Bernoulli property for homogeneous systems

Adam Kanigowski

Abstract

Let $G$ be a semisimple Lie group with Haar measure $\mu$ and let $\Gamma$ be an irreducible lattice in $G$. For $g \in G$, we consider left translation $L_g$ acting on $(G/\Gamma, \mu)$. We show that if $L_g$ is $K$ (which is equivalent to positive entropy of $L_g$) then $L_g$ is a Bernoulli automorphism. As a corollary, we also obtain analogous results for homogeneous flows.

Contents

1 Introduction .......................................................... 1
2 Basic definitions ......................................................... 4
  2.1 Bernoulli and very weak Bernoulli properties .................. 4
  2.2 Preliminaries on Lie groups and homogeneous spaces ........... 6
  2.3 Homogeneous systems .............................................. 7
  2.4 Sobolev norms .................................................... 7
  2.5 Homogeneous parallelograms, partitions of $M$ ................. 8
  2.6 Partitions of $M$ ................................................ 8
  2.7 Outline of the proof of Theorem 2.5 ............................ 9
3 Preliminary results on homogeneous spaces .......................... 9
  3.1 Basic lemmas in Lie theory ..................................... 9
  3.2 Tubular intersections, families of cubes ....................... 11
  3.3 Equidistribution estimates ...................................... 14
4 Proof of the VWB property ........................................... 15
  4.1 Good returns for most points .................................... 15
  4.2 Proof of Proposition 4.1 ....................................... 16
  4.3 Proof of Proposition 4.2 ....................................... 18
5 Appendix: Proof of Lemma 3.10 ..................................... 23

1 Introduction

In the theory of dynamical systems, the most random systems are so called Bernoulli shifts, coming from symbolic dynamics, and determined by independent stationary processes. Smooth systems, by their nature, are not of the above symbolic form. One of the main discoveries in the theory of smooth dynamical systems in the second half of the twentieth century is their possible randomness whose strongest manifestation is expressed
by a measure-theoretic isomorphism to a system determined by a stationary independent process. Such systems are called Bernoulli. Another important notion which “measures” chaoticity of a system is the $K$ property, introduced by Kolmogorov. By [30], $K$ property is equivalent to \emph{completely positive entropy}, i.e. every non-trivial factor of the system has \textit{positive entropy}. Since factors of Bernoulli shifts are Bernoulli, [31], it follows that Bernoulli systems enjoy the $K$ property. Kolmogorov conjectured that the converse should also be true. This however was disproved by Ornstein in [32]. Moreover, the work of Katok, [18], provided counterexamples in the smooth category. On the other hand, for many classes of smooth dynamical systems $K$ property implies Bernoullicity. Indeed, Katznelson, [19], proved that ergodic toral automorphisms are Bernoulli and this result was then extended by Lind, [23], to infinite dimensional toral automorphisms. Thomas and Miles, [24], [25] and independently Aoki, [1], showed that ergodic automorphisms of compact groups are Bernoulli. Bowen, [3], established the Bernoulli property for Axiom A diffeomorphisms. Ledrappier, [22], showed that quadratic maps with absolutely continuous invariant measure enjoy the Bernoulli property.

The above principle that “natural” smooth systems which are $K$ are in fact Bernoulli was confirmed in the class of flows. Ornstein and Weiss, [34], proved that geodesic flows on surfaces of constant negative curvature are Bernoulli. This result was extended by Pesin, [35] to higher dimensional manifolds (without focal points). Ratner, [37], established the Bernoulli property for $C^2$ Anosov flows. Then Ratner, [36], and Bunimovich, [5], showed that weakly mixing suspensions over Anosov automorphisms are Bernoulli. For non-uniformly hyperbolic maps and flows (with singularities) the Bernoulli property was proved by Chernov and Haskell, [6]. Moreover, Katok in [17] constructed smooth Bernoulli systems on any smooth surface; this was later extended to all manifolds in [4].

It is not hard to notice that a fundamental class of smooth systems, namely, the class of \textit{homogeneous systems} on quotients of semisimple Lie groups is not listed above. Indeed, the equivalence of $K$ and Bernoulli properties has not been known in it. S. G. Dani conjectured in the mid 1970’s that the $K$ and Bernoulli properties are equivalent for homogeneous systems (see also (1) in [33]). In [9], [10], Dani proved a special case of this conjecture, namely, assuming additionally that the \textit{adjoint operator} of the system is diagonalizable on the central space (we explain the details in Section 2.2). Since then, there has been no progress on the general case of Dani’s conjecture although it often reappeared in the literature, e.g. Conjecture 3.2. in [27], Section 2.3. b in [20], Problem 37 in [38], Question 11.11 in [13]. The main aim of this paper is to prove Dani’s conjecture in full generality. More precisely, we have the following result:

\textbf{Theorem 1.1.} Let $G$ be a semisimple Lie group, $\Gamma$ an irreducible lattice in $G$ and let $L_g$ be a left translation by $g \in G$ on $(G/\Gamma, \mu)$ that enjoys the $K$ property. Then $L_g$ is a Bernoulli automorphism.

Obviously, a necessary condition for homogeneous systems on irreducible quotients to be $K$ is \textit{positive entropy}. In [8], Dani showed that this necessary condition is also sufficient for the $K$ property. Therefore, \textbf{Theorem 1.1} has the following consequence on the Bernoulli property for homogeneous flows.

\textbf{Corollary 1.2.} Let $G$ be a semisimple Lie group, $\Gamma$ an irreducible lattice in $G$ and let $(\Psi_t)$ be a homogeneous flow on $(G/\Gamma, \mu)$ with positive entropy. Then $(\Psi_t)$ is a Bernoulli flow.
Theorem 1.1 and Corollary 1.2 by Ornstein’s theory [29] have the following consequence:

**Corollary 1.3.** Entropy is a full invariant of isomorphism in the class of positive entropy homogeneous systems on irreducible quotients of semisimple Lie groups.

We also have the following corollary for general homogeneous systems (without the irreducibility assumption on the lattice):

**Corollary 1.4.** Let \( L_g \) be a weakly mixing translation on \((G/\Gamma, \mu)\). Then either \( L_g \) is Bernoulli or it has a zero entropy homogeneous factor.

Indeed, notice that \((L_g, G/\Gamma, \mu)\) is a finite extension of the system 
\((L_{g_1} \times \ldots \times L_{g_k}, \prod_{i=1}^k G_i/\Gamma_i, \prod_{i=1}^k \mu_i)\), where \( G_i \) are semisimple Lie groups, \( \Gamma_i \) are irreducible in \( G_i \) and \( \mu_i \) is the Haar measure on \( G_i/\Gamma_i \). By Rudolph’s result, [40], it follows that if \( L_g \) is weakly mixing then \( L_g \) is Bernoulli if and only if all \( L_{g_i} \) are Bernoulli. This by Corollary 1.3 is equivalent to positive entropy of every \( L_{g_i} \).

Finally it follows from our result that a strong form of Pinsker’s conjecture [4] is true for homogeneous systems (although it is not true in general, [30]):

**Corollary 1.5.** Let \( L_g \) be a translation on \((G/\Gamma, \mu)\). Then \( L_g \) is isomorphic to a product of a Bernoulli system and a system of zero entropy.

Indeed, by [12] it follows that finite extension of a system which is a product of Bernoulli and zero entropy is also of that form. It remains to notice that by Corollary 1.3 for every \( i \in \{1, \ldots k\} \), every \( L_{g_i} \) is either Bernoulli or of zero entropy.

Let us give the simplest example for which Theorem 1.1 applies, while [9], [10] do not.

**Example 1.6.** Let 
\[ X = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \]
be the generators of, respectively the geodesic and the horocycle flow. Let \( L(x, y) = [\exp(X) \times \exp(U)](x, y) \Gamma \), where \( \Gamma \) is an irreducible lattice \( SL(2, \mathbb{R})^2 \). Then \( L \) on \( SL(2, \mathbb{R})^2/\Gamma \) is a \( K \) system by [8]. However, it is not diagonalizable on \( g^0 = sl(2, \mathbb{R}) \), so we cannot apply [9], [10]. On the other hand Theorem 1.1 shows that \( L \) is a Bernoulli automorphism. One can also consider the flow \( \Psi_t(x, y) = [\exp(tX) \times \exp(tU)](x, y) \Gamma \) and deduce the same result.

Theorem 1.1 yields examples of partially hyperbolic, Bernoulli systems for which the center direction is not an isometry (in fact, it may have polynomial growth). The only known such examples until now were ergodic toral automorphisms. Let us briefly describe the novelty of our approach compared to the ones used by [34], [36] and [9], [10] (in all these results, it was crucially used that the center direction is an isometry). One of the main tools in proving the Bernoulli property is the geometric method developed in [34]. The main difficulty in this method is to find a good matching between nearby pieces of local unstable manifolds. If the considered system is an isometry on the center space (recall that it contracts the stable space), then a good matching \( \theta \) is given by the central stable holonomy. It is crucial that the orbits of points \( x \) and \( \theta x \) stay together for all times. However, if the

---

1Pinsker’s conjecture states that every automorphism is isomorphic to a Cartesian product of a \( K \)-system and a system of entropy 0.
center is not an isometry, then the center-stable holonomy is not a good matching anymore, since points $x$ and $θx$ have different center components and therefore will split after some time. Our method is therefore different from the previously used methods. We use strong equidistribution properties of the unstable foliation (horospherical subgroup) and exact bounds on the growth on the center, to create a matching between local unstable manifolds. Finally, we emphasize that our proof is geometric (it uses effective equidistribution of the unstable foliation and controlled growth on the center). Therefore it has the potential of being generalized to other partially hyperbolic systems with non-isometric center.

Outline of the paper: In Section 2 we introduce some basic notation and definitions. In Section 3 we state some preliminary results on homogeneous spaces. In Section 3.3 we state a Lemma on equidistribution of the unstable foliation. This Lemma is proven in the appendix. In Section 4 we use results from Sections 3 to show the very weak Bernoulli property. One of the main results is Proposition 4.2 in which we construct a map between local unstable manifolds. We provide a general outline of the proof of Theorem 1.1 at the end of Section 2 and an outline of the proof of Proposition 4.2 at the beginning of Section 4.3.

2 Basic definitions

We will be always dealing with measure preserving actions of $R$ (flows) and $Z$ (automorphisms) on standard probability Borel spaces.

2.1 Bernoulli and very weak Bernoulli properties.

Let $A$ be a finite set and let $p = (p_i)_{i=1}^{|A|}$ be a probability vector, i.e. $p_i \geq 0$ for $i \in \{1, \ldots, |A|\}$ and $\sum_{i=1}^{|A|} p_i = 1$. A Bernoulli shift is a transformation $σ$ on $A^Z$ given by $σ((x_i)_{i∈Z}) = (x_{i+1})_{i∈Z}$. Notice that $σ$ preserves the measure $p^Z$.

An automorphism $T : (X, B, µ) → (X, B, µ)$ is called a Bernoulli system (or simply Bernoulli) if it is isomorphic to some Bernoulli shift (i.e. for some $A$ and some probability vector $p$.) We say that a flow $(T_t)$ on $(X, B, µ)$ is a Bernoulli flow (or simply Bernoulli) if for every $t_0 \in R \setminus \{0\}$ the automorphism $T_{t_0}$ is Bernoulli. We will use the notion of very weak Bernoullicity which we now define. First we need to recall some basic notation.

Let $P = (P_1, \ldots, P_k)$ and $Q = (Q_1, \ldots, Q_l)$ be two finite partitions of $(X, µ)$. Then $P ∨ Q$ is the smallest common refinement of $P$ and $Q$, i.e. the partition into sets of the form $P_i \cap Q_j$. For $A \subset X$, $P|_A$ denotes the induced partition of the space $(A, µ|_A)$, i.e. $P|_A := (P_1 \cap A, \ldots, P_k \cap A)$ (here $µ|_A(B) = µ(A ∩ B) / µ(A)$). Moreover if $k = l$, we can introduce the following distance on the space of partitions of $(X, µ)$:

$$\tilde{d}(P, Q) := \sum_{i=1}^k µ(P_i ∆ Q_i).$$

Now let $P^s = (P_1^s, \ldots, P_k^s)$, $s = 1, \ldots, S$ be a sequence of finite partitions of $(X, µ)$ and $Q^s = (Q_1^s, \ldots, Q_k^s)$, $s = 1, \ldots, S$ be a sequence of finite partitions of $(Y, ν)$. If additionally $(X, µ) = (Y, ν)$, then

$$\tilde{d}((P^s)_{s=1}^S, (Q^s)_{s=1}^S) := \frac{1}{S} \sum_{s=1}^S \tilde{d}(P^s, Q^s).$$
More generally, if \((P^s)_{s=1}^S\) and \((Q^s)_{s=1}^S\) are partitions of different spaces, we say that \(P^s \sim Q^s\) for \(s = 1, \ldots, S\) if \(\mu(P^s_i) = \nu(Q^s_i)\) for \(i = 1, \ldots, k\) and \(s = 1, \ldots, S\). We can then compare the distance between \((P^s)_{s=1}^S\) and \((Q^s)_{s=1}^S\) by setting
\[
\tilde{d}((P^s)_{s=1}^S, (Q^s)_{s=1}^S) = \inf_{\bar{Q}^s \sim Q^s, s=1,..,S} \tilde{d}((P^s)_{s=1}^S, (\bar{Q}^s)_{s=1}^S),
\]
where the infimum is taken over sequences of partitions \(\bar{Q}^s\) of \((X, \mu)\). We say that a property holds for \(\varepsilon\) a.e. atom of a partition \(\mathcal{P}\) if it holds for all atoms except a set of atoms whose union has measure less than \(\varepsilon\). We denote by \(T^n\mathcal{P}\) the partition given by \((T^n P_1, \ldots, T^n P_k)\).

**Definition 2.1** (very weak Bernoulli). Let \(T \in \text{Aut}(X, \mathcal{B}, \mu)\) and let \(\mathcal{P}\) be a finite partition of \(X\). Then \(\mathcal{P}\) is a very weak Bernoulli partition (VWB partition) if for every \(\varepsilon > 0\) there exists \(N_0 \in \mathbb{N}\) such that for every \(N' \geq N \geq N_0\) every \(S \geq 0\) and \(\varepsilon\) a.e. atom of \(\bigvee_{N}^{N'} T^i(\mathcal{P})\), we have
\[
\tilde{d}((T^{-i}\mathcal{P})_{s=0}^S, (T^{-i}\mathcal{P}|_{A})_{s=0}^S) < \varepsilon.
\]

The following classical theorem is a crucial tool in establishing Bernoullicity of a system (see e.g. \cite{[34]}). Recall that a sequence of partitions \((\mathcal{P}_k)_{k=1}^{+\infty}\) of \((X, \mathcal{B}, \mu)\) converges to partition into points if the smallest \(\sigma\)-algebra with respect to which all \(\mathcal{P}_k\) are measurable, is \(\mathcal{B}\).

**Theorem 2.2.** If \((\mathcal{P}_k)_{k=1}^{+\infty}\) is a sequence of partitions of \((X, \mathcal{B}, \mu)\) converging to partition into points and, for every \(k \geq 1\), \(\mathcal{P}_k\) is VWB partition for \(T \in \text{Aut}(X, \mathcal{B}, \mu)\), then \(T\) is a Bernoulli system.

We will now recall the main method of establishing VWB property, \cite{[34]}. For a partition \(\mathcal{P} = (P_1, \ldots, P_k)\) of \((X, \mu)\), an integer \(S \geq 1\) and \(x \in X\) the \(S, \mathcal{P}\)-name of \(x\) is sequence \((x^\mathcal{P}_i)_{i=0}^S \in \{1, \ldots, k\}^{S+1}\) given by the condition \(T^i(x) \in P^\mathcal{P}_x\). Let \(e : \mathbb{Z} \to \mathbb{Z}\) be given by \(e(0)=0\) and \(e(n) = 1\) for \(n \neq 0\). A map \(\theta : (X, \mu) \to (Y, \nu)\) is called \(\varepsilon\)-measure preserving if there exists a set \(E' \subset X\), \(\mu(E') < \varepsilon\) and such that for every \(A \in X \setminus E'\), we have
\[
\left| \frac{\nu(\theta(A))}{\mu(A)} - 1 \right| < \varepsilon.
\]

We have the following lemma:

**Lemma 2.3** (Lemma 1.3. in \cite{[34]}). Let \(T \in \text{Aut}(X, \mathcal{B}, \mu)\) and \(\mathcal{P}\) be a finite partition of \(X\). If for every \(\varepsilon > 0\) there exists \(N \in \mathbb{N}\) such that for every \(N' \geq N\), \(\varepsilon\) a.e. atom \(A \in \bigvee_{N}^{N'} T^i(\mathcal{P})\) and every \(S \geq 1\) there exists an \(\varepsilon\)-measure preserving map \(\theta = \theta(N, S, A) : (A, \mu|_A) \to (X, \mu)\) such that
\[
d(x, \theta(x)) := \frac{1}{S} \sum_{i=0}^{S-1} e\left( x^P_i - (\theta(x))^\mathcal{P}_i \right) < \varepsilon.
\]
then \(\mathcal{P}\) is a VWB partition.

Finally we recall the definition of Kolmogorov property (K property for short).
Definition 2.4. We say that $T \in \text{Aut}(X, B, \mu)$ has the $K$ property if there exists a generating partition $\mathcal{P}$ such that for every $B \in \mathcal{B}$ and every $\varepsilon > 0$ there exists $N_0 = N_0(\varepsilon, B)$ such that for every $N' \geq N \geq N_0$, $\varepsilon$ a.e. atom $A \in \bigvee_{N'} T^\mathcal{P}$ satisfies

$$|\mu|_A(B) - \mu(B)| < \varepsilon.$$ 

By [39] it follows that the $K$ property is equivalent to completely positive entropy: every factor of $T \in \text{Aut}(X, B, \mu)$ has positive entropy.

2.2 Preliminaries on Lie groups and homogeneous spaces

Let $G$ be a semisimple Lie group with Haar measure $\mu$, Lie algebra $\mathfrak{g}$ and denote the right invariant metric on $G$ by $d_G$. Let $\exp : \mathfrak{g} \to G$ be the exponential map with the local inverse (around $e$) given by $\log_{\text{alg}} : G \to \mathfrak{g}$. Let $[\cdot, \cdot] : \mathfrak{g} \to \mathfrak{g}$ denote the Lie bracket. For $g \in G$, let $C(g) : G \to G$ be given by $C(g)(h) := ghg^{-1}$ and let $\text{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ be the adjoint operator, i.e. derivative of $C(g)$ at $e \in G$. Let $\mathfrak{g}_C$ denote the complexification of $\mathfrak{g}$ and, for $\lambda \in \mathbb{C}$, let

$$V_\lambda := \{ U \in \mathfrak{g}_C : \exists j \in \mathbb{N} \ (\text{Ad}(g) - \lambda I)^j(U) = 0\}$$

be the generalized eigenspace of $\lambda$. Then $[V_\lambda, V_\mu] \subset V_{\lambda + \mu}$. Let

$$\mathfrak{g}_C^+ := \bigoplus_{|\lambda| > 1} V_\lambda, \quad \mathfrak{g}_C^0 := \bigoplus_{|\lambda| = 1} V_\lambda, \quad \mathfrak{g}_C^- := \bigoplus_{|\lambda| < 1} V_\lambda.$$

and let respectively $\mathfrak{g}^+, \mathfrak{g}^0, \mathfrak{g}^- \subset \mathfrak{g}$ be the real spaces corresponding to $\mathfrak{g}_C^+, \mathfrak{g}_C^0, \mathfrak{g}_C^-$. We can then decompose $\mathfrak{g}$:

$$\mathfrak{g} := \mathfrak{g}^+ \oplus \mathfrak{g}^0 \oplus \mathfrak{g}^-.$$  \hspace{1cm} (1)

It follows that for $i \in \{ +, 0, - \}$, $\mathfrak{g}^i$ is a subalgebra and hence $G^i := \exp(\mathfrak{g}^i)$ is a subgroup of $G$. It follows that $G^-$ is the horospherical subgroup for $L_g$, i.e.

$$G^- = \{ h \in G : g^i h g^{-i} \to 0, \text{ as } l \to +\infty\}.$$ 

Analogously, $G^+$ is the horospherical subgroup for $L_{-g}$. Moreover, $\text{Ad}(g)$ restricted to $\mathfrak{g}^0$ is called quasi-unipotent. Recall also that $\mathfrak{g}^+ \neq \{0\}$ iff $\mathfrak{g}^- \neq \{0\}$. From now on we fix $g \in G$ for which we assume that $\mathfrak{g}^+ \neq \{0\}$ (equivalently $\mathfrak{g}^- \neq \{0\}$).

For $i \in \{ +, 0, - \}$, let $\mathbf{V}^i = \{ V_j^i \}_{j=1}^{\dim \mathfrak{g}^i}$ be a basis of $\mathfrak{g}^i$. Then $\mathbf{V} := \mathbf{V}^+ \cup \mathbf{V}^0 \cup \mathbf{V}^-$ is a basis of $\mathfrak{g}$. For $i \in \{ +, 0, - \}$ and a vector $a = (a_j) \in \mathbb{R}^{\dim \mathfrak{g}^i}$, we denote $a \mathbf{V} := \sum_{j=1}^{\dim \mathfrak{g}^i} a_j V_j^i$. Then we define the norm on $\mathfrak{g}^i$ by setting

$$\|U\|_i = \max\{|a_j| : j \in \{1, \ldots, \dim \mathfrak{g}^i\}\},$$

where $U = a \mathbf{V}$.

For $i \in \{ +, 0, - \}$, $\text{Leb}^i$ denotes the Lebesgue measure on $\mathfrak{g}^i$. Moreover, $\text{Leb}$ denotes the Lebesgue measure on $\mathfrak{g}$.

\hspace{1cm} \footnote{This means that $\bigvee_{-\infty}^+ T^\mathcal{P} = B$.}
2.3 Homogeneous systems

Let \( \Gamma \subset G \) be a lattice in \( G \). We define the homogeneous space, \( M := G/\Gamma \). Then \( M \) is equipped with the measure given locally by \( \mu \) (we denote the measure on \( M \) also by \( \mu \)). For \( g \in G \), the left translation on \( M \) is given by

\[
\Psi(x\Gamma) = L_g(x\Gamma) := (gx)\Gamma.
\]

It follows that \( \Psi \) preserves \( \mu \). The right invariant metric on \( G \) induces the following metric on the homogeneous space:

\[
d_M(x, y) := \inf_{\gamma \in \Gamma} d_G(x, y\gamma).
\]

Our main theorem can be rephrased in the following way:

**Theorem 2.5.** Let \( g \in G \) be such that \( g^\perp \neq \{0\} \). If \( \Gamma \) is irreducible, then \( \Psi \) on \((M, \mu)\) is Bernoulli.

By Dani’s theorem, [9, 10] it follows that if \( g^\perp \neq \{0\} \) (which is equivalent to \( h_\mu(\Psi) > 0 \)) and \( Ad(g)|_{g^0} : g^0 \to g^0 \) is diagonalizable (all generalized eigenspaces are one-dimensional), then \( \Psi \) is Bernoulli. Hence our theorem addresses the case in which \( Ad(g)|_{g^0} : g^0 \to g^0 \) has non-trivial Jordan blocks.

2.4 Sobolev norms

We now briefly recall some basic facts on Sobolev norms. For more details see eg. [11] or [2]. Let \( C_c(M) \) denote the space of compactly supported functions on \( M \) with the topology of uniform convergence on compact sets. Let \( C_c^\infty(M) \subset C_c(M) \) denote the set of infinitely differentiable functions on \( M \). Recall that for \( Y \in g \) and \( \phi \in C_c^\infty(M) \) the Lie derivative in direction \( Y \) is given by

\[
\mathcal{L}_Y(\phi)(x) := \lim_{t \to 0} \frac{\phi(\exp(tY)x) - \phi(x)}{t}.
\]

Moreover, for a fixed basis \( V_1, \ldots, V_n \) of \( g \) and a multiindex \((m_1, \ldots, m_n)\), we define the operator of degree \( \deg W = \sum_{i=1}^n m_i \), by setting

\[
\mathcal{L}_W := \mathcal{L}^{m_1}U_1 \ldots \mathcal{L}^{m_n}U_n.
\]

Let \( r(x) \) denote the injectivity radius of \( x = g\Gamma \), i.e. the largest \( r > 0 \) such that the quotient map \( \pi : G \to M \) restricted to a closed \( d_G \) ball of radius \( r \) around \( g \) is injective. Then the Sobolev norms of order \( d \) are defined by

\[
S_d(\phi) := \left( \sum_{\deg W \leq d} \left\| r(\cdot)^d \mathcal{L}_W(\phi)(\cdot) \right\|_{L^2(M)}^2 \right)^{\frac{1}{2}}.
\]

Sobolev norms (of order \( d \)) measure the \( L^2 \) growth of derivatives of \( \phi \) up to level \( d \).

We recall the following lemma which establishes quantitative mixing for the action of \( G \) on \( M \) (recall that we assume that \( G \) is semisimple) and is based on the decay of matrix coefficients:

**Lemma 2.6.** There exist \( C, \zeta, d > 0 \) such that for any \( \phi, \psi \in C_c^\infty(M) \), we have

\[
\left| \int_M \phi(gx)\psi(x)d\mu - \int_M \phi(x)d\mu \int_M \psi(x)d\mu \right| \leq C e^{-\zeta d_G(g,e)} S_d(\phi)S_d(\psi).
\]

This results are based on work of several authors, Cowling [7], Howe [13], Moore [26], and Oh [28].
2.5 Homogeneous parallelograms, partitions of $M$

We now define a special class of sets that will be used in the proof of Theorem 2.5. Fix $g \in G$ and recall that for $i \in \{+,0,-\}$, $V^i$ denotes the basis of $g^i$. (see [1]). For $h \in \{g,g^+,g^0,g^-\}$ we define the $\delta$ cube in $h$ by setting

$$C(\delta,h) := \{V \in h : \|V\| \leq \delta\}.$$  \hfill (2)

Let $H \subset G$ be the Lie group of $h$. Any cube in $h$ defines a cube in $H$ by setting

$$C(\delta,H) := \{\{V \in h : \|V\| \leq \delta\}.$$  \hfill (3)

For $y \in M$, the $\xi$-local unstable space is defined by

$$W^u(\xi,y) := C(\xi,G^+)y \subset M.$$  \hfill (3)

An $HC$-set of size $\eta > 0$ around $x \in M$ is the set

$$D(\eta,x) := [C(\eta,G^+) \times C(\eta,G^0 \times G^-)]x \subset M.$$  \hfill (4)

Let $\eta \in (0,1)$. We say that $E \subset M$ intersects an $HC$-set $D(\eta,x)$ in an $g^+$-tubular subset if $y \in E \cap D$ implies that $W^u(1,y) \cap D \subset E \cap D$.

2.6 Partitions of $M$

Fix $\varepsilon > 0$ and let $K_\varepsilon \subset M$ be a subset such that $\mu(K_\varepsilon) \geq 1 - \varepsilon$. Let $r_\varepsilon = \min(\varepsilon, \text{sys}(K_\varepsilon)/10)$ (see Lemma 3.7). Consider the cover $K_\varepsilon \subset \bigcup_{x \in K_\varepsilon} B(x, r_\varepsilon)$ and let

$$K_\varepsilon \subset \bigcup_{j=1}^{N(\varepsilon)} B(x_j, r_\varepsilon)$$

be a finite subcover. We then make the sets $\{B(x_j, r_\varepsilon)\}_{j=1}^{N(\varepsilon)}$ disjoint by defining $P_j := B(x_j, r_\varepsilon) \setminus \bigcup_{t \subset j} B(x_t, r_\varepsilon)$. We call the new (disjoint) sets $P_1, ..., P_{m(\varepsilon)}$ moreover we set $P_0 := M \setminus \bigcup_{j=1}^{N(\varepsilon)} B(x_j, r_\varepsilon)$. We will be always considering partitions of the form $\{P_j\}_{j=0}^{m(\varepsilon)}$.

Notice that $\partial B(x_j, r_\varepsilon)$ is smooth (notice that by the definition on $r_\varepsilon$ it follows that the map $\pi : \pi^{-1}(B(x_j, r_\varepsilon)) \rightarrow M$ is injective). Since the sets $\{P_j\}_{j=0}^{m(\varepsilon)}$ are obtained from the sets $\{B(x_j, r_\varepsilon)\}_{j=1}^{N(\varepsilon)}$ by unions, intersections and differences (and there is finitely many of them), it follows that for every $\varepsilon' > 0$ there exists $\varepsilon'' > 0$ such that

$$\mu(V^{\varepsilon''}(\partial P)) \leq \varepsilon'/20,$$

where $\partial P = \bigcup_{j=0}^{m(\varepsilon)} \partial P_j$, and $V^{\varepsilon''}(\partial P)$ denotes the $\varepsilon''$ neighborhood if $\partial P$.

Notice that if $\varepsilon_n = \frac{1}{n}$ then the sequence $\{P_j\}_{j=1}^{m(\frac{1}{n})}$ for $n \in \mathbb{N}$ converges to partition into points and hence it is enough to establish the VWB property along $\{P_j\}_{j=1}^{m(\frac{1}{n})}$, $n \in \mathbb{N}$.

From now on we always assume that we have fixed an $\varepsilon_0 = \frac{1}{n_0}$ and $P$ always denotes the partition $\{P_j\}_{j=1}^{m(\frac{1}{n_0})}$.

To make the paper more readable, we will now present a general strategy of the proof of Theorem 2.5. 

Footnote 4: Acronym for homogeneous cube.
2.7 Outline of the proof of Theorem 2.5

We outline some standard reductions introduced in [34] in the problem of VWB-property for $K$-automorphisms. By Lemma 2.3 one needs to show existence of an $\varepsilon$-measure preserving map $\theta : (A, \mu|_A) \to (X, \mu)$ for $\varepsilon$ a.e. atom $A \in \bigvee_N \Psi^i(\mathcal{P})$ such that the orbits of $x$ and $\theta(x)$ are in one atom of $\mathcal{P}$ for most times. The first step is to use the $K$ property of $\Psi$ (see Definition 2.4) to reduce the general problem to a local one, i.e. it is enough to find an $\varepsilon$-measure preserving map $\bar{\theta}_D : (A \cap D, \mu|_{A \cap D}) \to (D, \mu|_D)$, where $D$ belongs to a family of small cubes in $M$ (smallness depending on $\varepsilon$) whose union has large measure and then use the $K$ property to extend the family of obtained $\varepsilon$ preserving maps to a global map $\theta$ (see Proposition 4.1). To construct the local map $\bar{\theta}_D$ one uses Lemma 3.8 to say that $\varepsilon$ a.e. atom $A$ intersects $D$ in an $g^+$ tubular subset. This by the local product structure (see the proof of Proposition 4.1) reduces the local problem to finding an $\varepsilon$-measure preserving map $\tilde{\theta} : W^u(\delta, z) \to W^u(\delta, z')$ between pieces of two local unstable manifolds in $D$ (with conditional measures $\mu^u_z$ and $\mu^u_{z'}$).

Finding the map $\tilde{\theta}$ is the most difficult and important part in proving VWB-property (it is called the Main Lemma in [34]). This is done in Proposition 4.2 (for most unstable pieces, for which we have good control of ergodic properties). This is the part on which most of the proof is devoted. The main obstruction in finding $\tilde{\theta}$ is non-trivial (polynomial) orbit growth on the center space (otherwise, if the center is isometry, one can take for $\tilde{\theta}$ the center stable holonomy). The idea is to use quantitative equidistribution of the local unstable space. At time $N$ we divide the space $M$ into a family of cubes of diameter $\frac{1}{N^r}$ (where $r$ is the exponent of the growth on the center space). If we can guarantee that unstable pieces (of most points) equidistribute at time $N$ for all the cubes of order $\frac{1}{N^r}$ one can construct the mapping $\bar{\theta}$ by mapping subsets of $W^u(\delta, z)$ that belong to the $i$-th cube at time $\varepsilon^3 N$ to subsets of $W^u(\delta, z')$ that belong to the $i$-th cube (using the holonomy inside each cube). Quantitative equidistribution estimates imply that such map is $\varepsilon$-measure preserving (for a more detailed outline of the construction of $\tilde{\theta}$ see the outline of the proof of Proposition 4.2 in Section 4.3). The above strategy shows, that one needs quantitative equidistribution for families of cubes. We state such result in Section 3.3 (see Lemma 3.10). The proof of Lemma 3.10 based on Proposition 2.4.8. from [21], is given in the appendix. Then in Section 4 we use the above outline to reduce the problem to finding $\varepsilon$-measure preserving map between pieces of unstable manifolds (see Proposition 4.2).

3 Preliminary results on homogeneous spaces

In this section we will recall some basic results from Lie groups, Lie algebras and homogeneous spaces.

3.1 Basic lemmas in Lie theory

We first state some classic results on norms of vectors in Lie algebra. In the lemma below $G$ is a Lie group with Haar measure $\mu$ and $\mathfrak{g}$ is the Lie algebra of $G$. Let $\text{Leb}$ denote the Lebesgue measure on $\mathfrak{g}$.

**Lemma 3.1.** The following hold:
(i) Let $g = e_1 \oplus \ldots \oplus e_k$ be any decomposition of $g$ into a sum of vector subspaces. If for every $g \in G$, $dG(g, e)$ is sufficiently small, then $g = \exp(E_1) \cdots \exp(E_k)$, with $E_i \in e_i$ for every $i \in \{1, \ldots, k\}$. Moreover such decomposition is unique.

(ii) there exist $\kappa_1, \kappa_2 > 0$ such that $\|(U_1, \ldots, U_n)\| < \kappa_2$, then

$$\left\| \log_{\text{alg}} (\exp(U_1) \cdots \exp(U_n)) - \sum_{i=1}^n U_i \right\| \leq \kappa_1 \max_{i,j \in \{1, \ldots, n\}} \|U_i\| \|U_j\|.$$  

(iii) let $B \subset g$. For every $\varepsilon' > 0$ there exists $\varepsilon'' > 0$ such that if $\sup_{V \in B} \|V\| < \varepsilon''$ then

$$(1 - \varepsilon') \text{Leb}(B) \leq \mu(\exp(B)) \leq (1 + \varepsilon') \text{Leb}(B).$$

(iv) for every $\varepsilon' > 0$ there exists $\varepsilon'' > 0$ such that for every $0 < \tau < \varepsilon''$, we have

$$C((1 - \varepsilon')\tau, g) \subset C(\tau, g^+) \times C(\tau, g^0 \oplus g^-) \subset C((1 + \varepsilon')\tau, g).$$

Proof. Properties (i) and (iii) are a consequence of the fact that $\exp: g \to G$ is a smooth function which satisfies $\exp'(0) = Id$, for a more detailed argument see e.g. [10], Lemma 3.7. Property (ii) is a consequence of the Taylor formula. Finally (iv) is a straightforward consequence of (ii).

Remark 3.2. We will be constructing $\varepsilon$- measure preserving maps between subsets of $M$ (with small diameter). Property (iii) in the above lemma gives an important tool in doing this (on the local level): it is enough to construct an $\varepsilon$-measure preserving map on the Lie algebra level (which is a vector subspace), and then send it by $\exp$ to $M$, where the above lemma implies that the new map will be $2\varepsilon$-measure preserving.

We now define the holonomy maps, which play an important role in the proof of theorem. Let $\text{hol}_s: g^+ \oplus [g^0 \oplus g^-] \to g^+$, $\text{hol}_{uc}: g^+ \oplus [g^0 \oplus g^-] \to g^0 \oplus g^-$ be given by

$$\exp(E) \exp(F) = \exp(\text{hol}_{uc}(E, F)) \exp(\text{hol}_s(E, F)).$$

(6)

The following result is an immediate consequence of (ii) in Lemma 3.1

Lemma 3.3. There exist $\kappa_1, \kappa_2 > 0$ such that for $(E, F) \in g^+ \oplus [g^0 \oplus g^-]$, with $\|(E, F)\| \leq \kappa_1$ the functions $\text{hol}_s(E)$ and $\text{hol}_{uc}(F)$ are well defined. Moreover

$$\max (\|\text{hol}_s(E, F) - E\|, \|\text{hol}_{uc}(E, F) - F\|) \leq \kappa_2 \|E\| \|F\|.$$  

Proof. It is enough to use (ii) in Lemma 3.1 for $(U_1, U_2) = (E, F)$ and then $(U_1, U_2) = (\text{hol}_s(E, F), \text{hol}_{uc}(E, F)).$  

Remark 3.4. The function $\text{hol}_s: g^+ \to g^+$ is called the stable holonomy and plays an important role in proving VWB Bernoulli property. The above lemma implies that if we consider the function $h_s$ on a small cube, then the image, up to a subset of small measure, remains inside the cube.

For a linear operator $\Psi$ on a vector space $(V, \|\cdot\|)$, we denote $\|\Psi\| = \sup_{v \in V \setminus \{0\}} \|\Psi(v)\| / \|v\|$. The following lemma describes the behavior of the action of $Ad(g): g \to g$. 

Lemma 3.5. There exist $c_0 > 0$, $\lambda_0 > 1$, $r' > 0$ such that

A. for every $t < 0$, $||\text{Ad}(g)||_\mathfrak{g}^+ \leq c_0 \lambda_0^t$;

B. for every $t > 0$, $||\text{Ad}(g)||_\mathfrak{g}^- \leq c_0 \lambda_0^{-t}$;

C. for every $t \in \mathbb{R}$, $||\text{Ad}(g)||_\mathfrak{g}^0 \leq c_0 |t|^{r'}$.

Proof. Notice that A. follows from the fact that for $\text{Ad}(g)$, we have $|\lambda| > 1$ on $\mathfrak{g}^+$ and analogously B. follows since $|\lambda| < 1$ on $\mathfrak{g}^-$. So we only need to show C. This however follows from the fact that the growth of the $\text{Ad}(g)$ on $\mathfrak{g}^0$ is at most polynomial, since $|\lambda| = 1$ on $\mathfrak{g}^0$.

We have the following immediate Lemma:

Lemma 3.6. There exists $r_0 \in \mathbb{N}$ such that for every $\varepsilon > 0$ there exists $N_\varepsilon > 0$ such that for every $x \in M$ and every $y \in C(N^{-r_0}, G^0 \times G^-) x \in M$ with $N \geq N_\varepsilon$, we have

$$d_M(\Psi^n x, \Psi^n y) < \varepsilon$$

for every $n \in [0, N]$.

Proof. By right invariance, we have

$$d_G(\Psi^n x, \Psi^n y) = d_G(e, g^n y x^{-1} g^{-n}).$$

Then the statement follows from B. and C. in Lemma 3.5: we have that the growth of $||\text{Ad}^n(g)||$ is at most polynomial in $n$ (actually it is bounded on $\mathfrak{g}^+$ and at most polynomial on $\mathfrak{g}^0$). Therefore, there exists $r_0 \in \mathbb{N}$ (for instance $r_0 = r' + 1$, where $r'$ comes from C.) and, for every $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that if $N \geq N_\varepsilon$, then we have

$$d_G(e, g^n y x^{-1} g^{-n}) \leq \frac{C}{N}.$$

This finishes the proof.

Finally, we recall a classical result which uses only the fact that $\Gamma$ is a discrete subgroup in $G$.

Lemma 3.7. For every compact set $K \subset M$, we have

$$\text{sys}(K) := \inf_{x \in K} \inf_{\gamma \in \Gamma \setminus \{e\}} d_G(x \gamma x^{-1}, e) > 0.$$
Proof. The proof follows the scheme of the proof of Lemma 2.1. in [34] and uses (exponential) contraction of the local unstable space for negative times. More precisely, define $G_k$ by setting $x \in G_k$ if there exists an atom $\Psi^k A \in \Psi^k P$ such that $x \in D \cap \Psi^k A$ but $W^u(\xi, x) \cap D$ is not a subset of $D \cap \Psi^k A$. Notice that this is equivalent, to $\Psi^{-k}(D \cap W^u(\xi, x))$ being $d_k$ close to the boundary of $A$, where, by A. in Lemma 3.5 and 3, 

$$d_k \leq c_0 \lambda_0^{-k}.$$ 

Indeed, for $\exp(B)x \in W^u(\xi, x)$, $B \in C(\xi, G^+)$, we have 

$$d_M(\Psi^{-k}x, \Psi^{-k}\exp(B)x) \leq d_G(\Psi^{-k}\exp(B)\Psi^k, e)$$ 

and we use A. Using this, the fact that $\mu$ is $\Psi$ and also using that $A$ has smooth boundaries it follows that 

$$\mu(G_k) \leq c'_0 \alpha^{-k},$$ 

for some constant $c'_0 > 0$. Let $N_{\varepsilon''}$ be such that if $G := \bigcup_{k \gg 0}^+ G_k$, then $\mu(G) < \varepsilon''^2$. Using Markov’s inequality, it follows that $\varepsilon''$ almost every atom $A \in \bigvee_{N_{\varepsilon''}}^\Psi \Psi^k P$, we have 

$$\mu(A \cap G) \leq \varepsilon'' \mu(G)(\text{otherwise by summing over this A would contradict } \mu(G) < \varepsilon''^2).$$ 

This implies that it is enough to define $E := A \cap G^c$, and then $E$ satisfies the assertion of the lemma. 

The following lemma will be needed in the proof of Theorem 2.5. This lemma is not needed in Dani’s paper, [10], i.e. if $Ad(g)$ is semisimple on $g^0$ but plays an important role otherwise, recall (4).

Lemma 3.9. For every $\varepsilon' > 0$ there exists $\varepsilon'' > 0$ such that for every $0 < \delta < \varepsilon''$ there exists a finite set $\{x_j\}_{j=1}^{N(\delta)} \subset M$ such that:

- **d1.** for $i \neq j$, we have $D(\delta, x_j) \cap D(\delta, x_i) = \emptyset$;
- **d2.** we have $\mu\left(\bigcup_{j=1}^{N(\delta)} D(\delta, x_j)\right) \geq 1 - \varepsilon'$;
- **d3.** there exists a compact set $K_{\varepsilon'}$ such that for every $\delta \in (0, \varepsilon''), \bigcup_{j=1}^{N(\delta)} D(\delta, x_j) \subset K_{\varepsilon'}$.

Proof. By $(iv)$ in Lemma 3.1 it follows that it is enough to show the existence of a family $\{C(\delta, G)x_j\}_{j=1}^{N(\delta)}$ satisfying d1., d2. and d3. (for $\varepsilon' = \varepsilon'/10$). Let $K_{\varepsilon'} \subset M$ be a compact set with $\mu(K_{\varepsilon'}) > 1 - \varepsilon'/10$. Similarly to section 2.6, we cover $K_{\varepsilon'} \subset \bigcup_{x \in K_{\varepsilon'}} B(x, r_{\varepsilon'})$ and let 

$$K_{\varepsilon'} \subset \bigcup_{j=1}^{m(\varepsilon')} B(x_j, r_{\varepsilon'})$$ 

be a finite subcover. Let then $K_{\varepsilon'}$ be the closure of $\bigcup_{j=1}^{m(\varepsilon')} B(x_j, r_{\varepsilon'})$, i.e.

$$K_{\varepsilon'} = \overline{\bigcup_{j=1}^{m(\varepsilon')} B(x_j, r_{\varepsilon'})}.$$ 

Notice that $\mu(K_{\varepsilon'}) \geq \mu(\overline{K_{\varepsilon'}}) \geq 1 - \varepsilon'/10$. Analogously to the construction of the sets $\{P_j\}_{j=1}^{m(\varepsilon')}$ in Section 2.6 let $R = \{R_k\}_{k=1}^{W(\varepsilon')}$ be any partition of $K_{\varepsilon'}$ such that $\mu(R_1) \leq \varepsilon'/10$
and every $R_k$, $k \geq 2$ has (piecewise) smooth boundary and $\text{diam}(R_k) < \min \left( \text{sys}(K,\varepsilon^2), \varepsilon^2 \right)$ (see Lemma 3.7). We will WLOG assume that $\text{sys}(\varepsilon') = \text{sys}(K,\varepsilon^2) < \varepsilon^2$ (if not, we consider $\text{sys}(\varepsilon) = \min(\text{sys}(\varepsilon), \varepsilon^2)$). We will show that for every $\delta$ small enough and every $k \in \{2, \ldots, W\}$, we can find a family $\{C(\delta,G)x_j^{\prime}\}_{j=1}^{p(k,\delta)}$ such that $C(\delta,G)x_j^{\prime} \subset R_k$ for every $j \in \{1, \ldots, p\}$. $C(\delta,G)x_j^{\prime} \cap C(\delta,G)x_{j'}^{\prime} = \emptyset$ for $j \neq j'$ and

$$
\mu \left( \bigcup_{j=1}^{p} C(\delta,G)x_j^{\prime} \right) \geq (1 - \varepsilon')\mu(R_k). \quad (7)
$$

The general result then follows by taking the union over all $k \in \{2, \ldots, W\}$ of the families $\{C(\delta,G)x_j^{\prime}\}_{j=1}^{p(k,\delta)}$.

Let $x$ be any point in $R_k$. and let $Gr := [2(1 + \varepsilon') \delta]2^{\text{dim}_G} \cap [0, \text{sys}(\varepsilon')] \dim_G$ be a grid in $[0, \text{sys}(\varepsilon')] \dim_G$ of size $2(1 + \varepsilon') \delta$. Let $Gr' \subset Gr$ be such that for any $a \in Gr'$, we have $x_a := \exp(aV)x \in R_k$ and also $C(\delta,G)x_a \subset R_k$. We then define the family

$$
\{C(\delta,G)x_a\}_{a \in Gr'}. \quad (8)
$$

Notice first that if $a, a' \in Gr'$ then

$$
C(\delta,G)x_a \cap C(\delta,G)x_{a'} = \emptyset. \quad (9)
$$

Indeed, if not then for some $\gamma \in \Gamma$ and $g, g' \in C(\delta,G)$, we have

$$
g \exp(aV)x = g' \exp(a'V)x_{\gamma}.
$$

If $\gamma = e$, then the above equation becomes

$$
g'^{-1}g = \exp(a'V) \exp(-aV).
$$

Since $g', g \in C(\delta,G)$, by $(ii)$ in Lemma 3.1 it follows that

$$
\| \log_{alg}(g^{-1}g) \| \leq 2\delta + O(\kappa_2 \delta^2) < 2(1 + \varepsilon'/10)\delta.
$$

Similarly, we write $a'V = aV + [a' - a]V$ and using $[aV, a'V] = 0$, the fact that $\|a' - a\| = 2(1 + \varepsilon')\delta$ and $a \leq \text{sys}(\varepsilon')/3$, by $(ii)$ in Lemma 3.1 we get

$$
\| \log_{alg}(\exp(a'V) \exp(-aV)) \| \geq 2(1 + \varepsilon)\delta + O(2(1 + \varepsilon) \kappa_2 \text{sys}(\varepsilon)/3\delta) > 2(1 + \varepsilon/3)\delta.
$$

The two last inequalities are contradictory and this finishes the proof if $\gamma = e$. On the other hand, if $\gamma \neq e$, then the above equation transforms to

$$
x_{\gamma}x^{-1} = \exp(-a'V)g'^{-1}g \exp(aV).
$$

Notice however that by triangle inequality, we have

$$
d_G(\exp(-a'V)g'^{-1}g \exp(aV), e) \leq 2\delta + 2\text{sys}(\varepsilon')/3,
$$

which contradicts the fact that $x \in K_\delta$ (since, by Lemma 3.7 $d_G(x\gamma x^{-1}, e) > \text{sys}(\varepsilon')$) if $\delta$ is small enough. So $(9)$ indeed holds. So it remains to show $(7)$ for the family in $(8)$. 
For this aim, we will first show that
\[
\{ y \in R_k : d_M(y, \partial R_k) > 4(1 + \epsilon')\delta \} \subset \bigcup \limits_{a \in Gr'} C((1 + 3\epsilon')\delta, G)x_a.
\] (10)

Indeed, notice first that by the definition of \(Gr'\) it follows that if \(d_M(x_a, \partial R_k) > 2\delta\), then \(C(\delta, G)x_a \subset R_k\). Take \(y \in R_k\) such that \(d_M(y, \partial R_k) > \delta\). Since \(\text{diam}(R_k) < \text{sys}(\epsilon)^2\) it follows that there exists \(\gamma \in \Gamma\) and \(q \in \left[0, \text{sys}(\epsilon)^2\right]^{\dim g}\) such that \(y = \exp(qV)x\gamma\). Let \(a'' \in Gr\) be a point which minimizes \(\|q - a\|\) over all \(a \in Gr\). By the definition of the set \(Gr\) it follows that \(\|q - a''\| \leq (1 + \epsilon')\delta\). Notice that
\[
y = \exp(qV)x\gamma = \exp([q - a)V + aV]x\gamma.
\]

By Lemma (ii) in Lemma 3.1 and since \([a''V, a''V]\) = 0 it follows that
\[
\|\log_{alg} [\exp([q - a''V])V + a''V]\exp(-a''V)]\| = q - a'' + O((1 + \epsilon')\delta|a''|) \leq (1 + \epsilon')\delta + \epsilon'^2(1 + \epsilon')\delta < \delta + 2\epsilon'\delta.
\]

Therefore \(y \in C((1 + 3\epsilon')\delta, G)x_{a''}\). It remains to show that \(a'' \in Gr'\) (which is equivalent to \(C(\delta, G)x_{a''} \subset R_k\)). Indeed, notice that
\[
d_M(x_{a''}, \partial R_k) \geq d_M(y, \partial R_k) - d_M(y, \exp(a''V)x) \geq 4(1 + \epsilon')\delta - d_G(\exp(qV), \exp(a''V)) \geq 4(1 + \epsilon')\delta - 2(1 + \epsilon')\delta = 2(1 + \epsilon')\delta.
\]

Consequently, \(C(\delta, G)x_{a''} \subset R_k\) and hence \(a'' \in Gr'\). This finishes the proof of (10). By (iii) in Lemma 3.1, 10 and (9), we have
\[
\frac{1 + \epsilon'}{1 - \epsilon}(1 + 3\epsilon')^{\dim g} \mu \left( \bigcup \limits_{a \in Gr'} C(\delta, G)x_a \right) = \sum \limits_{a \in Gr'} \frac{1 + \epsilon'}{1 - \epsilon}(1 + 3\epsilon')^{\dim g} \mu (C(\delta, G)x_a) \geq \sum \mu \left( C((1 + 3\epsilon')\delta, G)x_a \right) \geq \mu \left( \{ x \in R_k : d_M(x, \partial R_k) > 4(1 + \epsilon')\delta \} \right) \geq (1 - \phi(\delta))\mu(R_k),
\] (11)

where \(\lim_{\delta \to 0} \phi(\delta) = 0\) (since the boundary of \(R_k\) is piecewise smooth). So taking sufficiently small \(\delta\) (if necessary, we can make \(\delta\) smaller when \(k\) ranges over \(\{2, \ldots, W\}\)) and changing \(\epsilon' = \epsilon''\) so that \(1 - \epsilon'' = 1 + \frac{\epsilon''}{1 + \epsilon'}(1 + 3\epsilon')^{\dim g}\) proves (7) and hence finishes the proof of the lemma.

\[\square\]

### 3.3 Equidistribution estimates

In this section we will establish quantitative (polynomial) equidistribution estimates of the horospherical subgroup under the action of \(\Psi\). We will consider the family of cubes from Lemma 3.9 i.e. for \(\delta_k = (\log k)^{-r'}\), let
\[
\{ D(\delta_k, x_j) \}_{j=1}^{N_k}
\]
be the family from the statement of Lemma 3.9. Notice that by d1 and d2 it follows that \(N_k \leq (\log k)^{r'\dim g}\). Let moreover \(K_\epsilon\) be the compact set from d3. We have the following lemma:
Lemma 3.10. For every \( \varepsilon > 0 \) and every \( \xi \in (0, \varepsilon) \) there exists \( k_\varepsilon \) such that for every \( k \geq k_\varepsilon \), every \( N \geq \log k \) and every \( x \in K_\varepsilon \), we have

\[
(1 - \varepsilon^2) \text{Leb}^+(C(\xi, g^+)) \mu(D(\delta_k, x_j)) \leq \text{Leb}^+(P \in C(\xi, g^+) : \psi^N \exp(P)x \in D(\delta_k, x_j)) \leq (1 + \varepsilon^2) \text{Leb}^+(C(\xi, g^+)) \mu(D(\delta_k, x_j)).
\]

The result follows from Proposition 2.4.8. in [21] together with an approximation argument. We will give the proof in the appendix.

4 Proof of the VWB property

4.1 Good returns for most points

Recall that the partition \( \mathcal{P} \) is now fixed (see Subsection 2.6). We will use Lemma 2.3 to establish the VWB property for \( \Psi \). Fix \( \varepsilon > 0 \). Let \( D = \{D(\varepsilon^\prime, x_j)_{j=1}^{N(\varepsilon^\prime)} \} \) satisfying \( d_1 \), \( d_2 \) and \( d_3 \) (we wlog assume that \( \varepsilon^\prime \leq \varepsilon^{100} \)). Notice that since \( \varepsilon^\prime \) depends on \( \varepsilon \) only, by the \( K \)-property of \( \Psi \) (see Definition 2.4) with \( B_j = D(\varepsilon^\prime, x_j) \) and \( \varepsilon = \frac{\varepsilon^{\mu(D_j)}}{10N(\varepsilon^\prime)} \), we get that there exists \( N_j(\varepsilon, B_j) \) such that for every \( N' \geq N \geq N_j \) and for \( \varepsilon \) a.e. atom \( A \in \bigvee_{N=N_j}^{N(N_j)} \Psi^i(\mathcal{P}) \), we have

\[
\left| \frac{\mu(A \cap D_j)}{\mu(A)} - \mu(D_j) \right| \leq \frac{1}{50} \varepsilon \mu(D_j).
\]

Defining \( N'_2 = \max_{j \in \{1, \ldots, N(\varepsilon^\prime)\}} N_j \), by the above it follows that for every \( N' \geq N \geq N'_2 \) and for \( \varepsilon \) a.e. atom \( A \in \bigvee_{N=N'_2}^{N(N_j)} \Psi^i(\mathcal{P}) \), we have

\[
\left| \frac{\mu(A \cap D_j)}{\mu(A)} - \mu(D_j) \right| \leq \frac{1}{10} \varepsilon \mu(D_j),
\]

for all \( j \in \{1, \ldots, N(\varepsilon^\prime)\} \).

Let \( \partial \mathcal{P} = \bigcup_{P \in \mathcal{P}} \partial P \), and let \( V_{\varepsilon^\prime}(\partial \mathcal{P}) \) be the \( \varepsilon' \) neighborhood of \( \partial \mathcal{P} \), where \( \varepsilon' \) is such that \( \mu(V_{\varepsilon^\prime}(\partial \mathcal{P})) < \varepsilon/20 \) (see [5] in Subsection 2.6). Since \( \varepsilon' \) is an explicit function of \( \varepsilon \) we will below use \( \varepsilon' \) instead of \( \varepsilon' \). Let \( p = \dim g^+ \). The following proposition is important in the proof of Theorem 2.5.

Proposition 4.1. Let \( D \in \mathcal{D} \). There exists \( N_3 \) such that for every \( S, N \geq N_3 \) and \( \varepsilon/50 \) a.e. atom \( A \in \bigvee_{N=N_3}^{N(\varepsilon')} \Psi^i(\mathcal{P}) \) there exists an \( \varepsilon/50 \) -measure preserving map \( \theta_S : (A \cap D, \mu|_{A \cap D}) \to (D, \mu|_D) \) such that

\[
d(\Psi^k(x), \Psi^k(\theta_S x)) < \varepsilon/10 \quad \text{(for } 1 - \varepsilon^2 \text{ proportion of } k \in \{0, \ldots, S-1\})
\]

Before we prove Proposition 4.1 let us show how it implies Theorem 2.5.

Proof of Theorem 2.5. We will use Lemma 2.3. By ergodic theorem for the characteristic function of \( V_{\varepsilon'}(\partial \mathcal{P}) \) there exists a set \( \tilde{A}_\Psi, \mu(\tilde{A}_\Psi) > 1 - \varepsilon^8 \) and \( N_4 = N_4(\varepsilon) \), such that for every \( x \in \tilde{A}_\Psi \) and every \( N \geq N_4 \), we have

\[
\text{card} \{i \in \{0, \ldots, N-1\} : \Psi^i x \not\in V_{\varepsilon'}(\partial \mathcal{P})\} \geq (1 - \varepsilon/2)N
\]

(16)
Let $N = \max(N_2, N_3, N_4)$ where $N_2$ comes from (14) and $N_3$ comes from Proposition 4.1.

Let $B_1 = \bigcup_{D \in D, D' \subset D}$, where $D' \subset D$ is such that for $D \in D'$, we have $\mu(D \cap A_\psi) \leq (1 - \varepsilon)\mu(D)$. Since $\mu(A_\psi) \geq 1 - \varepsilon^8$, it follows that $\mu(B_1) \geq 1 - \varepsilon^3$.

We say that $A \in \bigvee_{N_2}^\infty \Psi^i(P)$ is good if $\mu(A \cap A_\psi) \geq (1 - \varepsilon)\mu(A)$ and (14) is satisfied for $A$. Analogously, since $\mu(A_\psi) \geq 1 - \varepsilon^8$, it follows that $\varepsilon^3$ a.e. atom is good. Fix a good atom $A$. If $D'' \subset D$ is such that for $D \in D''$, we have $\mu(A \cap A_\psi \cap D) \leq (1 - \varepsilon^2)\mu(A \cap D)$, then a similar reasoning shows that $\mu(\bigcup_{D \in D''} D) \leq \varepsilon/2$. Let $D''' = D' \cup D''$. For every $S \supset N$ we will construct a measure preserving map $\theta_S : A \to M$ such that the $S$-names of $x$ and $\theta_S(x)$ are the same except at most $\varepsilon$ proportion. For this aim we use Proposition 4.1. For a good atom $A$ and $D \in D, D'''$, let $\theta_S = \theta_{S,D} : A \cap D \to D$ be the $\varepsilon/50$ measure preserving map as in Proposition 4.1 so that (15) holds. Since $D \in D \setminus D'''$ it follows that $\mu(A \cap A_\psi \cap D) \in (1 - \varepsilon^2, 1]$. Therefore $\theta_S$ descends to an $\varepsilon/40$ measure preserving map $\theta_S : A \cap D \cap A_\psi \to D$ (recall that we need to rescale the measure on $A \cap D \cap A_\psi$). This by the definition and (14) means that for every $B \subset A \cap D \cap A_\psi$, we have

$$\frac{\mu(A)\mu(\theta_{S,D}(B))}{\mu(A \cap B \cap A_\psi \cap D)} = \frac{\mu(\theta_{S,D}(B))\mu(A \cap D \cap A_\psi)}{\mu(D)\mu(A \cap D \cap A_\psi B)} = \frac{\mu(A)\mu(D)}{\mu(A \cap D \cap A_\psi)} \in (1 - \varepsilon/10, 1 + \varepsilon/10),$$

(17)

since both fractions are $\varepsilon/30$ close to 1 – the first one since $\theta_{S,D}$ is $\varepsilon/40$ measure preserving and the second by (14) and the fact that $\frac{\mu(A \cap A_\psi \cap D)}{\mu(A \cap D)} \in (1 - \varepsilon^2, 1]$. We then naturally extend the maps $\theta_{S,D}$ to a map

$$\theta_S : A \cap A_\psi \cap \bigcup_{D \in D \setminus D''} D \to \bigcup_{D \in D \setminus D''} D.$$  

We will use (17) to show that $\theta_S$ is $\varepsilon/5$ measure preserving. Take $B \in A \cap A_\psi \cap \bigcup_{D \in D \setminus D''} D$, and let $B_j := A \cap A_\psi \cap D_j$, where $j$ is an indexing of the set $D \setminus D'''$.

Then $B = \bigcup_j B_j$ and, by (17), we have

$$\frac{\mu(\theta_S(B))}{\mu(A)} = \frac{\mu(A)\sum_j \mu(\theta_S(B_j))}{\mu(A)} = \frac{\mu(A)\sum_j \mu(B_j)}{\mu(A)\sum_j \mu(B_j)} \in (1 - \varepsilon/5, 1 + \varepsilon/5).$$

This shows that $\theta_S$ is $\varepsilon/5$ measure preserving on $\text{Dom}(\theta_S) := A \cap A_\psi \cap \bigcup_{D \in D \setminus D''} D$. Then, by the above choice of $A$, it follows that $\mu(\text{Dom}(\theta_S)) \geq (1 - \varepsilon)\mu(A)$. Notice that for $x \in A_\psi$, (15) and (16) imply that the $S$ names of $x$ and $\theta_S x$ match up to en error at most $\varepsilon S$. We define $\theta_S$ in an arbitrary way on $M \setminus \text{Dom}(\theta_S)$ to get an $\varepsilon/2$-measure preserving map $\theta_S : A \to M$. This finishes the proof.

So it remains to prove Proposition 4.1. We will do it in the next subsection.

### 4.2 Proof of Proposition 4.1

The following proposition is the most important part of the proof. It is a generalization of the „Main Lemma” in [33]. Recall that for $z \in M, W^+(\xi, z) = C(\xi, G^+) z$. For a set $\exp(S) z \in M$, where $S \in g^+$, we will use the following notation $\text{Leb}^+(\exp(S) z) = \text{Leb}^+(S)$, where $\text{Leb}^+$ denotes the $\rho$ dimensional Lebesgue measure on $g^+$.

We have the following:
Proposition 4.2 (Main Proposition). Let \( z, z' \in K_\varepsilon \) (see d3 in Lemma 3.9). There exists \( N_5 \) such that for every \( S \geq N_5 \) there exists an \( \varepsilon/100 - \text{Leb}^+ \) measure preserving map \( \tilde{\theta} := \tilde{\theta}_{z,z',S} : W^u(\varepsilon'', z) \to W^u(\varepsilon'', z') \) such that for every \( x \in W^u(\varepsilon'', z) \), we have
\[
d_M(\Psi^s x, \Psi^s \tilde{\theta} x) \leq \varepsilon \quad \text{for } (1 - \varepsilon^2) \text{ proportion of } s \in [0, S].
\] (18)

Before we prove Proposition 4.2 let us show how it implies Proposition 4.1.

Proof of Proposition 4.2. In the proof we will use that \( \mu \) is \( \varepsilon^2 \) close to \( \text{Leb}^+ \times \text{Leb}^0 \times \text{Leb}^- \) (see (iii) in Lemma 3.1) around \( e \in G \). Fix \( D \in D \) and \( A = \bigvee \psi_i(\mathcal{P}) \) satisfying the assumptions of Proposition 4.2. So by the definition of \( \tilde{\theta} \) it follows that
\[
\varepsilon = \mu(A) \geq (1 - \varepsilon^6)p \mu(A) \quad \text{and } E \text{ intersects } D \text{ in an } \mathfrak{g}^+ \text{ tubular subset. Since } E \text{ intersects } D \text{ in an arbitrary way on } S \text{ it follows that }
\[
E = C(\varepsilon'', \mathfrak{g}^+) \times [E \cap C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-)]x_D \subset M.
\]

Let \( \tilde{E} \subset E \cap C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-) \) be such that for every \( x \in \tilde{E} \) we have \( \text{Leb}^+(C(\varepsilon'', \mathfrak{g}^+)x \cap \psi_i) \geq 1 - \varepsilon^3p \). By Fubini’s theorem it follows that
\[
(\text{Leb}^0 \times \text{Leb}^-)(\tilde{E}) \geq (1 - \varepsilon^3)(\text{Leb}^0 \times \text{Leb}^-)(E \cap C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-))
\]
In what follows we will construct \( \varepsilon \) measure preserving maps on several domains. We always consider the induced measures on the domains considered. We will first construct a map \( \zeta : \tilde{E} \cap C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-) \to C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-) \) such that \( \zeta \) is \( \varepsilon^6 \) close \( \text{Leb}^0 \times \text{Leb}^- \) measure preserving. We will work in the Lie algebra \( \mathfrak{g}^0 \oplus \mathfrak{g}^- \). Notice that \( \tilde{E} \cap C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-) = \exp(E'') \) for some measurable subset \( E'' \subset \mathfrak{g}^0 \oplus \mathfrak{g}^- \). Let \( \zeta_0 : (E'', \text{Leb}^0 \times \text{Leb}^-_{E''}) \to C(\varepsilon'', \mathfrak{g}^0 \oplus \mathfrak{g}^-) \) be an \( \varepsilon^6 \) measure preserving map. Notice that such map always exists since \( E'' \) is a subset of Euclidean space, so we can approximate \( E'' \) by disjoint cubes \( \{C\mu_i E\}_{i=1}^{L(\varepsilon)} \) (up to \( \varepsilon^6 \) and then divide (up to \( \varepsilon^6 \) the cube \( C(\varepsilon'', \mathfrak{g}^0 \oplus \mathfrak{g}^-) \) into disjoint subcubes \( \{C\mu_i E\}_{i=1}^{L(\varepsilon)} \), and set \( C\mu_i \) affinely on \( C\mu_i \) for \( i = 1, \ldots, L(\varepsilon) \). Let then \( \zeta : \tilde{E}x_D \to C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-)x_D \) be the lift of the map \( \zeta_0 \) (then \( \zeta \) is also \( \varepsilon^6 \) \( \text{Leb}^0 \times \text{Leb}^- \) measure preserving). We now define the map \( \theta_S : C(\varepsilon'', \mathfrak{g}^+) \times \tilde{E}x_D \to C(\varepsilon'', \mathfrak{g}^+) \times C(\varepsilon'', \mathfrak{g}^0 \times \mathfrak{g}^-)x_D \)
\[
\theta_S(x,z) = \tilde{\theta}_{z,\zeta(z),S}(x),
\] (19)
we extend \( \theta_S \) to \( E \) in an arbitrary way on \( \tilde{E} \). Notice that by Fubini’s theorem, \( \theta_S \) is \( \varepsilon^2 \) measure preserving. Therefore, on \( C(\delta, \mathfrak{g}^+) \times \tilde{E}x_D \), we have that \( z, \zeta(z) \) satisfy the assumptions of Proposition 4.2. So by the definition of \( \theta_S \) (restricted to \( C(\delta, \mathfrak{g}^+) \times \tilde{E}x_D \) ) it follows that (18) holds which implies that (19) holds. This finishes the proof of Proposition 4.1.

So it only remains to prove Proposition 4.2. We will do this in a separate subsection.
4.3 Proof of Proposition 4.2

In this section we will give the proof of Proposition 4.2. Before we do that we will give an outline of the proof to explain the main ideas.

Outline of the proof. The main idea is to use effective equidistribution on cubes (see Lemma 3.10), i.e. for a fixed time $S$ let $k$ be such that $S$ is of order $\log k$. Let $z, z' \in K_\epsilon$. Then for a fixed $j \in \{1, \ldots, N_k\}$ we consider the set of all points from $x \in W^u(\varepsilon'', z)$ such that $\Psi^N x \in D_j$. By the invariance of foliations it follows that this set is a union of (exponentially small) cubes in $W^u(\varepsilon'', z)$ and each cube is mapped onto the unstable piece in $D_j$ as Figure 1 shows. The same holds for $z'$. Then there is a natural mapping $\theta$ between an exponentially small cube in $W^u(\varepsilon'', z)$ and $W^u(\varepsilon'', z')$, see Figure 1. It is the composition of the maps on the unstable pieces with the center-stable holonomy map in $D_j$. Notice that since the size of $D_j$ is a large power of $\frac{1}{\log k}$ (and the time $S$ is of order $\log k$), it follows that (on the two small cubes) $S$-orbits of points $\psi^N x$ and $\psi^N \theta x$ are $\varepsilon$ close – indeed, they lie on the same central stable piece (because of the holonomy map) and the growth on the central stable piece is at most polynomial in $S$ (see Lemma 3.6).

![Figure 1: The matching function between two small cubes in $W^u(\varepsilon'', z)$ and $W^u(\varepsilon'', z')$](image)

The key observation is that for each $j \in \{1, \ldots, N_k\}$ the number of small cubes in $W^u(\varepsilon'', z)$ and $W^u(\varepsilon'', z')$ is the same up to an error of order $\varepsilon^2$, see Figure 2. This is a consequence of quantitative equidistribution on $D_j$ (see Lemma 3.10). Moreover, if we consider the union over all $j$ of all small cubes in $W^u(\varepsilon'', z)$ (or $W^u(\varepsilon'', z')$), then again by Lemma 3.10 they cover $W^u(\varepsilon'', z)$ (or $W^u(\varepsilon'', z')$) up to a proportion of $\varepsilon^2$. Extending $\theta$ in a natural way over all $j \in \{1, \ldots, N_k\}$ yields a matching between $W^u(\varepsilon'', z)$ and $W^u(\varepsilon'', z')$. 
Proof of Proposition 4.2. Fix \( S \geq \max(\varepsilon^{-10}N_3, \varepsilon^{-10}k_\varepsilon) \). Let \( N = [\varepsilon^3 S] \). Let \( k \in \mathbb{N} \) be such that \( N \in [\log k, \log(k+1)] \). Let \( N_k \) and \( D_j = D(\delta_k, x_j), j = 1, \ldots, N_k \) be the family of cubes from Lemma 3.10 with \( \delta_k = \frac{1}{(\log k)^p} \). Fix \( j \in 1, \ldots, N_k \). By Lemma 3.10 with \( \xi = \varepsilon'' \), we have for \( w \in \{z, z'\} \subset K_\varepsilon \), we have

\[
\left| \frac{\text{Leb}^+(\{P \in C(\varepsilon'', g^+) : \Psi^N(\exp(P)w) \in D_j\})}{\text{Leb}^+(C(\varepsilon'', g^+))\mu(D_j)} - 1 \right| \leq \varepsilon^2. \tag{20}
\]

Let \( C_N \subset C(\delta, g^+) \) be given by

\[
C_N := C((1 - \lambda^{-N})\delta, g^+), \tag{21}
\]

where \( \lambda := \frac{1}{2} \min_{|\lambda| > 1} |\lambda| \). Notice that

\[
\text{Leb}^+(C(\delta, g^+) \setminus C_N) \leq [1 - (1 - \lambda^{-N})^p]\text{Leb}^+(C(\delta, g^+)). \tag{22}
\]

Recall that \( j \in 1, \ldots, N_k \) is now fixed. Take any \( P_j \in C_N \) such that \( \Psi^N(\exp(P_j)z) \in D_j \), i.e.

\[
\Psi^N \exp(P_j)z = \exp(V_j) \exp(W_j)x_j
\]

for some \( V_j \in C(\delta_k, g^+) \) and \( W_j \in C(\delta_k, g^0 \oplus g^-) \).

We now define a map \( f_{P_j} : C(\delta_k, g^+) \to C(\varepsilon'', g^+) \) which sends \( V_j \) to \( P_j \) and such that \( f(C(\delta_k, g^+)) \subset C(\varepsilon'', g^+) \) (in fact, as we will see below, the image of \( C(\delta_k, g^+) \) will be an exponentially small set in \( C(\varepsilon'', g^+) \)). Let \( R_Q : G \to G, R_Q(h) := h \exp(Q) \) and define\(^5\)

\[
f_{P_j}(U) := [\log_{\text{alg}} \circ R_{P_j} \circ \exp \circ \text{Ad}(g)^{-N} \circ \log_{\text{alg}} \circ R_{-V_j} \circ \exp]^\circ(U).
\]

\(^5\)Notice that if \( P = V = 0 \in g^+ \), then \( f_{P_j}(U) = \text{Ad}(g)^{-N}(U) \), which is the adjoint action on \( U \). In our case we want the image of \( V_j \) to be \( P_j \) and therefore we need to shift twice by \( \exp(P_j) \) and \( \exp(-V_j) \).
Moreover, since \( \log_{\text{alg}} \circ R_{V_j} \circ \exp \circ Ad(g)^N \circ \log_{\text{alg}} \circ R_{-P_j} \circ \exp \) (denoting \( \epsilon'' \) (in fact, \( \| V_j \| \leq \epsilon'' \)), we have
\[
|\det [\text{Jac}(\log_{\text{alg}} \circ R_{P_j} \circ \exp)(\cdot)]| \in (1 - \epsilon'''/3, 1 + \epsilon'''/3).
\]
Similarly,
\[
|\det [\text{Jac}((\log_{\text{alg}} \circ R_{-V_j} \circ \exp)(U))| \in (1 - \epsilon'''/3, 1 + \epsilon'''/3).
\]

Notice that in the proof we consider \( Ad(g)^{-N} \) restricted to the space \( \mathfrak{g}^+ \). To simplify the notation we use \( \overline{Ad(g)}^{-N} \) to denote \( Ad(g)_{\mathfrak{g}^+}^{-N} \). Since \( Ad(g)^{-N} \) is a linear map, we have (denoting \( c_g := \prod_{|\lambda| > 1} |\lambda| \dim V_\lambda > 1 \))
\[
\det \text{Jac}(\overline{Ad(g)}^{-N}(\cdot)) = |\det Ad(g)^{-N}| = c_g^{-N}.
\]

Summarizing, it follows that for \( U \in C(\delta_k, \mathfrak{g}^+) \), we have
\[
|\det [\text{Jac}(f_{P_j})](U)| \in [1 - \epsilon'''/10, 1 + \epsilon'''/10] c_g^{-N}.
\]

Moreover, since \( P_j \in C_N \), it follows that \( f_{P_j}(C(\delta_k, \mathfrak{g}^+)) \subset C(\epsilon'', \mathfrak{g}^+) \) (since \( C(\delta_k, \mathfrak{g}^+) \) is contracted exponentially by \( f_{P_1}, f_{P_j}(V_j) = P_j \) and \( P_j \) is separated from the boundary of \( C(\epsilon'', \mathfrak{g}^+) \) and \( N \) is large enough, see \( \text{(20)} \)).

Therefore, by the construction of the function \( f_{P_j} \) and since \( Ad(g)^N \) preserves the spaces \( C(\cdot, \mathfrak{g}^+) \) and the spaces \( \mathfrak{g}^+ \) and \( \mathfrak{g}^0 \oplus \mathfrak{g}^- \) are transversal, it follows that there exists \( \{ P_{j,s}^\epsilon(z) \}_{s=1}^{\ell_j(z)} \subset C_N \) such that
\[
\Psi^N \exp(P_{j,s}^\epsilon(z)) \exp(\{ V_{j,s}^\epsilon(z) \}) x_j
\]
for \( s = 1, \ldots, \ell_j(z) \) and
\[
C_N \cap \{ P \in C(\epsilon'', \mathfrak{g}^+) : \Psi^N \exp(P)z \in D_j \} \subset \bigcup_{s=1}^{\ell_j(z)} f_{P_{j,s}^\epsilon(z)}(C(\delta_k, \mathfrak{g}^+)) \subset \{ P \in C(\epsilon'', \mathfrak{g}^+) : \Psi^N \exp(P)z \in D_j \}.
\]

We get analogous sets \( \{ P_{j,s}^{\epsilon'}(z') \}_{s=1}^{\ell_j(z')} \) for \( z' \). Notice that by the bounds on the Jacobian, \( \text{(24)} \), it follows that
\[
\text{Leb}^+ \left( \bigcup_{s=1}^{\ell_j(z)} f_{P_{j,s}^{\epsilon}(z)}(C(\delta_k, \mathfrak{g}^+)) \right) \in \ell_j(z)[1 - \epsilon'''/10, 1 - \epsilon'''/10] c_g^{-N}.
\]

Moreover, since \((1 - (1 - \lambda^{-N})^p) \leq \epsilon^2 \mu(D_j) \) and \( N \geq \log k \), by \( \text{(20)} \) and \( \text{(22)} \) it follows that
\[
\text{Leb}^+ \left( C_N \cap \{ P \in C(\epsilon'', \mathfrak{g}^+) : \Psi^N \exp(P)z \in D_j \} \right) \geq (1 - \epsilon^2) \text{Leb}(C(\epsilon'', \mathfrak{g}^+)) \mu(D_j) - \text{Leb}^+(C(\epsilon'', \mathfrak{g}^+) \setminus C_N) \geq (1 - \epsilon^{3/2}) \text{Leb}^+(C(\epsilon'', \mathfrak{g}^+) \mu(D_j)
\]
This together with (26) and (27) implies that (since we assume that $\varepsilon' < \varepsilon^{100}$)
\[
\ell_j(z) \in [1 - \varepsilon^{4/3}, 1 + \varepsilon^{4/3}]C^N \log (C(\varepsilon', g^+))\mu(D_j)
\]
(29)
The same estimate holds for $\ell_j(\zeta)$. Let $\ell_j := \min(\ell_j(z), \ell_j(\zeta))$, $C(j, z, s) = f_{P_j,s}(C(\delta_k, g^+))$ and define
\[
C(j, z) := \bigcup_{s=1}^\infty C(j, z, s) \subset C(\varepsilon'', g^+).
\]
(30)

Since $\exp(W^\ast_{j,s}) \in C(\delta_k, g^0 \oplus g^-)$ and $\exp(W^\ast'_{j,s}) \in C(\delta_k, g^0 \oplus g^-)$ (see (25)) it follows that $\log_{\text{alg}} \left( \exp(W^\ast_{j,s}) \exp(-W^\ast'_{j,s}) \right)$ exists. Let $\text{hol}_{j,s}^\ast : g^+ \oplus \{ \log_{\text{alg}} \left( \exp(W^\ast_{j,s}) \exp(-W^\ast'_{j,s}) \right) \} \to g^+$ be the stable holonomy defined in (6) (since $z, z' \in M$ are fixed throughout the proof, we drop it from the notation of the holonomy map). By (6) and Lemma 3.3 it follows that for $R \in C(\delta_k, g^+)$, we have
\[
\exp(R) \exp(W^\ast_{j,s}) \exp(-W^\ast'_{j,s}) = \exp(W^\ast_{j,s}) \exp(\text{hol}_{j,s}(R)),
\]
where $W^\ast_{j,s} = \text{hol}_{j,s}^\ast \left( R, \log_{\text{alg}} \left( \exp(W^\ast_{j,s}) \exp(-W^\ast'_{j,s}) \right) \right) \in g^0 \oplus g^-$. This transforms to
\[
\exp(R) \exp(W^\ast_{j,s}) = \exp(W^\ast_{j,s}) \exp(\text{hol}_{j,s}(R)) \exp(W^\ast_{j,s})
\]
(31)

Moreover, by Lemma 3.3 it follows that
\[
\|W^\ast_{j,s} - \log_{\text{alg}} \left( \exp(W^\ast_{j,s}) \exp(-W^\ast'_{j,s}) \right) \| \leq k_2 \| R \| \| \log_{\text{alg}} \left( \exp(W^\ast_{j,s}) \exp(-W^\ast'_{j,s}) \right) \|.
\]
This together with the fact that $R \in C(\delta_k, g^+)$, $W^\ast_{j,s} \in C(\delta_k, g^0 \oplus g^-)$ and $W^\ast'_{j,s} \in C(\delta_k, g^0 \oplus g^-)$ implies, using $(ii)$ in Lemma 3.1 that
\[
\|W^\ast_{j,s}\| \leq 4 \delta_k, \text{ and } W^\ast_{j,s} \in g^0 \oplus g^-.
\]
(32)

By the same arguments and Lemma 3.3, we also have that if $\| R \| < (1 - \varepsilon^{100})\delta_k$, then $\| \text{hol}_{j,s}(R) \| \leq \delta_k$. By throwing away a subset of measure $\varepsilon^{100} \text{Leb}^+(C(\delta_k, g^+))$. Notice that by Lemma 3.3 we have $| \det \text{Jac} \text{hol}_{j,s} | \in (1 - \varepsilon^4, 1 + \varepsilon^4)$. We define the following matching map $\theta_{j,s}$ between $C(j, z, s)$ and $C(j, z', s')$.
\[
\theta_{j,s}(R) = f_{P_j,s} \circ \text{hol}_{j,s} \circ f_{P_j,s}^{-1}(R)
\]
(33)

We then naturally, using (30), extend the map to a map $\theta_j : C(j, z) \to C(j, z')$ and finally we extend it to a map $\theta_{z',z} : \bigcup_{j=1}^N C(j, z) \to \bigcup_{j=1}^N C(j, z)$, since the sets $\{ C(j, z) \}_{j=1}^N$ are pairwise disjoint (see the right inclusion in (26) and recall that the sets $\{ D_j \}_{j=1}^N$ are disjoint). We will show that $\theta_{z',z}$ is $\varepsilon/1000$ measure preserving. Notice first that by (29) it follows that $| \ell_j - \ell_j(z) | < 2 \varepsilon^{4/3} \ell_j(z)$. Therefore and using (28) and (26), we get that the domain of $\theta_j$ has measure at least
\[
\text{Leb}^+ \left( \bigcup_{s=1}^\infty C(j, z, s) \right) \geq \text{Leb}^+ \left( \bigcup_{s=1}^{\ell_j(z)} C(j, z, s) \right) \geq (1 - \varepsilon^{3/2} - 2\varepsilon^{4/3}) \text{Leb}^+(C(\delta, g^+))\mu(D_j).
\]
Therefore the domain of $\theta_{z,z'}$ has measure at least

$$\text{Leb}^+(\bigcup_{j=1}^{N_k} C(j, z)) \geq (1 - \epsilon^{3/2} - 2\epsilon^{3/3}) \text{Leb}^+(C(\delta, g^+)) \mu(\bigcup_{j=1}^{N_k} D_j) \geq (1 - \epsilon^{11/10}) \text{Leb}^+(C(\delta, g^+)),$$

since by definition of the family $D_j$, we have $\mu(\bigcup_{j=1}^{N_k} D_j) \geq (1 - \epsilon^2)$ (see d2 for $\epsilon' = \epsilon^2$ in Lemma (3.9)). Therefore $\theta_{z,z'}$ is defined on a proportion $1 - \epsilon^{11/10}$ of $C(\varepsilon'', g^+)$. Finally, we define

$$\tilde{\theta}_{z,z',S}(\exp(R)z) = \exp(\theta_{z,z'}(R))z', \quad (34)$$

We need to show that $\tilde{\theta}_{z,z',S}$ is $\varepsilon / 1000$-Leb$^+$ measure preserving and that (18) holds. To show the first part, it is enough to show that $\tilde{\theta}_{j,s}$ (see (33)) is $\varepsilon / 1000$-Leb$^+$ measure preserving. Note that the Jacobian of $\exp, \log_{\text{alg}}$ is $\varepsilon^5$ close to 1. Moreover, the Jacobian of $\text{hol}_s$ is also $\varepsilon^5$ close to 1. Finally by (24), we get that

$$|\det[\text{Jac}(f_{p_{j,s}})]| |\det[\text{Jac}(f_{p_{j,s}^{-1}})]| \in [(1 - \varepsilon)^2, (1 + \varepsilon)^2].$$

So by the chain rule we have that

$$|\det[\text{Jac}(\theta_{j,s})]| \in [1 - \varepsilon^2, 1 + \varepsilon^2].$$

This shows that $\theta_{j,s}$ (and hence $\theta_{z,z',S}$) is $\varepsilon / 1000$-measure preserving. It remains to show that $\tilde{\theta}_{z,z',S}$ satisfies (18). Take $x \in W^u(\varepsilon'', z)$, $x = \exp(U_x)z$ and such that $\tilde{\theta}_{z,z',S}$ is defined on $x$ (notice that the measure of such points is at least $1 - \epsilon^{11/10})\text{Leb}^+(C(\varepsilon'', g^+))$). Let $j \in \{1, \ldots, N_k\}$ and $s \in \{1, \ldots, \ell\}$ be such that $x \in C(j, z, s)$. Then by (34) and the definition of $\tilde{\theta}_{z,z',S}, \tilde{\theta}_{z,z',S}(\exp(U_x)z) = \exp(\theta_{j,s}(U_x))z'$. Denote $\theta_{j,s}(U_x) = \tilde{U}_x$.

We have

$$\Psi^N \exp(U_x)z = \Psi^N \exp(\log_{\text{alg}}(\exp(U_x) \exp(-P_{j,s}^z))) \Psi^-N \Psi^N \exp(P_{j,s}^z)z = \exp \left( \text{Ad}(g)^N \circ \log_{\text{alg}} \circ R_{-P_{j,s}^{-1}}(\exp(U_x)) \right) \exp(V^z_{j,s}) \exp(W^z_{j,s})x_j = \exp \left( \log_{\text{alg}} \left[ \exp(\text{Ad}(g)^N \circ \log_{\text{alg}} \circ R_{-P_{j,s}^{-1}}(\exp(U_x)) \exp(V^z_{j,s})) \right] \right) \exp(W^z_{j,s})x_j. \quad (35)$$

Hence the unstable component of $\Psi^N \exp(U_x)z$ equals (see (23))

$$\log_{\text{alg}} \circ R_{V^z_{j,s}} \circ \exp \circ \text{Ad}(g)^N \circ \log_{\text{alg}} \circ R_{-P_{j,s}^{-1}} \circ \exp(U_x) = f_{p_{j,s}^{-1}}(U_x) \quad (36)$$

Analogously,

$$\Psi^N \exp(\tilde{U}_x)z' = \Psi^N \exp(\log_{\text{alg}}(\exp(\tilde{U}_x) \exp(-P_{j,s}^z))) \Psi^-N \Psi^N \exp(P_{j,s}^z)z = \exp \left( \text{Ad}(g)^N \circ \log_{\text{alg}} \circ R_{-P_{j,s}^{-1}}(\exp(\tilde{U}_x)) \right) \exp(V^z_{j,s}) \exp(W^z_{j,s})x_j = \exp \left( \log_{\text{alg}} \left[ \exp(\text{Ad}(g)^N \circ \log_{\text{alg}} \circ R_{-P_{j,s}^{-1}}(\exp(\tilde{U}_x)) \exp(V^z_{j,s})) \right] \right) \exp(W^z_{j,s})x_j. \quad (37)$$
Hence the unstable component of $\Psi^N \exp(\tilde{U}_x)z$ equals (see (23))
\[
\log_{alg} \circ R_{\Psi_j,s} \circ \exp \circ \text{Ad}(g)^N \circ \log_{alg} \circ R_{-P_{j,s}} \circ \exp(\tilde{U}_x) = \frac{1}{P_{j,s}}(\tilde{U}_x).
\] (38)

Recall that $\tilde{U}_x = \theta_{j,s}(U_x)$. So by (33) it follows that
\[
f_{P_{j,s}}^{-1}(\tilde{U}_x) = \text{hol}_{j,s} \circ f_{P_{j,s}}^{-1}(U_x).
\] (39)

Therefore, by (35) and (36), we get
\[
\Psi^N \exp(U_x)z = \exp(f_{P_{j,s}}^{-1}(U_x)) \exp(W_{j,s}^z)x_j
\] (40)

and by (37), (38) and (39), we have
\[
\Psi^N \exp(\tilde{U}_x)z' = \exp\left(\text{hol}_{j,s} \left(f_{P_{j,s}}^{-1}(U_x)\right)\right) \exp(W_{j,s}^z)x_j.
\] (41)

Moreover by (31) for $R = f_{P_{j,s}}^{-1}$, for some $W''_{j,s} \in C(4\delta_k, g^0 \oplus g^-)$ (see (32)) we have
\[
\exp\left(\text{hol}_{j,s} \left(f_{P_{j,s}}^{-1}(U_x)\right)\right) \exp(W''_{j,s}x_j) = \exp(-W''_{j,s}) \exp(f_{P_{j,s}}^{-1}(U_x)) \exp(W_{j,s}^z)x_j.
\] (42)

It follows that (40), (41) and (42) give
\[
\Psi^N \exp(U_x)z = \exp(W''_{j,s}) \exp(\tilde{U}_x)z'.
\]

Therefore, for every $s \in [0, S - N]$, we have
\[
d_M(\Psi^{s+N} \exp(U_x)z, \Psi^{s+N} \exp(\tilde{U}_x)z') \leq d_G(\Psi^s(\exp(W''_{j,s})), \Psi^s(e)) \leq \varepsilon^2,
\]

since $s \leq S - N \leq S \leq \varepsilon^{-3} N \leq \varepsilon^{-3} \log(k + 1)$, $W''_{j,s} \in C(4\delta_k, g^0 \oplus g^-)$, $\delta_k = (\log k)^{-r'}$ and we use Lemma 3.6 remembering that $r' > r_0 + 1$ and $\varepsilon^{-3}$ is small compared to $\log k$. This finishes the proof of Proposition 4.2. \hfill \square

5 Appendix: Proof of Lemma 3.10

Let $m_+$ denote the Haar measure on the expanding subgroup $G^+$. We will use the following modification of Proposition 2.4.8. in [21]:

**Proposition 5.1.** There exists $\lambda, d, d', q > 0$ such that for any $f \in C_c^\infty(G^+)$, supp$f \subset B^+(\varepsilon, 1)$, any $\phi \in C_c^\infty(M)$ and any compact set $K \subset M$, there exists a constant $C(K)$ such that for every $n \geq 0$ and every $x \in K$, we have
\[
\left| \int_{G^+} f(g)\phi(\Psi^n(gx))dm_+ - \int_{G^+} f dm_+ \int_M \phi d\mu \right| \leq C(K)[\max_{x \in M} \nabla \phi]^d S_d(f)S_d(\phi)q^ne^{-\lambda n}.
\] (43)

**Proof.** The proof of the above proposition is analogous to the proof of Proposition 2.4.8. in [21]. There the authors consider a flow $(g_t)$ instead of an automorphism $\Psi$ but the important property used is the exponential mixing of $(g_t)$, which also holds in our case for $\Psi$ (see Lemma 2.6). Moreover, the statement of Proposition 2.4.8. is weaker, i.e. the
authors claim the existence of a constant $C(f, \phi, K)$ such that (43) holds with $C(f, \phi, K)$ instead of $C(K)|\max_{x \in \Omega} \nabla \phi|^2 S_d(f) S_d(\phi)$. But the proof of Proposition 2.4.8. gives an explicit constant. First, by changing the constant (but only depending on $K$), one can assume that the projection maps $\pi_x : G \to M$, $\pi_x(g) = gx$ are injective on $\text{supp} f$ for every $x \in K$. If this is the case, equation 2.4.8 in [21] and the estimates below explicitly give the constant $C(f, \phi, K)$ in the last line of the proof.

Then Lemma 3.10 follows from approximating characteristic functions of $D(\delta_k, x_i)$ with smooth compactly supported functions, with good control on the Sobolev norms. Notice also, that close to $\epsilon \in G$ the measures $\text{Leb}^+$ and $m_+$ are close (see (iii) in Lemma 3.1).

We will prove the following lemma:

For $\epsilon > 0$ and $B \subset M$, we denote the $\epsilon$ neighborhood $V_\epsilon(B) := \{ x \in M : d_M(x, B) < \epsilon \}$. Analogously we define $\epsilon$ neighborhood (in $G$) of a set $B \subset G$. Let $r' = 100r_0^2$, where $r_0$ comes from Lemma 3.6. For $\epsilon > 0$ let $K_\epsilon$ be the set from $d3$ in Lemma 3.9. Let $\text{sys}(\epsilon) = s_{\text{sys}}(K_\epsilon)$ (see Lemma 3.7) and let

$$s(\epsilon) := \min((\text{sys}(\epsilon))^2, \epsilon^4).$$

First we state the following approximation lemma.

**Lemma 5.2.** Let $\delta_k := (\log k)^{-\nu}$ and denote $\tilde{C}_k := C(\delta_k, G^+) \times C(\delta_k, G^0 \times G^-) \subset G$. For every $\epsilon > 0$ there exists $k_\epsilon$ such that for every $k \geq k_\epsilon$ there are function $\phi_k^+, \tilde{\phi}_k^+ \in C_c^\infty(G)$ such that the following holds:

V0. $\phi_k^- \leq \chi \tilde{C}_k \leq \phi_k^+$

V1. for $w \in \{+,-\}$, $\int_G |\phi_k^w - \chi \tilde{C}_k| d\mu_G < \epsilon^3 \mu_G(\tilde{C}_k)$;

V2. for $w \in \{+,-\}$, $\phi_k^w \equiv 0$ outside $V_{s(\epsilon)}(\tilde{C}_k)$;

V3. there exists a global constant $R = R(d) > 0$ such that $S_d(\phi_k^w) \leq (\log k)^R$ for $w \in \{+,-\}$.

**Proof.** We will only show how to construct $\phi_k^+$, the construction of $\phi_k^-$ is analogous.

Let $\eta_\epsilon^f \in C_c^\infty(\mathbb{R}^\ell)$ be a **smoothing function** in $\mathbb{R}^\ell$, such that

$$\eta_\epsilon^f(\mathbf{v}) = 1 \text{ if } \mathbf{v} \in [-1, 1]^\ell, \ \eta_\epsilon^f(\mathbf{v}) = 0 \text{ if } \mathbf{v} \notin [-1 - s(\epsilon)^3, 1 + s(\epsilon)^3]^\ell$$

and the derivatives satisfy: for every $m \in \mathbb{N}$ and for every $i \in \{1, \ldots, \ell\}$,

$$\left| \frac{\partial^m(\eta_\epsilon^f)}{\partial^m x_i} \right| \leq s(\epsilon)^{-66m}$$

Let then $\eta_{\epsilon, \ell}(x) := \eta_\epsilon^f \left( \frac{x}{\pi_k \epsilon} \right)$ (since $\epsilon$ is fixed, we will drop it from the notation). If we pick $k \geq k_\epsilon$, where $k_\epsilon$ is such that $\delta_k^{-1} \geq \log k_\epsilon > s(\epsilon)^{6\ell}$, we have

$$\left| \frac{\partial^m(\eta_{\epsilon, \ell})}{\partial^m x_i} \right| \leq \delta_k^{-126m}, \text{ for every } m \in \mathbb{N} \text{ and } i \in \{1, \ldots, \ell\}. \quad (44)$$
Let $\ell_1 = \dim g^+$ and $\ell_2 = \dim g^0 + \dim g^-$. Define\[\tilde{\phi}(\exp(U) \exp(H)) := \exp(\eta_1^\ell(U) V_+) \exp(\eta_2^\ell(H) V_{0,-})\]
if $\exp(U) \exp(H) \in V_{s(\varepsilon)}(\tilde{C}_k)$ and set $\tilde{\phi}(g) = 0$ otherwise. With this definition of $\tilde{\phi}$, properties $\text{V0}$ and $\text{V2}$, hold automatically. Moreover by the bounds on the derivatives of $\eta_1^\ell$ and $\eta_2^\ell$ (see (44)) and the definition of Sobolev norms (see Section 2.4) we get that there exist $R' = R'(d) > 0$ and $R = rR'$ such that
\[S_d(\tilde{\phi}) \leq \delta_k^R' = (\log k)^{rR'} \leq \log_k^R,\]
this gives $\text{V3}$. For $\text{V1}$ notice that $\tilde{f}(g) = 1$ on $\tilde{C}_k$ and hence
\[\int_G \tilde{\phi} - \chi_{\tilde{C}_k} d\mu(G) < \mu \left( V_{s(\varepsilon)}(\tilde{C}_k) \setminus \tilde{C}_k \right) \leq \varepsilon^3 \mu(\tilde{C}_k),\]
where the last inequality follows from $(iii)$ in Lemma 3.1 for $V_{s(\varepsilon)}(\tilde{C}_k)$ and $\tilde{C}_k$. This finishes the proof.

Using Lemma 5.2 and Proposition 5.1 we can prove Lemma 3.10.

Proof. Let $\tilde{C}_k = C(\delta_k, G^+) \times C(\delta_k, G^0 \times G^-) \in G$. Then by definition, $D (\delta_k, x_j) = \tilde{C}_k x_j \Gamma$. Let
\[\tilde{\phi}_j^+ : V_{s(\varepsilon)}(\tilde{C}_k) x_j \to \mathbb{R}, \quad \tilde{\phi}_j^+(g) := \tilde{\phi}^+(g x_j^{-1}),\]
where $\tilde{\phi}^+$ comes from Lemma 5.2 for $\tilde{C}_k$. Notice that $\tilde{\phi}_j^+ \equiv 0$ outside $V_{s(\varepsilon)}(\tilde{C}_k) x_j$. Moreover, for every $\gamma \in \Gamma \setminus \{e\}$, we have
\[V_{s(\varepsilon)}(\tilde{C}_k) x_j \cap V_{s(\varepsilon)}(\tilde{C}_k) x_j \gamma = \emptyset.\]
Indeed, if $z = g_1 x_j = g_2 x_j \gamma$, with $g_1, g_2 \in V_{s(\varepsilon)}(\tilde{C}_k)$, then since $x_j \in K_{z'}$ (see $d_3$ in Lemma 3.9), we have
\[0 = d_G(g_1 x_j, g_2 x_j \gamma) \geq d(x_j \gamma, x_j^{-1}) - d(g_1, e) - d(g_2, e) \geq \text{sys}(\varepsilon) - 2 \varepsilon - 2 \delta_k > 0,\]
since $s(\varepsilon) < \text{sys}(\varepsilon)^2$ and $k$ is large enough. Since $\tilde{\phi}_j^+ \equiv 0$ outside $V_{s(\varepsilon)}(\tilde{C}_k) x_j$ and by (45), we get that $\tilde{\phi}_j^+$ extends to a $\Gamma$-periodic function $\tilde{\phi}_j^+$ such $\tilde{\phi}_j^+ = \tilde{\phi}_j^+$ in $V_{c}(\tilde{C}_k) x_j$. So $\tilde{\phi}_j^+$ descends to a $C^\infty$ function on $M$. By the construction of $\tilde{\phi}_j^+$ and Lemma 5.2 it follows that $\text{V1}$, $\text{V2}$, $\text{V3}$ hold with $G$, $\tilde{\phi}$, $\mu_G$ and $\tilde{C}_k$ replaced respectively by $M$, $\tilde{\phi}_j$, $\mu$, $D (\delta_k, x_j)$. For $\xi \in (0, \varepsilon)$ let $f_{+\xi} \in C^\infty_c(G^+)$ be a function such that
\begin{itemize}
  \item [W0.] $\chi_{\exp(C(\xi, g^+))} \leq f_{+\xi}$,
  \item [W1.] $\int_{G^+} |f_{+\xi} - \chi_{\exp(C(\xi, g^+))}| d\mu_+ < \varepsilon^3 \text{Leb}^+(C(\xi, g^+))$;
  \item [W2.] $f_{+\xi} \equiv 0$ on $V_{s(\varepsilon)}(\exp(C(\xi, g^+)))$;
\end{itemize}
\footnote{For a vector $U \in g^+$ and a function $\eta : \mathbb{R}^{\dim g^+} \to \mathbb{R}$, we denote $\eta(U) := \eta(a_1, \ldots, a_{\dim g^+})$, where $U = \sum a_i V_i$. We use analogous notation for $g^0 \oplus g^-$.}
W3. $S_d(f_+\xi) \leq \xi^{-2d}$.

Then, since $m_+$ and $\text{Leb}^+$ are almost the same around $e$,

$$\int_{C(\xi, \phi^+)} \chi_{D(\delta_k, x_j)}(\Psi^N \exp(H)x) d\text{Leb}^+ \leq \int_{C(\xi, \phi^+)} f_+\xi(\exp(H))\tilde{\phi}^+_j(\Psi^N \exp(H)x) d\text{Leb}^+ \leq (1 + \varepsilon^3) \int_{G^+} f_+\xi(g)\tilde{\phi}^+_j(\Psi^N gx) dm_+.$$

But by Proposition 5.1 and Lemma 5.2 and W3., for $x \in K_\varepsilon$ and $N \geq \log k$, we have

$$\left| \int_{G^+} f_+\xi(g)\tilde{\phi}^+_j(\Psi^N gx) dm_+ - \int_{G^+} f_+\xi dm_+ \int_G \tilde{\phi}^+_j d\mu \right| \leq C(K_\varepsilon) [\max_{x \in \Omega} \nabla \tilde{\phi}^+_j]^d S_d(f_+\xi) S_d(\tilde{\phi}^+_j) t^\alpha e^{-\lambda N} \leq C(K_\varepsilon) [\log k]^{2R} \xi^{-2d} k^{-\lambda} \leq \varepsilon^{10} \int_{G^+} f_+\xi dm_+ \int_G \tilde{\phi}^+_j d\mu,$$

for $k$ large enough (largeness depending on $\varepsilon$ and $\xi$). It remains to notice that by V1., we have

$$\int_{G^+} f_+\xi dm_+ \int_G \tilde{\phi}^+_j d\mu \leq (1 + \varepsilon^3) \text{Leb}^+(C(\xi, g^+)) \mu(D(\delta_k, x_j)).$$

Combining the above inequalities, we get:

$$\text{Leb}^+(H \in C(\xi, g^+) : \psi^N \exp(H)x \in D(\delta_k, x_j)) \leq (1 + \varepsilon^3)^2 (1 + \varepsilon^{10}) \text{Leb}^+(C(\xi, g^+)) \mu(D(\delta_k, x_j)).$$

This gives the LHS inequality in Lemma 3.10. Analogously by considering $f_-\xi$ and $\tilde{\phi}^-_j$ we prove the right inequality. This finishes the proof of Lemma 3.10.

References

[1] N. Aoki, *Ergodic Automorphisms of Compact Metric Groups are Isomorphic to Bernoulli Shifts*, Publications mathématiques et informatique de Rennes, no. S4 (1975), p. 1-10.

[2] M. Björklund, M. Einsiedler, A. Gorodnik, *Quantitative multiple mixing*, to appear in JEMS.

[3] R. Bowen, *Bernoulli equilibrium states for Axiom A diffeomorphisms*, Math. System Theory 8 (1975), 289-294.

[4] M. Brin, J. Feldman, A. Katok, *Bernoulli diffeomorphisms and group extensions of dynamical systems with non-zero characteristic exponents*, Ann. of Math., 113 (1981), 159-179.

[5] L. Bunimovich, *On a class of special flows*, Math. USSR Izv., 8 (1974), 219-232.

[6] N. Chernov, C. Haskell, *Nonuniformly hyperbolic K-systems are Bernoulli*, Vol. 16, Issue 1 (1996), 19-44.
[7] M. Cowling, *Sur les coefficients des représentations unitaires des groupes de Lie simples*, Analyse harmonique sur les groupes de Lie, II, Springer, Berlin (1979), 132-178.

[8] S. G. Dani, *Kolmogorov automorphisms on homogeneous spaces*, Amer. J. Math. 98 (1976), no. 1, 119-163.

[9] S. G. Dani, *Dynamical systems on homogeneous spaces*, Bulletin of the AMS, Vol. 82 (6), 1976.

[10] S. G. Dani, *Bernoulli Translations and Minimal Horospheres on Homogeneous Spaces*, Journal of the Indian Mathematical Society, Volume 40, Issue 1-4, 1976.

[11] M. Einsiedler, G. Margulis, A. Venkatesh, *Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces*. Invent. Math. 177 (2009), no. 1, 137-212.

[12] P. Gabriel, *unpublished manuscript*.

[13] B. Hasselblatt, *Problems in dynamical systems and related topics raised in connection with the Clay Mathematics Institute/ Mathematical Sciences Research Institute, Workshop on “Recent progress in dynamics”*.

[14] R. Howe, *On a notion of rank for unitary representations of the classical groups*, Harmonic analysis and group representations, 223-331, Liguori, Naples, 1982.

[15] A. Kanigowski, K. Vinhage, D. Wei, *Slow entropy of some parabolic systems*, submitted.

[16] A. Kanigowski, K. Vinhage, D. Wei, *Kakutani equivalence of unipotent flows*, submitted.

[17] A. Katok, *Bernoulli diffeomorphisms on surfaces*, Ann. of Math., 110 (1979), 529-547.

[18] A. Katok, *Smooth non-Bernoulli K-automorphisms*, Invent. Math. 61 (1980), 291-300.

[19] Y. Katznelson, *Ergodic automorphisms of $T^n$ are Bernoulli shifts*, Israel J. Math., 10 (1971), 186-195.

[20] D. Kleinbock, N. Shah, A. Starkov, *Dynamics of subgroup actions on homogeneous spaces of lie groups and applications to number theory*, Chapter 11, Handbook of Dynamical Systems Volume 1, Part A, 2002, 813-930.

[21] D. Kleinbock, G. Margulis *Bounded orbits of nonquasiunipotent flows on homogeneous spaces*, Amer. Math. Soc. Transl. 171 (1996) 141-172.

[22] F. Ledrappier, *Some properties of absolutely continuous invariant measures on an interval*, Ergod. Th. Dynam. Syst., 1 (1981), 77-94.

[23] D. A. Lind, *Ergodic automorphisms of the infinite torus are Bernoulli*, Israel J. Math., 17, (1974), 162-168.

[24] G. Miles, R. K. Thomas, *Generalized torus automorphisms are Bernoullian*, Adv. in Math Suppl. Stud., 2, Academic Press, New York-London, 231-249.
[25] G. Miles, R. K. Thomas, The breakdown of automorphisms of compact topological groups, Adv. in Math Suppl. Stud., 2, Academic Press, New York-London, 207-218.

[26] C. Moore, Exponential decay of correlation coefficients for geodesic flows, Group representations, ergodic theory, operator algebras, and mathematical physics (Berkeley, Calif., 1984), 163-181, Math. Sci. Res. Inst. Publ., 6, Springer, New York, 1987.

[27] D. W. Morris, Dani’s work on dynamical systems on homogeneous spaces, Amer. Math. Soc., Providence, RI, 2015, 131-144.

[28] H. Oh, Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants, Duke Math. J. 113 (2002), no. 1, 133-192.

[29] D. Ornstein, Bernoulli shifts with the same entropy are isomorphic, Advances in Math., Vol. 4 (1970), 337-352.

[30] D. Ornstein, A mixing transformation for which Pinsker’s conjecture fails, Adv. in Math., 10, 1973, 103-120.

[31] D. Ornstein, Factors of Bernoulli shifts are Bernoulli, Adv. Math. 5 (1961), 349-364.

[32] D. Ornstein, An example of a Kolmogorov automorphism that is not a Bernoulli shift Advances in Math. (10) 1973, 49-62.

[33] D. Ornstein, A survey of some recent results in ergodic theory, Studies in probability theory, MAA Stud. Math. 18, Math. Assoc. America, Washington, D. C. 1978.

[34] D. Ornstein, B. Weiss, Geodesic flows are Bernoullian, Israel J. Math., 14 (1973), 184-198.

[35] Y. B. Pesin, Geodesic flows on closed Riemannian manifolds without focal points, Math. USSR Izv., 11 (1977), 1195-1228.

[36] M. Ratner, Bernoulli flows over maps of the interval, Israel J. Math., 31 (1978), 298-314.

[37] M. Ratner, Anosov flows with Gibbs measures are also Bernoullian, Israel J. Math., 17 (1974), 380-391.

[38] F. Rodriguez-Hertz, M. Rodriguez-Hertz, R. Ures, A survey of partially hyperbolic dynamics, Fields Inst. Commun. 51 (2007),35-88.

[39] V. A. Rohlin, Y. G. Sinai, The structure and properties of invariant measurable partitions, Dokl. Akad. Nauk SSSR, 141, (1961),1038-1041.

[40] D. Rudolph, If a finite extension of a Bernoulli shift has no finite rotation factors, it is Bernoulli, Israel J. Math., 30 (1978), 193-206.