_{e^{-s}r}a_su_{a_su_{e^{-s}r}(t)} \in \mathcal{K} \}) \geq (1 - \eta)\mu_{xa_s}^{PS}(B_U(e^8)) \]
and
\[ \mu_{xa_s}^{PS}(\{ t \in B_U(e^8) : xu_{e^{-s}r}a_su_{e^{-s}r}u_t \notin \mathcal{K} \}) \geq (1 - \eta)\mu_{xa_s}^{PS}(B_U(e^8)). \]

For all $t \in B_U(e^8)$, $\text{Jac}(\beta_r(t)) = 1 + O(\varepsilon')$, and hence (6.11) implies that
\[ \mu_{xa_s}^{PS}(\{ t \in B_U(e^8) : xu_{e^{-s}r}a_su_{a_su_{e^{-s}r}(t)} \notin \mathcal{K} \}) \leq \mu_{xa_s}^{PS}(B_U(e^8 + \varepsilon')) \]
\[ \leq (c_1\eta) \cdot \mu_{xa_s}^{PS}(B_U(e^8)) \]
where $c_1$ is given by Lemma 4.2. This is equivalent to saying that
\[ \mu_{xa_s}^{PS}(\{ t \in B_U(e^8) : xu_{e^{-s}r}a_su_{a_su_{e^{-s}r}(t)} \notin \mathcal{K} \}) \leq (c_1\eta) \cdot \mu_{xa_s}^{PS}(B_U(e^8)) \]
(6.12)
Note that \((x a_s u_{x,t}^- u_{β,s}(t))^+ = (x a_s u_t)^+\) since \(g_{x,t} \in AMU^-.\) It follows from the definition of the PS-measure \(μ_{x,s}^{PS}(t) = e^{δβ(yu,t) - (α,yu)} ν_o((yu)^+)\) and the fact \(|e^{-s}r| = O(e^{-s}e')\) that
\[
dμ_{x,s}^{PS}(u_{β,s}(t)) = (1 + O(e^{-s}e'))dμ_{x,s}(u_t)
\]
on \(B_U(e^s)\).

Therefore (6.12) implies that
\[
μ_{x,s}^{PS}(t) \in B_U(e^s) : xu^-a_s u_{β,s}(t) \notin \mathcal{K} \leq (c_2η) \cdot μ_{x,s}^{PS}(B_U(e^s))
\]
for some \(c_2 > 0\) depending only on \(\mathcal{K}\). Hence this together with (6.10) implies the claim (6.9).

Put
\[
Θ_{x,s}(t) = \min\{d(v_i(xu^-a_s)u_{x,t}^-)u_t, Ψ(xa_s)u_t^2, 1\}.
\]

Then
\[
\frac{1}{μ_{x,s}^{PS}(B_U(e^s))} \int_{B_U(e^s)} Θ_{x,s}(t)dμ_{x,s}(t)
\]
\[
\leq c_0η + \frac{1}{μ_{x,s}^{PS}(B_U(e^s))} \int_{\mathcal{K}x,s} Θ_{x,s}(t)dμ_{x,s}(t), \quad \text{by (6.9)}
\]
\[
\leq c_0η + ε, \quad \text{by (6.8)}.
\]

Putting \(ζ = c_0η + ε\), it follows that there exist \(1 < k(i) ≤ l\) and a subset \(J(s) \subseteq B_U(e^s)\) with \(|J(s)| ≥ \frac{|B_U(e^s)|}{l}\) so that
\[
d(v_i(xu^-a_s)u_{x,t}^-)u_t, v_{k(i)}(xa_s)u_t^2 ≤ 2ζ/C \quad \text{for all } t \in J(s).
\]

Recall now that if \(y,z \in X\) are two point then for all \(t \in \mathbb{F}\) so that \(d(xu_t,yu_t)\) is sufficiently small, the map \(t \mapsto d(xu_t,yu_t)^2\) is governed by a polynomial of bounded degree, see [29]. Therefore the above and the fact that polynomials of a bounded degree are \((C,α)-\text{good}\) with respect to the Haar measure, see (6.3), imply that
\[
|B_U(e^s)|/l ≤ |J(s)| ≤ Cε^α (\sup_{|t|≤e^s} Θ_{x,s})^{-α}|B_U(e^s)|
\]
and hence
\[
d(v_i(xu^-a_s)u_{x,t}^-)u_t, v_{k(i)}(xa_s)u_t^2 ≤ O_1(ζ) \quad \text{for all } t \in B_U(e^s)
\]
where the implied constant depends only on \(ζ\). Therefore for every \(1 ≤ i ≤ l\) and \(s\) as above, if we define \(g_{s,i} \in G\) by the following
\[
v_i(xu^-a_s)u_{x,t}^- = v_{k(i)}(xa_s)g_{s,i},
\]
then \(g_{s,i}\) is contained in an \(O(ζ)\) neighborhood of \(e\).

If we write \(g_{s,i} = \begin{pmatrix} x_s & y_s \\ z_s & w_s \end{pmatrix}\), then
\[
u_{x,t}^{-1}g_{s,i}u_t = \begin{pmatrix} x_s + yst & y_s \\ z_s + (w_s - x_s)t - yst^2 & w_s - yst \end{pmatrix}.
\]
Therefore, (6.14) and \( \det g_{s,i} = 1 \) imply that
\[
|z_s| \leq O(\zeta), \quad |1 - x_s| \leq O(\zeta) e^{-s}, \quad |1 - w_s| \leq O(\zeta) e^{-s}, \quad |y_s| \leq O(\zeta) e^{-2s},
\]
in particular,
\[
d(e, a_s g_{s,i} a_{-s}) = O(e^{-s}). \tag{6.15}
\]
For every \( 1 \leq i \leq l \) we have shown that
\[
\nu_i (x u^- r a_s) u^- e^{-r} a_{-s} \in \Upsilon(x a_s) g_{s,i} a_{-s} = \Upsilon(x) a_s g_{s,i} a_{-s}. \tag{6.16}
\]
In view of the \( A \)-equivariance, this implies that
\[
\Upsilon(x u^-) u^- r \subseteq \Upsilon(x) \cdot \{ g \in G : d(e, g) = O(e^{-s}) \}.
\]
This establishes (6.6) and finishes the proof.

7. JOINING CLASSIFICATION

We keep the notation for \( G, U, MAU^{-}, \) etc. from section 6; so \( G = \text{PSL}_2(\mathbb{F}) \) where \( \mathbb{F} = \mathbb{R}, \mathbb{C} \). We will use the notation \( \Delta \) for the diagonal embedding map of \( G \) into \( G \times G \); so \( \Delta(g) = (g, g) \) for \( g \in G \) and \( \Delta(S) = \{(s, s) : s \in S\} \) for a subset \( S \) of \( G \).

7.1. Construction of a polynomial like map. Recall \( P = MAU^{-} \); so \( P \) is the subgroup consisting of all upper triangular matrices. We fix a rational cross-section \( \mathcal{L} \) for \( \Delta(U) \) in \( G \times G \) as follows:
\[
\mathcal{L} = \{(e) \times G\} \cdot \Delta(P) = P \times G.
\]
Then \( \mathcal{L} \cap \Delta(U) = \{e\} \) and the product map from \( \mathcal{L} \times \Delta(U) \) to \( G \) defines a diffeomorphism onto a Zariski open dense subset of \( G \times G \). We will use \( \mathcal{L} \) as the transversal direction to \( \Delta(U) \) in \( G \times G \).

We observe that \( N_{G \times G}(\Delta(U)) = \Delta(AM) \cdot (U \times U) \) and
\[
N_{G \times G}(\Delta(U)) \cap \mathcal{L} = \Delta(AM) \cdot \{(e) \times U\}
\]
Suppose that we are given a sequence \( h_k = (h_k^1, h_k^2) \in G \times G \) such that \( h_k \notin N_{G \times G}(\Delta(U)) \) with \( h_k \to e \) as \( k \to \infty \). Associated to \( \{h_k\} \), we will construct a quasi-regular map
\[
\varphi : \Delta(U) \to \mathcal{L} \cap N_G(\Delta(U))
\]
following [20, Section 5]. Via the identification \( \mathbb{F} = \Delta(U) \) given by \( t \mapsto \Delta(u_t) \), we will define the map \( \varphi \) on \( \mathbb{F} \), which will save us some notation. Accordingly, we will write \( t \in B_U(1) \) to mean that \( u_t \in B_U(1) \), etc.

For \( g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G \) and for \( t \neq -ab^{-1} \), define
\[
\alpha_g(t) = \frac{c + dt}{a + bt}.
\]
We denote the pole of \( \alpha_g \) by \( R(g) \) and put \( R(g) = \infty \) if \( \alpha_g \) is defined everywhere. That is, \( R(g) = -ab^{-1} \) if \( b \neq 0 \) and \( \infty \) otherwise.
Explicit construction of

We choose a norm $\|\cdot\|$ on $\mathcal{Y}$ so that

$$h_k\Delta(u_t) = \Delta(u_{\alpha_k(t)})\varphi_k(t) \in \Delta(U)\mathcal{L}.$$  

We renormalize these maps $\varphi_k$ using a representation corresponding to $\Delta(U)$. Recall that by a theorem of Chevalley there is a finite dimensional representation $(\rho, \Upsilon)$ of $G \times G$, where $G \times G$ acts from the right on $\Upsilon$ and a unit vector $q \in \Upsilon$ so that

$$\Delta(U) = \{ h \in G \times G : q\rho(h) = q \}.$$  

Then

$$N_{G \times G}(\Delta(U)) = \{ h : q\rho(h)\rho(\Delta(u_t)) = q\rho(h) \text{ for all } u_t \in U \}.$$  

We choose a norm $\|\cdot\|$ on $\mathcal{Y}$ so that $B(q, 2) \cap q\rho(G \times G) \subseteq q\rho(G \times G)$. 

Now for each $k$, define $\tilde{\phi}_k : \mathbb{F} \to \mathcal{Y}$ by

$$\tilde{\phi}_k(t) = q\rho(h_k\Delta(u_t));$$

$\tilde{\phi}_k$ is a polynomial map of degree bounded in terms of $\rho$ and $\tilde{\phi}_k(0) = q$.

**Explicit construction of $\tilde{\phi}_k$:** Consider the following representation: let $G \times G$ act on $\mathcal{Y} = \mathbb{F}^2 \oplus \mathbb{F}^2 \oplus M_2(\mathbb{F})$ by

$$(g_1, g_2), (v_1, v_2, Q) = (v_1 g_1, v_2 g_2, g_1^{-1} Q g_2).$$

Then the stabilizer of $q := (e_1, e_1, I_2)$ is precisely $\Delta(U)$. If we write $h_k = \left( \begin{array}{cc} a_k & b_k \\ c_k & d_k \end{array} \right)$, then, up to an additive constant vector, say, $q_k \in \mathcal{Y}$ we have

$$\tilde{\phi}_k(t) = q_k + \left( b_k^1 t, 0, b_k^2 t, 0, \begin{pmatrix} -A_k t & 0 \\ A_k t^2 & B_k t & A_k t \end{pmatrix} \right)$$

where $A_k = b_k^1 d_k^2 - b_k^2 d_k^1$ and $B_k = a_k^1 d_k^2 + b_k^1 c_k^2 - b_k^2 c_k^1 - a_k^2 d_k^1$. Hence $\tilde{\phi}_k$ is a polynomial of degree at most $2$.

Let $T_k > 0$ be the infimum of $T > 0$ such that

$$\sup_{t \in B(\mathcal{Y})(1)} \|\tilde{\phi}_k(t) - q\| = 1.$$  

Since $h_k \notin N_{G \times G}(\Delta(U))$, we get $T_k \neq \infty$, moreover, in view of our assumption $h_k \to e$ we have $T_k \to \infty$ as $k \to \infty$.

By normalizing $\tilde{\phi}_k$ by

$$\phi_k(t) := \tilde{\phi}_k(T_k t),$$

we obtain a sequence of equicontinuous polynomials $\phi_k$. Hence, after passing to a subsequence, $\phi_k$ converges to $\phi$ where

- $t \mapsto \phi(t)$ is a non-constant polynomial of degree at most 2.
\( \sup_{t \in B_U(1)} \| \phi(t) - q \| = 1 \) and \( \phi(0) = q \),
\( \{ \phi(t) : t \in B_U(1) \} \subseteq q \rho(G \times G) \)
- the convergence is uniform on compact subsets of \( F \).

Put
\[ \varphi = (\rho_L)^{-1} \circ \phi \]
where \( \rho_L \) is the restriction to \( L \) of the orbit map \( g \mapsto q \rho(g) \). Then \( \varphi : F \to L \) is a rational map defined on a Zariski open dense subset \( W \subseteq F \) containing zero and \( \varphi(0) = e \). We have
\[ \varphi(t) = \lim_k \varphi_k(T_k t) \]
and the convergence is uniform on compact subsets of \( W \).

Note also that for any \( t_0 \in F \), we have
\[ \phi(t) \rho(\Delta(u_{t_0})) = \lim_k \tilde{\phi}_k(T_k t) \rho(\Delta(u_{t_0})) = \lim_k q \rho(h_k \Delta(u_{t_0})) \rho(\Delta(u_{t_0})) = \lim_k q \rho(h_k \Delta(u_{t_0}(t + t_0/T_k))) = \phi(t). \]
Therefore \( \varphi(t) \in N_{G \times G}(\Delta(U)) \cap L \).

The following observation will be important in our application:

**Proposition 7.1.** \( 0 \) is an isolated point in \( \varphi^{-1}(e) \); i.e., there is \( \eta > 0 \) such that \( \varphi(t) = e \) and \( t \in B_U(\eta) \) implies \( t = 0 \).

**Proof.** Since \( \phi \) is a non-constant polynomial of degree at most 2, the set \( \{ t \in F : \phi(t) = q \} \) consists of at most two points. \( \square \)

### 7.2. \( \Delta(U) \)-recurrence for the pull back function \( \Psi \).

In the rest of this section, we let \( \Gamma_1 \) and \( \Gamma_2 \) be geometrically finite, and Zariski dense subgroups of \( G \) and set \( X_i := \Gamma_i \backslash G \). As before, we normalize \( |m_{\Gamma_i}^{\text{BMS}}| = 1 \) for \( i = 1, 2 \). For the sake of simplicity, we will often omit the subscript \( \Gamma_i \) in the notation of \( m_{\Gamma_i}^{\text{BMS}} \) and \( m_{\Gamma_i}^{\text{BR}} \).

We fix a non-negative function \( \psi \in C_c(X_1) \) with \( m^{\text{BR}}(\psi) > 0 \) and set
\[ \Psi := \psi \circ \pi_1 \in C(X_1 \times X_2). \]
We also fix a compact subset \( \Omega \subset X_1 \) such that \( m^{\text{BR}}(\Omega) > 0 \) and that for all \( x_1 \in \Omega, x_1 \in \Lambda_r(\Gamma) \). Let
\[ 0 < r_\Omega := r(\frac{1}{2}, \Omega) < 1 \]
be as in Theorem 5.3.

Since \( \Psi \) is defined as the pull-back of a function on \( X_1 \), we can transfer the recurrence properties of \( U \)-orbits in \( X_1 \) to statements about \( \Delta(U) \)-recurrence properties with respect to \( \Psi \). The rest of this subsection is devoted to recording these properties of \( \Psi \).
Lemma 7.2. For any $x \in X_1 \times X_2$ with $\pi_1(x)^{-1} \in \Lambda_r(\Gamma)$, we have
\[
\int_{B_U(T)} \Psi(x\Delta(u_t))dt \to \infty \quad \text{as } T \to \infty.
\]

Proof. For such an $x$, we have, by Theorem 4.6,
\[
\int_{B_U(T)} \Psi(x\Delta(u_t))dt = \int_{B_U(T)} \psi(\pi_1(x)u_t)dt \sim m^{PS}_{\pi_1(x)}(B_U(T)) \cdot m^{BR}(\psi).
\]
Since $m^{BR}(\psi) > 0$, we have $\int_{B_U(T)} \Psi(x\Delta(u_t))dt \to \infty$. \qed

Lemma 7.3. There exists $T_0 = T_0(\psi, \Omega) > 1$ such that for any $x \in X_1 \times X_2$ with $\pi_1(x) \in \Omega$ and any $T > T_0$,
\[
\int_{B_U(rT)} \Psi(x\Delta(u_t))dt \leq \frac{1}{2} \int_{B_U(T)} \Psi(x\Delta(u_t))dt
\]
for any $0 \leq r \leq r_\Omega$.

Proof. For $x \in X_1 \times X_2$ with $\pi_1(x) \in \Omega$, we have
\[
\int_{B_U(T)} \Psi(x\Delta(u_t))dt - \int_{B_U(rT)} \Psi(x\Delta(u_t))dt \\
= \int_{B_U(T)} \psi(\pi_1(x)u_t)dt - \int_{B_U(rT)} \psi(\pi_1(x)u_t)dt \\
\geq \frac{1}{2} \int_{B_U(T)} \psi(\pi_1(x)u_t)dt \\
= \frac{1}{2} \int_{B_U(T)} \Psi(x\Delta(u_t))dt
\]
where the inequality follows by Theorem 5.3 for all $0 \leq r \leq r_\Omega$. \qed

7.3 Joining measure. In the rest of this section, we let $\mu$ be a $\Delta(U)$-invariant ergodic joining of $(X_1 \times X_2, m^{BR}_{\Gamma_1} \times m^{BR}_{\Gamma_2})$. In particular, if $\pi_i : X_1 \times X_2 \to X_i$ denotes the projection onto $i$-th coordinate, then
\[
(\pi_i)_*\mu = m^{BR}_{\Gamma_i}.
\]

We will fix a non-negative function $\psi \in C_c(X_1)$ with $m^{BR}(\psi) > 0$, and set
\[
\Psi := \psi \circ \pi_1 \in C(X_1 \times X_2).
\]
Note that since $(\pi_1)_*\mu = m^{BR}$, we have
\[0 < \mu(\Psi) < \infty; \text{ in particular, } \Psi \in L^1(\mu)\]
We also fix the following:
- a compact subset $\Omega \subset X_1$ such that for all $x_1 \in \Omega$, we have $x_1^- \in \Lambda_r(\Gamma)$ and Theorem 4.6 holds for $\psi$ uniformly for all $x_1 \in \Omega$ (such $\Omega$ exists in view of Remark 4.8);
- $r := 1/4 \cdot r(1/2, \Omega) < 1$ as in Theorem 5.3;
- $c_0 := c_0(1/10, \Omega, \psi)$ and $T_0 = T_0(\Omega, \psi)$ as in Lemma 7.8;
• a compact subset
\[ Q = \Omega_1 \times \Omega_2 \subset X_1 \times X_2 \]
such that \( \Omega_1 \subset \Omega \), \( \mu(Q) > 0 \) and for all \( x \in Q \) and for all \( f \in C_c(X_1 \times X_2) \),
\[
\lim_{T \to \infty} \frac{\int_{B_U(T)} f(x\Delta(u_t))dt}{\int_{B_U(T)} \Psi(x\Delta(u_t))dt} = \frac{\mu(f)}{\mu(\Psi)}.
\]

(7.2)

Lemma 7.4. The following set has a full \( \mu \)-measure in \( X_1 \times X_2 \):
\[
\{ (x_1, x_2) \in X_1 \times X_2 : x_i^- \in \Lambda_r(\Gamma_i) \text{ for each } i = 1, 2 \}.
\]

Proof. Since \( \Lambda_r(\Gamma_i) \) has the full Patterson-Sullivan measure in \( \Lambda(\Gamma) \) by Sullivan [41], \( Y_i := \{ x_i \in X_i : x_i^- \in \Lambda_r(\Gamma_i) \} \) has a full \( m^{BR} \)-measure. Since \( \pi_i^* \mu = m^{BR} \), we have \( \pi_1^{-1}(Y_1) \cap \pi_2^{-1}(Y_2) \) has a full \( \mu \)-measure. \( \square \)

In this section we will use ergodic theorems and polynomial like behavior of unipotent orbits, the construction of a polynomial-like map in § 7.1 in order to produce extra quasi-invariance for the measure. Recall we say \( \mu \) is \( h \) quasi-invariant means that \( h_\cdot \mu \) is a positive multiple of \( \mu \) where \( h_\cdot \mu(f) := \mu(h^{-1} \cdot f) \).

We begin with the following.

Lemma 7.5. Let \( Y \subseteq X_1 \times X_2 \) be a subset such that for all \( y \in Y \),
\[
(1) \lim_{T \to \infty} \int_{B_U(T)} \Psi(y\Delta(u_t))dt = \infty;
(2) \text{ for all } f \in C_c(X_1 \times X_2),
\]
\[
\lim_{T \to \infty} \frac{\int_{B_U(T)} f(y\Delta(u_t))dt}{\int_{B_U(T)} \Psi(y\Delta(u_t))dt} = \frac{\mu(f)}{\mu(\Psi)}.
\]

If \( h \in N_{G \times G}(\Delta(U)) \) is so that \( Y \cap Yh \neq \emptyset \), then \( \mu \) is \( h \) quasi-invariant.

Proof. Let \( h \in N_{G \times G}(\Delta(U)) = \Delta(AM) \cdot (U \times U) \) and \( y \in Y \) such that \( yh \in Y \). Since, under conjugation, \( \Delta(AM) \) acts on \( \Delta(U) \) by homothety composed with rotations,
\[
h^{-1}\Delta(u_t)h = \Delta(u_{\beta_h(t)})
\]
where \( \beta_h : U \to U \) is a homothety composed with a rotation.

Let \( B'_U(T) \) denote the Euclidean ball of radius \( T \). If the \( \Delta(A) \)-component of \( h \) is \( (a_s, a_s) \), then \( \beta_h(B'_U(T)) = B'_U(e^sT) \) and the Jacobian of \( \beta_h \) is a constant \( e^{(n-1)s} \) where \( (n-1) \) is the dimension of \( U \) as a real vector space.
For any \( f \in C_c(X_1 \times X_2) \) and for any all large \( T \gg 1 \),
\[
\left| \frac{\mu(h.f)}{\mu(h.\Psi)} - \frac{\mu(f)}{\mu(\Psi)} \right|
\leq \frac{\mu(h.f)}{\mu(h.\Psi)} \left( \int_{B_U^c(T)} f(y \Delta(u_t)h) dt + \int_{B_U(T)} f(yh \Delta(u_t)) dt - \mu(f) \right) \left( \int_{B_U^c(T)} \Psi(y \Delta(u_t)h) dt + \int_{B_U(T)} \Psi(yh \Delta(u_t)) dt - \mu(\Psi) \right).
\]
Since both \( y \) and \( yh \) belong to \( Y \), it follows that each term in the right-hand side tends to zero as \( T \to \infty \). Hence
\[
\mu(h.f) = \frac{\mu(h.\Psi)}{\mu(\Psi)} \cdot \mu(f),
\]
finishing the proof. \( \square \)

For any Borel function \( f \) on \( X_1 \times X_2 \), all \( T > 0 \) and any \( x \in X_1 \times X_2 \) put
\[
D_T f(x) = \int_{I_r(T)} f(x \Delta(u_t)) dt
\]
where \( I_r(T) := B_U(T) - B_U(rT) \).

**Corollary 7.6.** For any \( f \in C_c(X_1 \times X_2) \), and any \( x \in Q \), we have
\[
\frac{D_T f(x)}{D_T \Psi(x)} \to \frac{\mu(f)}{\mu(\Psi)} \text{ as } T \to \infty.
\]

**Proof.** This follows from Lemma 7.3 and (7.2). \( \square \)

Fix \( \varepsilon > 0 \) which is small enough so that \( \frac{1}{1+\varepsilon} \leq (1 - 2\varepsilon)^2 \). Choose \( \eta > 0 \) small enough so that \( \mu(Q\{g : |g| \leq \eta\}) \leq (1 + \varepsilon)\mu(Q) \). We fix such \( \eta = \eta(Q) \) and put
\[
Q^+ = Q\{g : |g| \leq \eta/4\}, \text{ and } Q^{++} = Q\{g : |g| \leq \eta\}.
\]
Set
\[
\mathcal{F} := \{\chi_Q, \chi_{Q^+}, \chi_{Q^{++}}\}.
\]

Now using Corollary 7.6 and Egorov’s theorem we can find a compact subset
\[
Q_\varepsilon \subset Q
\]
with \( \mu(Q_\varepsilon) > (1 - \varepsilon)\mu(Q) \) with the following property: for any \( f \in \mathcal{F} \) and any \( \theta > 0 \), there exists some \( T_0 = T_0(f, \theta) \) so that if \( T \geq T_0 \), then
\[
\left| \frac{D_T f(x)}{D_T \Psi(x)} - \frac{\mu(f)}{\mu(\Psi)} \right| \leq \theta \text{ for all } x \in Q_\varepsilon.
\]
Such a subset will be referred to as a set of uniform convergence for the family \( \mathcal{F} \), see also [22, Lemma 7.5].

The following lemma will be used to compare ergodic averages along two nearby orbits.
Lemma 7.7. Let \( \{ R_k \} \) be a sequence tending to infinity as \( k \to \infty \) and \( 0 < \sigma < 1/2 \) be a fixed small number. Suppose that for each \( k \), \( \alpha_k : F \to \mathbb{F} \) is a rational map with no poles on \( B_{R_k}(F) \) and satisfying that for all \( t \in B_{R_k}(F) \),

\[
1 - \sigma \leq |\text{Jac}(\alpha_k)(t)| \leq 1 + \sigma.
\]

Then there exist \( c_1 > 1 \) and \( T_1 = T_1(\Omega, \psi, F) > 1 \) such that for all \( T_1 < T < R_k/4 \), \( x \in Q_\varepsilon \) and for each \( f \in F \), we have

\[
c_1^{-1} \cdot \mathcal{D}_T f(x) \leq \int_{I_r(T)} f(x \Delta(u_{\alpha_k(t)})) dt \leq c_1 \cdot \mathcal{D}_T f(x).
\]

Proof. First note that

\[
\int_{I_r(T)} f(x \Delta(u_{\alpha_k(t)})) dt = \int_{\alpha_k(I_r(T))} f(x \Delta(u_t)) |\text{Jac}(\alpha_k)(t)| dt. \tag{7.5}
\]

Setting \( r_\sigma^+ = (1 + \sigma) r \) and \( T_\sigma^+ = (1 + \sigma) T \), note that

\[
I_{r_\sigma^+}(T_\sigma^-) \subset \alpha_k(I_r(T)) \subset I_{r_\sigma^-}(T_\sigma^+).
\]

Now for all \( T > 2T_0(\Omega, \psi) \), where \( T_0 \) is as in Sublemma 7.8, we have

\[
\int_{I_r(T)} f(x \Delta(u_{\alpha_k(t)})) dt \geq (1 - \sigma) \int_{I_{r_\sigma^+}(T_\sigma^-)} f(x \Delta(u_t)) dt
\]

Sublemma 7.8
\[
\geq (1 - \sigma) c_0 \int_{B_U(T)} \Psi(x \Delta(u_t)) dt + a(T)
\]

where \( a(T) \) and \( a'(T) \) tend to 0. Since \( \mathcal{D}_T f(x) \to \infty \) as \( \mathcal{D}_T \Psi(x) \) does, we have

\[
(1 - \sigma) c_0 \mathcal{D}_T f(x) + a'(T) \geq c_1^{-1} \mathcal{D}_T f(x)
\]

for some \( c_1 > 1 \) and for all \( T \) bigger than some fixed \( T_1 > 1 \). The other inequality can be proved similarly.

Sublemma 7.8. Let the notation be as in Lemma 7.7 and the proof of that lemma, in particular let \( 0 < \sigma < 1/2 \). There exist some \( T_0 = T_0(\Omega, \psi) \) and \( c_0 = c_0(\Omega, \psi) > 0 \) such that for all \( T > T_0 \) and for all \( x \in Q_\varepsilon \) we have

\[
\int_{I_{r_\sigma^+}(T_\sigma^-)} \Psi(x \Delta(u_t)) dt \geq c_0 \int_{B_U(T)} \Psi(x \Delta(u_t)) dt.
\]
Proof. Note that our assumption that Theorem 4.6 holds for \( \psi \) uniformly for all \( \pi_1(x) \in \Omega \) implies that Theorem 5.4(2) also holds for \( \psi \) uniformly for all \( \pi_1(x) \in \Omega \) (see 5.5). Therefore, for any \( x \in \mathcal{Q}_\varepsilon \) we have
\[
\int_{I_{\psi}(T_\sigma)} \Psi(x\Delta(u_t))dt = \int_{I_{\psi}(T_\sigma)} \psi(\pi_1(x)u_t)dt \\
\geq \frac{1}{2}m^{\text{BR}}(\psi) \left( \mu^{\text{PS}}_{\pi_1(x)}(B_U(T_\sigma^-)) - B_U(r_\sigma^+T^-) \right)
\]
for all \( T \) bigger than some \( T_0 > 1 \), since Theorem 5.4(2) holds for \( \psi \), uniformly for all \( \pi_1(x) \in \Omega \).

On the other hand, by Lemma 5.2 and Lemma 4.2, there exist \( T_1 > T_0 \) and \( c'_0 > 0 \), depending only on \( \Omega \) such that if \( \pi_1(x) \in \Omega \),
\[
\mu^{\text{PS}}_{\pi_1(x)}(B_U(T_\sigma^-)) - B_U(r_\sigma^+T^-) \geq \frac{1}{2}\mu^{\text{PS}}_{\pi_1(x)}(B_U(T_\sigma^-)) \geq c'_0 \mu^{\text{PS}}_{\pi_1(x)}(B_U(T))
\]
for some constant \( c'_0 > 0 \). Since Theorem 4.6 holds for \( \psi \), uniformly for all \( \pi_1(x) \in \Omega \), we have, for all sufficiently large \( T \gg 1 \),
\[
m^{\text{BR}}(\psi)\mu^{\text{PS}}_{\pi_1(x)}(B_U(T)) \geq \frac{1}{2} \int_{B_U(T)} \psi(\pi_1(x)u_t)dt.
\]

Therefore,
\[
\int_{I_{\psi}(T_\sigma)} \Psi(x\Delta(u_t))dt \geq \frac{c'_0}{4} m^{\text{BR}}(\psi) \mu^{\text{PS}}_{\pi_1(x)}(B_U(T)) \\
\geq \frac{c'_0}{4} \int_{B_U(T)} \psi(\pi_1(x)u_t)dt = \frac{c'_0}{4} \int_{B_U(T)} \Psi(x\Delta(u_t))dt.
\]

Recall our notation: for a subset \( S \subseteq G \times G \), we denote by \( \langle S \rangle \) the minimal connected subgroup of \( G \times G \) containing \( S \).

**Lemma 7.9.** Let \( h_k \in G \times G - N_{G \times G}(\Delta(U)) \) be a sequence tending to \( e \) as \( k \to \infty \). Suppose that \( \mathcal{Q}_\varepsilon h_k \cap \mathcal{Q}_\varepsilon \neq \emptyset \) for all \( k \). Then \( \mu \) is quasi-invariant under a nontrivial connected subgroup of \( \langle \text{Im}\varphi \rangle \) where \( \varphi \) is the map constructed in section 7.1 with respect to \( \{h_k\} \).

**Proof.** We use the notation used in the construction of the map \( \varphi \). By our assumption we have: there are points \( y_k \in \mathcal{Q}_\varepsilon \) so that \( x_k = y_k h_k \in \mathcal{Q}_\varepsilon \). Recalling the maps \( \varphi_k \) and \( \alpha_k \) from above, we have
\[
x_k \Delta(u_t) = y_k h_k \Delta(u_t) = y_k \Delta(u_{\alpha_k(t)})\varphi_k(t).
\]
Now let
\[
\tau'_k := \sup_{t \in B_U(\gamma)} d(e, \varphi_k(t)) = \eta/4, \quad \tau_k = \min\{\tau'_k, R_k\},
\]
and
\[
R_k = \sup\{0 < R < \infty : 0.9 \leq \text{Jac}(\alpha_k)|_{B_{R}(e)} \leq 1.1\}.
\]
Note that
\[ \sup_{t \in B_{U}(r'_{k})} d(e, \varphi^{-1}_{k}(t)) = \eta/4. \]

Note that \( \Theta_{k} = \tau_{k}/T_{k} \) is bounded away from 0; in particular, \( \tau_{k} \to \infty \). Passing to a subsequence we may and will assume that \( \theta_{k} \to \Theta \).

By the definition of \( Q_{\varepsilon} \), we have, for all large enough \( T \),
\[ \left| \frac{\int_{I_{r}(T)} f_{1}(z_{k}(u_{t})) dt}{\int_{I_{r}(T)} f_{2}(z_{k}(u_{t})) dt} - \frac{\mu(f_{1})}{\mu(f_{2})} \right| \leq \theta \]
for \( f_{1}, f_{2} \in F = \{ \chi_{Q}, \chi_{Q^{+}}, \chi_{Q^{++}} \} \) and \( z_{k} = x_{k}, y_{k} \).

With this notation, the above implies: for all large enough \( k \) and all \( T_{0} \leq T \leq T_{k} \) we have
\[ \{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q \} \neq \emptyset. \tag{7.6} \]
To see this, let \( k \) be large and let \( T_{0} \leq T \leq T_{k} \). Then
\[ \{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q \} \subseteq \{ t \in I_{r}(T) : y_{k}(u_{t}) \in Q^{+} \} \]
\[ \subseteq \{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q^{++} \}. \]

On the other hand we have
\[ |\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q \}| \geq (1 - \varepsilon)|\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q^{++} \}| \tag{7.7} \]
where \(| \cdot | \) denotes the Lebesgue measure on \( F \). From these two and Lemma 7.7 we get
\[ |\{ t \in I_{r}(T) : y_{k}(u_{t}) \in Q \}| \geq c_{1}|\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q^{++} \}| \tag{7.8} \]
\[ \geq c_{1}(1 - \varepsilon)|\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q^{+} \}| \tag{7.9} \]
\[ \geq c_{1}(1 - \varepsilon)|\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q \}| \]
\[ \geq c_{1}(1 - \varepsilon)|\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q^{+} \}| \]
\[ \geq c_{2}^{2}(1 - \varepsilon)|\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q^{++} \}|. \]

Now (7.6) follows by applying (7.7) and (7.8), in view of the choice of \( \varepsilon \) and the fact that by Corollary (7.6) we have
\[ |\{ t \in I_{r}(T) : x_{k}(u_{t}) \in Q^{++} \}| > 0 \text{ for all large enough } T. \]

For each \( k \) let \( m_{k} \geq 0 \) be the maximum integer so that \( r^{m_{k}}\tau_{k} \geq T_{0} \). Then for any \( \ell \geq 0 \) and all large enough \( k \) we have \( \ell \leq m_{k} \). Let \( \ell \geq 0 \) and apply (7.6) with \( T_{k,\ell} = r^{\ell}\tau_{k} \). Then for each \( k \) we find \( t \in I_{r}(T_{k,\ell}) \) so that \( z_{k,\ell} = y_{k}(u_{\alpha_{k}(t)}) \) satisfies \( z_{k,\ell} \in Q \) and \( z_{k,\ell}(\varphi_{k}(t)) \in Q \). Passing to a subsequence we get: there exist some \( z_{\ell} \in Q \) and some \( s \in B_{U}(\Theta) - B_{U}(r\Theta) \) so that \( z_{\ell}(\varphi(s)) \in Q \). Therefore by Lemma 7.5 we have \( \mu \) is \( \varphi(s) \) quasi-invariant. Now if we choose \( \ell \) large enough, then \( \varphi(s) \neq e \) in view of Proposition 7.1 however, it can be made arbitrary close to the identity by choosing large \( \ell \)'s. This implies the claim. \( \Box \)
7.4. **Infinite joining measure cannot be invariant by** $\{e\} \times V$. We recall some basic facts about dynamical systems. Consider an action of

one-parameter subgroup $W = \{w_t\}$ on a separable, $\sigma$-compact and locally compact topological space $X$ with an invariant Radon measure $\mu_0$. A Borel subset $E \subset X$ is called *wandering* if $\int_E \chi_{E}(x\tilde{v}_t) < \infty$ for almost all $x \in E$. The Hopf decomposition theorem says that $X$ is a disjoint union of invariant subsets $D(W)$ and $C(W)$ where $D(W)$ is a countable union of wandering subsets which is maximal in the sense that any wandering subset is contained in $D(W)$ up to null sets (see [17]). The sets $D(W)$ and $C(W)$ are called the dissipative part, and the conservative part of $X$ respectively. If $D(W)$ (resp. $C(W)$) is a null set, this action is called *conservative* (resp. *dissipative*). If the $W$-action is ergodic, then it is either conservative or dissipative. The following is well known for a single transformation (e.g. [1]), but we could not find a reference for a flow; so we provide a proof for the sake of completeness.

**Lemma 7.10.** If $\mu_0$ is conservative, ergodic and infinite, then for any non-negative $f \in L^1(\mu_0)$,

$$\frac{\int_{-T}^{T} f(xw_t)dt}{2T} \to 0$$

for almost all $x \in X$.

**Proof.** Since $X$ is $\sigma$-compact it suffices to prove that for any compact subset $K \subset X$ and almost all $x \in K$ the above holds. Let $K \subset X$ be a compact subset. We will show that for any $\varepsilon > 0$, the set $\{x \in K : \frac{\int_{-T}^{T} f(xw_t)dt}{2T} \to 0\}$ has co-measure less than $\varepsilon > 0$. Write $X = \bigcup_{N=1}^{\infty} \Omega_N$ as an increasing union of compact subsets with $\mu_0(\Omega_1) > 0$. By the Hopf ratio theorem and Egorov’s theorem we can find a subset $K_\varepsilon$ of $K$ with co-measure at most $\varepsilon$ such that the following convergence is uniform for all $x \in K_\varepsilon$ and for all $N$:

$$\frac{\int_{-T}^{T} f(xw_t)dt}{\int_{-T}^{T} \chi_{\Omega_N}(xw_t)dt} \to \frac{\mu_0(f)}{\mu_0(\Omega_N)}.$$ 

Hence for any $\eta > 0$, there exists $T_\eta$ such that for all $x \in K_\varepsilon$, $T > T_\eta$,

$$\limsup_{N} \left| \frac{\int_{-T}^{T} f(xw_t)dt}{\int_{-T}^{T} \chi_{\Omega_N}(xw_t)dt} - \frac{\mu_0(f)}{\mu_0(\Omega_N)} \right| \leq \eta.$$

Since $\mu_0(\Omega_N) \to |\mu_0| = \infty$ and $\int_{-T}^{T} \chi_{\Omega_N}(xw_t)dt = 2T$ for all large $N$, it follows that

$$\frac{\int_{-T}^{T} f(xw_t)dt}{2T} \leq \eta,$$

for all $T > T_\eta$, and hence $\frac{\int_{-T}^{T} f(xw_t)dt}{2T} \to 0$ for all $x \in K_\varepsilon$.

**Remark 7.11.** We recall that if $\Gamma$ is not a lattice, then the BR measure is an infinite measure; this was proved in [25] using Ratner’s measure classification theorem.
We take this opportunity to present an alternative argument. To see this, we note that if the BR-measure were a finite measure, it would have to be $A$-invariant, since $|a_s m_{BR}| = e^{(2-\delta)s}|m_{BR}|$ for all $s$ and hence $\delta = 2$. For $\Gamma$ geometrically finite, this implies $\Gamma$ is a lattice. In the general case, one can utilize facts from entropy to prove a similar result as we now explain. Indeed by the Mautner phenomenon, any $AU$-invariant finite ergodic measure on $\Gamma \setminus G$ is $A$-ergodic so we may reduce to the ergodic case. Now we have an $A$-ergodic measure which is $U$ invariant; in particular it has maximum entropy. This implies the entropy contribution from $U^-$ has to be maximum as well which implies the measure is also $U^-$ invariant, see [20, Theorem 9.7] for a more general statement. This implies $\Gamma \setminus G$ has a finite $G$ invariant measure, finishing the proof.

We need the following lemma which says almost all ergodic components of $m_{BR}$ is infinite for any one-parameter subgroup of $U$; our proof of this lemma uses Ratner’s classification theorem for finite invariant measures for unipotent flows.

**Lemma 7.12.** Let $\Gamma$ be a Zariski dense, discrete subgroup of $G$. Suppose $\Gamma$ is not a lattice. Let $V$ be a one-parameter subgroup of $U$, and let $m_{BR} = \int_Y \eta_y d\sigma(y)$ be the ergodic decomposition with respect to $V$. Then for $\sigma$-a.e. $y$ we have $\eta_y$ is an infinite measure.

**Proof.** We will use the fact that the set $\Lambda_p(\Gamma)$ of parabolic limit points is a null set for the Patterson-Sullivan measure since $\Lambda_p(\Gamma)$ is a countable set and that a proper Zariski closed subset of $G$ is a null set for the $\tilde{m}_{BR}$-measure, since $\Gamma$ is Zariski dense. Assume the contrary, that is: the set $Y_0 = \{ y \in Y : \eta_y$ is a finite measure $\}$ has positive measure.

It follows from Ratner’s measure classification theorem [32]: that for all $y \in Y_0$, we have one of the following holds

1. $\text{supp } \eta_y = xV$ for some compact orbit $xV$;
2. $\text{supp } \eta_y = xU$ for some compact orbit $xU$.
3. there exists $H$ which is locally isomorphic to $\text{PSL}_2(\mathbb{R})$ so that for some $g \in G$, $V \subset g^{-1}Hg$, $\eta_y$ is a $g^{-1}Hg$ invariant (finite) measure on a closed orbit $\Gamma Hg$.
4. $\eta_y$ is $\text{PSL}_2(\mathbb{C})$ invariant.

In both (1) and (2) above we get $x^{-1}$ is a parabolic fixed point and these form a measure zero subset of $m_{BR}$. The conclusion in (4) cannot hold on a positive measure set as it would imply $\Gamma$ is a lattice, contrary to our assumption. Therefore for $\sigma$-a.e. $y \in Y_0$ the conclusion (3) above holds.

We first note that the collection of $H$ so that (3) holds is countable, see [32, 6]. Indeed similar reductions are possible using Hopf argument in more general settings.
Theorem 1.1] or [6, Proposition 2.1]. Therefore if (3) holds there exists some \( H \) with \( \Gamma H \) a closed orbit (with finite volume) so that

\[
m_{\text{BR}}^{\Gamma} \{ g \in G : gV \subset Hg \} > 0.
\]  
(7.10)

Since \( \{ g \in G : gV \subset Hg \} \) is a proper Zariski closed subset, this yields a contradiction. \( \square \)

Lemma 7.13. Suppose \( \mu \) is an infinite joining measure for \((m_{\Gamma_1}^{\text{BR}}, m_{\Gamma_2}^{\text{BR}})\). Then \( \mu \) is not invariant under \( \{ e \} \times V \) for any non-trivial connected subgroup \( V \) of \( U \).

Proof. Without loss of generality, we assume \( m_{\Gamma_2}^{\text{BR}} \) is infinite. It suffices to prove the claim when \( V \) is one dimensional subgroup of \( U \). Set \( V = \{ e \} \times v_t : t \in \mathbb{R} \} \) and \( \tilde{V} = \{ e \} \times v_t : t \in \mathbb{R} \} \).

By the choice of \( \Psi \), \( \Psi \in L^1(\mu) \), and for any \( x = (x_1, x_2) \in X_1 \times X_2 \) and any \( T \geq 1 \), we have

\[
\int_{-T}^{T} \Psi(x\tilde{v}_t)dt = \psi(x_1).
\]  
(7.11)

Also note that every element of the sigma algebra

\[
\Xi = \{ B \times X_2 : B \subset X_1 \text{ any Borel set} \}
\]

is \( \tilde{V} \) invariant. In particular, \( \tilde{V} \)-ergodic components of \( \mu \) are supported on atoms of \( \Xi \) which are of the form \( \{ x^1 \} \times X_2 \) for \( x^1 \in X_1 \). Let \( \mu' \) be a probability measure on \( X_1 \times X_2 \) in the same measure class of \( \mu \) and let \( \int_Y \eta_y d\sigma(y) \) be an ergodic decomposition of \( m_{\Gamma_2}^{\text{BR}} \) for the action of \( \tilde{V} \). Then

\[
\mu = \int_{X_1 \times X_2} \int_Y \eta_y d\sigma(y) d\mu'(x)
\]
gives an ergodic decomposition of \( \mu \) for the action of \( \tilde{V} \).

By Lemma 7.12 a. e. \( \eta_y \) is an infinite measure. Therefore (7.11) and Lemma 7.10 imply that almost all ergodic components of \( \mu \) are dissipative, and hence \( \mu \) is dissipative; so \( \mathcal{C}(\tilde{V}) \) is a null set.

Therefore we can find a Borel subset \( E \subset \pi_1^{-1}(\text{supp}(\psi)) \cap \mathcal{D}(\tilde{V}) \) such that \( \mu(E) > 0 \) and that \( R := \sup_{x \in E} \int_{\mathbb{R}} \chi_E(x\tilde{v}_t)dt < \infty \). Then

\[
\int_{E} \int_{\mathbb{R}} \Psi(x\tilde{v}_t) dt d\mu(x) = \int_{X_1 \times X_2} \Psi(x) \int_{\mathbb{R}} \chi_E(x\tilde{v}_t) dt d\mu(x)
\]

\[\leq R \cdot \int \Psi d\mu < \infty.
\]

This is contradiction to (7.11). \( \square \)

7.5. For the rest of the section, we now suppose that \( \mu \) is an infinite measure

that is, at least one of \( m_{\Gamma_i}^{\text{BR}} \) is infinite. Without loss of generality we will assume \( m_{\Gamma_2}^{\text{BR}} \) is an infinite measure.
Corollary 7.14. The joining measure $\mu$ is quasi-invariant under a subgroup $\{am, uam^{-1} : am \in A'\}$ where $A'$ is a one-parameter unbounded connected subgroup of $AM$ and $u \in U$.

Proof. Let the notation be as in Lemma 7.9. In particular, $Q$ is a compact subset with $\mu(Q) > 0$ and $Q_c \subseteq Q$ with $\mu(Q_c) \geq (1 - \varepsilon)\mu(Q)$.

With this notation let $x_k, y_k \in Q_c$; suppose $x_k = y_kh_k$, with $h_k \rightarrow e$ as $k \rightarrow \infty$. Since $\pi_{is}(\mu) = m^{BR}$, we can choose $x_k$ so that for at least one of $i = 1, 2$, $\pi_i(h_k) \notin N_G(U)$. This, in particular, implies that $h_k \notin N_G(\Delta(U))$.

Now apply Lemma 7.9 with $\{x_k\}$ and $\{y_k\}$. We get a map

$$\varphi : \Delta(U) \rightarrow N_{G \times G}(\Delta(U)) \cap L = \Delta(AM) \cdot \{(e) \times U\}$$

so that $\mu$ is quasi-invariant under a non-trivial connected subgroup, $L$ say, of $\langle \text{Im}(\varphi) \rangle$.

Note that by Lemma 7.13 we have $\{(e) \times V\}$ is not contained in $L$ for any nontrivial subgroup $V$ of $U$. Indeed since $\{(e) \times V\}$ is a unipotent subgroup, quasi-invariance implies invariance.

The conclusion now follows as $L$ is an unbounded connected Lie subgroup of the group $\Delta(AM) \cdot \{(e) \times U\}$, and hence must contain a one-parameter unbounded subgroup of $AM$, up to conjugation by an element of $\{(e) \times U\}$. \[\square\]

Let $\mathcal{P}(X_i)$ denote the space of probability Borel measures. By the standard disintegration theorem (cf. [1, 1.0.8]), for each $i = 1, 2$, there exists an $m_i^{BR}$ co-null set $X_i' \subset X_i$ and a measurable function $x^i : X_i' \rightarrow \mu_{x^i}^{-1}(\Delta^\varepsilon) = \mu_{x^i}^{-1}(\Delta^\varepsilon) \cap \mathcal{P}(X_i)$ such that for any Borel subsets $Y \subset X_1 \times X_2$ and $Z \subset X_i$,

$$\mu(Y \cap \pi_i^{-1}(Z)) = \int_Z \mu_{x^i}(Y) \ dm^{BR}(x^i).$$

The measure $\mu_{x^i}$ is called the fiber measure over $\pi_i^{-1}(x^i)$.

Theorem 7.15. There exists a positive integer $l > 0$ and an $m_i^{BR}$ conull subset $X_i' \subset X_i$ so that $\pi_i^{-1}(x^i)$ is a finite set with cardinality $l$ for all $x^i \in X_i'$. Furthermore,

$$\mu_{x^i}^{-1}((x^1, x^2)) = 1/l$$

for any $x^1 \in X_i'$ and $(x^1, x^2) \in \pi_i^{-1}(x^i)$.

Proof. We first prove that for a.e. $x^1 \in X_1$, the fiber measure $\mu_{x^1}^{-1}$ is fully atomic. Assuming the contrary, we will show that $\mu$ is invariant under $\{(e) \times V\}$ for some non-trivial connected subgroup of $U$, which will be a contradiction by Lemma 7.13.

Put $B = \{x^1 \in X_1 : \mu_{x^1}^{-1}$ is not fully atomic$\}$, and suppose that $m_i^{BR}(B) > 0$. For any $x^1 \in B$ we write

$$\mu_{x^1}^{-1} = (\mu_{x^1}^{-1})^a + (\mu_{x^1}^{-1})^c$$

where $(\mu_{x^1}^{-1})^a$ and $(\mu_{x^1}^{-1})^c$ are respectively the purely atomic and the continuous part of the fiber measure [13]. Let

$$B' = \{(x^1, x^2) : x^1 \in B, \ x^2 \in \text{supp}((\mu_{x^1}^{-1})^c)\}.$$
We take $Q \subset B'$ and $Q_\epsilon \subset Q$ be as in section 7.3 for each small $\epsilon > 0$; In particular, (7.4) holds for $Q$.

Let now $x = (x^1, x^2) \in Q_\epsilon$ be so that there exists a sequence $\{x_k = (x^1_k, x^2_k)\} \subset Q_\epsilon$ so that $x_k \rightarrow x$. Such $x$ exists since $Q \subset B'$. We write

$$x_k = (x^1_k, x^2_k) = (x^1_k, x^2_k)(e, g_k)$$

where $g_k \neq e$ and $g_k \rightarrow e$. There are two possibilities to consider: Recall that

$$N_{G \times G}(\Delta(U)) \cap (\{e\} \times G) = \{e\} \times U.$$

**Case 1.** For all large enough $k$, we have $g_k \in U$, and hence $(e, g_k) \in N_{G \times G}(\Delta(U))$. Since $(x^1, x^2), (x^1, x^2 g_k) \in Q_\epsilon$, Lemma 7.5 implies that $\mu$ is quasi-invariant under $((e, g_k))$. Since $g_k \rightarrow e$ and $U$ is a unipotent group, we get $\mu$ is invariant by $\{e\} \times V$ for some non-trivial connected subgroup $V \subset U$, which is a contradiction by Lemma 7.13.

**Case 2.** By passing to a subsequence, we have $g_k \notin U$, that is, $h_k := (e, g_k) \notin N_{G \times G}(\Delta(U))$. Then we use the construction of $\varphi$ in section 7.3. Note that in view of the choice of $h_k$, the image of $\varphi$ is contained in $N_{G \times G}(\Delta(U)) \cap (\{e\} \times G) = \{e\} \times U$. Indeed, $\varphi_k(t) = (e, u^{-1} g_k u t)$ and $\alpha_k(t) = t$. Thus by Lemma 7.9, we get $\mu$ is quasi-invariant, and hence invariant, by a non-trivial connected subgroup $\{e\} \times V$ with $V > U$. This is again a contradiction by Lemma 7.13. This shows that almost all fiber measures are atomic. Set

$$\Sigma = \{(x^1, x^2) \in X_1 \times X_2 : \mu_{x^1, x^2}^y = \max_{y \in \pi_2^{-1}(x^2)} \mu^y_{x^2}(y)\}. $$

Then $\Sigma$ is a $\Delta(U)$-invariant set. Since almost all fiber measures are atomic, we have $\mu(\Sigma) > 0$. Therefore, in view of the $\Delta(U)$-ergodicity of $\mu$, we have $\Sigma$ is conull. We thus conclude that for $\mu$-almost every point, the fiber measures are uniform distribution on $\ell$-points. \hfill \Box

### 7.6. Reduction to the rigidity of measurable factors.

By Corollary 7.15, we have: $\mu$-a.e fibers of $\pi_1$ have cardinality $\ell$ for some fixed $\ell \in \mathbb{N}$.

We put

$$\Upsilon(x^1) = \pi_1^{-1}(x^1).$$

Then there exist a $U$-invariant BR conull subset $X' \subset X_1$ and $\ell$ measurable maps

$$v_1, \ldots, v_\ell : X' \rightarrow X_2$$

so that $\Upsilon(x^1) = \{v_1(x^1), \ldots, v_\ell(x^1)\}$ for all $x^1 \in X'$. Furthermore,

1. the set map $\Upsilon$ is $U$-equivariant; for all $x^1 \in X'$ and every $u_t \in U$, we have

$$\Upsilon(x^1 u_t) = \Upsilon(x^1) u_t.$$
Recall from Corollary 7.14 that there exists some \( u \in U \) so that \( \mu \) is quasi invariant under some one parameter unbounded (Lie) subgroup \( A' \) of \( AM \). Replacing \( v_i \) by \( u^{-1} v_i \) for all \( 1 \leq i \leq \ell \), and replacing \( X' \) by a BR conull subset, which we continue to denote by \( v_i \) and \( X' \), we have
\[
\Upsilon(x^1 am) = \Upsilon(x^1 am) \quad \text{for all } x^1 \in X' \text{ and all } am \in A'.
\]
Moreover, by the following lemma 7.17,

(ii) the set map \( \Upsilon \) is \( AM \)-equivariant; for all \( x^1 \in X' \) and every \( am \in AM \), we have
\[
\Upsilon(x^1 am) = \Upsilon(x^1 am) \quad \text{(7.13)}
\]

We will use the following Lemma 7.16.

There exists a compact subset \( C \subset X_1 \) with the following property: for any \( x \in X_1 \) with \( x - \in \Lambda_r(\Gamma_1) \), there exists a sequence \( s_i \to +\infty \) so that \( xa_{-s_i} \in C \).

**Proof.** We use notations from subsection 5.1. If \( x - \in \Lambda_r(\Gamma) \), then
\[
d_{\text{hyp}}(\mathcal{C}(\Gamma_1), xa_{-s}) \to 0
\]
as \( s \to \infty \). Take \( C \) to be the 1-neighborhood of the thick part \( \mathcal{C}_0(\Gamma_1) \) of the convex core of \( \Gamma_1 \). Suppose now that the claim does not hold, i.e., if \( xa_s \) does not intersect \( C \) for all sufficiently large \( s \), \( xa_s \) must remain in one of the thin components of \( \mathcal{C}(\Gamma_1) \) for all sufficiently large \( s \). However this implies \( x^- \) is a parabolic limit point of \( \Gamma_1 \), which is a contradiction. \( \Box \)

**Lemma 7.17.** The measure \( \mu \) is \( \Delta(M) \)-invariant.

**Proof.** If \( F = \mathbb{R} \), then \( M = \{e\} \); so we assume \( G = \text{PSL}_2(\mathbb{C}) \) for this lemma. The proof is similar to the proof of Theorem 6.1. Let \( C \) be as in Lemma 7.16. Let \( K \supset C \) be an \( M \) invariant compact subset of \( X' \) with positive BR-measure and \( m^{BR}(\partial K) = 0 \). For \( \varepsilon > 0 \), since \( \Upsilon \) is measurable, there is a compact subset \( K' \) of \( K \) with \( m^{BR}(K') > (1 - \varepsilon) \cdot m^{BR}(K) \) on which the map \( \Upsilon \) is continuous and, in particular bounded. Let \( Y \) be a set where Lemma 7.5 holds true with \( \Psi = \pi_{1^{-1}}(\psi) \) as before. It suffices to find a dense subset \( M' \subset M \) so that for all \( m \in M' \) there is some \( y \in Y \) with \( y\Delta(m) \in Y \).

Replacing \( X' \) with a conull set, we may and will assume the Hopf ratio theorem holds with \( \psi \) in the denominator and for all functions \( c_c(X_1) \) as well as for the characteristic functions \( \chi_K \) and \( \chi_{K'} \) in the numerator. Since \( Y \) is \( \Delta(U) \) invariant, replacing \( X' \) with a conull subset, one more time, we may and will assume
\[
\{(x^1, x^2) : x^1 \in X', x^2 \in \Upsilon(X')\} \subset Y.
\]
For every \( x^1 \in K' \), put
\[
M_{x^1} = \{m : x^1 m \in K'\}.
\]
Since \( m^{BR} \) is \( M \) invariant, we have
\[
m^{BR}\{x^1 \in K' : |M_{x^1}| \geq 1/2\} > 0.8 m^{BR}(K') \quad \text{(7.15)}
\]
where \(|M_{x^1}|\) means the measure of \(M_{x^1}\) with respect to the probability Haar measure on \(M\). Fix some \(x^1 \in K'\) with \(|M_{x^1}| \geq 1/2\) and let \(m \in M_{x^1}\). We claim
\[
\Upsilon(x^1m) = \Upsilon(x^1)m.
\]
Recall that \(m^{BR}(K') \geq (1 - \varepsilon)m^{BR}(K)\). Note that \(x^1mu_t = x^1u_{m,t}m\) for all \(t \in \mathbb{R}\). Since the Hopf ratio theorem holds for every \(x^1 \in K\), the \(M\)-invariance of \(K\) implies that for all large \(T\), we have
\[
|(\{t \in B_U(T) : x^1u_t, x^1u_tm \in K'\}| \geq (1 - 4\varepsilon)\int_{B_U(T)} \chi_K(x^1u_t)dt \quad (7.16)
\]
Fixing a small \(\eta > 0\) and choosing \(m \in M\) small enough, we may and will assume that if both \(x^1u_t\) and \(x^1u_tm\) belong to \(K'\), then
\[
d(\Upsilon(x^1u_t), \Upsilon(x^1m)m^{-1}u_t) \leq \eta,
\]
\(\Upsilon(x^1m)m^{-1}u_t = \Upsilon(x^1m)(m^{-1}u_t)m^{-1} = \Upsilon(x^1u_t)m^{-1}\), \(\Upsilon\) is uniformly continuous on \(K'\), and hence \(d(\Upsilon(x^1u_t)m^{-1}, \Upsilon(x^1u_t)) \leq \eta\). Put
\[
F(t) = \min\{d(\Upsilon(x^1u_t), \Upsilon(x^1m)m^{-1}u_t)^2, 1\}.
\]
In view of (7.16) we have
\[
\frac{\int_{B_U(T)} \chi_K(x^1u_t)F(t)dt}{\int_{B_U(T)} \chi_K(x^1u_t)dt} \leq 4\varepsilon + \eta.
\]
By Lemma 7.16 and since \(C \subset K\) there is a sequence \(T_k \to \infty\) so that \(x_{a_{-1}} \log T_k \in K\). Therefore, the above and Lemma 6.2 imply
\[
\sup_{B_U(2RT_k)} F(t) \leq C \cdot (4\varepsilon + \eta) \quad (7.17)
\]
for some constant \(C > 0\).

Let us write \(\Upsilon(x^1m) = \{v_1(x^1)g_1, \ldots, v_l(x^1)g_l\}\) and for each \(1 \leq j \leq l\) put
\[
f_j(t) = u_{-1}g_jm^{-1}u_t.
\]
Then (7.17) implies \(f_j\) is a constant function for all \(1 \leq j \leq l\). That is, \(g_jm^{-1}\) must be in the centralizer of \(U\), which is \(U\). Hence for each \(j\), there exists some \(u_j \in U\) so that \(g_jm^{-1} = u_j\).

All together we get \((x^1, v_j(x^1)) \in Y\) and \((x^1m, v_j(x^1)u_jm) \in Y\) for all \(1 \leq j \leq l\). Therefore, by Lemma 7.5 we get \(\mu\) is quasi invariant under \((m, u_jm)\). Since \(u_jm \in MU\), \(u_jm = u'_jm(u'_j)^{-1}\) for some \(u'_j \in U\). Since the group \(MU \times MU\) does not admit any non-trivial character, we get that the measure is actually invariant under such elements.

In the case at hand \(M\) is one dimensional and \(m \in M_{x^1}\) can be chosen from a positive measure case. Therefore, we get that there exists \(u_0 \in U\) so that \(\mu\) is invariant under a subgroup \(R := \{(m, u_0m^{-1}) : m \in M\}\).

We claim \(u_0 = e\); this finishes the proof. Suppose \(u_0 \neq e\). Let \(A'\) be as in (ii) above, let \(am_a \in A'\) and \(m \in M\). Consider the commutator of elements
of $A(A')$ and $(m, u_0 m u_0^{-1}) \in R$. Since the first component is a commutator among elements of $AM$ which is commutative, the first component is necessarily $e$. So for any $a m_d \in A'$ and $(m, u_0 m u_0^{-1}) \in R$, set

$$(e, v) := (a m_d, a m_d)(m, u_0 m u_0^{-1})(a^{-1} m_d^{-1}, a^{-1} m_d^{-1})(m^{-1}, u_0 m^{-1} u_0^{-1}) = (e, a m_d u_0 m u_0^{-1} a^{-1} m_d^{-1} u_0 m^{-1} u_0^{-1}).$$

Note that $U$ is the commutator subgroup of $AMU$ and hence $v \in U$. Moreover $v \neq e$ if $a m_d \neq e, m \neq e$, since $u_0 \neq e$. Observe that $v$ can be made arbitrarily close, but not equal, to $e$ by choosing these parameters close but not equal to the identity.

Recall now that $A(A')$ leaves $\mu$ quasi-invariant and $\{(m, u_0 m u_0^{-1}) : m \in M\}$ leaves $\mu$ invariant. From the above we get: if $u_0 \neq e$, then there is some $v_i \in U$ with $v_i \to e$, but not equal to $e$, so that $(e, v_i)$ leaves the measure invariant for all $i \geq 1$. Therefore, there is a one parameter subgroup $V \subset U$ so that $\mu$ is invariant under $\{e\} \times V$. This contradicts Lemma 7.13 and finishes the proof.  \hfill $\square$

Let us recall that the BMS measure and the $BR$-measure on $X_1$ have product structures, and have the “same transversal measures”. To be more precise, let $\psi \in C_c(X)$ and further assume that supp$(\psi) \subseteq y P \varepsilon U_\varepsilon$ with $P = MAU^-$. Then

$$m^{BR}(\psi) = \int \int \psi(y p u) d\nu(y p) d\mu_{Leb}(u), \text{ and } m^{BMS}(\psi) = \int \int \psi(y p u) d\nu(y p) d\mu_{PS}(u) \quad (7.18)$$

where $d\nu$ is the transversal measure on $P$.

**Proposition 7.18.** Let $\Upsilon : X_1 \to X_2$ be as above, in particular, it satisfies that $\Upsilon(xg) = \Upsilon(x) g$ for all $x \in X'$ and all $g \in MAU$. Then there exists $q_0 \in G$ such that $\Gamma_1 \cap q_0^{-1} \Gamma q_0$ has finite index in $\Gamma_1$ so that if we put $\Gamma_2 q_0 \Gamma_1 = \bigcup_{1 \leq j \leq l} \Gamma_2 q_0 \gamma_j$ for $\gamma_j \in \Gamma_1$, then

$$\Upsilon(\Gamma_1 g) = \{\Gamma_2 q_0 \gamma_j g : 1 \leq j \leq l\}$$

on a BR conull subset of $X_1$. Moreover $\mu$ is a $A(U)$-invariant measure supported on $\{(x^1, \Upsilon(x^1)) : x^1 \in X_1\}$ and hence a finite cover self-joining (see Def. (1.2)).

**Proof.** By Theorem 6.1, we have $\Upsilon(x g) = \Upsilon(x) g$ for all $x \in X'$ and all $g \in U^- MAU$ where $X'$ is given by loc. cit. Fixing $x_0 \in X'$, define

$$\Upsilon(x_0 g) = \Upsilon(x_0) g$$

for all $g \in G$. Then $\Upsilon$ and $\Upsilon'$ coincide with each other on $Y' := x_0(U^- MAU)$. Since $U^- MAU$ is a Zariski open subset of $G$, $m^{BMS}(Y') = 1$, and hence $\Upsilon = \Upsilon'$ on a BMS-conull subset. Hence we may assume without loss of generality

$$\Upsilon(x g) = \Upsilon(x) g \quad (7.19)$$
for all $g \in G$ and $x \in Y'$ with $xg \in Y'$. Let $\Gamma_1 g_0 \in Y'$ and write $\Upsilon(\Gamma_1 g_0) = \{\Gamma_2 h_1, \ldots, \Gamma_2 h_\ell\}$. Then for every $g \in G$ such that $\Gamma_1 g_0 g \in Y'$ we have

$$\Upsilon(\Gamma_1 g_0 g) = \{\Gamma_2 h_1 g, \ldots, \Gamma_2 h_\ell g\}.$$ 

Note that for all $\gamma \in \Gamma_1$ we have $\Gamma_1 g_0 (g_0^{-1} \gamma g_0) = \Gamma_1 g_0 \in Y'$. Therefore, applying the $G$-equivariance (7.19) of $\Upsilon$ to $g_0^{-1} \gamma g_0 \in g_0^{-1} \Gamma_1 g_0$, we get the set $\{\Gamma_2 h_1 g, \ldots, \Gamma_2 h_\ell g\}$ is right invariant under $g_0^{-1} \Gamma_1 g_0$. It follows that $\Gamma_2 \setminus \Gamma_2 h_i (g_0^{-1} \Gamma_1 g_0)$ is finite for each $i$. Putting $q_i := h_i g_0^{-1}$, we have $\Gamma_2 \setminus \Gamma_2 q_i \Gamma_1$ is finite. Let $q_1, \ldots, q_r$ be such that the corresponding cosets $\Gamma_2 \setminus \Gamma_2 q_i \Gamma_1$ are distinct and $\cup_{1 \leq i \leq r} \Gamma_2 \setminus \Gamma_2 q_i \Gamma_1 = \cup_{1 \leq i \leq r} \Gamma_2 \setminus \Gamma_2 h_i g_0^{-1} \Gamma_1$. Thus, if for each $1 \leq i \leq r$ we put $\Gamma_2 q_i \Gamma_1 = \cup_{1 \leq j \leq i} \Gamma_2 q_i \gamma_{ij}$ for $\gamma_{ij} \in \Gamma_1$, then

$$\Upsilon(\Gamma_1 g) = \{\Gamma_2 q_1 \gamma_{11} g, \ldots, \Gamma_2 q_1 \gamma_{1s} g, \ldots, \Gamma_1 g_r \gamma_{r1} g, \ldots, \Gamma_2 q_r \gamma_{r1} g\} \quad (7.20)$$

on a BMS conull subset of $X_1$.

In particular we get $q_i^{-1} \Gamma_1 q_i$ is commensurable with a subgroup of $\Gamma_2$. Repeating the argument with $\Gamma_2$ we get, up to a conjugation, $\Gamma_1$ and $\Gamma_2$ are commensurable with each other.

Let us note now that in view of (7.18) we have that the $U$ invariant set $X'' = Y' \cap X' U$ has full BR measure. Let now $g \in G$ be so that $\Gamma_1 g \in X''$ then we can write $g = g'u$ where $\Gamma_1 g' \in X'$ and $u \in U$. Now property (7.12) of $\Upsilon$ and (7.20) imply

$$\Upsilon(\Gamma_1 g) = \Upsilon(\Gamma_1 g') u$$

$$= \{\Gamma_2 q_1 \gamma_{11} g' u, \ldots, \Gamma_2 q_1 \gamma_{1s} g' u, \ldots, \Gamma_2 q_r \gamma_{r1} g' u, \ldots, \Gamma_2 q_r \gamma_{r1} g' u\}$$

$$= \{\Gamma_2 q_1 \gamma_{11} g, \ldots, \Gamma_2 q_1 \gamma_{1s} g, \ldots, \Gamma_2 q_r \gamma_{r1} g, \ldots, \Gamma_2 q_r \gamma_{r1} g\}.$$ 

Now for every $1 \leq i \leq r$ we have that the set

$$\{(x^1, x^2) : x^1 = \Gamma_1 g \in X'', x^2 \in \{\Gamma_2 q_1 \gamma_{11} g, \ldots, \Gamma_2 q_r \gamma_{r1} g\}\} \quad (7.21)$$

is $\Delta(U)$ invariant and has positive $\mu$ measure. Therefore, by the ergodicity of $\mu$, we get $r = 1$ and $\mu$ is a $\Delta(U)$-invariant measure supported on the set (7.21). This implies that $\mu$ is a finite self-joining as defined in the introduction. This finishes the proof. \hfill \Box

Remark 7.19. Note that for a general discrete non-elementary subgroup $\Gamma$, the Patterson-Sullivan density always exists, although it may not be a unique $\Gamma$ conformal density of dimension $\delta_\Gamma$, and hence the BR-measure is well-defined on $\Gamma \setminus G$. Given this, the above proof yields a stronger version of Corollary 1.4 in the introduction where $\Gamma_2$ is not assumed to be geometrically finite. Namely, we have: if $\Gamma_1$ is geometrically finite and Zariski dense and $\Gamma_2$ is a Zariski dense (not necessarily geometrically finite) subgroup of $G$ with infinite co-volume, then a $U$-joining on $\Gamma_1 \setminus G \times \Gamma_2 \setminus G$ with respect to the pair $(\mu_1^{BR}, \mu_2^{BR})$ exists only when $\Gamma_1$ is commensurable with a subgroup of $\Gamma_2$, up to conjugation.
References


Department of Mathematics, The University of Texas at Austin, Austin, TX 78750

E-mail address: amir@math.utexas.edu

Mathematics department, Yale university, New Haven, CT 06511 and Korea Institute for Advanced Study, Seoul, Korea

E-mail address: hee.oh@yale.edu