POINTWISE EQUIDISTRIBUTION AND TRANSLATES OF MEASURES ON HOMOGENEOUS SPACES

OSAMA KHALIL

Abstract. Let $(X, \mathfrak{B}, \mu)$ be a Borel probability space. Let $T_n : X \to X$ be a sequence of continuous transformations on X. Let ν be a probability measure on X such that $\frac{1}{N}\sum_{n=1}^{N}(T_n)_*\nu \to \mu$ in the weak-$*$ topology. Under general conditions, we show that for ν almost every $x \in X$, the measures $\frac{1}{N}\sum_{n=1}^{N}\delta_{T_n x}$ get equidistributed towards μ if N is restricted to a set of full upper density. We present applications of these results to translates of closed orbits of Lie groups on homogeneous spaces. As a corollary, we prove equidistribution of exponentially sparse orbits of the horocycle flow on quotients of $SL(2,\mathbb{R})$, starting from every point in almost every direction.

1. Introduction

Many problems in number theory and geometry can be recast in terms of the equidistribution of translates of appropriate measures on quotients of certain Lie groups. The general set up of these results is a Borel probability space $(X, \mathfrak{B}, \mu)$, a probability measure ν on X (usually singular with respect to μ) and a sequence of transformations $T_n : X \to X$ such that

$$\frac{1}{N}\sum_{n=1}^{N}(T_n)_*\nu \xrightarrow{N\to\infty} \mu \tag{1.1}$$

where $(T_n)_*\nu$ is the pushforward of ν under T_n and the convergence is in the weak-$*$ topology. A natural question is to what extent can one extend such results to describe the behavior of measures of the form

$$\frac{1}{N}\sum_{n=1}^{N}\delta_{T_n x} \tag{1.2}$$

for ν-almost every x, where δ_y denotes the dirac delta measure at a point y.

Recently, an analogous question for flows was addressed by Chaika and Eskin [CE] in the context of flat surfaces. In that setting, X is some affine submanifold of the moduli space of flat structures on a surface, μ is a natural affine $SL(2,\mathbb{R})$ invariant measure, ν is the measure supported on an orbit of $SO(2)$ and the transformations are of the form $a(t) = diag(e^t, e^{-t})$. They show that for ν almost every x, the empirical measures analogous to those in (1.2) get equidistributed towards μ.

In the context of homogeneous spaces, Shi [Shi] explored this question for translates of measures supported on (pieces of) orbits of certain horospherical subgroups of Lie groups by one parameter diagonalizable subgroups. Here X is a homogeneous space for a Lie group G, ν is a measure on an orbit of a certain horospherical subgroup, and $T_n = T^n$, where T is an

2010 *Mathematics Subject Classification.* 37A30, and 22F30.

Key words and phrases. Pointwise Equidistribution, Birkhoff's Theorem, Ratner's Theorem.

Ad-diagonalizable element of G. Equidistribution of empirical measures towards the natural G-invariant Haar measure is proven ν almost everywhere.

In [KSW], an effective version of this result is obtained via different methods. The convergence of empirical measures analogous to (1.2) is proven for general dynamical systems under the hypothesis of some form of exponential mixing of the transformation T with respect to the non-invariant measure ν.

In all three cases, equidistribution was obtained by exploiting specific properties of the system at hand, while not directly utilizing the fact that (1.1) holds.

1.1. **Statement of Results.** In this article, we approach this question in the general context of continuous measure preserving transformations, assuming (1.1) only. We obtain equidistribution results of measures in (1.2) under general conditions yet only along subsequences of full upper density. Recall that the upper density of a subset $A \subseteq \mathbb{N}$, denoted by $\overline{d}(A)$ is defined to be

$$\overline{d}(A) = \limsup_{N\to\infty} \frac{\#(A \cap [1,N])}{N}$$

In what follows, X will be a locally compact, second countable topological space and $\mathcal{B}$ is its Borel σ-algebra. A pair $(X, \mathcal{B})$ will be called a standard Borel space. The following is our first main result for the case of translations by powers of a single transformation.

Theorem 1.1. *Suppose $(X, \mathcal{B}, \mu)$ is a standard Borel probability space. Let T be a continuous ergodic measure preserving transformation of X. Assume ν is a probability measure on X satisfying*

$$\frac{1}{N}\sum_{n=0}^{N-1} T_*^n \nu \xrightarrow[N\to\infty]{weak-*} \mu$$

Then, for ν-almost every $x \in X$, there exists a sequence $A(x) \subseteq \mathbb{N}$, of upper density 1, such that for all $\psi \in C_c(X)$,

$$\lim_{\substack{N\to\infty \\ N\in A(x)}} \frac{1}{N}\sum_{n=0}^{N-1} \psi(T^n x) = \int \psi \, d\mu$$

Our next result concerns the more general situation of translating by sequences of transformations. In this generality, we assume more structure on the possible limit points of the empirical measures.

Theorem 1.2. *Suppose $(X, \mathcal{B}, \mu)$ be a standard Borel probability space. Let $(T_n)_n$ be a sequence of continuous transformations of X. Let $S : X \to X$ be a continuous ergodic μ preserving transformation. Let ν be a probability measure on X. Assume the following holds:*

(1) $\frac{1}{N}\sum_{n=1}^{N} (T_n)_*\nu \xrightarrow{N\to\infty} \mu$ in the weak-$*$ topology.
(2) For ν-almost every $x \in X$, any limit point of the sequence of measures $\frac{1}{N}\sum_{n=1}^{N} \delta_{T_n x}$ is S-invariant.
(3) There exists a Borel measurable, σ-compact set $Z \in \mathcal{B}$ such that $\mu(Z) = 0$ and for all ergodic S invariant measures $\lambda \neq \mu$, $\lambda(Z) = 1$.

Then, for ν-almost every $x \in X$, there exists a sequence $A(x) \subseteq \mathbb{N}$, of upper density equal to 1, such that for all $\psi \in C_c(X)$,

$$\lim_{\substack{N\to\infty \\ N\in A(x)}} \frac{1}{N}\sum_{n=1}^{N} \psi(T_n x) = \int \psi \, d\mu$$

We remark that the maximal inequality used in the proof of Theorem 1.2 does not require the hypotheses of the Theorem.

We now discuss some applications of these results. Theorem 1.1 is broadly applicable to general dynamical systems and thus we present some applications of Theorem 1.2 within homogeneous dynamics where we demonstrate that its hypotheses are verified.

1.2. **Sparse Equidistribution and Translates of Orbits of Lie Groups.** Our motivation for studying this question comes from the problem of sparse equidistribution of unipotent flows on homogeneous spaces. This was conjectured by Shah in [Sha]. Recent progress was achieved in [Ven] for the horocycle flow on compact quotients of $SL_2(\mathbb{R})$ along sequences of the form $n^{1+\gamma}$ for small values of γ. See also [TV, Zhe, FFT] for more results in this direction and the work of Sarnak and Ubis on the equidistribution along the primes [SU] on $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$. However, the question in full generality remains open.

With the help of Theorem 1.2, we obtain an equidistribution result for exponentially sparse orbits of unipotent flows, of which the horocycle flow is an example.

In order to apply Theorem 1.2 in this setup, we introduce the notion of Ratner Sequences. We say a sequence g_n of elements of G is a **Ratner Sequence** with respect to a probability measure ν on G/Γ if there exists a non-trivial Ad-unipotent element $u \in G$ such that for ν almost every $x \in G/\Gamma$, any limit point of the sequence of empirical measures $\frac{1}{N}\sum_{n=1}^{N}\delta_{g_n x}$ is invariant by u. Our main theorem in this set up is the following.

Theorem 1.3. *Suppose G is a connected semisimple Lie group and Γ is an irreducible lattice in G. Assume g_n is a Ratner Sequence with respect to a probability measure ν on G/Γ satisfying*

$$\frac{1}{N}\sum_{n=1}^{N}(g_n)_*\nu \xrightarrow{N\to\infty} \mu_{G/\Gamma} \tag{1.3}$$

where $\mu_{G/\Gamma}$ denotes the unique G-invariant Haar probability measure on G/Γ. Then, for ν almost every $x \in G/\Gamma$, there exists a sequence $A(x) \subseteq \mathbb{N}$ of upper density 1 such that

$$\lim_{\substack{N\to\infty \\ N\in A(x)}} \frac{1}{N}\sum_{n=1}^{N} \delta_{g_n x} = \mu_{G/\Gamma}$$

Apart from Theorem 1.2, a key ingredient in the proof of Theorem 1.3 is Ratner's measure classification theorem [Rat].

The assumption on the equidistribution of the translates of a probability measure ν by a sequence of elements g_n is satisfied in numerous examples in homogeneous dynamics. For example, when ν is the Haar measure supported on a closed orbit of a symmetric subgroup of G, it was shown in [EM] that (1.3) is satisfied as soon as g_n tends to infinity in G/H. Recall that H is said to be a symmetric subgroup if it is the fixed point set of an involution of G.

Using Ratner's theorem, this result was extended in [EMS] to include translates of maximal reductive subgroups that is, subgroups which are only invariant by an involution of G. Thus, the applicability of Theorem 1.3 boils down to the existence of Ratner sequences which is the subject of the next section.

1.2.1. ***Existence of Ratner Sequences***. We prove a general criterion on the existence of sparse Ratner Sequences inside 1-parameter unipotent subgroups. The proof relies on the representation theory of embedded copies of $SL_2(\mathbb{R})$ inside G in addition to a generalization of a technique developed by Chaika and Eskin [CE] in the setting of strata of Abelian differentials.

For a subgroup $H \leqslant G$, let $Z_G(H)$ denote the centralizer of H inside G. For 2 sequences of positive real numbers a_n and b_n, we use $a_n \asymp b_n$ to mean their ratio is uniformly bounded from above and below by positive constants for all n. The following is the main statement.

Theorem 1.4. *Suppose G is a semisimple Lie group and Γ is a discrete subgroup of G. Let $H \leqslant G$ be a closed connected subgroup and let μ_H denote the H-invariant probability measure supported on a closed orbit of H on G/Γ. Assume $U = \{u_t : t \in \mathbb{R}\}$ is a 1-parameter Ad-unipotent subgroup of G such that $U \not\subset Z_G(H)$. Then, for every sequence $t_n > 0$ satisfying $t_n \asymp e^{\lambda n}$ for some constant $\lambda > 0$, u_{t_n} is a Ratner sequence for μ_H.*

Unipotent invariance is deduced using a law of large numbers argument which is similar in spirit to Breiman's law of large numbers. However, the relevant random variables are weakly dependent and exponential growth is used to guarantee a sufficiently fast rate of decay of correlations. This leaves open the question of whether Theorem 1.4 holds for polynomially growing sequences.

In the appendix, using different methods, we prove a more crude criterion for the existence of Ratner sequences when H is a symmetric subgroup of G, but we drop the restriction on the unipotence of the elements g_n.

We show that any sequence g_n satisfying an exponential growth condition similar to the one in Theorem 1.4 contains a Ratner sequence as a subsequence. See Theorem A.1 for the precise statement. Note that we only require H to be closed and connected in Theorem 1.4.

In the particular instance when $G = SL_2(\mathbb{R})$ and the sequence g_n comes from the action of the horocycle flow, Theorem 1.3 takes the following form:

Corollary 1.5. *Let $G = SL(2,\mathbb{R})$, $\Gamma \subset G$ a lattice and let $H = SO(2)$. Let $\lambda > 0$ and for $n \in \mathbb{N}$, let $t_n = e^{\lambda n}$. Let*

$$g_n = \begin{pmatrix} 1 & t_n \\ 0 & 1 \end{pmatrix}, k_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Then, for every $x \in G/\Gamma$ and for almost every $\theta \in [0, 2\pi]$, there exists a sequence $A(\theta) \subseteq \mathbb{N}$ of upper density 1 such that

$$\lim_{\substack{N\to\infty \\ N \in A(\theta)}} \frac{1}{N} \sum_{n=1}^{N} \delta_{g_n k_\theta x} = \mu_{G/\Gamma}$$

Moreover, if G/Γ is compact then $A(\theta) = \mathbb{N}$.

Proof. Since H is its own centralizer in G, by Theorem 1.4, g_n is a Ratner sequence for H. By the work of Eskin and McMullen [EM], one has that $g_n\mu_H \to \mu_{G/\Gamma}$. Thus, the corollary follows from Theorem 1.3. When G/Γ is compact, the Haar measure is uniquely ergodic under the action of unipotent elements on G/Γ. Since g_n is a Ratner sequence, the claim follows. □

The main point of the Corollary is that it holds for every x. This is not guaranteed by any general theorem on sparse equidistribution almost everywhere. We remark that even density of the sparse orbits considered in Theorem 1.3 is not known in this generality.

The paper is organized as follows. In § 2, we prove an analogue of the maximal ergodic theorem in our set up. We use this to prove Theorems 1.1 and 1.2 in § 3 and § 4 respectively. In § 5, we discuss the existence of Ratner sequences and prove Theorem 1.4. In § 6, we provide a proof of Theorem 1.3.

Acknowledgements. I would like to thank my advisor, Nimish Shah, for numerous valuable discussions.

2. An Analogue of the Maximal Inequality

The following proposition is an extension of the classical maximal ergodic theorem to the set up involving sequences of transformations and a non-invariant measure. The key observation is to convert the classical statement of the maximal inequality into one concerning finite shifted orbit averages with quantitative control on the size of the shift in comparison with the length of the orbit segment.

Proposition 2.1. *Let $(X, \mathcal{B})$ be a standard Borel probability space. Let T_n be a sequence of continuous transformations on X. Let ν, μ be probability measures on X such that*

$$\frac{1}{N}\sum_{n=0}^{N-1}(T_n)_*\nu \xrightarrow[N\to\infty]{weak-*} \mu$$

Let $\psi \in C_c(X)$ and let $\alpha > 0$ and $\beta \in (0,1)$. For every $j, N \geq 1$, define the set

$$E^{\psi}_{\alpha,N,j} = \left\{ x \in X : \sup_{1\leq M\leq N} \left| \frac{1}{M}\sum_{k=j}^{j+M-1} \psi(T_k x) \right| > \alpha \right\}$$

Then, for all sufficiently large $N \in \mathbb{N}$, depending on ψ and satisfying $\beta \leq 1 - 1/N$, there exists some $0 \leq j_N < \beta N$, such that

$$\alpha\beta\nu(E^{\psi}_{\alpha,N,j_N}) \leq 12||\psi||_{L^1(\mu)}$$

We will deduce this proposition from the classical maximal inequality for $l^1(\mathbb{Z})$ which is a consequence of Vitali's covering lemma.

Lemma 2.2 (Lemma 2.29, [EW])**.** *Let $\phi \in l^1(\mathbb{Z})$. Define the following maximal function, for $a \in \mathbb{Z}$:*

$$\phi^*(a) = \sup_{N\geq 1}\left|\frac{1}{N}\sum_{i=0}^{N-1}\phi(i+a)\right|$$

Let $\alpha > 0$ and define

$$E_\alpha = \{a \in \mathbb{Z} \mid \phi^*(a) > \alpha\}$$

Then,

$$\alpha\,|E_\alpha| \leq 3||\phi||_{l^1(\mathbb{Z})}$$

Proof of Proposition 2.1. Let $\psi \in C_c(X)$ and let $\alpha > 0$, $\beta \in (0,1)$. Let $N \geq 1$ be such that $\beta \leq 1 - 1/N$ and let $E^{\psi}_{\alpha,N,j}$ be as in the statement. Let $x \in X$ and let $J > N$ be a parameter to be determined later. Define the following function

$$\phi(j) = \begin{cases} \psi(T_j x) & 0 \leq j \leq J \\ 0 & \textit{otherwise} \end{cases}$$

Then, clearly $\phi \in l^1(\mathbb{Z})$. For $a \in \mathbb{Z}$, define the following two functions

$$\phi^*(a) = \sup_{1 \leq M} \left| \frac{1}{M} \sum_{k=0}^{M-1} \phi(k+a) \right|, \qquad \phi^*_N(a) = \sup_{1 \leq n \leq N} \left| \frac{1}{n} \sum_{k=0}^{n-1} \phi(k+a) \right|$$

Define the corresponding exceptional sets

$$E^{\phi}_{\alpha} = \{a \in \mathbb{Z} \mid \phi^*(a) > \alpha\}, \qquad E^{\phi}_{\alpha,N} = \{a \in [0, J-N-1] \mid \phi^*_N(a) > \alpha\}$$

By Lemma 2.2 applied to ϕ, we have

$$\alpha \left| E^{\phi}_{\alpha,N} \right| \leq \alpha \left| E^{\phi}_{\alpha} \right| \leq 3 ||\phi||_{l^1(\mathbb{Z})} \tag{2.1}$$

Note that for $a \in [0, J-N-1]$, we have

$$\phi^*_N(a) = \sup_{1 \leq n \leq N} \left| \frac{1}{n} \sum_{k=0}^{n-1} \psi(T_{k+a} x) \right| \tag{2.2}$$

Let $\chi_{\alpha,N,j}$ denote the indicator function of $E^{\psi}_{\alpha,N,j}$. Thus, for $j \in [0, J-N-1]$,

$$\chi_{\alpha,N,j}(x) = 1 \text{ if and only if } j \in E^{\phi}_{\alpha,N} \tag{2.3}$$

Thus, combining (2.1), (2.2) and (2.3) along with the definition of ϕ, we get

$$\alpha \sum_{j=0}^{J-N-1} \chi_{\alpha,N,j}(x) = \alpha \left| E^{\phi}_{\alpha,N} \right| \leq 3 \sum_{j=0}^{J} |\phi(j)| = 3 \sum_{j=0}^{J} |\psi(T_j x)|$$

Integrating both sides of the above with respect to ν yields

$$\alpha \sum_{j=0}^{J-N-1} \nu(E^{\psi}_{\alpha,N,j}) \leq 3 \sum_{j=0}^{J} \int |\psi(T_j x)| \, d\nu(x) \tag{2.4}$$

Taking $J = (1+\beta)N$ in (2.4), dividing both sides by $J - N$ and noting that $(1+\beta)N + 1 \leq 2N$, we get the following.

$$\begin{aligned} \frac{\alpha}{\beta N} \sum_{j=0}^{\beta N - 1} \nu(E^{\psi}_{\alpha,N,j}) &\leq 3 \frac{(1+\beta)N+1}{\beta N} \frac{1}{(1+\beta)N+1} \sum_{j=0}^{(1+\beta)N} \int |\psi(T_j x)| \, d\nu(x) \\ &\leq \frac{6}{\beta} \frac{1}{(1+\beta)N+1} \sum_{j=0}^{(1+\beta)N} \int |\psi(T_j x)| \, d\nu(x) \end{aligned} \tag{2.5}$$

Now, by assumption,

$$\frac{1}{(1+\beta)N+1} \sum_{j=0}^{(1+\beta)N} \int |\psi(T_j x)| \, d\nu(x) \to \int |\psi| \, d\mu = ||\psi||_{L^1(\mu)}$$

Thus, for all N sufficiently large, depending on ψ, we have

$$\frac{1}{(1+\beta)N+1}\sum_{j=0}^{(1+\beta)N}\int |\psi(T_j x)|\, d\nu(x) \leq 2||\psi||_{L^1(\mu)}$$

Combining this with (2.5), we get for all N sufficiently large,

$$\frac{1}{\beta N}\sum_{j=0}^{\beta N-1}\nu(E^{\psi}_{\alpha,N,j}) \leq \frac{12||\psi||_{L^1(\mu)}}{\alpha\beta}$$

Thus, there must exist some $j = j(N) \in [0, \beta N - 1]$ for which the conclusion of the Proposition holds.

□

3. An Analogue of Birkhoff's Ergodic Theorem

This section is dedicated to the proof of Theorem 1.1. With the maximal inequality for the non-invariant measure ν in place (Proposition 2.1), our proof will follow Bourgain's approach in deducing pointwise convergence from the mean ergodic theorem (cf. [Bou], Section 2-C).

Recall that for a sequence of sets A_n, the *limsup* of these sets is the set of elements which belong to A_n for infinitely many n. More precisely,

$$\limsup_{n\to\infty} A_n = \bigcap_{n\geq 1}\bigcup_{k\geq n} A_k$$

The following simple observation will be used repeatedly in what follows.

Lemma 3.1. *Let X be a standard Borel space and let μ be a probability measure on X. Let $A_n \subseteq X$ be a sequence of measurable sets such that $\mu(A_n) \geq \alpha$ for some $\alpha \in [0,1]$. Then,*

$$\mu\left(\limsup_{n\to\infty} A_n\right) \geq \alpha$$

Proof. This follows from the definition of $\limsup_{n\to\infty} A_n$ as a decreasing intersection and the continuity of the measure μ. □

The following Lemma is the main step in the proof of Theorem 1.1.

Lemma 3.2. *Let $(X, \mathcal{B}, \mu)$ be a standard Borel probability space. Let T be an ergodic measure preserving transformation on X. Let ν be a probability measure on X such that*

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n_*\nu \xrightarrow[N\to\infty]{weak-*} \mu$$

Let $f_1, \dots, f_n \in C_c(X)$. Then, for ν-almost every $x \in X$, there exists a sequence $A \subseteq \mathbb{N}$, of upper density 1, depending on x and the functions $f_1, \dots, f_n$, such that for all $k = 1, \dots, n$,

$$\lim_{\substack{N\in A\\ N\to\infty}} \frac{1}{N}\sum_{n=0}^{N-1} f_k(T^n x) = \int f_k \mu$$

Let us deduce Theorem 1.1 from this Lemma first.

3.1. **Proof of Theorem 1.1.** Let $\mathcal{F} = \{f_k \in C_c(X) : k \in \mathbb{N}\}$ be an enumeration of a countable set of continuous functions which are dense in $C_c(X)$ in the uniform norm. Then, it suffices to show that for ν almost every $x \in X$, there exists a sequence $A(x) \subseteq \mathbb{N}$, of full upper density such that for all $f \in \mathcal{F}$,

$$\lim_{\substack{N\to\infty \\ N\in A(x)}} \frac{1}{N}\sum_{k=1}^{N} f(T^k x) = \int f \, d\mu \tag{3.1}$$

For each $n \in \mathbb{N}$, let $\mathcal{F}_n = \{f_1, \dots, f_n\} \subset \mathcal{F}$. By Lemma 3.2, for each n, there exists a set X_n with $\nu(X_n) = 1$, such that for all $x \in X_n$, there exists a sequence $A(x, \mathcal{F}_n) \subseteq \mathbb{N}$, along which the limit in (3.1) holds for all $f \in \mathcal{F}_n$.

Let $Y = \cap_n X_n$. Then, $\nu(Y) = 1$. Let $y \in Y$. We will build a sequence $A(y)$ by induction from the sequences $A(y, \mathcal{F}_n)$. For each $n \in \mathbb{N}$, let $N_n \in \mathbb{N}$ be such that for all $N \geq N_n$ with $N \in A(y, \mathcal{F}_n)$, and all $f \in \mathcal{F}_n$,

$$\left| \frac{1}{N}\sum_{k=1}^{N} f(T^k y) - \int f \, d\mu \right| \leq \frac{1}{n} \tag{3.2}$$

Let $M_1 = N_1$. If M_j has been defined, let M_{j+1} be such that the following holds

$$\frac{|A(y, \mathcal{F}_j) \cap [1, M_{j+1}]|}{M_{j+1}} \geq 1 - \frac{1}{j}$$
$$\frac{M_j}{M_{j+1}} \leq \frac{1}{j}$$
$$M_{j+1} \geq N_{j+1}$$

Note that the above implies that

$$\frac{|A(y, \mathcal{F}_j) \cap [M_j, M_{j+1}]|}{M_{j+1}} \geq 1 - \frac{2}{j}$$

Now, define the sequence $A(y)$ as follows:

$$A(y) = \bigcup_{j=1}^{\infty} A(y, \mathcal{F}_j) \cap [M_j, M_{j+1}]$$

Thus, by construction, the upper density of $A(y)$ is equal to 1. Now, let $f \in \mathcal{F}$. Then, $f \in \mathcal{F}_n$ for all $n \geq n_0$, for some $n_0 \in \mathbb{N}$. Thus, for all $N \geq M_{n_0}$ such that $N \in A(y)$, there exists $j \geq n_0$, such that $N \in A(y, \mathcal{F}_j) \cap [M_j, M_{j+1}]$. Thus, since $M_j \geq N_j$, by (3.2), the conclusion follows.

3.2. **Proof of Lemma 3.2.** For any function ψ and for every $N \geq 1$, let $\mu(\psi) = \int \psi \, d\mu$ and let

$$A_N(\psi) = \frac{1}{N}\sum_{n=0}^{N-1} \psi \circ T^n$$

Let $\varepsilon \in (0, 1/2)$. By Von Neumann's mean ergodic theorem, for all k,

$$A_N(f_k) \xrightarrow{L_1(\mu)} \mu(f_k)$$

Hence, we can find some $M \gg 1$, for all $k \leq n$,

$$\int |A_M(f_k) - \mu(f_k)| \, d\mu < \frac{\varepsilon^3}{n}$$

Let $\beta = \frac{\varepsilon}{C}$, where

$$C = \max_{1 \leq k \leq n} 2||f_k||_{L^\infty} + 1$$

Let $g_k = A_M(f_k) - \mu(f_k)$. Note that $||g_k||_\infty \leq C$. For all k and for every $N \in \mathbb{N}$, define

$$E^k_{\varepsilon,N} = \left\{ x \in X : \sup_{1 \leq m \leq N} \left| \frac{1}{m} \sum_{l=0}^{m-1} g_k(T^l x) \right| > \varepsilon \right\}$$

Note that since $\beta \leq \varepsilon < 1/2$, for all $N \geq 2$, one has that $\beta \leq 1 - 1/N$. Thus, by the analogue of the maximal inequality, Proposition 2.1, applied to g_k, the sequence of transformations $T_l = T^l$ and $E^{g_k}_{\varepsilon,N,j} = T^{-j} E^k_{\varepsilon,N}$, for all $N \geq 2$ sufficiently large, depending on ε, there exists $j_{N,k} \in [0, \beta N]$ such that

$$\nu(T^{-j_{N,k}} E^k_{\varepsilon,N}) \leq \frac{12||g_k||_{L^1(\mu)}}{\varepsilon\beta} \leq \frac{12C\varepsilon}{n} \tag{3.3}$$

Let $G^\varepsilon_{N,k} = X \setminus T^{-j_{N,k}} E^k_{\varepsilon,N}$ and let $G^\varepsilon_N = \bigcap_{k=1}^n G^\varepsilon_{N,k}$. Thus, by (3.3) and Lemma 3.1,

$$\nu\left(\limsup_{N \to \infty} G^\varepsilon_N \right) \geq 1 - 12C\varepsilon \tag{3.4}$$

Now, let $y \in G^\varepsilon_N$ and let $Q \in [\sqrt{\varepsilon}N, N] \cap \mathbb{N}$. Then, for all $k = 1, \ldots, n$, by definition of $E^k_{\varepsilon,N}$ and our choice of β,

$$\begin{aligned} \left|A_{\lceil Q+\beta N \rceil}(g_k)(y)\right| &\leq \left| \frac{Q}{Q+\beta N} \frac{1}{Q} \sum_{l=j_{N,k}}^{j_{N,k}+Q-1} g_k(T^l y) \right| + \frac{\beta N ||g_k||_{L^\infty}}{Q+\beta N} \\ &\leq \left|A_Q(g_k)(T^{j_{N,k}} y)\right| + \frac{C\beta}{\sqrt{\varepsilon}} \\ &\leq \varepsilon + \sqrt{\varepsilon} \leq 2\sqrt{\varepsilon} \end{aligned} \tag{3.5}$$

where for any $R \in \mathbb{R}$, $\lceil R \rceil$ denotes the least integer greater than R.

Hence, in particular, for any $y \in \limsup_N G^\varepsilon_N$, there exists a sequence $N_i \to \infty$ for which (3.5) holds for all $Q \in [\sqrt{\varepsilon}N_i, N_i] \cap \mathbb{N}$ and for all $k = 1, \ldots, n$. Define the following sequence for $y \in \limsup_N G^\varepsilon_N$

$$A(y, \varepsilon) = \bigcup_{N_i : y \in G^\varepsilon_{N_i}} [(\sqrt{\varepsilon} + \beta)N_i, (1+\beta)N_i] \cap \mathbb{N} \tag{3.6}$$

Now, a simple computation shows that for all N, and any function ψ,

$$\begin{aligned} A_N(A_M(\psi)) &= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \psi \circ T^{n+m} \\ &= A_N(\psi) + O_M\left(\frac{||\psi||_{L^\infty}}{N} \right) \end{aligned} \tag{3.7}$$

Combining 3.5 and 3.7 implies that for every $y \in \limsup_N G_N^\varepsilon$, all $k \leq n$ and for all $Q \in A(y,\varepsilon)$ such that $Q \gg M$,

$$|A_Q(f_k)(y) - \mu(f_k)| \leq 3\sqrt{\varepsilon} \tag{3.8}$$

Note that the choice of ε in the above considerations was arbitrary. Hence, we may consider sets $B^{\varepsilon_i} = X \setminus \limsup_N G_N^{\varepsilon_i}$, where $\varepsilon_i = 1/i^2$. Then, by (3.4) and the Borel-Cantelli lemma applied to B^{ε_i},

$$\nu\left(X \setminus \limsup_{i\to\infty} B^{\varepsilon_i}\right) = 1 \tag{3.9}$$

Let $y \in X \setminus \limsup_i B^{\varepsilon_i}$. We claim that there exists a sequence $A(y) \subseteq \mathbb{N}$ of full upper density, such that for all $k \leq n$,

$$\limsup_{\substack{N\to\infty \\ N\in A(y)}} |A_N(f_k)(y) - \mu(f_k)| = 0 \tag{3.10}$$

By (3.9), this will conclude the proof.

Since $y \in \limsup_N G_N^{\varepsilon_i}$ for all i sufficiently large, we have sequences $A(y,\varepsilon_i)$ as before for all ε_i sufficiently small. Without loss of generality, we may assume this holds for all ε_i. Note that by (3.6), the upper density of $A(y,\varepsilon_i)$ is at least $\frac{1-\sqrt{\varepsilon_i}}{1+\varepsilon_i/C}$.

We build the sequence $A(y)$ from $A(y,\varepsilon_i)$ by induction as follows. Let $N_i \in A(y,\varepsilon_i)$ be such that (3.8) holds for all $k \leq n$ and all $Q \geq N_i$. Let $M_1 = N_1$. If M_j is defined, let $M_{j+1} \in \mathbb{N}$ be such that

$$\frac{|A(y,\varepsilon_j) \cap [1, M_{j+1}]|}{M_{j+1}} \geq \frac{1-\sqrt{\varepsilon_j}}{1+\varepsilon_j/C} - \frac{1}{j}$$
$$\frac{M_j}{M_{j+1}} \leq \frac{1}{j}$$
$$M_{j+1} \geq N_{j+1}$$

This in particular, implies that

$$\frac{|A(y,\varepsilon_j) \cap [M_j, M_{j+1}]|}{M_{j+1}} \geq \frac{1-\sqrt{\varepsilon_j}}{1+\varepsilon_j/C} - \frac{2}{j}$$

Now, define the sequence $A(y)$ as follows:

$$A(y) = \bigcup_{j=1}^{\infty} (A(y,\varepsilon_j) \cap [M_j, M_{j+1}])$$

Thus, since $\varepsilon_j \to 0$, the upper density of $A(y)$ is equal to 1. Moreover, by (3.8) and by choice of M_j, we have that (3.10) holds as desired.

4. An Analogue of Birkhoff's Theorem for Sequences of Transformations

In this section, we prove Theorem 1.2. We will use similar ideas to those used in the proof of Theorem 1.1 by applying the weak-type maximal inequality to a carefully chosen set of continuous functions capturing the structure of the ergodic invariant measures under the transformation S.

First, we make some standard reductions. Note that since X is locally compact and second countable, by passing to the one point compactification and extending all the transformations on X trivially to the point at infinity, we may assume that X is in fact compact.

The set Z in the assumption will then be enlarged to include the point at infinity since the Dirac measure at that point will be an ergodic invariant measure for S. Also, since X is now assumed compact, the space of probability measures on it is weak-$*$ compact and thus we can always find limit points of infinite sequences.

Proof of Theorem 1.2. Let $\varepsilon \in (0, 1/16)$ be fixed and let $\varepsilon_n = \varepsilon^2/4^n$ for $n \in \mathbb{N}$. Write $Z = \cup_n K_n$, where K_n is compact and $K_n \subseteq K_{n+1}$ for each n. By regularity of the measure μ, since $\mu(Z) = 0$, there exists an open set U_n containing Z, such that $\mu(U_n \setminus K_n) < \varepsilon_n$, for each n.

Moreover, by Urysohn's lemma, we can find a continuous function $0 \leq f_n \leq 1$ such that $f_n|_{K_n} \equiv 1$ and $f_n \equiv 0$ on $X \setminus U_n$. Thus,

$$||f_n||_{L^1(\mu)} = \int f_n(x)\, d\mu(x) < \varepsilon_n$$

Let n be fixed. For each $j \in \mathbb{N}$, $k \leq n$ and $\alpha \in \mathbb{R}$, define the following set

$$E^k_{\alpha,N,j} = \left\{ x \in X : \sup_{1 \leq M \leq N} \frac{1}{M} \sum_{m=j+1}^{j+M} f_k(T_m x) > \alpha \right\}$$

Applying the analogue of the maximal inequality, Proposition 2.1, with f_k, $\alpha_k = \beta_k = \varepsilon_k^{1/4} < 1/2$, we get that for all $N \geq 2$ sufficiently large, depending on f_k, there exists $j_{N,k} \in [0, \beta_k N]$ such that

$$\nu\left(E^k_{\alpha_k,N,j_{N,k}}\right) \leq \frac{12||f_k||_{L^1(\mu)}}{\alpha_k \beta_k} \ll \varepsilon_k^{1/2} \tag{4.1}$$

for each $k \leq n$.

Let $G_{N,k} = X \setminus E^k_{\alpha_k,N,j_{N,k}}$. Let $y \in G_{N,k}$ and let $Q \in [\varepsilon_k^{1/8} N, N] \cap \mathbb{N}$. Then, by definition of $E^k_{\alpha_k,N,j_N}$,

$$\begin{aligned} \frac{1}{Q + \beta_k N} \sum_{l=1}^{Q+\beta_k N} f_k(T_l y) &\leq \frac{Q}{Q + \beta_k N} \frac{1}{Q} \sum_{l=j_{N,k}+1}^{j_{N,k}+Q} f_k(T_l y) + \frac{\beta_k N ||f_k||_{L^\infty}}{Q + \beta_k N} \\ &\leq \alpha_k + \frac{\beta_k}{\varepsilon_k^{1/8}} \leq 2\varepsilon_k^{1/8} \end{aligned} \tag{4.2}$$

Now, for each $N \gg 1$, depending on n, define the following set

$$V_{N,n} = \bigcap_{k=1}^{n} G_{N,k} \tag{4.3}$$

and let $W_n = \limsup_N V_{N,n}$. The sets W_n have the following properties:

- By (4.1), since $\nu(G_{N,k}) \geq 1 - \varepsilon/2^k$, $k = 1, \ldots, n$, we have $\nu(V_{N,n}) \geq 1 - \varepsilon$ for all $N \gg 1$. In particular,

$$\nu(W_n) = \nu\left(\limsup_{N\to\infty} V_{N,n}\right) \geq 1 - \varepsilon \tag{4.4}$$

- For each $y \in W_n$, by (4.2) and (4.3), and noting that $\varepsilon > \varepsilon_k$, there exists a sequence $A(y,n) \subseteq \mathbb{N}$ defined by

$$A(y,n) = \bigcup_{N_i : y \in V_{N_i,n}} [(\varepsilon^{1/4} + \varepsilon^{1/8})N_i, N_i] \cap \mathbb{N} \tag{4.5}$$

such that for all $k = 1, \ldots, n$ and for all $Q \in A(y,n)$,

$$\frac{1}{Q} \sum_{l=1}^{Q} f_k(T_l y) \leq 2\varepsilon_k^{1/8} \tag{4.6}$$

Let $W = \limsup_n W_n$. Then, by (4.4),

$$\nu(W) \geq 1 - \varepsilon \tag{4.7}$$

Let $y \in W$. Then, there exists a sequence $n_i \to \infty$, such that $y \in W_{n_i}$ for all i. We will construct a sequence $A(y)$ from the sequences $A(y, n_i)$ defined in (4.5) as follows. Let

$$\eta = \varepsilon^{1/4} + \varepsilon^{1/8}$$

First, we define a sequence N_i by induction. Let $N_0 = 1$. If N_j has been defined, let N_{j+1} be such that the following holds

$$\frac{|A(y, n_{j+1}) \cap [1, N_{j+1}]|}{N_{j+1}} \geq 1 - 2\eta$$

$$\frac{N_j}{N_{j+1}} \leq \eta$$

This is possible since the sequences $A(y,n)$ have upper density at least $1-\eta$. These conditions imply that

$$\frac{|A(y, n_{j+1}) \cap [N_j, N_{j+1}]|}{N_{j+1}} \geq 1 - 3\eta \tag{4.8}$$

Now, define the sequence $A(y)$ as follows:

$$A(y) = \bigcup_{j=0}^{\infty} A(y, n_{j+1}) \cap [N_j, N_{j+1}] \tag{4.9}$$

Thus, by (4.8), we get

$$\limsup_{N \to \infty} \frac{|A(y) \cap [1, N]|}{N} \geq 1 - 3\eta \tag{4.10}$$

We claim that

$$\lim_{\substack{N \to \infty \\ N \in A(y)}} \frac{1}{N} \sum_{n=1}^{N} \delta_{T_n y} = \mu \tag{4.11}$$

Let λ_∞^y be any weak-$*$ limit of the sequence $\frac{1}{N}\sum_{n=1}^{N} \delta_{T_n y}$, $N \in A(y)$. First, we claim that $\lambda_\infty^y(Z) = 0$. Suppose otherwise. Then, since $Z = \bigcup_i K_i$ and $K_i \subseteq K_{i+1}$ for all i, there exists some i_0 such that for all $i > i_0$:

$$\lambda_\infty^y(K_i) \geq \lambda_\infty^y(K_{i_0}) > 0$$

Fix some $i > i_0$. By definition of the functions f_i, $\lambda_\infty^y(f_i) \geq \lambda_\infty^y(K_i)$. Then, for all $n_j \geq i$ and for all $N \in A(y, n_{j+1}) \cap [N_j, N_{j+1}] \subset A(y)$, by (4.6), we get

$$\lambda_\infty^y(f_i) \leq 2\varepsilon_i^{1/8}$$

In particular, we get that $\lambda_\infty^y(K_{i_0}) \leq 2\varepsilon_i^{1/8}$. But, this applies to $i > i_0$. Thus, since $\varepsilon_i \to 0$, we get that $\lambda_\infty^y(K_{i_0}) = 0$, a contradiction.

Next, by our hypothesis, (after possibly intersecting W with a set of full measure), λ_∞^y is S-invariant. However, all the ergodic S-invariant measures different from μ live on Z to which λ_∞^y assigns 0 mass. Thus, by the ergodic decomposition, we get that $\lambda_\infty^y = \mu$. Hence, the sequence $\frac{1}{N}\sum_{n=1}^N \delta_{T_n y}$, $N \in A(y)$ has μ as its only weak-$*$ limit point as desired.

Thus far, we proved that for all $y \in W$, a set of ν measure at least $1 - \varepsilon$, there exists a sequence $A(y)$ of upper density at least $1 - 3\eta$ such that (4.11) holds. Since ε was arbitrary, the conclusion of the theorem holds ν almost everywhere as desired.

□

5. Existence of Ratner Sequences

In this section, we prove a general criterion for the existence of Ratner sequences, Theorem 1.4. We fix some notation which will be used throughout the section. G is a semisimple Lie group, Γ is a discrete subgroup of G, and H is a closed connected subgroup. We assume that there exists some $x \in G/\Gamma$ such that the orbit Hx is closed in G/Γ and supports an H-invariant probability measure, which we denote by μ_H. By replacing Γ by a conjugate subgroup, we may assume that x is the identity coset.

Let $Z_G(H)$ denote the centralizer of H in G. We use $\mathfrak{g}$ and $\mathfrak{h}$ to denote the Lie algebras of G and H respectively. For $g \in G$, $Ad(g)$ denotes the linear transformation on $\mathfrak{g}$ induced by the adjoint action of g. We also fix some norm $\|\cdot\|$ on $\mathfrak{g}$.

Recall that a sequence of elements $g_n \in G$ is said to be a Ratner sequence for μ_H if there exists a one-parameter unipotent subgroup $W < G$ such that for μ_H-almost every $x \in G/\Gamma$, any limit point of the empirical measures $\frac{1}{N}\sum_{n=1}^N \delta_{g_n x}$ is invariant by W.

5.1. $SL_2(\mathbb{R})$ Representations and Unipotent Invariance.

The first key step in the proof of Theorem 1.4 is Proposition 5.1 below. It relies on the representation theory of embedded copies of $SL_2(\mathbb{R})$ inside semisimple Lie groups and its proof is inspired by Ratner's H-principle appearing in the proof of her measure classification theorem.

The statement roughly says that it is possible to change the starting point of a unipotent orbit of a group U in a direction parallel to H so that the 2 unipotent orbits differ roughly by a unipotent element in the centralizer of U. The following is the precise statement.

Proposition 5.1. *Let $U = \{u_t : t \in \mathbb{R}\}$ be an Ad-unipotent 1-parameter subgroup of G such that $U \not\subset Z_G(H)$. Then, for every sequence $t_n \to \infty$, there exists a sequence $v_n \in \mathfrak{h} = Lie(H)$ satisfying the following: for all n sufficiently large*

$$\|v_n\| \ll \frac{1}{\|u_{t_n}|_{\mathfrak{h}}\|}, \qquad \exp(Ad(u_{t_n})(v_n)) \xrightarrow{n\to\infty} u$$

where $u \in Z_G(U)$ is a non-trivial Ad-unipotent element of G.

Proof. Let $X \in \mathfrak{g}$ be such that $u_t = \exp(tX)$ for all $t \in \mathbb{R}$. Since u_t is Ad-unipotent, X is an ad-nilpotent element of $\mathfrak{g}$. That, $ad(X)$ is a nilpotent linear transformation of $\mathfrak{g}$.

Since $\mathfrak{g}$ is a semisimple Lie algebra, by the Jacobson-Morozov theorem, X may be extended to an $\mathfrak{sl}_2$-triple. That is there exists $Y, h \in \mathfrak{g}$ such that the following relations hold:

$$[h, X] = 2X, \qquad [h, Y] = -2Y, \qquad [X, Y] = h$$

Hence, the Lie subalgebra $\mathfrak{f}$ generated by X, Y and h is isomorphic to the Lie algebra $\mathfrak{sl}_2(\mathbb{R})$. Since $\mathfrak{f}$ is semisimple, $\mathfrak{g}$ decomposes into irreducible representations under the adjoint action of $\mathfrak{f}$ as follows:

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_s$$

For $1 \leq i \leq s$, let $\pi_i : \mathfrak{g} \to V_i$ denote the associated projections and note that $Ad(u_t)$ commutes with π_i for all i. Let $v \in \mathfrak{h}$.

Let $1 \leq i \leq s$ be such that $\pi_i(v)$ is not fixed by $Ad(u_t)$. Let $n_i \in \mathbb{N}$ be such that

$$\dim(V_i) = n_i + 1$$

By the standard description of irreducible $\mathfrak{sl}_2(\mathbb{R})$ representations, V_i decomposes into 1 dimensional eigenspaces for the action h as follows:

$$V_i = W_0^{(i)} \oplus W_1^{(i)} \oplus \cdots \oplus W_{n_i}^{(i)}$$

so that for each $0 \leq l \leq n_i$ and every $w \in W_l^{(i)}$, we have

$$h \cdot w = (n_i - 2l)w$$

Let $q_l : V_i \to W_l^{(i)}$ denote the associated projections. Let $\left\{ w_l^{(i)} : 0 \leq l \leq n_i \right\}$ denote a basis of V_i consisting of eigenvectors of h and write

$$\pi_i(v) = \sum_{l=0}^{n_i} c_l^{(i)} w_l^{(i)}$$

Note that for each l, we have that

$$Ad(u_t) \cdot w_l^{(i)} = \sum_{k=0}^{l} \binom{l}{k} t^{l-k} w_k^{(i)}$$

In particular, we get the following

$$q_0(\pi_i(Ad(u_t)(v))) = q_0(Ad(u_t)(\pi_i(v))) = \sum_{k=0}^{n_i} c_k^{(i)} t^k w_0^{(i)} \tag{5.1}$$

Note also that the degree of the polynomial appearing in the coefficient of $q_l(\pi_i(Ad(u_t)(v)))$ for any $l > 0$ is strictly less than the degree of the polynomial in (5.1). Let

$$d_i(v) = \max \left\{ 0 \leq k \leq n_i : c_k^{(i)} \neq 0 \right\}$$

And, define the following natural number

$$d_{\mathfrak{h}} = \max \left\{ d_i(v) : 1 \leq i \leq s, v \in \mathfrak{h} \right\}$$

By assumption, we have that $U \not\subset Z_G(H)$. Thus, $d_{\mathfrak{h}} \neq 0$. In particular, we can find $v \in \mathfrak{h}$ and $1 \leq i \leq s$ such that

$$d_i(v) = d_{\mathfrak{h}}$$

Now, let $t_n \in \mathbb{R} \setminus \{0\}$ be a sequence tending to infinity and define v_n to be

$$v_n = \frac{1}{t_n^{d_\mathfrak{h}}} v$$

Thus, for each i with $d_i(v) = d_\mathfrak{h}$, by (5.1), we get that

$$q_0(\pi_i(Ad(u_{t_n})(v_n))) = (c_{d_\mathfrak{h}}^{(i)} + O(1/t_n))w_0^{(i)}$$

And, for each $l > 0$, we have that

$$q_l(\pi_i(Ad(u_{t_n})(v_n))) \xrightarrow{n\to\infty} 0$$

In particular, we get that

$$Ad(u_{t_n})(v_n) \xrightarrow{n\to\infty} \sum_{\substack{1\leq i\leq s \\ d_i(v)=d_\mathfrak{h}}} c_{d_\mathfrak{h}}^{(i)} w_0^{(i)} \neq 0 \tag{5.2}$$

Next, since $v_n \to 0$, $\exp(v_n)$ converges to the identity element in G. Hence, all the eigenvalues of the linear transformation $Ad(\exp(v_n))$ tend to 1. Thus, since conjugation doesn't change eigenvalues, we get that $Ad(u_{t_n})(\exp(v_n))$ converges to an Ad-unipotent element of G which is non-trivial by (5.2). Since $Ad(u_t)$ fixes $w_0^{(i)}$ for all i, by (5.2), the limiting element belongs to the centralizer of U.

Finally, note that equation (5.1) and the definition of $d_\mathfrak{h}$ imply that

$$\|Ad(u_t)|_\mathfrak{h}\| = O(t^{d_\mathfrak{h}}) \tag{5.3}$$

as t tends to ∞. This implies the bound on $\|v_n\|$ and completes the proof. ☐

The following example demonstrates the above Proposition in the concrete set up of $G = SL_2(\mathbb{R})$.

Example 5.2. Let $G = SL(2,\mathbb{R})$ and $H = K = SO(2)$. Let $t_n \to +\infty$ be a sequence. Let g_n be the following sequence:

$$g_n = \begin{pmatrix} 1 & t_n \\ 0 & 1 \end{pmatrix}$$

Let $k_\theta \in K$. Then,

$$g_n k_\theta g_n^{-1} = \begin{pmatrix} \cos(\theta) - t_n \sin(\theta) & (t_n^2+1)\sin(\theta) \\ -\sin(\theta) & \cos(\theta) + t_n\sin(\theta) \end{pmatrix}$$

Let $\alpha \in (0,1)$ be a fixed real number. For all large n, let θ_n be such that $t_n^2 \sin(\theta_n) = \alpha$. Then, as $n \to \infty$, $\theta_n \to 0$ and $t_n \sin(\theta_n) \to 0$. Hence, we get

$$g_n k_{\theta_n} g_n^{-1} \to u(\alpha) = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \neq id$$

5.2. **Decay of Correlations.** Let Γ, H and G be as above and define

$$g_n = u_{t_n}$$

Let $\varphi \in C_c^\infty(G/\Gamma)$. For each n, define the following function on $H\Gamma/\Gamma$:

$$f_n(h\Gamma) = \varphi(g_n \exp(v_n)h\Gamma) - \varphi(g_n h\Gamma) \tag{5.4}$$

where $v_n \in Lie(H)$ is as in the conclusion of Proposition 5.1 applied to the sequence t_n. The reason for defining such functions is the following

Proposition 5.3. *To prove Theorem 1.4, it suffices to show that for μ_H almost every $x \in H\Gamma/\Gamma$, the following holds:*

$$\frac{1}{N}\sum_{n=1}^{N} f_n(x) \to 0$$

Proof. Let $h_n = \exp(v_n)$. By Proposition 5.1, we have $g_n h_n g_n^{-1} \to u$, where u is a non-trivial Ad-unipotent element. Since, φ is uniformly Lipschitz, we have

$$\varphi(ug_nx) - \varphi(g_nh_nx) = \varphi(ug_nx) - \varphi(g_nh_ng_n^{-1}g_nx) = O\left(d(u, g_nh_ng_n^{-1})\right)$$

where $d(.,.)$ is the right invariant metric on G. Hence, $|\varphi(ug_nx) - \varphi(g_nh_nx)| \to 0$ as $n \to \infty$.

Hence, by assumption,

$$\left|\frac{1}{N}\sum_{n=1}^{N}(\varphi(ug_nx) - \varphi(g_nx))\right| \leq \frac{1}{N}\sum_{n=1}^{N}|\varphi(ug_nx) - \varphi(g_nh_nx)| + \left|\frac{1}{N}\sum_{n=1}^{N} f_n(x)\right| \longrightarrow 0$$

But, φ was an arbitrary function. Thus, any limit point must be invariant by the group generated by u.

□

In order to use Proposition 5.3, we need to control the correlations between the functions f_n. This is established in the following lemma which is an analogue of [CE, Lemma 3.3] in our setting. We shall need the following definition.

Definition 5.4. For any $x \in H\Gamma/\Gamma \cong H/(H \cap \Gamma)$, the injectivity radius at x, denoted by inj_x is defined to be the infimum over all $r > 0$ such that the map $h \mapsto hx$ is injective on the ball of radius r around identity in H.

Lemma 5.5. *For all $n \geq m \geq 1$ such that $\|Ad(g_m)|_{\mathfrak{h}}\| / \|Ad(g_n)|_{\mathfrak{h}}\|$ is sufficiently small, the following holds*

$$\int f_n(h\Gamma)f_m(h\Gamma)\, d\mu_H(h\Gamma) = O\left(\left(\frac{\|Ad(g_m)|_{\mathfrak{h}}\|}{\|Ad(g_n)|_{\mathfrak{h}}\|}\right)^{1/2}\right) \tag{5.5}$$

Proof. Let v_n be as in the definition of the functions f_n and let $h_n = \exp(v_n)$. Let $d_n = \|Ad(g_n)|_{\mathfrak{h}}\|$ and $d_m = \|Ad(g_n)|_{\mathfrak{h}}\|$. Define

$$r = \left(\frac{1}{d_m d_n}\right)^{1/2} \tag{5.6}$$

Let $B_H(e,r)$ denote the ball of radius r around the identity in H. By abuse of notation, we'll use μ_H to denote the Haar measure on H and on $H\Gamma/\Gamma$.

Let $\psi : H\Gamma/\Gamma \to \mathbb{R}$ be any integrable function. Then, by Fubini's theorem and left H-invariance of μ_H,

$$\int_{H\Gamma/\Gamma} \psi(x)\, d\mu_H(x) = \int_{H\Gamma/\Gamma} \frac{1}{\mu_H(B_H(e,r))} \int_{B_H(e,r)} \psi(hx)\, d\mu_H(h)\, d\mu_H(x)$$

Define the following set

$$Thick_r = \{x \in H\Gamma/\Gamma : inj_x \geq r\}$$

where inj_x denotes the injectivity radius at x in $H/(H \cap \Gamma)$. Let $Thin_r = H\Gamma/\Gamma - Thick_r$. Using the structure of Siegel sets, one can show (Lemma 11.2, [BO]) that $\mu_H(Thin_r) \ll r^p$, for some $p > 0$ as $r \to 0$. Hence, it suffices to prove for all $x \in Thick_r$,

$$\frac{1}{\mu_H(B_H(e,r))} \int_{B_H(e,r)} f_n(hx) f_m(hx)\, d\mu_H(h) = O\left(\left(\frac{\|Ad(g_m)|_{\mathfrak{h}}\|}{\|Ad(g_n)|_{\mathfrak{h}}\|} \right)^{1/2} \right) \tag{5.7}$$

Let $w \in Thick_r$ be fixed. Let B_r denote the ball of radius r around the w in $H\Gamma/\Gamma$ in the metric induced by the metric on G. Then, for every $x \in B_r$, there exists some $l \in B_H(e,r)$ such that $x = lw$. Since $\varphi \in C_c^\infty(G/\Gamma)$, φ is uniformly Lipschitz. Thus, we get

$$\varphi(g_m h_m x) - \varphi(g_m h_m w) = \varphi(g_m h_m l w) - \varphi(g_m h_m w) = O\left(d(g_m h_m l h_m^{-1} g_m^{-1}, e)\right)$$

Since the sequences d_n, d_m are tending to infinity, for all n, m sufficiently large, r will be small enough so that the exponential map is a diffeomorphism from a neighborhood of 0 in $\mathfrak{h} = Lie(H)$ onto $B_H(e,r)$.

Thus, we can write $l = \exp(v)$ for some $v \in \mathfrak{h}$. So, we have

$$\|Ad(g_m h_m)(v)\| \leq \|Ad(g_m)|_{\mathfrak{h}}\| \cdot \|Ad(h_m)\| \cdot \|v\|$$

But, since $h_m \to id$ as $m \to \infty$ and since the norm is continuous, for all m sufficiently large, we have $\|Ad(h_m)\| \ll 1$.

Moreover, since the differential of the exponential map at 0 is the identity, its Jacobian is 1 at 0 and hence, when r is sufficiently small, we have $\|v\| \ll d(l,e) \leq r$. Combining these estimates, we get for all $x \in B_r$,

$$\|Ad(g_m h_m)(v)\| = O\left(\|Ad(g_m)|_{\mathfrak{h}}\|\, r\right) = O\left(\left(\frac{\|Ad(g_m)|_{\mathfrak{h}}\|}{\|Ad(g_n)|_{\mathfrak{h}}\|} \right)^{1/2} \right)$$

But, as before, the exponential map is nearly an isometry near identity. Hence, when $\|Ad(g_n)|_{\mathfrak{h}}\|$ is sufficiently larger than $\|Ad(g_m)|_{\mathfrak{h}}\|$, $Ad(g_m h_m)(v)$ will be sufficiently close to 0 so that $d(\exp(Ad(g_m h_m)(v)), e) \ll \|Ad(g_m h_m)(v)\|$ up to absolute constants. Thus, we get for all $x \in B_r$,

$$\varphi(g_m h_m x) - \varphi(g_m h_m w) = O\left(\left(\frac{\|Ad(g_m)|_{\mathfrak{h}}\|}{\|Ad(g_n)|_{\mathfrak{h}}\|} \right)^{1/2} \right)$$

Similarly, we get the same estimate for $\varphi(g_m x) - \varphi(g_m w)$ for all $x \in B_r$. Thus, by definition of f_m, we get

$$f_m(x) - f_m(w) = O\left(\|Ad(g_m)|_{\mathfrak{h}}\|\, r\right)$$

Thus, we get that

$$\frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x) f_m(x)\, d\mu_H(x) = \frac{f_m(w)}{\mu_H(B_r)} \int_{B_r} f_n(x)\, d\mu_H(x) + O\left(\|Ad(g_m)|_{\mathfrak{h}}\|\, r\right) \tag{5.8}$$

Next, note that by definition of f_n and left-invariance of μ_H,

$$\begin{aligned} \frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x)\, d\mu_H(x) &= \frac{1}{\mu_H(B_r)} \int_{h_n B_r} \varphi(g_n x)\, d\mu_H(x) - \frac{1}{\mu_H(B_r)} \int_{B_r} \varphi(g_n x)\, d\mu_H(x) \\ &= O\left(\frac{\mu_H(h_n B_r \triangle B_r)}{\mu_H(B_r)} \right) \end{aligned}$$

Since $w \in Thick_r$, B_r isometric to $B_H(e,r)$. Hence, for all r sufficiently small, we may apply Proposition 5.6 below to get

$$\frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x)\, d\mu_H(x) = O\left(\frac{d(h_n,e)}{r}\right) \tag{5.9}$$

Note that Proposition 5.6 requires that $h_n \in B_H(e,r)$. To get around this assumption, observe that for n sufficiently large, $h_n = \exp(v_n)$ will be sufficiently close to identity and hence, we have

$$d(h_n,e) \ll \|v_n\|$$

But, by Proposition 5.1, we have $\|v_n\| \ll 1/\left\|Ad(g_n)|_{\mathfrak{h}}\right\|$ for all n sufficiently large. But, since $\|Ad(g_n)|_{\mathfrak{h}}\| \geq \|Ad(g_m)|_{\mathfrak{h}}\|$ by assumption, we have $\|Ad(g_n)|_{\mathfrak{h}}\| \geq 1/r$. Thus, in particular, h_n will be contained in a ball of radius comparable to r for all large n, which doesn't affect our estimate.

Moreover, this observation, along with (5.9), imply that

$$\frac{1}{\mu_H(B_r)} \int_{B_r} f_n(x)\, d\mu_H(x) = O\left(\frac{1}{r\left\|Ad(g_n)|_{\mathfrak{h}}\right\|}\right) \tag{5.10}$$

Combining this estimate with (5.7) and (5.8) gives

$$\int_{Thick_r} f_n(h\Gamma) f_m(h\Gamma)\, d\mu_H(h\Gamma) = O\left(\left(\frac{\|Ad(g_m)|_{\mathfrak{h}}\|}{\|Ad(g_n)|_{\mathfrak{h}}\|}\right)^{1/2}\right) \tag{5.11}$$

and the conclusion of the lemma follows. □

5.2.1. *A measure estimate.* The following estimate was used in the proof of Lemma 5.5.

Proposition 5.6. *Let H be a Lie group and let B_r denote a ball of radius $r > 0$ around the identity in H. Then, for all $r > 0$ sufficiently small and all $h \in B_r$,*

$$\frac{\mu_H(hB_r \triangle B_r)}{\mu_H(B_r)} = O\left(\frac{d(h,e)}{r}\right)$$

where μ_H denotes a left-invariant Haar measure on H and $d(.,.)$ denotes a right invariant metric.

Proof. Let $\mathfrak{h} = Lie(H)$. Fix a norm on $\mathfrak{h}$ inducing the metric d. Let $r > 0$ be small enough such that the exponential map is a diffeomorphism from a ball around 0 in $\mathfrak{h}$ onto B_r. Since the differential of the exponential is the identity at 0, such ball will have a radius comparable to r, denote it by $B^{\mathfrak{h}}_{r'}$

Let $g \in B_r$. Let $X, Y \in \mathfrak{h}$ be such that $h = \exp(X)$ and $g = \exp(Y)$. Then, if r is sufficiently small, by the Campell-Baker-Hausdorff formula, there exists some $Z \in \mathfrak{h}$ so that $hg = \exp(Z)$ and

$$Z - Y = X + o(\|X\|)$$

In particular, there is some $C \geq 1$ such that $hB_r \subseteq \exp(B^{\mathfrak{h}}_{r'} + CX)$. And, hence, we get that

$$hB_r \triangle B_r \subseteq \exp((B^{\mathfrak{h}}_{r'} + CX) \triangle B^{\mathfrak{h}}_{r'})$$

Let Leb denote the Lebesgue measure on $\mathfrak{h}$. It is then a standard fact from convex euclidean geometry that

$$Leb((B^{\mathfrak{h}}_{r'} + CX) \triangle B^{\mathfrak{h}}_{r'}) \ll \|X\|\, r^{dimH-1} \tag{5.12}$$

where the implicit constants are absolute and depend only on the dimension (see for example [Gro]). Here we are using that a ball in the norm on $\mathfrak{h}$ is equivalent to a standard euclidean ball of comparable radius.

Again, since the differential of the exponential is the identity at 0, the Haar measure on H near identity is comparable up to absolute constants with the pushforward of the Lebesgue measure under the exponential map.

In particular, one has $\mu_H(B_r) \ll r^{dimH}$. Combining this with (5.12) gives the desired conclusion. ☐

5.3. **Law of Large Numbers.** This section is dedicated to the proof of Proposition 5.7 below. By Proposition 5.3, this concludes the proof of Theorem 1.4.

Proposition 5.7. *Under the same hypotheses of Theorem 1.4, if $\varphi \in C_c^\infty(G/\Gamma)$, then for μ_H almost every $x \in H\Gamma/\Gamma$,*

$$\frac{1}{N}\sum_{n=1}^{N} f_n(x) \to 0$$

where f_n is defined by (5.4).

Proof. For $x \in H\Gamma/\Gamma$ and $N \in \mathbb{N}$, let $S_N(f)(x) = \sum_{n=1}^N f_n(x)$. As noted in equation (5.3), there exists a natural number $d \geq 1$, depending only on H and U, such that

$$\|Ad(g_n)|_{\mathfrak{h}}\| = \|Ad(u_{t_n})|_{\mathfrak{h}}\| = O(\left(t_n^d\right)$$

as n tends to infinity. Hence, by assumption, there exists $\lambda > 0$ such that for all $n \geq 1$, we have

$$\frac{\|Ad(g_m)|_{\mathfrak{h}}\|}{\|Ad(g_n)|_{\mathfrak{h}}\|} \ll e^{d\lambda(m-n)} \tag{5.13}$$

Then, we have

$$\begin{aligned}\int |S_N(f)(x)|^2 \, d\mu_H(x) &= \sum_{1\leq n,m\leq N} \int f_n(x) f_m(x) \, d\mu_H(x) \\ &= O\left(N^{3/2}\right) + \sum_{|n-m|\geq N^{1/2}} \int f_n(x) f_m(x) \, d\mu_H(x)\end{aligned}$$

Here we estimated the number of pairs (m, n) with $|m-n| < N^{1/2}$ using the area between the 2 lines $m \pm n = N^{1/2}$ in the square $[0, N]^2$.

But, by (5.13), when $N \gg 1$, for $n \geq m$ such that $|n-m| \geq N^{1/2}$, we have that $\|Ad(g_m)|_{\mathfrak{h}}\| \,/\, \|Ad(g_n)|_{\mathfrak{h}}\|$ will be sufficiently small so that Lemma 5.5 applies. This implies that for all $N \gg 1$:

$$\begin{aligned}\frac{1}{N^2}\int |S_N(f)(x)|^2 \, d\mu_H(x) &= O\left(N^{-1/2}\right) + O\left(e^{\frac{-d\lambda N^{1/2}}{2}}\right) \\ &= O\left(N^{-1/2}\right)\end{aligned}$$

Let $\varepsilon > 0$. Then, by the Chebyshev-Markov inequality,

$$\mu_H\left(\left\{x : \left|\frac{S_N(f)(x)}{N}\right| > \varepsilon\right\}\right) \ll \frac{N^{-1/2}}{\varepsilon^2}$$

For all $k \in \mathbb{N}$, let $N_k = k^4$. Thus, the above observation shows that the sequence $N_k^{-1/2}$ is summable. Hence, by the Borel-Cantelli lemma, we have

$$\mu_H \left(x : \left| \frac{S_{N_k}(f)(x)}{N_k} \right| > \varepsilon \text{ for infinitely many } k \right) = 0$$

Since ε was arbitrary, by taking a countable sequence ε_i decreasing to 0, we conclude that for μ_H almost every x,

$$\lim_{k\to\infty} \frac{S_{N_k}(f)(x)}{N_k} = 0$$

We are left with bootstrapping this conclusion to all sequences, for which we use a standard interpolation argument. Let $M_i \to \infty$ be a sequence. Observe that for each $M_i \in \mathbb{N}$, there exists some $k_i \in \mathbb{N}$ such that $N_{k_i} \leq M_i \leq N_{k_i+1}$.

Moreover, we have $N_{k_i+1} - N_{k_i} = O(k_i^3)$. Thus, we get that

$$\left| \frac{S_{M_i}(f)(x)}{M_i} \right| \leq \frac{N_{k_i}}{M_i} \left| \frac{S_{N_{k_i}}(f)(x)}{N_{k_i}} \right| + O(k_i^{-1}) \xrightarrow{i\to\infty} 0$$

as desired. □

6. Proof of Theorem 1.3

This section is dedicated to the proof of Theorem 1.3. Let the notation be the same as in § 5.

Proof of Theorem 1.3. Let g_n be a Ratner sequence for H. Let W denote the one-parameter unipotent subgroup generated by an Ad-unipotent element u as in the definition of Ratner sequences above.

We will apply Theorem 1.2 with $X = G/\Gamma$, $\mu = \mu_{G/\Gamma}$, $\nu = \mu_H$ and $T_n = g_n$. Let us verify the hypotheses. By assumption, we have that

$$\frac{1}{N} \sum_{n=1}^{N} (g_n)_* \nu \to \mu$$

In particular, this implies condition (1) of Theorem 1.2. Moreover, by definition of Ratner sequences, condition (2) is satisfied, with the transformation S being multiplication by the unipotent element u.

Let $\mathcal{L}$ denote the collection of proper analytic subgroups L of G such that $L \cap \Gamma$ is a lattice. Then, $\mathcal{L}$ is a countable set [Rat].

For $L \in \mathcal{L}$, define the following set

$$N(L, W) = \left\{ g \in G : g^{-1} W g \subseteq L \right\}$$

Let $\pi : G \to G/\Gamma$ denote the natural projection. The set Z appearing in the hypotheses of Theorem 1.2 will be defined to be

$$Z = \bigcup_{L\in\mathcal{L}} \pi \left(N(L, W) \right)$$

Then, since Z is a countable union of analytic subvarieties of G/Γ [Rat], Z admits a filtration by compact sets. Moreover, since $\mathcal{L}$ is countable, and $\mu_{G/\Gamma}(\pi(N(L, W))) = 0$ for all $L \in \mathcal{L}$, we have

$$\mu_{G/\Gamma}(Z) = 0$$

Finally, by Ratner's measure rigidity theorem [Rat], any ergodic W invariant probability measure $\lambda \neq \mu_{G/\Gamma}$ is supported on $N(L,W)$ (in fact supported on a single closed orbit of a conjugate of L) for some $L \in \mathcal{L}$. Thus, all the hypotheses of Theorem 1.2 are verified and hence the conclusion of Theorem 1.3 follows.

□

Appendix A. Symmetric Groups and Ratner Sequences

Throughout this section, G is a connected semisimple Lie group with finite center and H is symmetric subgroup of G. We let Γ be an irreducible lattice in G and assume the orbit $H\Gamma$ is closed in G/Γ and supports an H-invariant probability measure.

We prove a general criterion for a sequence of elements of a Lie group G to contain a Ratner sequence as a subsequence with respect to H. The precise statement is Theorem A.1 below. The proof of this theorem follows the same lines as the proof of Theorem 1.4. The only difference being Proposition A.3 below which acts as a replacement for Proposition 5.1 in the proof of Theorem 1.4. The arguments in sections 5.2 and 5.3 carry over verbatim to this setting.

A.1. **Structure of Affine Symmetric Spaces.** Our main tool will be the structure of affine symmetric spaces which we recall here. We follow the exposition in [EM] closely for the material in this section. Let $\sigma : G \to G$ be an involution such that H is the fixed point set of σ. Then, G/H is an affine symmetric space. By abuse of notation, let σ also denote the differential of σ at identity. Let $\mathfrak{g}$ denote the Lie algebra of G. Then, we have

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$$

where $\mathfrak{h}$ is the eigenspace corresponding to the eigenvalue 1 of σ and $\mathfrak{p}$ corresponds to the -1 eigenspace, and $\mathfrak{h}$ is the Lie algebra of H.

It is well known (Proposition 7.1.1, [Sch]) that one can find a Cartan involution θ of G commuting with σ. Let θ also denote its differential at identity. Then, similarly $\mathfrak{g}$ splits as

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{q}$$

where $\mathfrak{k}$ (resp. $\mathfrak{q}$) is the $+1$ (resp. -1) eigenspace of θ. Since θ is a Cartan involution, $\mathfrak{k}$ is the Lie algebra of a maximal compact subgroup, denote it by K.

Now, let $\mathfrak{a}$ be a maximal abelian subspace of $\mathfrak{p} \cap \mathfrak{q}$. Then, $\mathfrak{a}$ is the Lie algebra of a maximal abelian subgroup A and the exponential map $\mathfrak{a} \to A$ is a diffeomorphism.

Recall that G admits a decomposition of the form $G = KAH$ (See [Sch], Proposition 7.1.3 or [EM], Proposition 4.2). Elements of the fiber of the map $(k, a, h) \mapsto kah$ have the form $(kl, a, l^{-1}h)$ for some element $l \in K \cap H$. In particular, the fiber lies in a compact group.

Consider the adjoint action of $\mathfrak{a}$ on $\mathfrak{g}$. There exists a finite subset $\Sigma \subset \mathfrak{a}^*$ of non-zero elements of the dual of $\mathfrak{a}$ such that $\mathfrak{g}$ splits as

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Sigma} \mathfrak{g}_\alpha$$

such that for all $X \in \mathfrak{a}$ and all $Z \in \mathfrak{g}_\alpha$,

$$ad_X(Z) = \alpha(X)Z$$

And, for $Z \in \mathfrak{g}_0$, $ad_X(Z) = 0$. Recall that the subspaces

$$\{X \in \mathfrak{a} : \alpha(X) = 0\}$$

for $\alpha \in \Sigma$ divide $\mathfrak{a}$ into a finite collection of cones, called Weyl Chambers.

Let $\mathfrak{C}$ be one such Weyl chamber. Let Σ^+ denote the set of $\alpha \in \Sigma$ such that $\alpha(X) > 0$ for all $X \in \mathfrak{C}$. We call Σ^+ the set of positive roots relative to $\mathfrak{C}$. Then, $\mathfrak{g}$ splits as follows:

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{g}_0 \oplus \mathfrak{n}^+$$

where $\mathfrak{n}^- = \bigoplus_{\alpha \in \Sigma - \Sigma^+} \mathfrak{g}_\alpha$ and $\mathfrak{n}^+ = \bigoplus_{\alpha \in \Sigma^+} \mathfrak{g}_\alpha$.

A.2. **Unipotent Invariance.** We need to fix some notation before stating the main theorem. Fix some norm on $Lie(G)$ inducing the metric on G and denote it by $\|\cdot\|$. This norm induces a matrix norm for the Adjoint maps. Let $\|Ad(g)\|$ denote such matrix norm.

The following is the main theorem of this section.

Theorem A.1. *In the notation above, if $g_n \in G$ is a sequence tending to infinity in G/H and satisfying the following growth condition:*

(1) There exists a constant $\lambda > 0$ such that for all $n \geq 1$,

$$\|Ad(g_n)\| = O\left(e^{\lambda n}\right)$$

(2) Writing $g_n = k_n a_n h_n$, we have that $\|Ad(h_n^{-1})\|$ is uniformly bounded for all n.

Then, g_n contains a Ratner sequence for H as a subsequence.

Remark A.2.
(1) Passage to a subsequence in the conclusion of Theorem A.1 is needed to insure invariance in the limit by a *single* one-parameter unipotent subgroup. This is very important for applying Theorem 1.2 to prove Theorem 1.3.
(2) The second growth condition in Theorem A.1 makes sense, since the element h_n in the decomposition of g_n is unique up to left multiplication by elements inside the compact group $H \cap K$.
(3) The growth rate of $\|Ad(g_n)\|$ required by this theorem is not the most general one which works with our techniques. It is possible to obtain the same conclusions assuming there exist constants $\lambda, c > 0$ such that $\|Ad(g_n)\| = O\left(e^{\lambda n^c}\right)$.

A.3. **Expansion Properties of the Adjoint Action.** We shall need the following lemma regarding the Adjoint action of G. This lemma exploits the relationship between diagonalizable elements and their associated horospherical subgroups. We also make use of the structure of affine symmetric spaces.

Proposition A.3. *Let g_n be as in Theorem A.1. Then, there exists a sequence $v_n \to 0 \in Lie(H)$ satisfying the following for all n,*

$$\|v_n\| \ll \frac{1}{\|Ad(g_n)\|}$$

and such that after passing to a subsequence of the g_n's, we have

$$g_n \exp(v_n) g_n^{-1} \to u \neq id$$

where u is an Ad-unipotent element in G.

We will need the following fact for the proof of Proposition A.3.

Lemma A.4. *(Lemma 3, [Moz]) If G is semisimple over $\mathbb{R}$ with finite center, then the Adjoint representation $Ad : G \to GL(\mathfrak{g})$ is a proper map.*

We are now ready for the proof.

A.3.1. *Proof of Proposition A.3.* Write $g_n = k_n a_n h_n$. Then, by passing to a subsequence of g_n, we may assume that there exists a single Weyl chamber $\mathfrak{C}$ such that $a_n = \exp(X_n)$ and $X_n \in \mathfrak{C}$ for all n. Let Σ^+ be a set of positive roots associated with $\mathfrak{C}$.

First, we'll assume that $g_n \in KA$ and write $g_n = k_n a_n$. Note that the map $Ad : G \to GL(\mathfrak{g})$ is proper by Lemma A.4. In particular, by assumption, since $g_n \to \infty$, we have

$$\|Ad(g_n)\| \to \infty$$

We claim that the image of $\mathfrak{h}$ under the projection onto $\mathfrak{n}^+$ is non-zero. To see this, note that given $X, Y \in \mathfrak{p}$, we have that

$$\sigma(ad_X(Y)) = ad_{\sigma(X)}(\sigma(Y)) = ad_X(Y)$$

Hence, $ad_X(Y) \in \mathfrak{h}$. On the other hand, if $X \in \mathfrak{a} \subseteq \mathfrak{p}$ and $Y \in \mathfrak{g}_\alpha$ for some $\alpha \neq 0 \in \Sigma$, then $ad_X(Y) \in \mathfrak{g}_\alpha$. In particular, this implies

$$\mathfrak{n}^+ \cap \mathfrak{p} = \{0\}$$

Thus, given any $X \neq 0 \in \mathfrak{n}^+$, the element $X + \sigma(X) \neq 0$ and is σ invariant and hence belongs to $\mathfrak{h}$.

Next, note that since $\sigma(X) = -X$ for all $X \in \mathfrak{a}$, we have $\sigma(\mathfrak{n}^+) = \mathfrak{n}^-$. Thus, in particular, for any $v \in \mathfrak{n}^+$,

$$\|Ad(g_n)(v)\| \to \infty, Ad(g_n)(\sigma(v)) \to 0 \tag{A.1}$$

Let $V = \{v_\alpha \in \mathfrak{g}_\alpha : \|v_\alpha\| = 1, \alpha \in \Sigma^+\}$ be fixed. For each n, let $v_{\alpha_n} \in V$ be such that

$$\alpha_n(X_n) = \max\left\{\alpha(X_n) : \alpha \in \Sigma^+\right\}$$

where $X_n \in \mathfrak{C}$ was such that $a_n = \exp(X_n)$. Now, for each n, let

$$v_n = \frac{v_{\alpha_n} + \sigma(v_{\alpha_n})}{\|Ad(g_n)\|} \tag{A.2}$$

Then, for all n, $v_n \neq 0$ in $\mathfrak{h}$ and satisfies

$$\|v_n\| \ll \frac{1}{\|Ad(g_n)\|}$$

Moreover, by the standard identity $Ad(\exp) = \exp(ad)$ and by A.1,

$$Ad(g_n)(v_n) = \frac{e^{\alpha_n(X_n)}}{\|Ad(g_n)\|} Ad(k_n)(v_{\alpha_n}) + o(1)$$

By compactness of K, we have $\|Ad(g_n)\| \ll \|Ad(a_n)\|$. But, by our choice of α_n, $e^{\alpha_n(X_n)}$ is the largest eigenvalue of $Ad(a_n)$ and $Ad(a_n)$ is diagonalizable. Thus, $e^{\alpha_n(X_n)} / \|Ad(a_n)\| = O(1)$ and so we get

$$\|v_{\alpha_n}\| \ll \|Ad(g_n)(v_n)\| \leq \|Ad(g_n)\| \, \|v_n\| \ll \|v_{\alpha_n}\|$$

Hence, by passing to a subsequence, we get that

$$g_n \exp(v_n) g_n^{-1} = \exp(Ad(g_n)(v_n)) \to u \neq id$$

Since $v_n \to 0$, we have that $\exp(v_n) \to id$. Hence, all the eigenvalues of $Ad(\exp(v_n))$ converge to 1. Since conjugation doesn't change eigenvalues, we get that u must be an Ad-unipotent element, which finishes the proof in the case $g_n \in KA$.

For the general case, by KAH decomposition, we write $g_n = k_n a_n h_n$. Then, we can find $v_n \in \mathfrak{h}$ as above such that $Ad(k_n a_n)(\exp(v_n)) \to u \neq id$. Thus, the elements $w_n =$

$h_n^{-1} v_n h_n$ will satisfy $Ad(g_n)(\exp(w_n)) = Ad(k_n a_n)(\exp(v_n)) \to u$. By our assumption on the boundedness of $\|Ad(h_n^{-1})\|$, we get

$$\|w_n\| \leq \left\|Ad(h_n^{-1})\right\| \|v_n\| \ll \frac{1}{\|Ad(g_n)\|}$$

Hence, the sequence w_n satisfies the conclusion of the Proposition.

References

[BO] Yves Benoist and Hee Oh. Effective equidistribution of S-integral points on symmetric varieties. *Annales de l'institut Fourier* **62**(2012), 1889–1942.

[Bou] Jean Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publications Mathmatiques de l'IHS* **69**(1989), 5–41.

[CE] Jon Chaika and Alex Eskin. Every flat surface is birkhoff and oseledets generic in almost every direction. *Journal of Modern Dynamics* **9**(2015), 1–23.

[EW] Manfred Einsiedler and Thomas Ward. *Ergodic Theory*. Springer Nature, 2011.

[EM] Alex Eskin and Curt McMullen. Mixing, counting, and equidistribution in Lie groups. *Duke Math. J.* **71**(07 1993), 181–209.

[EMS] Alex Eskin, Shahar Mozes, and Nimish Shah. Unipotent Flows and Counting Lattice Points on Homogeneous Varieties. *Annals of Mathematics* **143**(1996), 253–299.

[FFT] Livio Flaminio, Giovanni Forni, and James Tanis. Effective equidistribution of twisted horocycle flows and horocycle maps. *Geometric and Functional Analysis* **26**(2016), 1359–1448.

[Gro] H. Groemer. On the symmetric difference metric for convex bodies. *Beitrge zur Algebra und Geometrie* **41**(2000), 107–114.

[KSW] Dmitry Kleinbock, Ronggang Shi, and Barak Weiss. Pointwise equidistribution with an error rate and with respect to unbounded functions. *Mathematische Annalen* **367**(2017), 857–879.

[Moz] Shahar Mozes. Mixing of all orders of Lie groups actions. *Inventiones mathematicae* **107**(1992), 235–242.

[Rat] Marina Ratner. On Raghunathan's Measure Conjecture. *Annals of Mathematics* **134**(1991), 545–607.

[SU] Peter Sarnak and Adrián Ubis. The horocycle flow at prime times. *Journal de Mathématiques Pures et Appliquées* **103**(feb 2015), 575–618.

[Sch] Henrik Schlichtkrull. *Hyperfunctions and Harmonic Analysis on Symmetric Spaces (Progress in Mathematics)*. Birkhuser, 1984.

[Sha] Nimish A. Shah. Limit distributions of polynomial trajectories on homogeneous spaces. *Duke Math. J.* **75**(09 1994), 711–732.

[Shi] Ronggang Shi. Pointwise equidistribution for one parameter diagonalizable group action on homogeneous space, 2014.

[TV] James Tanis and Pankaj Vishe. Uniform Bounds for Period Integrals and Sparse Equidistribution. *International Mathematics Research Notices* **2015**(2015), 13728.

[Ven] A. Venkatesh. Sparse equidistribution problems, period bounds and subconvexity. *Annals of Math.* **172**(2010), 989–1094.

[Zhe] Cheng Zheng. Sparse equidistribution of unipotent orbits in finite-volume quotients of $\mathrm{PSL}(2, \mathbb{R})$. *Journal of Modern Dynamics* **10**(2016), 1–21.

Department of Mathematics, The Ohio State University, Columbus, OH 43210
E-mail address: khalil.37@osu.edu