Kakutani equivalence of unipotent flows

Adam Kanigowski Kurt Vinhage∗ Daren Wei†

Abstract

We study Kakutani equivalence in the class of unipotent flows acting on compact quotients of semisimple Lie groups. For every such flow we compute the Kakutani invariant of M. Ratner, the value of which being explicitly given by the Jordan block structure of the unipotent element generating the flow. This, in particular, answers a question of M. Ratner. Moreover, it follows that the only standard unipotent flows are given by
\[
\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \times \text{id}
\]
acting on \( SL(2, \mathbb{R}) \times G'/\Gamma' \), where \( \Gamma' \) is an irreducible lattice in \( SL(2, \mathbb{R}) \times G' \) (with the possibility that \( G' = \{ e \} \)).

1 Introduction

Classical ergodic theory studies representations of a group \( G \) as automorphisms of measure spaces: \( g \mapsto T_g \). For such group actions there is a natural notion of isomorphism: two group actions \( (T_g)_{g \in G} \) on \( (X, \mu) \) and \( (S_g)_{g \in G} \) on \( (Y, \nu) \) are (measure theoretically) isomorphic if there exists a (measure preserving, invertible) map \( R : (X, \mu) \to (Y, \nu) \) such that
\[
R \circ T_g = S_g \circ R, \quad \text{for all } g \in G.
\]

In what follows, we will consider the classical setting, where the acting group is \( \mathbb{Z} \) or \( \mathbb{R} \), corresponding to the iteration of a single automorphism, or flow along a one-parameter measurable family of automorphisms, respectively. Classifying \( \mathbb{Z} \) or \( \mathbb{R} \) actions up to isomorphism is too difficult of a problem in full generality (see e.g. [2], [12], [11]). A much weaker equivalence relation than isomorphism is that of orbit equivalence. We recall, that \( (T_g)_{g \in G} \) and \( (S_g)_{g \in G} \) are called orbit equivalent if there exists a measure preserving, invertible map \( R : (X, \mu) \to (Y, \nu) \) such that \( R \circ T_g = S_g \circ R \), for all \( g \in G \).

In what follows, we will consider the classical setting, where the acting group is \( \mathbb{Z} \) or \( \mathbb{R} \), corresponding to the iteration of a single automorphism, or flow along a one-parameter measurable family of automorphisms, respectively. Classifying \( \mathbb{Z} \) or \( \mathbb{R} \) actions up to isomorphism is too difficult of a problem in full generality (see e.g. [2], [12], [11]). A much weaker equivalence relation than isomorphism is that of orbit equivalence. We recall, that \( (T_g)_{g \in G} \) and \( (S_g)_{g \in G} \) are called orbit equivalent if there exists a measure preserving, invertible map \( R \) taking \( (T_g)_{g \in G} \) orbits to \( (S_g)_{g \in G} \) orbits (as sets). This notion is however too weak as according to Dye’s theorem, [6], [7] it follows that for \( G = \mathbb{Z} \) (or \( \mathbb{R} \)), any two ergodic measure preserving actions are orbit equivalent.

For \( \mathbb{Z} \) and \( \mathbb{R} \) actions, an equivalence relation weaker than isomorphism but stronger than orbit equivalence was introduced by S. Kakutani [13]. Following [13], we say that two \( \mathbb{Z} \) actions \( T \) and \( S \) are Kakutani equivalent if there exist measurable sets \( A \subset X \) and \( B \subset Y \) such that \( (T|_A, A, \mu_A) \) and \( (S|_B, B, \nu_B) \) are isomorphic, where \( T|_A \) and \( S|_B \) denote the corresponding induced isomorphisms and \( \mu_A \) and \( \nu_B \) denote the induced measures.

Analogously, we say that two \( \mathbb{R} \)-actions \( (T_t)_{t \in \mathbb{R}} \) and \( (S_t)_{t \in \mathbb{R}} \) are Kakutani equivalent if there exists an \( L^1(X, \mu) \) time change of \( (T_t)_{t \in \mathbb{R}} \) which is isomorphic with \( (S_t)_{t \in \mathbb{R}} \) (see Definition

∗K. V. was supported by the National Science Foundation under Award DMS 1604796
†D. W. was partially supported by the NSF grant DMS-16-02409
Kakutani originally called this notion weak isomorphism, but as this terminology was later used in a different context, modern treatments use Kakutani equivalence instead.

By Abramov’s formula, \[1\], it follows that Kakutani equivalence preserves the class of zero-entropy, finite entropy systems and infinite entropy systems. In the present paper we study the zero entropy case.

A. Katok, \[15\], showed that any two ergodic actions with discrete spectrum are Kakutani equivalent. In particular, we call an automorphism \( T \) (a flow \( (T_t)_{t \in \mathbb{R}} \)) standard or loosely Bernoulli of zero entropy if it is Kakutani equivalent to an irrational rotation (to a linear flow on \( \mathbb{T}^2 \)) \footnote{Notice that standardness implies ergodicity. By the above result of A. Katok, all irrational rotations (linear flows on \( \mathbb{T}^2 \)) are Kakutani equivalent.}. Kakutani originally conjectured that all zero entropy systems were standard (although he did not use this terminology) \[13\]. It turns out that the class of standard systems is quite broad, it contains all systems of local rank one \[10\] and is closed under factors, inverse limits and compact extensions, \[15\], \[20\], \[3\]. Hence, all distal systems are standard, and, in particular, all nil-systems are standard.

The first non-standard system of zero entropy was constructed by J. Feldman, \[9\], by the cutting and stacking method. Later, A. Katok, \[15\], and D. Ornstein, D. Rudolph, and B. Weiss, \[20\], independently, constructed uncountably many non-Kakutani equivalent zero entropy systems. However, these systems were manufactured to be non-standard and were not systems of general interest. Instead, they were created via certain combinatorial constructions which were later shown to have smooth models. In fact, until now, Kakutani classification of smooth zero-entropy systems which were not created solely for this purpose, has only been answered in a few special cases by M. Ratner. Namely, in \[21\], it is shown that horocycle flows \( (h_t)_{t \in \mathbb{R}} \) acting on finite volume quotients of \( SL(2, \mathbb{R}) \) are standard. Then, in \[22\], it was shown that \( h_t \times h_t \) acting on \( SL(2, \mathbb{R})/\Gamma \times SL(2, \mathbb{R})/\Gamma \), the cartesian square of these systems, is not standard, for any (hyperbolic) cocompact lattice \( \Gamma \). Finally, in \[23\], it was shown that the product of \( k \)-copies of \( (h_t) \) is not Kakutani equivalent to the product of \( l \)-copies with \( k \neq l \). The method in \[23\] was to introduce, for a general flow \( (T_t) \), an invariant of Kakutani equivalence, which was called the Kakutani invariant and denoted by \( e((T_t), \log) \), which then was estimated to be different for \( (h_t)^k \) and \( (h_t)^l \).

Notice that these examples come from a very specific class: unipotent flows on quotients of semisimple Lie groups. The study of the Kakutani invariant for these flows was suggested by M. Ratner (see Problem 1, \[25\]). In this class, all previous methods require the use of certain properties of the lattice action. As a result, results were limited to the very restricted class of products of \( SL(2, \mathbb{R}) \) with reducible lattices. For many years, the study of the Kakutani equivalence for unipotent flows had no progress in view of these limitations. In fact, since the work of M. Ratner in the 1980s, no progress was made on the question of Kakutani equivalence for any naturally defined systems. The results of the present paper represent the first major step forward in over thirty years for our understanding of Kakutani equivalence of algebraic actions. We show that for every unipotent flow on a semisimple Lie group quotient, the Kakutani invariant can be explicitly computed from the Jordan block structure of the unipotent element that generates the flow (see Definition \[3.8\]).

There is a remarkable difference between the semisimple and nilpotent cases: for the semisimple case, there is a nontrivial but explicit formula for the Kakutani invariant in terms of the slow entropy. In particular, by Corollary \[1.2\] there are very few unipotent flows
which are standard. In the nilpotent case, the slow entropy can be arbitrarily large, but the resulting systems are always standard. At first glance this may be quite a surprise, since the local behavior of unipotent flows on quotients of semisimple groups and nilmanifolds are very similar. In fact, there is a unified argument that shows that the slow entropy of these systems does not see the global structure of these groups (see [17]). However, in the nilmanifold case, the directions in which the maximal divergence is seen are not mixed with the directions which cause divergence, even after recurrence. These directions are central in the group and descend tori on the nilmanifold. But in the semisimple case, the directions in which divergence are seen are mixed with the remaining directions which cause divergence after they recur.

The proof we implement here is not an adaptation of Ratner’s argument in [22] and [23], which uses specific properties of hyperbolic lattices in $PSL(2, \mathbb{R})$ and their boundary actions. Instead, we replace it by using multi-scale analysis, which controls orbits on intermediate scales, combined with a polynomial divergence property, which generalizes the divergence properties of horocycle flows and was first observed in [28]. This has many advantages over previously used technology, as it works for arbitrary semisimple groups. In particular, we use only very coarse properties of these groups such as exponential volume growth, estimates on the number of lattice points in balls, and the existence of certain renormalizing flows which interact in special ways with the unipotent flows (see Section 3.3.1).

1.1 Statement of Main Results

In what follows $G$ is a semisimple linear Lie group, and $\Gamma$ is a cocompact lattice in $G$. Let moreover $\mathfrak{g} = \text{Lie}(G)$ denote the Lie algebra of $G$. A flow $(\phi_t)$ on $G/\Gamma$ is called unipotent, if $\phi_t$ is the left translation action by $\exp(tU)$, where $U \in \mathfrak{g}$ is such that $\text{ad}_U^k = 0$ for some $k$, where $\text{ad}_U \in \text{End}(\mathfrak{g})$ is the adjoint operator, $\text{ad}_U(X) = [U,X]$. The flow $(\phi_t)$ preserves Haar measure $\mu$ on $G/\Gamma$. We may also associate a list of numbers $(m_1, \ldots, m_n)$ called the chain structure of $U$ which are the sizes of the Jordan blocks for $\text{ad}_U$ (see Definition 3.8). Then, we have the following invariant which is the growth rate or slow entropy of $(\phi_t)$:

$$GR(U) := \frac{1}{2} \sum_{i=1}^{n} m_i (m_i + 1).$$

As shown in [17], the number $GR(U)$ describes the asymptotic orbit growth (both in the topological and metric category). Moreover (see Section 3.4), it follows that $GR(U) \geq 3$. The main theorem is the following (see Definition 2.3):

**Theorem 1.1.** Let $G$ be a semisimple linear Lie group and $(\phi_t) = L_{\exp(tU)}$ a unipotent flow on $G/\Gamma$, where $\Gamma$ is cocompact. We have

$$e((\phi_t), \log) = GR(U) - 3.$$ 

Moreover, if $GR(U) = 3$, then $(\phi_t)$ is standard.

By a direct computation, one gets $e((h_t)^k, \log) = 3k - 3$. This, in particular, generalises M. Ratner’s result, [23] to any co-compact lattice on $SL(2, \mathbb{R})^k$. Theorem 1.1 allows one to deduce the following immediately from Lemma 3.9.

---

2 We do not lose much generality in assuming that $G$ is a linear Lie group, as any Lie group is a discrete cover of some matrix group
Corollary 1.2. The only ergodic unipotent flows on compact quotients of linear semisimple Lie groups which are standard are of the form \( \phi_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \times \text{id} \) acting on \((SL(2, \mathbb{R}) \times G')/\Gamma\), where \( \Gamma \) is irreducible.

Theorem 1.1 gives a solution to M. Ratner’s Problem 1 in [25] (see also [18]):

Corollary 1.3. Let \( G \) be a linear semisimple Lie group with \( \dim G > 3 \), and \( G/\Gamma \) be a compact homogeneous space of \( G \).

(i) There are ergodic unipotent flows on \( G/\Gamma \) which are not standard.

(ii) If \( G \) is simple, no unipotent flow on \( G/\Gamma \) is standard.

(iii) If \( G \) has real rank at least two, there are two unipotent flows on \( G/\Gamma \) (which are not identity, but not necessarily ergodic) which are not Kakutani equivalent.

(iv) If \( G \cong SL(d, \mathbb{R}) \), then there are at least \( d - 1 \) flows on \( G/\Gamma \) which are pairwise non-Kakutani equivalent.

In fact, we expect the number of pairwise non-Kakutani equivalent flows on \( SL(d, \mathbb{R})/\Gamma \) to grow on the order of \( d^3 \) (see Remark 3.10). A proof of Corollary 1.3 is given in Section 3.4.

Moreover, our main result also allows one to construct algebraic examples which answer negatively the following question by A. Katok, [15]: if \( T \circ S = S \circ T \) (i.e. \( S \in C(T) \)) and \( T \) is standard, does it follow that \( S \) is standard? The first such counterexamples were constructed by de la Rue in [5]. However, these examples were Gaussian systems which are not known to have smooth finite dimensional models.

Corollary 1.4. Let \( T = h_1 \times \text{id} \) and \( S = h_1 \times h_1 \) acting on \( SL(2, \mathbb{R})^2/\Gamma \) with \( \Gamma \) irreducible. Then \( S \) and \( T \) are ergodic and commute, with \( T \) standard and \( S \) non-standard.

We finish the introduction with the following questions:

**Question 1.** When is the flow \( (\phi_t) = L^{\exp(u)} \) acting on \( G/\Gamma \) Kakutani equivalent with its action on \( G'/\Gamma' \)?

Notice that if \( \Gamma \) and \( \Gamma' \) are conjugated, then the actions are isomorphic and hence Kakutani equivalent. Therefore the interesting case is to consider question 1 for \( \Gamma \) and \( \Gamma' \) which are not algebraically related.

The above question is a particular case of the following general question:

**Question 2.** Let \( U \in \mathfrak{g} \) and \( U' \in \mathfrak{g}' \). Is it true that if \( GR(U) = GR(U') \) then the flows \( (\phi_t) = L^{\exp(u)} \) and \( (\phi'_t) = L^{\exp(u')} \) are Kakutani equivalent?

A positive answer to Question 2 would mean that the Kakutani invariant is a full invariant in the class of unipotent flows (the same way as Kolmogorov-Sinai entropy is a full invariant for Bernoulli shifts, [19]). Notice also that Question 1 is a special case of Question 2.

Notice that in Theorem 1.1 we use the log function to compute the Kakutani invariant \( e((\phi_t), \log) \). In general (see [26]), one may consider any \( u : [0, +\infty) \rightarrow [0 + \infty) \) such that \( \lim_{t \rightarrow +\infty} \frac{u(at)}{u(t)} = 1 \), for any \( a > 0 \). We have the following general problem:

**Problem 1:** For any function \( u \) as above construct a flow \( (T_t) \) such that \( 0 < e((T_t), u) < +\infty \).
Notice that it is much harder to construct systems with a prescribed Kakutani invariant than with the Hamming one (this invariant is called slow entropy in [16]). Indeed, it follows from [16] that (for natural systems such as unipotent flows) slow entropy behaves well under taking products, which is not the case for the Kakutani invariant, as is demonstrated by considering \((h_t)\) on \(SL(2, \mathbb{R})/\Gamma\) and \((h_t \times h_t)\), first considered in [22].

**Acknowledgements** The authors would like to thank Anatole Katok for suggesting this problem and his encouragements in its development. The authors are also grateful to Federico Rodriguez-Hertz, and Jean-Paul Thouvenot for offering their insight on the subject, as well as Dmitry Dolgopyat and Mariusz Lemańczyk on their useful comments on a preliminary version of the paper.

### 1.2 A Reader’s Guide

We write the paper with readers from two distinct fields in mind: ergodic theory and measurable invariants, and Lie groups and homogeneous dynamics. We therefore include a section to describe some standard tools from each (Sections 2 and 3). In Section 4 we combine ideas from each of these fields to make definitions which allow us to analyze the decay rate of Kakutani balls. Section 5 contains some algebraic lemmas which are applied in Sections 6 and 7. Since some proofs have a clear main idea but are technical, we have included outlines of each important reduction (before its proof) to explain what the technicalities mean intuitively.

The key technique of the paper is developing new counting results for the Kakutani invariant. The main idea is that if two points are Kakutani close (which, in general, is very hard to control), then they are algebraically close on a long block. In particular, the first reduction of Theorem 1.1 is in Section 6 (Theorem 6.1), which relates Kakutani balls to “Bowen-like” balls (Definition 4.2). These are algebraically, and not dynamically, defined and we can obtain good estimates on their decay rates. Therefore, the main purpose of Theorem 6.1 is to relate the dynamically defined Kakutani balls with a more algebraic definition.

The main difficulty with the Kakutani invariant is that the dynamical criterion for being in a Kakutani ball does not give us control over the full orbit. Therefore, the strength of Theorem 6.1 is the guarantee that this does happen: not only for a large proportion of time do we have closeness of orbits, but also for a very long interval.

The remainder of the paper is dedicated to the proof of Theorem 6.1. A series of further reductions to prove Theorem 6.1 are made in Section 7. The main idea is the following: to guarantee a long interval in which orbits are close, and not just a large proportion of time as guaranteed by the Kakutani condition, one must show that orbits cannot align, separate, and realign in a negligible amount of time on large scales. Proposition 7.1 is a way to guarantee that this cannot happen: for any matching of orbits, the smaller segments of matching times cannot take up a large portion of the matching interval.

Let us point out that our technique is different from Ratner’s methods from [22], [23]. Indeed, the methods in [22] and [23] are crucially based on the fact that the lattice is a product of hyperbolic lattices in \(SL(2, \mathbb{R})\). Our method is based on controlling the algebraic (polynomial) divergence of the unipotent flow and not on controlling the behaviour of the return times using finer properties of the lattice. The details will be explained more in future sections.
Remark 1.5. The only place were we use the fact that $G$ is linear is the computation in Appendix, where we compute products of elements from the $\mathfrak{sl}(2,\mathbb{R})$-triple in $G$. If $G$ is linear, it follows that the homomorphism $\phi : \mathfrak{sl}(2,\mathbb{R}) \rightarrow \mathfrak{g}$ lifts to a homomorphism $\Phi : SL(2,\mathbb{R}) \rightarrow G$ (and not just its universal cover, which is all that is guaranteed from general Lie theory), see Lemma 3.4. This allows us to make computations in $SL(2,\mathbb{R})$ and conclude things about the corresponding products in $G$.

2 Preliminaries on Kakutani Equivalence

In this section we will introduce some basic definitions. We first recall the definition of Kakutani equivalence. For a flow $(T_t)$ on $(X, \mathcal{B}, \nu)$ and a function $\alpha \in L^1_+(X, \mathcal{B}, \nu)$, the flow $(T_t^\alpha)$ is called a time change of $(T_t)$ (along $\alpha$) if

$$T_t^\alpha(x) = T_{u(t,x)}(x),$$

where $u(t, x)$ is a (unique) solution to

$$\int_0^{u(t,x)} \alpha(T_s x) ds = t.$$

it follows that $(T_t^\alpha)$ preserves measure $d\nu := \alpha(\cdot)d\nu$.

Definition 2.1 (Kakutani equivalence, [14]). Two ergodic measure preserving flows $(T_t)$ on $(X, \mathcal{B}, \nu)$ and $(S_t)$ on $(\tilde{X}, \mathcal{C}, \tilde{\nu})$ are Kakutani equivalent, if $(S_t)$ is isomorphic to $(T_t^\alpha)$ for some $\alpha \in L^1_+(X, \mathcal{B}, \nu)$.

Following [26], we will introduce the Kakutani invariant for an ergodic flow $(T_t)$ acting on a Lebesgue space $(X, \mathcal{B}, \nu)$. For a finite measurable partition $\mathcal{P}$ of $X$ and an element $x \in X$, we denote by $\mathcal{P}(x)$ the atom of $\mathcal{P}$ containing $x$ and let $I_R(x) := \{T_s x : s \in [0, R] \}$. Let $l$ denote the Lebesgue measure on $[0, R]$.

Definition 2.2 ($(\varepsilon, \mathcal{P})$-matchable, [26]). For $x, y \in X$, $\varepsilon > 0$ and $R > 1$, $I_R(x)$ and $I_R(y)$ are called $(\varepsilon, \mathcal{P})$-matchable if there exists a subset $A = A(x, y) \subset [0, R]$, $l(A) > (1-\varepsilon)R$ and an increasing absolutely continuous map $h = h(x, y)$ from $A$ onto $A' = A'(x, y) \subset [0, R]$, $l(A') > (1-\varepsilon)R$ such that $\mathcal{P}(T_t x) = \mathcal{P}(T_{h(t)} y)$ for all $t \in A$ and the derivative $h' = h'(x, y)$ satisfies

$$|h'(t) - 1| < \varepsilon \text{ for all } t \in A.$$  \hspace{1cm} (2)

We call $h$ an $(\varepsilon, \mathcal{P})$-matching from $I_R(x)$ onto $I_R(y)$.

The Kakutani invariant is defined based on the above definition.

Definition 2.3 (Kakutani invariant, [26]). Define

$$f_R(x, y, \mathcal{P}) = \inf\{\varepsilon > 0 : I_R(x) \text{ and } I_R(y) \text{ are } (\varepsilon, \mathcal{P})-\text{matchable}\}.$$ 

Then denote $B_R(x, \varepsilon, \mathcal{P}) = \{y \in X : f_R(x, y, \mathcal{P}) < \varepsilon\}$ as $(R, \mathcal{P})$-ball of radius $\varepsilon > 0$ centered at $x \in X$, $R > 1$. A family $\alpha_R(\varepsilon, \mathcal{P})$ of $(R, \mathcal{P})$-balls of radius $\varepsilon > 0$ is called $(\varepsilon, R, \mathcal{P})$-cover of $X$ if $\nu(\cup \alpha_R(\varepsilon, \mathcal{P})) > 1 - \varepsilon$. Denote $K_R(\varepsilon, \mathcal{P}) = \inf |\alpha_R(\varepsilon, \mathcal{P})|$ where $|A|$ denotes the cardinality of $A$ and infimum is taken over all $(\varepsilon, R, \mathcal{P})$-covers of $X$. Let $\mathcal{F}$
denote the family of all nondecreasing functions from $\mathbb{R}^+$ onto itself, converging to $+\infty$. For $u \in F$, we denote,

$$\beta(u, \varepsilon, P) = \liminf_{R \to \infty} \frac{\log K_R(\varepsilon, P)}{u(t)};$$
$$e(u, P) = \limsup_{\varepsilon \to 0} \beta(u, \varepsilon, P);$$
$$e((T_t), u) = \sup_{P} e(u, P).$$

We also recall the following theorems, the first one is the generator theorem.

**Theorem 2.4** (Generator theorem, [26]). Let $(T_t)$ be an ergodic measure-preserving flow on $(X, \mathcal{B}, \nu)$ and let $\mathcal{P}_1 \leq \mathcal{P}_2 \leq \ldots$ be an increasing sequence of finite measurable partitions of $X$ such that $\bigvee_{n=1}^{\infty} \mathcal{P}_n$ generates the $\sigma$–algebra $\mathcal{B}$. Then $e((T_t), u) = \sup_{P} e(u, \mathcal{P}_m)$ for all $u \in F$.

The following theorem shows that the above quantity is an invariant of Kakutani equivalence.

**Theorem 2.5** ([26]). Let $(T_t)$ and $(S_t)$ be two ergodic Kakutani equivalent measure preserving flows on $(X, \mathcal{B}, \nu)$ and $(\tilde{X}, \tilde{\mathcal{B}}, \tilde{\nu})$. Then $e((T_t), u) = e((S_t), u)$ for all $u \in F$ with $\lim_{t \to \infty} u(at) = 1$ for all $a > 0$.

Moreover, we have the following theorem (see e.g. [26]):

**Theorem 2.6.** A zero-entropy ergodic measure preserving flow $(T_t)$ is standard if and only if $e((T_t), u) = 0$ for all $u \in F$.

We will also use the following definition of matching balls:

**Definition 2.7.** Fix $\varepsilon > 0$, let $x, y \in M$ be $(\varepsilon, P)$-matchable (see Definition 2.2) and let $h : A(x, y) \to A'(x, y)$ be an $(\varepsilon, P)$-matching. For $u \in A(x, y)$ and $L > 0$ let

$$B(u, L) := \{ r \in A(x, y) : r \geq u, r - u \leq L \}.$$ 

denote the matching ball around $(u, h(u))$.

Finally, we give a simple general remark, which we will use in the proof of Theorem 1.1.

**Remark 2.8.** If there exists a set $D \subset X$, such that for every $y \in D$, we have

$$\mu(B_R(y, \varepsilon, P) \cap D) \leq a(R, \varepsilon),$$

for $\varepsilon < \mu(D)$, then $K_R(\varepsilon/5, P) \geq \frac{1}{a(R, \varepsilon)}$.

On the other hand if for every $y \in D_\varepsilon$, $\mu(D_\varepsilon) \geq 1 - \varepsilon$, we have

$$\mu(B_R(y, \varepsilon, P)) \geq b(R, \varepsilon),$$

then $K_R(5\varepsilon, P) \leq \frac{1}{a(R, \varepsilon)}$.

We recall also that $f_1(\cdot, \cdot, \mathcal{P})$ does not define a metric (triangle inequality fails), however it is close to a metric: if $x \in B_R(y, \varepsilon, \mathcal{P})$ and $y \in B_R(z, \varepsilon, \mathcal{P})$, then $x \in B_R(z, 5\varepsilon, \mathcal{P})$. 

3 Preliminaries on Homogeneous Spaces

In this section, we recall some basic facts from the theory of Lie groups and homogeneous spaces. Throughout the paper $G$ will denote a semisimple Lie group with Lie algebra $\mathfrak{g}$. Given $g \in G$, let $L_g, R_g : G \to G$ denote the left and right translations on $G$. Let $\exp : \mathfrak{g} \to G$ denote the exponential mapping of the Lie algebra $\mathfrak{g}$ onto $G$. Then $\exp$ has a local inverse $\log$ sending a neighborhood of $e \in G$ to a neighborhood of $0 \in \mathfrak{g}$.

3.1 Metrics and Measures on Homogeneous Spaces

Let $\Gamma \subset G$ be a (discrete) subgroup. We introduce a metric on the homogeneous space $\Gamma \backslash G$ by first introducing a right-invariant metric on $G$. Fix an inner product $\langle \cdot, \cdot \rangle_0$ on $\mathfrak{g}$, and define for $v, w \in T_g G$:

$$\langle v, w \rangle = \langle dR_g^{-1}v, dR_g^{-1}w \rangle_0$$

By construction, $\langle \cdot, \cdot \rangle$ is right-invariant, so it induces a Riemannian metric on the space $\Gamma \backslash G$. The Riemannian metric also has an associated exponential mapping $\exp_{\text{geom}} : \mathfrak{g} \to G$, which is $C^\infty$ and satisfies

$$d_0 \exp_{\text{geom}} = \text{id}. \quad (4)$$

Like the algebraic exponential, there is a local inverse of $\exp_{\text{geom}}$ which we will denote by $\log_{\text{geom}}$. The following is immediate from the definition of the inner product.

**Lemma 3.1.** The Riemannian volume is a (right) Haar measure on $G$. In particular, it is independent of the metric $\langle \cdot, \cdot \rangle_0$ when determining a probability measure on a homogeneous space.

3.2 The Adjoint Representation

$G$ acts on itself by conjugation $C_g : h \mapsto g^{-1}hg$, and taking the derivative at the identity in the coordinate $h$ gives the adjoint representation of $G$ on $\mathfrak{g} = T_e G$, $\text{Ad} : G \to GL(\mathfrak{g})$. Taking the derivative of this map in the $\mathfrak{g}$ coordinate yields the Adjoint representation of the Lie algebra $\mathfrak{g}$, $\text{ad} : \mathfrak{g} \to \text{End}(\mathfrak{g})$, which coincides with the Lie bracket: $\text{ad}(X)Y = [X,Y]$. The following are standard tools from the theory of Lie groups, which we write as a Lemma to reference.

**Lemma 3.2.** If $X, Y \in \mathfrak{g}$,

$$\exp_{\text{alg}}(-X) \exp_{\text{alg}}(Y) \exp_{\text{alg}}(X) = \exp_{\text{alg}}(\text{Ad}(\exp_{\text{alg}}(X))Y)$$

$$\exp(\text{ad}(X)) := \sum_{k=0}^{\infty} \frac{\text{ad}(X)^k}{k!} = \text{Ad}(\exp_{\text{alg}}(X))$$
3.3 Decompositions and Subgroups of Semisimple Groups

3.3.1 $\mathfrak{sl}(2, \mathbb{R})$ triples

Let $V = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $X = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $U = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ be the standard generators of $\mathfrak{sl}(2, \mathbb{R})$.

Let $G$ be a simply-connected semisimple Lie group of rank $r$, and $\mathfrak{g} = \text{Lie}(G)$. We abusively let $U \in \mathfrak{g}$ denote an arbitrary unipotent Lie algebra element (i.e., an element such that $0$ is the only eigenvalue of $U$). This is because given any unipotent element, there exists a homomorphism $\varphi : \mathfrak{sl}(2, \mathbb{R}) \to \mathfrak{g}$ such that $\varphi(U)$ is this given element. While this homomorphism is not unique, it is unique up to automorphism of $\mathfrak{g}$ fixing $U$. We therefore identify $\mathfrak{sl}(2, \mathbb{R})$ with its image under $\varphi$.

Given a subalgebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$ of $\mathfrak{g}$, we may consider the action $\text{ad} : \mathfrak{sl}(2, \mathbb{R}) \to \text{End}(\mathfrak{g})$ which maps $X \mapsto \text{ad}_X$. Since it is a subalgebra, this is a representation of $\mathfrak{sl}(2, \mathbb{R})$. Since $\mathfrak{sl}(2, \mathbb{R})$ is a semisimple algebra, this representation splits as a sum of irreducible representations. The irreducible representations of $\mathfrak{sl}(2, \mathbb{R})$ classified up to isomorphism, with classes indexed by $\mathbb{N}$. Let $E_n$ be an $(n + 1)$-dimensional real vector space generated by vectors $X_{2k-n}$, $k = 0, \ldots, n$. Then there exist nonzero constants $a_{n,k}$ such that:

\[
\begin{align*}
\pi_n(U)X_{2k-n} &= X_{2k-n+2} \\
\pi_n(X)X_{2k-n} &= (2k - n)X_{2k-n} \\
\pi_n(V)X_{2k-n} &= a_{n,k}X_{2k-n-2}
\end{align*}
\]

where we assume $U$ sends $X_n$ to $0$ and $V$ sends $X_{-n}$ to $0$. Note that the first three are special cases: $\pi_0$ is the trivial representation, $\pi_1$ is the standard representation and $\pi_2$ is the adjoint representation. Given elements $Y_1, \ldots, Y_n \in \mathfrak{g}$, let $C(Y_1, \ldots, Y_n)$ denote the common centralizer of the $Y_i$. That is:

\[C(Y_1, \ldots, Y_n) = \{ H \in \mathfrak{g} : \text{ad}_Y(H) = 0 \text{ for all } i = 1, \ldots, n \}\]

The following is a straightforward finite-dimensional version of the Howe-Moore theorem:

Lemma 3.3. If $A \in C(U, X)$, then $A \in C(U, V, X)$.

Proof. Pick a basis \( \{ X^j_{2i-m_j} : i = 0, \ldots, m_j \text{ and } j = 1, \ldots, n \} \) of $\mathfrak{g}$ such that $X^j_{2i-m_j}$ span a representation $\pi_{m_j}$ of $\mathfrak{sl}(2, \mathbb{R})$ with relations determined by (5)-(7). We may write $A = \sum a_{i,j} X^j_{2i-m_j}$, and notice that $\text{ad}_X(A) = \sum a_{i,j} (2i - m_j) X^j_{2i-m_j,j}$. If $\text{ad}_X(A) = 0$, then $a_{i,j} = 0$ unless $2i = m_j$. Furthermore, $\text{ad}_Y(A) = \sum a_{i,j} X^j_{2i-m_j+2,j}$. This implies that $a_{i,j} = 0$ unless $2i - m_j = m_j$. If $a_{i,j} \neq 0$, the first condition implies $i = m_j/2$ and the second implies $i = m_j$. The only way this occurs is when $i = m_j = 0$. That is, $A$ must be a sum of vectors spanning trivial representations, and $V$ must act trivially as well.

The following lemma allows us to make computations in $SL(2, \mathbb{R})$ directly for the corresponding elements of $G$:

Lemma 3.4. Let $H \subset GL(N, \mathbb{R})$ be a Lie group and $\phi : \mathfrak{sl}(2, \mathbb{R}) \to \text{Lie}(H)$ be a homomorphism. Then there exists a unique $\tilde{\phi} : SL(2, \mathbb{R}) \to H$ such that $d\tilde{\phi} = \phi$.
Proof. Notice that \( \phi \) induces a representation of \( \mathfrak{sl}(2, \mathbb{R}) \) on \( \mathbb{R}^N \) which we denote by the same name, since \( \phi(X) \) is a matrix in \( \text{Lie}(H) \subset \mathfrak{g}(N, \mathbb{R}) \) by definition. Then we may decompose \( \mathbb{R}^N \) as a direct sum of irreducible subrepresentations \( \mathbb{R}^N = \bigoplus_{i=1}^n E_i \). But it is known that each of the irreducible finite-dimensional representations of \( \mathfrak{sl}(2, \mathbb{R}) \) lifts to a unique representation of \( SL(2, \mathbb{R}) \). Therefore, we may lift \( \phi \) by lifting in each subspace \( E_i \), and taking the corresponding direct sum of representations.

3.3.2 \( SL(2, \mathbb{R}) \) and Hyperbolic Geometry

Recall that the group \( PSL(2, \mathbb{R}) = SL(2, \mathbb{R})/\{ \pm \text{id} \} \) is isomorphic to \( \text{Isom}(\mathbb{H}^2) \), which gives a canonical action of \( SL(2, \mathbb{R}) \) on \( \mathbb{H}^2 \).

Lemma 3.5. There exists \( C > 0 \) with the following property: If \( 0 < \varepsilon < 1 \) and \( S \subset SL(2, \mathbb{R}) \) has \( S \subset B(e, R) \), then the minimal number of \( \varepsilon \)-balls in \( SL(2, \mathbb{R}) \) required to cover \( S \) is less than \( C\varepsilon^{-3}e^{2CR} \).

Proof. Notice that \( \mu(B(x, \varepsilon)) = \mu(B(y, \varepsilon)) > C_1\varepsilon^3 \) for the standard hyperbolic measure \( \mu \) and all \( x, y \in \mathbb{H}^2 \), since \( \varepsilon \) is sufficiently small. Notice also that \( \mu(B(x_0, R)) \leq C_2e^{h_G R} \), where \( h_G \) is larger the exponential volume growth rate for balls in \( G \). Therefore, one has at most \( C_2e^{h_G R}/\mu(B(e, \varepsilon/2)) = C_1^{-1}C_2(\varepsilon/2)^{-3}e^{2h_G R} \) disjoint balls in \( B(e, R) \). Any such set which is maximally chosen will also cover \( B(e, R) \) when taking \( \varepsilon \)-balls, so the result holds.

The following Lemma gives estimates on distances of the horocycle flow on \( SL(2, \mathbb{R}) \):

Lemma 3.6. There exists \( C > 0 \), \( t_0 > 1 \) such that \( d(e, \exp(tU)) \leq C \log t \) if \( t \geq t_0 \).

Proof. Write \( \exp(tU) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \) as \( \exp(tU) = k_1ak_2 \), where \( k_1, k_2 \in SO(2, \mathbb{R}) \) are rotation matrices, and \( a = \exp(sX) \) is a diagonal matrix. Observe that \( d(e, \exp(tU)) \leq d(e, k_1ak_2) \leq d(e, k_1) + d(k_1, k_1a) + d(k_1a, k_1ak_2) \leq 2D + d(e, a) \), since the metric on \( SL(2, \mathbb{R}) \) is right-invariant, and left-invariant under \( SO(2, \mathbb{R}) \). But \( d(e, a) = s \), and we may compute \( s \) by finding the eigenvalues of \( \exp(tU) \) \( \exp(tU)^T = k_1a^Tk_2^T \) \( k_1^T = k_1a^2k_1^{-1} \).

Notice that:

\[ \exp(tU) \exp(tU)^T = \begin{pmatrix} 1 + t^2 & t \\ t & 1 \end{pmatrix} \]

which has top eigenvalue \( \frac{1}{2} \left( t^2 + \sqrt{t^2 + 4t} + 2 \right) \). Therefore, the distance from \( e \) to \( \exp(sX) \) is \( \frac{1}{2} \log \left( \frac{1}{2} \left( t^2 + \sqrt{t^2 + 4t} + 2 \right) \right) \leq 2 \log t \). Therefore, by choosing \( t_0 \) large enough we get that \( d(\exp(tU), e) \leq 2 \log t + 2D \leq C \log t \).

3.3.3 Presentation of Group Elements

Let \( G \) be a Lie group, \( \mathfrak{g} = \text{Lie}(G) \), and \( \mathfrak{g} = \mathfrak{e}_1 \oplus \mathfrak{e}_2 \oplus \cdots \oplus \mathfrak{e}_n \) be a vector subspace decomposition of \( \mathfrak{g} \). We do not require that the subspaces \( \mathfrak{e}_1 \) are subalgebras or that they commute with one another. The following is an easy adaptation of the classical lemma that \( \exp : \mathfrak{g} \rightarrow G \) is a local diffeomorphism at 0.
Lemma 3.7. If \( g \in G \) is sufficiently close to \( e \in G \), then there exists unique \( X_i \in \mathfrak{e}_i \) close to 0 such that \( g = \exp(X_1) \exp(X_2) \ldots \exp(X_n) \).

Proof. Let \( \varphi : \mathfrak{g} \to G \) be the map defined via \( \varphi(X) = \exp(X_1) \exp(X_2) \ldots \exp(X_n) \), where \( X = \sum X_i \) and \( X_i \in \mathfrak{e}_i \). One can easily check that \( \varphi'(0) = \text{id} \), and hence \( \varphi \) has a local inverse at \( e = \varphi(0) \) by the inverse function theorem.

3.4 Properties of unipotent flows

The following definition is important for describing the orbit growth of a unipotent flow (see [17]).

Definition 3.8. Let \( \mathfrak{g} \) be a Lie algebra and \( U \in \mathfrak{g} \) be a unipotent element. A chain in \( \mathfrak{g} \) with respect to \( U \) of depth \( m \) is a linearly independent set \( \{ X_i : 0 \leq j \leq m \} \) such that \( X_0 \) is in the centralizer of \( U \) and:

\[
\text{ad}_U(X_i) = X_{i-1} \quad \text{for all } 1 \leq i \leq m.
\]

A chain basis of \( \mathfrak{g} \) with respect to \( U \) is a basis of chains. The sequence of depths \( (m_1, \ldots, m_n) \) of chains is called the chain structure of \( U \). We will denote the chain basis by \( \{ X_i^{(1)} \}_{i=0}^{m_1}, \{ X_i^{(2)} \}_{i=0}^{m_2}, \ldots, \{ X_i^{(n)} \}_{i=0}^{m_n} \). We will often denote chains using the notation:

\[
X_n \mapsto X_{n-1} \mapsto \cdots \mapsto X_1 \mapsto X_0
\]

While every unipotent element \( U \) has a chain basis, we will use special structures associated to semisimple groups to construct a canonical one. In particular, notice that the weight spaces for the representations of the \( \mathfrak{sl}(2, \mathbb{R}) \) triple can be taken as the chain basis by [3]. We reindex them replacing the index \( n - 2k \) by \( i \) for convenience. Therefore, the basis element \( X_i^j \) is an eigenvector for \( \text{ad}_X \) with eigenvalue \( m_j - 2i \). The elements \( V \mapsto X \mapsto U \) may be taken as a chain, so there is always at least one chain of depth 2. We call this the Jacobson-Morozov chain. This implies that any unipotent flow in a semisimple homogeneous space has \( GR(U) \geq 3 \). Call any chain of depth 0 a trivial chain. Note that trivial chains span trivial subrepresentations of \( \text{ad} \).

Lemma 3.9. Let \( \phi_t(g \Gamma) = \exp(tU)g \Gamma \) act on \( G/\Gamma \) ergodically. The following are equivalent:

1. \( GR(U) = 3 \).
2. The only nontrivial subrepresentation of \( \text{ad} \) is the Jacobson-Morozov representation
3. \( \dim G - \dim C(X) \leq 3 \)
4. \( \mathfrak{g} \cong \mathfrak{sl}(2, \mathbb{R}) \oplus \mathfrak{g}' \), \( \Gamma \) is irreducible and under this isomorphism, \( U = \left( \begin{pmatrix} 0 & 1 \\ \ast & 0 \end{pmatrix}, 0 \right) \in \mathfrak{sl}(2, \mathbb{R}) \oplus \mathfrak{g}' \).

Proof. We show that 1. \( \implies \) 2. \( \implies \) 3. \( \implies \) 2. \( \implies \) 4. \( \implies \) 1. That 1. \( \implies \) 2. is a direct consequence of the definition of \( GR \) and the fact that the Jacobson-Morozov representation has depth 2. Now assume 2. As discussed in Section 3.3.1 the trivial chains
span trivial subrepresentations for \( \text{ad} |_{\mathfrak{sl}(2,\mathbb{R})} \) and therefore \( \text{ad}_X \) also acts trivially. That is, the remaining chain basis elements are in \( C(X) \).

Now suppose 3. We claim that this implies \( \dim G - \dim C(X) = 2 \). Notice that we have the lower inequality since \( U, V \in \mathfrak{g} \), but neither \( U \) nor \( V \) commute with \( X \). We must therefore rule out the case of 3. This implies that there exists exactly one more linearly independent element which fails to commute with \( X \). But by considering each representation \( \pi_n \) as described in Section 3.3.1 we see that there are always an even number of linearly independent elements which fail to commute with \( X \) in each chain basis. In particular, we have 2.

Now assume 2. We claim that in this case \( \mathfrak{sl}(2,\mathbb{R}) \) is an ideal in \( \mathfrak{g} \). Indeed, all basis elements which are not from the Jacobson-Morozov representation act trivially on \( V \), \( X \) and \( U \) (since the Lie bracket is anti-commutative), and \( \mathfrak{g}, X, V, U \) act on each other by the standard \( \mathfrak{sl}(2,\mathbb{R}) \) relations. Since \( \mathfrak{g} \) is semisimple and \( \mathfrak{sl}(2,\mathbb{R}) \) is an ideal, there exists a complementary subalgebra \( \mathfrak{g}' \). Since the flow must be ergodic, \( \Gamma \) must be irreducible. That is, we have 4.

One can see 4. \( \implies \) 1. by direct computation. \( \square \)

**Proof of Corollary 1.3.** Let \( G \) be a semisimple Lie group of dimension at least 4. Then \( \mathfrak{g} = \text{Lie}(G) \neq \mathfrak{sl}(2,\mathbb{R}) \). If \( \mathfrak{g} \neq \mathfrak{sl}(2,\mathbb{R}) \oplus \mathfrak{g}' \), by Lemma 3.9 any unipotent \( U \) has \( GR(U) > 3 \). In particular, any ergodic unipotent flow on the quotient of a simple Lie group other than covers of \( SL(2,\mathbb{R}) \) are non-standart (this proves (ii)). If \( \mathfrak{g} = \mathfrak{sl}(2,\mathbb{R}) \oplus \mathfrak{g}' \), then \( \mathfrak{g}' \) is also semisimple (since \( \mathfrak{g} \) is semisimple). Therefore, each simple factor has a Cartan subalgebra, with associated roots, and in particular, has unipotent elements of the root spaces. Take \( U' = U + \sum U_{\alpha_i} \), where \( U_{\alpha_i} \) is an element from a root space in each simple factor of \( \mathfrak{g}' \). By the Howe-Moore ergodicity theorem, the action of \( U' \) is ergodic, and \( GR(U') > 3 \) since each \( U_{\alpha_i} \) has its own \( \mathfrak{sl}(2,\mathbb{R}) \) triple in its semisimple factor (the \( \mathfrak{sl}(2,\mathbb{R}) \) triple for \( \mathfrak{g} \) will be the sum of the elements from each factor). Therefore, we have produced an ergodic unipotent flow which is non-standart (this proves (i)).

Finally, assume that \( \text{rank}_\mathbb{R}(G) \geq 2 \), and let \( \mathfrak{a} \) denote a Cartan subalgebra of \( \mathfrak{g} \). We have shown \( G \) carries an ergodic unipotent flow, so if there is one which is not ergodic, we produce two flows which are not Kakutani equivalent. Therefore, we need to produce elements \( U \) and \( U' \) such that \( GR(U) \neq GR(U') \). If \( G \) is not simple, we take some \( U \) in one factor and \( V \) in another, and set \( U' = U + V \). This clearly yields two flows with different values for \( GR \). If \( G \) is simple and has rank at least 2, we may decompose \( \mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha \) into a root space decomposition for some \( \mathfrak{a} \subset \mathfrak{g}_0 \subset \mathfrak{g} \). Here, \( \mathfrak{g}_0 \) is the centralizer of \( \mathfrak{a} \), a split Cartan subalgebra of \( \mathfrak{g} \), and \( \Delta \) is a set of \( \mathbb{R} \)-valued functionals on \( \mathfrak{a} \) such that if \( X \in \mathfrak{a} \) and \( Y \in \mathfrak{g}_\alpha \), then \([X,Y] = \alpha(X)Y \). Now take any two roots \( \alpha_1, \alpha_2 \in \Delta \) which are non-proportional, and let \( U_{\alpha_i} \in \mathfrak{g}_{\alpha_i} \) be elements of the root spaces. We set \( U = U_{\alpha_1} \) and \( U' = U_{\alpha_1} + U_{\alpha_2} \).

Let \( X_m \mapsto \cdots \mapsto X_1 \mapsto X_0 \mapsto 0 \) be any chain for \( U' \). We will show that from this we may produce chains for \( U \), hence we may choose a chain for \( U \) as chains smaller than that of \( U' \). If we show that there is at least one chain that is broken up into two smaller ones, then we get strict inequality, as desired. Write \( X_i \) as a sum of root spaces \( X_i = \sum_\beta X_i^\beta \), then \( \text{ad}_{U_{\alpha_1}}(X_i) = \sum_\beta \alpha(X_i^\beta) = \sum_\beta X_i^{\beta + \alpha_1} \). We may choose \( \alpha_1 \) and \( \alpha_2 \) as simple roots, and since all roots are integral linear combinations of the simple roots, every \( \beta \in \Delta \) has uniquely determined integer coefficients for \( \alpha_1 \) and \( \alpha_2 \), call them \( l_1(\beta) \) and \( l_2(\beta) \). Let
\(\mathfrak{g}_k = \bigoplus_{\beta: l(\beta) = k} \mathfrak{g}_\beta\). Then \(\text{ad}_U\) and \(\text{ad}_{U'}\) both map \(\mathfrak{g}_k\) to \(\mathfrak{g}_{k+1}\). Hence we may choose chain bases for both so that \(X_i^j \in \mathfrak{g}_k\) for some \(k\). Notice that any chain for \(U'\) which starts at \(\mathfrak{g}_k\) can be decomposed into chains for \(U\), by starting a new chain for \(U\) at \(\text{ad}_{U_{\alpha_2}}(X_i^j)\), if \(X_i^j\) were a chain basis element for \(U'\). In particular, the Jordan blocks for \(U\) are shorter than the corresponding ones for \(U'\). Finally, we need to find at least one chain for \(U'\) which is broken into shorter ones for \(U\). Consider the \(\mathfrak{sl}(2, \mathbb{R})\)-triple for \(U_{\alpha_2}\), giving \(V_{\alpha_2} \in \mathfrak{g}_{-\alpha_2}\) and \(X_{\alpha_2} \in \mathfrak{g}_0\). Since \(\alpha_1\) and \(\alpha_2\) are both simple, \(\alpha_1 - \alpha_2 \notin \Delta\) (the integral coefficients for the simple roots are either all positive or all negative). In particular, \(\text{ad}_U(V_{\alpha_2}) = 0\). But \(\text{ad}_{U'}(V_{\alpha_2}) = X_{\alpha_2}\), showing that at least one chain is shorter for \(U\) than for \(U'\). This proves (iii).

If \(G \cong SL(d, \mathbb{R})\), an explicit formula for \(GR(U)\) can be found. If \(U_l \in \mathfrak{sl}(d, \mathbb{R})\) has one Jordan block of size \(l\), then:

\[
GR(U_l) = \frac{1}{6}l(4l + 1)(l - 1) + l(d - l)(l - 1)
\]

Since it acts via the representation \(\pi_{l-1}\) on the off-diagonal blocks, with the main term coming from Corollary 1.13 of [17]. One easily confirms that these are distinct numbers for \(l = 2, \ldots, d\) by computing \(GR(U_{l+1}) - GR(U_l) = l(2d - l)\), giving \(d - 1\) flows with different Kakutani invariant. This proves (iv).

\[\square\]

**Remark 3.10.** In fact, the leading term of the general formula for \(GR(U)\) obtained in [17] is cubic, and \(GR(U_2)\) (in the notation above) grows linearly in \(d\). We expect most numbers interpolating the cubic and linear growth to be possible by taking more involved Jordan block structures, and therefore expect the number of pairwise non-Kakutani equivalent flows on \(SL(d, \mathbb{R})\) to grow like \(d^3\).

### 4 Some additional definitions

Let \(G\) be a semisimple Lie group and \(\Gamma \subset G\) be a cocompact lattice as considered in Section 3 with projection \(\pi: G \to G/\Gamma\). We fix a compact fundamental domain \(F \subset G\) for \(G/\Gamma\). Then every point of \(G/\Gamma\) has at least one lift to \(F\), and there is a unique lift on an open, dense subset of \(G/\Gamma\) (corresponding to the interior of \(F\)).

Notice that \(\pi: F \to G/\Gamma\) is a measurable isomorphism, so we can think of the left action \(G\) on \(G/\Gamma\) equivalently as an action on \(F\). In particular, if \(x, y \in G/\Gamma\), then they are cosets \(x = \tilde{x}\Gamma\) and \(y = \tilde{y}\Gamma\) for some unique (except for points on the boundary of \(F\)) \(\tilde{x}, \tilde{y} \in F\). Then \(gx = y\) is equivalent to \(g\tilde{x} = \tilde{y}\gamma\) for some \(\gamma \in \Gamma\). This allows us to consider the homogeneous flow \((\phi_t)\) on \(G/\Gamma\), as we shall in Section 7 which preserves the Haar measure \(\mu_F\). We may lift the metric on \(G/\Gamma\) to \(F/\sim\) by setting \(d_F(x, y) := \inf_{\gamma \in F} d_G(\tilde{x}, \tilde{y}\gamma)\), where \(F/\sim\) is the topological quotient space of \(F\) by the usual relation.

For a point \(y \in G/\Gamma\) let \(\text{inj}(y)\) denote the injectivity radius of \(y\), i.e.

\[
\text{inj}(y) := \sup \{r \geq 0 : B_G(y, r) \cap B_G(y, r)\gamma = \emptyset \text{ for all } \gamma \neq e\}.
\]

For a set \(K \subset G/\Gamma\) let \(\text{inj}(K) = \inf_{y \in K} \text{inj}(y)\). We have the following classical lemma which we state here for reference.
Lemma 4.1. $\text{inj}(G/\Gamma) > 0$

4.1 Kakutani-Bowen Balls

Given a unipotent element $U \in \mathfrak{g}$, we may use the results of Sections 3.3.1 and 3.4 to obtain a basis $\{V, X, U\} \cup \{X^j_i : i = 1, \ldots, m_j, j = 1, \ldots, n\}$, where $\{V, X, U\}$ generates a subalgebra of $\mathfrak{g}$ and satisfy the standard relations for the $\mathfrak{sl}(2, \mathbb{R})$-triple, and the remaining elements are chains for $U$. We may therefore apply Lemma 3.7 to write elements of $G$ sufficiently close to $e$ as

$$g = \exp(a_V(g)V) \exp(a_X(g)X) \exp \left( a_U(g)U + \sum a_{ij}(g)X^j_i \right) \quad (8)$$

Let $\tau(g) = \exp \left( a_U(g)U + \sum a_{ij}(g)X^j_i \right)$ be the standard chain component of $g$. The following definition combines dynamical and algebraic features, which is critical to our analysis of the Kakutani balls $B_R(x, \varepsilon, \mathcal{P})$ (see Theorem 6.1).

Definition 4.2 (Kakutani-Bowen Balls). For $\varepsilon > 0$, $R > 0$ let

$$\text{Bow}(R, \varepsilon, e) := \{g \in G : d_G(\exp(rU)g\exp(-rU), e) < \varepsilon, \text{ for every } r \in [0, R]\}$$

be the Bowen ball of $e \in G$ for $U$. If $x, y \in G/\Gamma$ we say that $x \in \text{Kak}(R, \varepsilon, y)$ if and only if $x = gy$ and

(a) $|a_V(g)| < \varepsilon/R$
(b) $|a_X(g)| < \varepsilon$
(c) $\tau(g) \in \text{Bow}(R, \varepsilon, e)$

where $a_V(\cdot)$, $a_X(\cdot)$ and $\tau(\cdot)$ are as in (8).

We will often consider Kakutani balls as subsets of $F$, since points in $F$ are in one-to-one correspondence with points in $G/\Gamma$ (except for those on the boundary). Furthermore, if $\varepsilon$ is sufficiently small, depending only on $\text{inj}(G/\Gamma)$, if $\tilde{y}$ is any point of $G$ such that $\tilde{y}\Gamma = y$, $\text{Kak}(R, \varepsilon, y)$ lifts uniquely to a neighborhood of $\tilde{y}$. The definition of $\text{Kak}(R, \varepsilon, y)$ has the following explanation. We will see that polynomials that differ in the direction $X^k_j$ will see polynomial divergence in the direction $X^k_j$ with degree $i - j$ for $j < i$. Since $\{V, X, U\}$ is taken as a chain for $U$, points that differ in $V$ direction split with quadratic speed in the direction of $U$ and with linear speed in the direction of $X$. Such points can be easily matched with the $f_t$-metric (even though they can not by $\tilde{d}$-Bowen metric) as long as we don’t see divergence in $X$ direction, since we are allowed to correct by the flow. This is the reason why in the definition of Kak we take $|a_V| \leq \frac{\varepsilon}{R}$. Similarly, points differences in $X$ direction yield linear divergence, but only in the $U$ direction and hence the control required on the $X$ coefficient does not grow.

Since the condition for being in a Kakutani ball does not tell us the direction of divergence, we make the following additional definitions which allow us to distinguish points whose divergence is first seen in the $\mathfrak{sl}(2, \mathbb{R})$-triple components, and those which see divergence in the other chains first.
Fix $\delta > 0$, and set

$$\text{Kak}^{1,\delta}(R, \varepsilon, y) := \text{Kak}(R, \varepsilon, y) \cap \left\{ x \in G/\Gamma : |a_V(g)| < R^{-(1+10\delta)} \right\} \text{ in Definition 4.2}$$ \hspace{1cm} (9)

and

$$\text{Kak}^{2,\delta}(R, \varepsilon, y) := \text{Kak}(R, \varepsilon, y) \setminus \text{Kak}^{1,\delta}(R, \varepsilon, y).$$ \hspace{1cm} (10)

For sufficiently large $R$, Kak$^{1,\delta}$ are exactly points in which the divergence is guaranteed to be seen first in the other chains (due to the increased control on $a_V$). Kak$^{2,\delta}$ is therefore the points in which we are guaranteed to see some divergence in the $X$ direction, since in this case $|a_V| \in [R^{-(1+10\delta)}, \varepsilon/R]$.

**Definition 4.3** (Splitting time). For $x, y \in G/\Gamma$ define the splitting time of $x, y$

$$S(x, y, \varepsilon) := \sup \{ R \geq 0 : x \in \text{Kak}(R, \varepsilon, y) \}. \hspace{1cm} (11)$$

The following observation is a straightforward consequence of continuity of $(\phi_t)$: there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that $f(m) \to +\infty$ as $m \to +\infty$ and

$$S(x, y, \varepsilon) > f(m) \text{ as } d_{G/\Gamma}(x, y) < m^{-1}. \hspace{1cm} (12)$$

We also have the following general definition which establishes a useful notation when dealing with matching of $x, y$. A priori, given a matching between $x$ and $y$, the points may have long periods of matching, diverge for a small amount of time, and realign to have another long period of matching. This is exactly what happens for the classical horocycle flow. The following definition identifies that maximal interval on which the matching could be extended before seeing divergence and waiting for another realignment.

**Definition 4.4.** Fix a partition $\mathcal{P}$ of $G/\Gamma$ and $x, y \in G/\Gamma$ which are $(\eta, \mathcal{P})$-matchable (with matching function $h$). For $u \in A(x, y)$ denote $x_u = \phi_u x \in G/\Gamma$, $y_u = \phi_{h(u)} y \in G/\Gamma$ and let for $\varepsilon > 0$

$$S(u, \varepsilon) = S(x_u, y_u, \varepsilon).$$
4.2 Summary of notations

| Symbol | Definition |
|--------|------------|
| $G$ | A fixed semisimple linear Lie group |
| $\mathfrak{g}$ | The Lie algebra of $G$ |
| $\Gamma$ | A cocompact lattice in $G$ |
| $x, y, z$ | Points of $G/\Gamma$ |
| $\tilde{x}, \tilde{y}, \tilde{z}$ | Lifts of $x, y, z$ to $G$ |
| $g, h$ | Elements of $G$ |
| $\text{inj}(G/\Gamma)$ | The largest number such that if $d_{G/\Gamma}(x, y) < \text{inj}(G/\Gamma)$, $x = gy$ for a unique $g \in B_G(\epsilon, \text{inj}(G/\Gamma))$ |
| $\{V, X, U\}$ | Fixed generators of a subalgebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$ |
| $\phi_t$ | The left translation action by $\exp(tU)$ |
| $GR(U)$ | The polynomial slow entropy of $\phi_t$ (i.e., the growth rate for the number of Bowen balls to cover $M$) |
| $\{X^j_i\}_{i=1}^m, j = 1, \ldots, n$ | Vectors generating $\mathfrak{g}$ together with $\{V, X, U\}$, having certain relations with $\{V, X, U\}$ (see Section 3.3.1) |
| $\text{Bow}(R, \epsilon, y)$ | The Bowen ball around $y$ of radius $\epsilon$ up to time $R$ |
| $B_R(y, \epsilon, \mathcal{P})$ | The Kakutani ball around $y$, i.e., the set of all points $x$ which are $(\epsilon, \mathcal{P})$-matchable with $y$ |
| $\text{Kak}(R, \epsilon, y)$ | Intuitively, points which stay $\epsilon$-close to $y$ after lifting to $G$, allowing correction of $x$ by the flow (see Section 4.1) |
| $\text{Kak}^{1,2}(R, \epsilon, y)$ | A set of points in $\text{Kak}(R, \epsilon, y)$ which see non-orbit divergence in directions other than $X$ first |
| $\text{Kak}^{2,3}(R, \epsilon, y)$ | The remaining points of $\text{Kak}(R, \epsilon, y)$ |
| $S(x, y, \epsilon)$ | The first time $x, y$ split and cannot be made close by applying $\phi_t$ to $y$ in the universal cover |
| $S(u, \epsilon)$ | With a fixed matching of $x$ and $y$, $S(x_u, y_u, \epsilon)$, where $x_u$ and $y_u$ is the matching at time $u$ |

5 Orbit divergence estimates

In this section we state results on orbit divergence for unipotent flows. These results play an important role in the proofs of Theorem 1.1 and 6.1.

We recall that the following lemmas, which will be used in the proof:

**Lemma 5.1.** Let $p(t) = \sum_{k=0}^d a_k t^k$ be a polynomial of degree $d$. There exists $C(d)$ such that if $|p(t)| < \epsilon$ for all $t \in [0, T]$, then $|a_k| < C(d)T^{-k}\epsilon$ for all $k = 0, \ldots, d$. Conversely, if $|a_k| < C(d)^{-1}T^{-k}\epsilon$ for all $k$, then $|p(t)| < \epsilon$ for all $t \in [0, T]$.

Let $g = \exp(a_U(g)U + \sum_{j=1}^n \sum_{i=0}^{m_j} a_{ij}(g)X^j_i)$. The following formulas are important for computing divergence rates:

\[
\exp(sX)\exp(tU)g\exp(-tU)\exp(-sX) = \exp\left(e^{2s}a_U(g)U + \sum_{j=1}^n \sum_{i=0}^{m_j} \left(\sum_{k=0}^{m_j-i} e^{(m_j-2i)k} \frac{k!}{k!} a_{(k+i)j}(g)\right) X^j_i\right)
\]  (13)
\[
\exp(tU) \exp(sX)g \exp(-sX) \exp(-tU) = \exp\left( e^{2sa_U(g)}U + \sum_{j=1}^{m_j} \sum_{i=0}^{m_j-i} e^{(m_j-2(k+i))(s)} \frac{t k}{k!} a_{(k+i)j}(g) X_i^j \right)
\] (14)

Equations (13) and (14) follow from Lemma 3.2 and the choice of chain basis made in section 3.3.1. The proof of the following lemma is almost identical to that of [17, Proposition 3.7], so we provide only a sketch.

Lemma 5.2. There exists \( \varepsilon_0 \) such that for every \( y \in G/\Gamma \) and every \( \varepsilon < \varepsilon_0 \), we have

\[
C(\varepsilon_0) R^{-GR(U)+2} \geq \mu_F(\text{Kak}(R, \varepsilon, y)) \geq C(\varepsilon_0) R^{-GR(U)+2}. \tag{15}
\]

Sketch of Proof. Set \( \varepsilon < \varepsilon_0 := \text{inj}(G/\Gamma)/3 \). It follows that the projection \( \pi : G \to G/\Gamma \) is injective on \( \text{Kak}(R, \varepsilon, y) \subset B(y, 3\varepsilon) \). Write \( x = gy \), and note that the coefficients of \( X_k^i \) for \( \log(\exp(tU)\tau(g)\exp(-tU)) \) are all polynomials of degree \( i \) by (13). Therefore, by Lemma 5.1 and (13) with \( s = 0 \), a sufficient condition for \( \tau(g) \) is that \( |a_{ik}| \leq C'(\varepsilon) R^{-i} \), and a necessary one is that \( |a_{ik}| \leq C'(\varepsilon)^{-1} R^{-i} \) (by shrinking \( \varepsilon_0 \) if necessary to absorb the constant \( C(d) \)). Let \( L : U \to \mathbb{R}V \oplus \mathbb{R}X \oplus \left( \mathbb{R}U \oplus \bigoplus_{i,k} \mathbb{R}X_k^i \right) \) be the inverse function provided by Lemma 3.7. Then we have shown that

\[
[-\varepsilon/R, \varepsilon/R] \times [-\varepsilon, \varepsilon] \times \left[ -\varepsilon, \varepsilon \times \prod_{i,k} \left[ -\varepsilon \frac{\varepsilon}{C'(\varepsilon) R^{i}}, \frac{\varepsilon}{C'(\varepsilon) R^{i}} \right] \right] \subset L(\text{Kak}(R, \varepsilon, y))
\]

\[
\subset [-\varepsilon/R, \varepsilon/R] \times [-\varepsilon, \varepsilon] \times \left[ -\varepsilon, \varepsilon \times \prod_{i,k} \left[ -\frac{C'(\varepsilon) \varepsilon}{R^{i}}, \frac{C'(\varepsilon) \varepsilon}{R^{i}} \right] \right]
\]

Notice that the hypercubes which contain and are contained in \( L(\text{Kak}(R, \varepsilon, y)) \) decay with the rate prescribed. Since the Jacobian of \( L \) is bounded above and below in a neighborhood of \( e \) in \( G \), we get the desired decay rate.

In the study of slow entropy, i.e. the covering rate for \( G/\Gamma \) via Bowen balls, the result analogous to Lemma 5.2 is sufficient to estimate the number of Bowen balls to cover the space. However, Kakutani balls have a more complicated behavior, since we only insist that the points are close for a large proportion of times. The remaining lemmas help to show that if points stay together for a certain interval, then that amount of time can be quantified, and that each such interval has a long interval afterwards in which the points diverge, but in a controlled way as to avoid recurrence.

The following lemma allows us to explicitly describe an optimal matching function when \( x \) and \( y \) are sufficiently close.
Lemma 5.3. There exists $\varepsilon_1 > 0$ such that if we let $h = \exp(a_V V) \exp(a_X X) \in G$ and
\[ \psi(t) := te^{ax} / (e^{-ax} - ae^{ax}t) \] with $|a_X| < \varepsilon_1$, then for every $|t| \in [0, \varepsilon_1a_V^{-1}]$, we have
\[ \exp(\psi(t)U)h\exp(-tU) = \exp(\alpha t V) \exp(\beta t X), \] (16)
with $|\alpha_t| \leq 2|a_V|$ and $|\beta_t| \leq 2(|a_X| + |a_V||t|)$.

Proof. By Lemma 5.2 we may make computations in $SL(2, \mathbb{R})$ and conclude the relevant relations in $G$. Abusing the notation in this proof slightly, we let $V, X, U \in \mathfrak{sl}(2, \mathbb{R})$ denote the generators of the opposite horocycle flow, geodesic flow, and horocycle flow respectively. By a direction computation and the definition of $\psi(\cdot)$, we have
\[ \exp(\psi(t)U) \exp(a_V V) \exp(a_X X) \exp(-tU) = \begin{pmatrix} e^{ax} + \psi(t)aVe^{ax} & 0 \\ aVe^{ax} & e^{-ax} - aVe^{ax}t \end{pmatrix}. \] (17)

Let $\alpha_t$ and $\beta_t$ be defined so that
\[ \exp(\alpha_t V) \exp(\beta_t X) = \exp(\psi(t)U) \exp(a_V V) \exp(a_X X) \exp(-tU). \]

Direct computation shows that
\[ \alpha_t e^{\beta_t} = aVe^{ax}, \]
\[ e^{-\beta_t} = e^{-ax} - aVe^{ax}t. \] (18)

Thus we have
\[ \alpha_t = aVe^{ax}(e^{-ax} - aVe^{ax}t), \]
\[ \beta_t = -\log(e^{-ax} - aVe^{ax}t). \] (19)

Since $|a_X| < \varepsilon_1$ and $|t| \in [0, \varepsilon_1a_V^{-1}]$, we have
\[ |\alpha_t| < 2|a_V|, |\beta_t| \leq 2(|a_X| + |a_V||t|). \]

This finishes the proof. \qed

Remark 5.4. We will use the following property of $\psi(\cdot)$ which follows by a direct computation: under the above assumptions if additionally $|a_X| < \varepsilon^2$, then for every $t \in [0, \varepsilon^2a_V^{-1}]$ we have
\[ \psi'(t) \in (1 - \varepsilon, 1 + \varepsilon). \]

Lemma 5.5. There exists $\varepsilon_1 > 0$ such that for every $\varepsilon \in (0, \varepsilon_1]$ and for every $x, y \in G / \Gamma$ if $x \in \text{Kak}(R, e^3, y)$, then for every $|L| \in [0, R/3]$ there exists $|t| < R$ such that
\[ \phi_t(x) \in \text{Kak}(R/2, \varepsilon, \phi_L y). \]

Proof. By Definition 4.2 it follows that we may write
\[ x = \exp(a_V V) \exp(b X) g y \]
where $|a_V| \leq \varepsilon^3 R^{-1}$, $|b| \leq \varepsilon^3$ and $g \in \text{Bow}(R, e^3, e)$, with $g$ having no $V$ or $X$ component. Therefore, for every $\ell \geq 0$
\[ \exp(\ell U)x = \left( \exp(\ell U) \exp(a_V V) \exp(b X) \exp(-LU) \right) \left( \exp(LU) \exp(-LU) \right) \exp(LU) y \] (20)
Moreover by Lemma 5.3 if we define $\ell := \psi(L)$ (by the bound on $a_V$ this is well defined), then $|\ell| \leq R/2$ (see Remark 5.4), and we have

$$\exp(\ell U) \exp(a_V V) \exp(bX) \exp(-LU) = \exp(b_L V) \exp(c_L X),$$

where $|b_L| \leq \varepsilon R^{-1}$, and $|c_L| < \varepsilon$. Notice also that if $m_\ell = \exp(LU)z \exp(-LU)$ then for every $t \in [0, R - L]$, we have

$$d_G(\exp(tU)m_\ell \exp(-tU), e) = d_G(\exp((t + L)U)g \exp(-(t + L)U), e) < \varepsilon^3$$

(21)

the last inequality since $g \in \text{Bow}(R, \varepsilon^3, e)$. Therefore $m_\ell \in \text{Bow}(R/2, \varepsilon^3, e)$ and the $V$ and $X$ coordinates of $m_\ell$ are zero since the spaces generated by each chain are invariant for $\text{ad}(tU)$ (see Definition 4.2). Therefore

$$\phi_\ell x = \exp(\ell U)x = \exp(\ell U)\exp(a_V V)\exp(bX)y = \exp(b_L V)\exp(c_L X)m_\ell \phi_{LY}$$

which by Definition 4.2 and (21) implies that $\phi_\ell x \in \text{Kak}(R/2, \varepsilon, \phi_{LY})$. This finishes the proof.

The following Lemma quantifies the renormalization phenomenon related to the relation $[X, U] = 2U$. Recall Definition 4.2 and equation (8).

**Lemma 5.6.** Let $g \in \text{Bow}(R, \varepsilon, e)$ be such that $a_V(g) = a_X(g) = 0$. There exists $C > 0$ (independent of $\varepsilon$) such that for every $\varepsilon \in (0, \varepsilon_1]$, every $\delta' \in [0, C^{-2})$, every $R > 0$ and every $s \in [0, \frac{1}{2}(1 + \delta') \log(R)]$, we have

$$\exp(-sX)g \exp(sX) \in \text{Bow}(R^{1/2-C\delta'}, \varepsilon^{1/3}, e)$$

and $a_V(\exp(-sX)g \exp(sX)) = a_X(\exp(-sX)g \exp(sX)) = 0$.

**Proof.** Since $a_V(g) = a_X(g) = 0$, we can write $g = \exp\left(a_V(g)U + \sum_{i,j} a_{ij}(g)X^j_i\right)$, where $\{X^1_i\}_{i=1}^{m_0}, \ldots, \{X^n_i\}_{i=1}^{m_n}$ are standard chains, each of which span a finite-dimensional representation of $\mathfrak{sl}(2, \mathbb{R})$ (see Definition 3.3). Note that the claims that

$$a_V(\exp(-sX)g \exp(sX)) = a_X(\exp(-sX)g \exp(sX)) = 0$$

follow from the fact that each chain $\left\{X_i^j\right\}$ spans a finite-dimensional representation of $\mathfrak{sl}(2, \mathbb{R})$ and that $U$ is an eigenvector for both $\text{ad}_X$ and $\text{ad}_U$. Since $g \in \text{Bow}(R, \varepsilon, e)$ it follows that for $0 \leq t \leq R$, $d_G(\exp(tU)g \exp(-tU), e) \leq \varepsilon$. Since $d_G(\exp(tU)g \exp(-tU), e) < \varepsilon$ for $0 \leq t \leq R$ it follows by Lemma 5.1 applied to the $X^j_i$ terms of (13) with $s = 0$ that

$$|a_{ij}| \leq \frac{C(d, \varepsilon)\varepsilon}{R^i}$$

(22)

for all $j$, where $C(d, \varepsilon)$ is determined by Lemma 5.1 and the norm of $d(\log_{\text{geom}} \circ \exp)$ on $B(e, \varepsilon_1)$ (since $d(e, \exp(Y)) = ||\log_{\text{geom}}(\exp(Y))||$).

Let $R_s(g) = \exp(-sX)g \exp(sX)$. We need to show that $R_s(g) \in \text{Bow}(R^{1/2-C\delta'}, \varepsilon^{1/3}, e)$. Hence we only need to show that for every $0 \leq t \leq R^{1/2-C\delta'}$, 

$$d_G(\exp(tU)R_s(g) \exp(-tU), e) < \varepsilon^{1/3}. $$
Notice that by (14),
\[ a_{ij}(s, t) = \sum_{k=0}^{m_j-i} \frac{t^k}{k!} e^{-(m_j-2k+i)s} a_{(i+k)j}(g) \]
denotes the coefficient of \( X_j^2 \) for \( \exp(tU)R_s(g)\exp(-tU) \) at time \( t \). We will control each coefficient \( a_{ijk} = \frac{1}{k!} e^{-(m_j-2k+i)s} a_{(k+i)j}(g) \).

Our bound on \( a_{ij}(g) \), (22), and the condition that \( 0 \leq s \leq \frac{1}{2}(1 + \delta') \log(R) \) gives:
\[ |a_{ij}| \leq \frac{C(d, \varepsilon_1)\varepsilon}{k!} R^{-\frac{1}{2}(m_j-2k+i)(1+\delta')-(k+i)} \leq C(d, \varepsilon_1)\varepsilon R^{-\frac{1}{2}m_j+(k+i-\frac{1}{2})m_j}\delta' \]
Therefore, by Lemma 5.1, since \( k+i \leq m_j \) and since the maximal power of \( t \) by \( a_{ij}(s, t) \) is \( m_j - i \) (and therefore \( t^k \leq R^{1/2-C\delta'}(m_j-i) \), we have
\[ \sup_{t \in [0, R^{1/2-C\delta'}]} a_{ij}(s, t) \leq m_j C_1(d, \varepsilon_1)\varepsilon R^{1/2-C\delta'-(m_j-i)-\frac{1}{2}m_j(1-\delta')} = m_j C_1(d, \varepsilon_1) R^{-1/2-C\delta'-(m_j-i)+1/2m_j}\delta' \]
Notice that \(-1/2i - C\delta'(m_j - i) + 1/2m_j\delta'\leq 0 \).

Therefore, if \( \varepsilon \) is sufficiently small, we can guarantee that \( \sup_{t \in [0, R^{1/2-C\delta'}]} a_{ij}(s, t) \) can be made less than \( \varepsilon^{1/3}/C'' \) for arbitrary \( C'' \). So by choosing \( \varepsilon \) sufficiently small, we may guarantee \( R_s(g) \in \text{Bow}(R^{1/2-C\delta'}, \varepsilon^{1/3}, e) \).

**Lemma 5.7.** There exist constants \( C_2, \varepsilon_1, R_0 > 0 \) such that for every \( \varepsilon \leq \varepsilon_1 \), every \( R \geq R_0, s \leq \frac{1}{2}\log(R) \), and for every \( g = \exp(Y) \in \text{Bow}(R, \varepsilon, e) \),
\[ \exp(-sX)g \exp(sX) \in \text{exp}(Y_C)B_{G}(e, C_2\varepsilon R^{-\frac{1}{2}}) \]
for some \( Y_C \in C(U) \). Moreover, for \( s = \frac{1}{2}\log(R) \), the same holds for some \( Y_C \in C(U, X) \).

**Proof.** Recall, (see (13) and (22)) that if \( g = \exp \left( \sum_{i,j} a_{ij}(g)X_i^j \right) \in \text{Bow}(R, \varepsilon, e) \), then
\[ |a_{ij}(g)| \leq \frac{C(d)\varepsilon}{R^i} \]
and
\[ \exp(-sX)g \exp(sX) = \exp \left( \sum_{j=1}^{n} \sum_{i=0}^{m_j} e^{-s(m_j-2i)} a_{ij}(g)X_i^j \right) \] (23)
Therefore, if \( s \leq \frac{1}{2}\log R \), we get the following bound on the \( X_i^j \) coefficient of \( \exp(-sX)g \exp(sX) \) (denoting \( s = c_s \log(R) \), so that \( c_s < \frac{1}{2} \))
\[ |e^{-s(m_j-2i)} a_{ij}(g)| \leq C(d)\varepsilon R^{-i-c_s(m_j-2i)} \]
Moreover, if \( m_j - 2i \geq 0 \), then \(-i - c_s(m_j - 2i) \leq -i \) and if \( m_j - 2i < 0 \), then \(-i - c_s(m_j - 2i) \leq -i - \frac{1}{2}(m_j - 2i) = -\frac{1}{2}m_j \) and recall that \( m_j \geq i \). Therefore, if \( i > 0 \), then for every \( j \), the coefficient by \( X_i^j \) in (23) is at most \( C(d)\varepsilon R^{-1/2} \). Notice that by Definition
Corollary 5.10. There exists $C_2(d) > 0$ and $\varepsilon_3 > 0$ such that for every $\eta > 0$ and $\varepsilon_3 > \varepsilon > 0$ if $g \notin \text{Bow}(R, \varepsilon, e)$ and $d_G(g, e) < C(d)^{-1}\varepsilon$ (recall $C(d)$ is defined in Lemma 5.1) then

$$\left| w \in [0, R^{1+\eta}] : d_G(\exp(wU)g \exp(-wU), e) < 10\varepsilon \right| \leq C_2(d)R^{1+\eta-\eta/d}.$$
Proof. Recall that if \( g = \exp(\sum_{i,j} a_{ij} X_i^j) \), then \( \exp(wU)g \exp(-wU) \) will have the coefficients of \( X_i^j \) as polynomials in \( w \) by (13). Let \( p_{ij}(w) \) denote the polynomial for \( X_i^j \) and let \( R \geq w \geq 0 \) be the largest number such that \( g \in \text{Bow}(w, \varepsilon, e) \), then analogously to (22) it follows that there exists \( j \) such that
\[
|p_{0j}(w)| \geq \varepsilon.
\]
Indeed, this follows from the fact that if \( \varepsilon \) is small enough (depending only on \( U \)), then \( |p_{ij}(w)| \leq |p_{0j}(w)| \).

Therefore using Lemma 5.9 we have
\[
|w| \in [0, R^{1+\eta}] : |p_{0j}(w)| \leq 10\varepsilon \leq C_1(d)R^{(1+\eta-\eta/d)}.
\]
Since \( d_G(\exp(wU)g \exp(-wU), e) < 10\varepsilon \) implies in particular that every coordinate is less than \( 10\varepsilon \) (by taking log and since Jacobian is close to 1 around \( e \)). Let \( C_2(d) = C_1(d) \).

This finishes the proof.

Denote \( L_U = \max\{m_j : j = 1, \ldots, n\} \) to be the depth of the longest chain for \( U \).

Lemma 5.11. For every \( \frac{1}{2L_U^2 + 2L_U + 1} > \eta' > 0 \), every \( \varepsilon \in [0, \varepsilon_0] \) there exist \( R_{\varepsilon, \eta'} \) such that for every \( R \geq R_{\varepsilon, \eta'} \), every \( g \in \text{Bow}(R, \varepsilon, e) \) such that \( a_V(0) = a_X(0) = 0 \) and \( d_G(g, e) < \varepsilon/C(d) \) and every \( t \in [0, R^{1+\eta'}] \) for which
\[
d_G(\exp(tU)g \exp(-tU), e) > 10\varepsilon,
\]
there exists \( |s| \in [0, 2(L_U + 1)\eta' \log |R|] \) and \( C_t := a_U(t)U + \sum_{j=0}^n c_j(t)X_0^j \) such that \( \max_j |c_j(t)| \geq \varepsilon/n \) and
\[
d_G(\exp(-sX) \exp(tU)g \exp(-tU) \exp(sX), \exp(C_t)) < R^{-\eta'},
\]
Proof. Write \( g = \exp \left( a_U(0)U + \sum_{j=1}^n \sum_{i=0}^{m_j} a_{ij}(g)X_i^j \right) \), and let
\[
p_{ij}(t, s) = \sum_{k=0}^{m_j-i} \frac{t^k}{k!} e^{-\eta k} a_{(k+i), j}
\]
be the coefficient of \( X_i^j \) for \( \exp(-sX) \exp(tU)g \exp(-tU) \exp(sX) \) as determined by (13). The condition that \( g \in \text{Bow}(R, \varepsilon, e) \) implies that \( |a_{k,j}| < C(d)\varepsilon k! / R^k \) by Lemma 5.1. Thus, if \( t \leq R^{1+\eta'} \),
\[
t^k / k! a_{(k+i), j} \leq \frac{R^{(1+\eta')k}}{R^{k+1}} C(d)\varepsilon k! = R^{\eta' k} C(d)\varepsilon k! < R^{1/2 - \varepsilon/(m_j + 1)}
\]
if \( R \) is sufficiently large.

Thus, if \( i \geq 1 \), \( |p_{ij}(t, 0)| < R^{-1/2} \varepsilon \). Similarly for sufficiently large \( R \), \( |p_{0j}(t, 0)| \leq (m_j + 1)C(d)\varepsilon R^{m_j \eta'} \leq C(d)^2 \varepsilon R^{\eta' \eta'} < 2R^{(L_U + 1)\eta'} \varepsilon / n \) for every \( j \).

By assumption, \( t \) is such that \( d_G(\exp(tU)g \exp(-tU), e) > 10\varepsilon \). Therefore, there exists some \( j \) such that \( p_{0j}(t, 0) \geq 8\varepsilon / n \). Let \( D = 2(L_U + 1) \), and consider the function \( \zeta : s \mapsto (p_{01}(t, s), \ldots, p_{0n}(t, s)) \in \mathbb{R}^n \). Since \( p_{0j}(t, 0) \geq 8\varepsilon / n \) for some \( j \), it follows that \( \zeta(0) \notin [0, 7\varepsilon / n]^n \). But since \( |p_{0j}(t, 0)| < 2R^{(L_U + 1)\eta'} \varepsilon / n \), we get
such that $p_{0j}(t, D_{\eta'} \log R) = R^{-m_j D_{\eta'}} |p_{0,j}(t, 0)| < 2R^{(L_U+1)(1-2m_j)\eta'} \leq 2\varepsilon/n$

for every $j$, provided $m_j \geq 1$. If $m_j = 0$, the only term appearing is $X'_{ij}$ which is constant in both $t$ and $s$. Therefore its coefficient is bounded by $\varepsilon/C(d)$ by the assumption that $d(z, e) < \varepsilon/C(d)$. Therefore, $||\zeta(D_{\eta'} \log R)|| \leq 2\varepsilon$.

By continuity of $\zeta$, we may therefore choose $s \in [0, D_{\eta'} \log R]$ such that $||\zeta(s)|| = 2\varepsilon$. Then, since $|p_{ij}(t, 0)| < R^{-1/2\varepsilon}$, we get that $|p_{ij}(t, s)| \leq R^{-(m_j-2\eta')D_{\eta'}-1/2\varepsilon} < R^{-2\eta'}$ for $i \geq 1$ if $\eta'$ is sufficiently small. Setting $C_t = e^{-2\eta' \log(z)}U + \sum_{j=1}^n p_{0j}(t, s)X'_{ij}$ gives that:

$$d_G(\exp(-sX) \exp(tU)z \exp(-tU) \exp(sX), \exp(C_t)) \leq \ell \left\| \sum_{j=1}^n \sum_{i=1}^{m_j} p_{ij}(t, s)X'_{ij} \right\| \leq \ell R^{-2\eta'} \sum_{j=1}^n m_j < R^{-\eta'}$$

Here, $\ell$ is the Lipschitz constant for $\log: B_G(e, 100\varepsilon) \to \mathfrak{g}$ and the last inequality holds if $R$ is chosen sufficiently large. Finally, notice that since $\zeta(s) = 2\varepsilon$, there exists some $j$ such that $p_{ij}(t, s) > \varepsilon/n$. $\square$

6 Proof of Theorem 1.1

In this section we will prove Theorem 1.1. The proof is rather technical and consists of several steps, which we will divide into subsections to improve readability. The following Theorem is a crucial step in the proof of Theorem 1.1 since it shows that for most points, being Kakutani close implies that there exists a long block on which they are close in the metric on $G/\Gamma$ (see Definition 1.2).

Sequence of Partitions. Let $P_m$ be a partition of $G/\Gamma$ such that the atoms of $P_m$ are sets with diameter in $[\frac{1}{2m}, \frac{1}{m}]$ (with smooth boundaries). It is clear that $(P_m)_{m \in \mathbb{N}}$ is generating (and we can use Theorem 2.1). Let $\varepsilon' > 0$ be a small constant fixed from now on, in particular $\varepsilon' < \min(\varepsilon_1, \text{inj}(G/\Gamma))$. We have the following theorem:

**Theorem 6.1.** Let $(\phi_t)$ be the flow generated by $U$ and assume that $U$ does not satisfy the conditions of Lemma 3.9. There exists $\delta_0 > 0$ such that for every $0 < \delta \leq \delta_0$ there exists a set $E = E_\delta \subset M$, $\mu(E_\delta) > 99/100$ and $m_\delta, R_\delta \in \mathbb{N}$ such that for every $m \geq m_\delta$, $R \geq R_\delta$ and every $x, y \in B$ and $(\phi_t)_m$-matchable there exists $u \in A(x, y)$ such that (see Definition 4.4)

$$S(u, \varepsilon') \geq R^{1-\delta}.$$

The proof of Theorem 1.1 can be deduced from Theorem 6.1. We will first give a conditional proof of Theorem 1.1 (assuming that Theorem 6.1 holds) and then prove Theorem 6.1 in a separate section. The proof of Theorem 1.1 is divided into two parts: (i) the upper bound on the number of balls and (ii) lower bound on the number of balls. It follows that we only need Theorem 6.1 for (ii). In the proof we will use Remark 2.8.

**Outline of the proof:** The proof shows first that $e((\phi_t), \log) \leq GR(U) - 3$ and then that $e((\phi_t), \log) \geq GR(U) - 3$. For the proof of the upper bound, we will apply Lemma 5.2. To do so, we need to relate $\text{Kak}(R, \varepsilon, y)$ and $B_R(y, \varepsilon, P)$. The key idea is this: $\text{Kak}(R, \varepsilon, y)$ is defined in such a way that all coordinates are controlled in such a way
that if \( x \in \text{Kak}(R, \varepsilon, y) \), for each \( u \in [0, R] \), we can apply \( \phi_t \) to bring \( \phi_u(x) \) close to \( \phi_u(y) \), with \( t \) well-controlled. The decay rate of \( \text{Kak} \), however, is 1 away from the claimed decay rate for \( B_R(y, \varepsilon, \mathcal{P}) \). This is because the definition of \( \text{Kak} \) assumes that \( x \) and \( y \) are close initially, while for \( B_R(y, \varepsilon, \mathcal{P}) \), we only require a forward orbit of \( x \) (which is small relative to \( R \), but can depend linearly on \( R \)) to be close to \( y \). This accounts for the slower decay rate of \( B_R \) versus \( \text{Kak} \).

**Claim A** is the relationship between \( \text{Kak} \) and \( B_R \): if \( x \in \bigcup_{p=0}^{3R} \phi_{-p}(\text{Kak}(R, \varepsilon^5, y)) \), then they are in one Kakutani ball, i.e. \( x \in B_R(y, \varepsilon, \mathcal{P}_m) \). Then **Claim B** shows that the forward orbits of \( \text{Kak}(R, \varepsilon^5, y) \) do not overlap giving the desired rate with Lemma \[5.2\] holds. This together with the upper bound in Lemma \[5.2\] implies the lower bound.

**Proof of Theorem** \[6.1\]. We will separately prove the upper bound which will follow from general estimates on asymptotic divergence of orbits and the lower bound which is a consequence of Theorem \[6.1\]. Fix \( m \in \mathbb{N} \) (this also fixes the partition \( \mathcal{P}_m \)) and let \( \varepsilon > 0 \), \( \varepsilon < \min \{ m^{-3}, \text{inj}(G/\Gamma) \} \).

**Upper bound on the number of balls.** Let \( V^m_\varepsilon \) be the \( \varepsilon^2 \) neighborhood of the boundary of \( \mathcal{P}_m \). Since the boundaries are smooth, it follows that \( \mu(V^m_\varepsilon) = O(m^2\varepsilon^2) \). Applying the ergodic theorem to \( (\phi_t) \) and the set \( \chi_{V^m_\varepsilon} \), we obtain a set \( D_\varepsilon \) such that \( \mu(D_\varepsilon) > 1 - \varepsilon^2 \) and a number \( N_\varepsilon > 0 \) such that for every \( \tilde{R} \geq N_\varepsilon \) and every \( y \in D_\varepsilon \), we have

\[
|\{ t \in [0, R] : \phi_t(y) \in V^m_\varepsilon \}| \leq \frac{\varepsilon R}{2}. \tag{24}
\]

The upper bound will follow from the following two claims:

**Claim A.** For every \( y \in D_\varepsilon \), every \( R \geq N_\varepsilon \) and every \( x \in F \) if there exists \( p \in [0, \varepsilon^3 R] \) such that if

\[
\phi_p x \in \text{Kak}(R, \varepsilon^5, y)
\]

then \( x \in B_R(y, \varepsilon, \mathcal{P}_m) \) (see Definition \[2.3\]).

**Claim B.** For every \( y \in D_\varepsilon \) and every \( p, q \in [0, \varepsilon^3 R] \), with \( |p - q| \geq 1 \), we have

\[
\phi_{-p}(\text{Kak}(R, \varepsilon^5, y)) \cap \phi_{-q}(\text{Kak}(R, \varepsilon^5, y)) = \emptyset.
\]

Before we prove the claims, let us show how they imply the upper bound.

Take \( y \in D_\varepsilon \). By **Claim A** it follows that

\[
\bigcup_{p \in [0, \varepsilon^3 R]} \phi_{-p}(\text{Kak}(R, \varepsilon^5, y)) \subset B_R(y, \varepsilon, \mathcal{P}_m).
\]

Therefore by **Claim B.** and Lemma \[5.2\] we have

\[
\mu(B_R(y, \varepsilon, \mathcal{P}_m)) \geq \mu \left( \bigcup_{p \in [0, \varepsilon^3 R]} \phi_{-p}(\text{Kak}(R, \varepsilon^5, y)) \right) \geq \varepsilon^d R \mu(\text{Kak}(R, \varepsilon^5, y)) \geq \varepsilon^{d+3} R^{-G(R(U)+3)}.
\]
Since this holds for every \( y \in D_\varepsilon \) and \( \mu(D_\varepsilon) \geq 1 - \varepsilon \) it follows that for some \( C(\varepsilon) > 0 \) depending on \( \varepsilon \) only and by Remark 2.3 we have
\[
K_R(5\varepsilon, \mathcal{P}) \leq C(\varepsilon)R^{GR(U) - 3}. 
\] (25)

Therefore \( \beta((\log 5\varepsilon), \mathcal{P}_m) \leq GR(U) - 3 \) and so by Theorem 2.3, \( e((\phi_t), \log) \leq GR(U) - 3 \).

Notice moreover, that if \( GR(U) = 3 \), then by (25) it follows that the number of balls does not depend on \( R \). Therefore \( e((\phi_t), u) = 0 \) for every function \( u \in \mathcal{F} \). By Theorem 2.0 it follows that if \( GR(U) = 3 \), then \( (\phi_t) \) is standard.

So it remains to prove Claim A and Claim B.

**Proof of Claim A.** Take \( y \in D_\varepsilon \) and let \( p \in \{0, \varepsilon^3 R\} \) be such that \( \exp(pU)x \in \text{Kak}(R, \varepsilon^5, y) \).

This by Definition 4.2 implies that for some \( |b| < \varepsilon^5 \), \( |a| < \varepsilon^5 \), and \( g \in \text{Bow}(R, \varepsilon^5, e) \) satisfying \( aV(g) = aX(g) = 0 \), we have
\[
x = \exp(-pU)\exp(aV)\exp(bX)gy. 
\] (26)

Let \( \psi(t) := \frac{t aV}{\exp(tU)\exp(aV)} \) be as in Lemma 5.3, \( h(t) = \psi(t) + p, A(x, y) := \{t \in [0, R] : \phi_t(y) \notin V_{m}^{y}\} \) (notice that by (24) we have \( |A(x, y)| \geq (1 - \varepsilon)R \)). Moreover (see Remark 5.4), \( |h'(t) - 1| < \varepsilon \) for every \( t \in [0, R] \) and hence \( h \) satisfies the condition to be a \((\varepsilon, \mathcal{P}_m)\)-matching function. We will show that for every \( t \in [0, R] \), we have
\[
d_F(\phi_t y, \phi_t h(x)) \leq \varepsilon^3. 
\] (27)

This by right invariance, the definition of \( h(\cdot) \) and (26) follows by showing
\[
d_G(e, \exp(\psi(t)U)(\exp(aV)\exp(bX)\exp(-tU))\exp(\exp(-tU))) \leq \varepsilon^3. 
\]

Since \( g \in \text{Bow}(R, \varepsilon^5, e) \) it follows that \( d_G(\exp(\exp(-tU))g\exp(-tU), e) < \varepsilon^5 \) for \( t \in [0, R] \). Moreover, by Lemma 5.3 we have
\[
d_G(\exp(\psi(t)U)\exp(aV)\exp(bX)\exp(-tU), e) < \varepsilon^4. 
\]

The above two inequalities finish the proof of (27). By (27), for every \( t \in A(x, y) \) (see (24)), we have \( \mathcal{P}_m(\phi_t y) = \mathcal{P}_m(\phi_t h(x)) \). Since \( |A(x, y)| \geq (1 - \varepsilon)R \), it follows that \( x \in B_R(y, \varepsilon, \mathcal{P}_m) \). This finishes the proof of Claim A.

**Proof of Claim B.** We will argue by contradiction assuming that there exists \( x \in \phi_{q-p}(\text{Kak}(R, \varepsilon^5, y)) \cap \text{Kak}(R, \varepsilon^5, y) \), with \( \varepsilon^3 R \geq |p - q| \geq 1 \) and \( y \in D_\varepsilon \). This, by the Definition 4.2 in particular means (denoting \( r = p - q \)) that
\[
x = \exp(aV)\exp(bX)gy,
\]
and,
\[
\exp(-rU)x = \exp(a'V)\exp(b'X)g'y,
\]
where \( g, g' \in \text{Bow}(R, \varepsilon^5, e) \) satisfy \( aV(\cdot) = aX(\cdot) = 0, |a|, |a'| \leq \frac{\varepsilon^5}{R} \) and \( |b|, |b'| \leq \varepsilon^5 \). Choose lifts \( \tilde{x}, \tilde{y} \in G, x, y \in G/T \) minimizing \( d_G(\tilde{x}, \tilde{y}) \). In particular since \( \varepsilon < \text{inj}(G/T) \),
\[
\tilde{x} = \exp(aV)\exp(vX)g\tilde{y} \quad \exp(-rU)\tilde{x} = \exp(a'V)\exp(b'X)g'\tilde{y}\gamma 
\] (28)
for some $\gamma \in \Gamma$. Therefore using the second equality to express $\gamma \gamma$ and the first to express $\gamma^{-1}$, we get

$$\gamma \gamma^{-1} = g^{-1} \exp(-b'X) \exp(-a'V) \exp(rU) \exp(aV) \exp(bX)g.$$ 

Notice that since $|r| \geq 1$ and all the other terms are $\varepsilon$ small, it follows that $\gamma \neq e$. Multiplying on the left by $\exp(-sX)$ and on the right by $\exp(sX)$ with $s = \frac{1}{2} \log \varepsilon^2 R$ gives

$$(\exp(-sX)\gamma) (\exp(-sX)\gamma)^{-1} = \exp(-sX)g^{-1} \exp(sX) \exp(-b'X).$$

By the definition of $s$ it follows that if $\gamma \neq e$, then (25)), it follows that if $\gamma \neq e$, then

$$(\exp(e^2a'V) \exp(-2s rU) \exp(e^{2s}aV) \exp(bX) \exp(-sX)g \exp(sX), e)$$

$$d_G(\exp(-sX)w \exp(sX), e) < \varepsilon^{5/3}.$$ 

Therefore, the RHS of (29) is $\varepsilon^2$ close to $e$. But since we have chosen $\varepsilon < \inj(G/\Gamma)$, $d_G((\exp(-sX)\gamma)(\exp(-sX)\gamma)^{-1}, e) \geq \varepsilon$. This contradiction finishes the proof of the upper bound.

**Lower bound on the number of balls.** Notice that from the upper bound estimates (in particular, (25)), it follows that if $GR(U) = 3$, then $(\phi_t)$ is standard. Hence in what follows we assume that $GR(U) > 3$, which, by Lemma 3.9, is equivalent to $\dim G - \dim(C(X)) - 3 > 0$ and we can use Theorem 6.1.

Fix $\delta > 0$, $0 < \varepsilon < 1/100$, and $m \geq m_\delta$ satisfying $1/m^2 < \varepsilon' / C(d)$ and $R \geq R_\delta$ and assume that $x \in G/\Gamma$ and $y \in E_\delta$ are $(\varepsilon, P_m)$-matchable. Using Theorem 5.1 and Definition 1.4 it follows that there exists $p, q \in [0, R]$ (in fact $p = u$ and $q = h(u)$ where $h$ is the matching function) such that

$$\phi_p x \in \Kak(R_{1-\delta}, \varepsilon', \phi_q y).$$

This implies that

$$x \in \phi_{-p} \left( \Kak(R^{1-\delta}, \varepsilon', \phi_q y) \right).$$

Therefore,

$$B_R(y, \varepsilon, P_m) \cap E_\delta \subset \phi_{-p} \left( \Kak(R^{1-\delta}, \varepsilon', \phi_q y) \right).$$

Let $(R_i)_{i=1}^{R^{2\delta}}$ be given by $R_i = i R^{1-2\delta}$. We will show that any $x, y \in E_\delta$ which are $(1/100, P_m)$-matchable have to satisfy

$$x \in \bigcup_{p \in [-2R, 2R]} \phi_{-p} \left( \bigcup_{i=1}^{R^{2\delta}} \Kak(R^{1-4\delta}, \varepsilon'^{1/3}, \phi_R, y) \right)$$

Before we give the proof of (30), let us show how it implies the lower bound. By (30), we have

$$B_R(y, \varepsilon, P_m) \cap E_\delta \subset \bigcup_{p \in [-2R, 2R]} \phi_{-p} \left( \bigcup_{i=1}^{R^{2\delta}} \Kak(R^{1-4\delta}, \varepsilon'^{1/3}, \phi_R, y) \right).$$
Since $\epsilon' < \text{inj}(G/\Gamma)$, by Lemma \ref{lem:matching_set}, for some $c(\epsilon') > 0$ (depending on $\epsilon'$ only)

$$
\mu(B_R(y, \varepsilon, P_m) \cap E_\delta) \leq 4R^{1+2\delta} \max_i \mu(\text{Kak}(R^{1-4\delta}, \varepsilon'^{1/3}, \phi_{R_i} y)) \leq c(\epsilon') R^{1+2\delta} R^{-(1-4\delta)(GR(U)-2)}.
$$

This gives

$$
\mu(B_R(y, \varepsilon, P_m) \cap E_\delta) \leq c(\varepsilon) R^{U(\delta)},
$$

where $U(\delta) = 1+2\delta - (1-4\delta)(GR(U)-2)$. Hence the number of balls needed to cover $1-\varepsilon$ of space, i.e. $K_R(\varepsilon, P_m)$ is at least $C(m) R^{-U(\delta)}$ (for some constant depending on $m$ only). Therefore $\beta(\log, \varepsilon, P_m) \geq -U(\delta)$ and since the sequence $(P_m)$ is generating $e((\phi_t), \log) \geq -U(\delta)$. The proof is finished by taking limit as $\delta$ goes to 0, since $-U(0) = GR(U) - 3$. So it remains to show (30).

Using Theorem \ref{thm:main} and Definition \ref{defn:matching} it follows that there exists $p, q \in [0, R]$ (in fact $p = u$ and $q = h(u)$ where $h$ is the matching function) such that

$$
\phi_p x \in \text{Kak}(R^{1-\delta}, \varepsilon', \phi_q y).
$$

Let $R_j$ be the number minimizing $|g - R_i|$ (over all $i$). To finish the proof of (30) it is enough to show that there exists $\ell = \ell(p, q)$, $|\ell| \leq R$ such that (since $|p+\ell| < 2R$)

$$
\phi_{p+\ell} x \subset \text{Kak}(R^{1-2\delta}, \varepsilon'^{1/3}, \phi_{R_j}(y)).
$$

This however follows by Lemma \ref{lem:matching_set} with $\varepsilon'^3 = \varepsilon'$, $z = \phi_p x$, $y = \phi_q y$, $L = R_j - q$. This finishes the proof of Theorem \ref{thm:matching}.

So it remains to prove Theorem \ref{thm:main}.

## 7 Proof of Theorem \ref{thm:main}

**Outline of the proof:** Assume that $x, y \in G/\Gamma$ are $(1/100, P_m)$-matchable. Then every matching arrow (of $\phi_u x$ and $\phi_{h(u)} y$) can be parametrized by $j = j(u) \in \mathbb{N}$ and $w \in \{1, 2\}$:

- $j \in \mathbb{N}$ measures the splitting time of $\phi_u x$ and $\phi_{h(u)} y$ in exponential scale (see the set $C_{j,R,m}(x, y)$ below).
- Moreover $w \in \{1, 2\}$ gives the direction which is responsible for the splitting, i.e. if $w = 1$ then the splitting is definitely produced by directions different than $V$ and if $w = 2$ then the splitting might be (but not necessarily has to be) produced by $V$ (see \ref{prop:splitting} and \ref{prop:splitting2}).

Propositions \ref{prop:splitting} and \ref{prop:splitting2} are purely of combinatorial nature (no dynamics involved, just a counting argument).

**Proposition \ref{prop:splitting}** states, that if for every $j$ (sufficiently large) the measure of arrows with label $j, w$, for $w \in \{1, 2\}$ is exponentially small (see \ref{lem:small_measure}) than the total measure of the matching is also small (since the series is summable over $j$). Proposition \ref{prop:splitting2} states that if in every window of size $2^{j(1+\delta)}$ for $c = 2$ (in (A)) and $c = 40$ (in (B)) the relative measure of arrows with label $j$ is exponentially small, then the total measure of arrows with label $j$ has to be small.

Fix $\delta > 0, m > 0$ and $R, j \in \mathbb{R}_+$. Assume that $x, y \in M$ are $(1/100, P_m)$-matchable. Let $A(x, y) \subset [0, R]$ denote the matching set and $h : A(x, y) \rightarrow [0, R]$ the matching function.
We define two sets which will play a crucial role in the proof. Recalling Definitions and 4.4 and 4.3, we define

\[ C_{j,R,m}(x,y) := \{u \in A(x,y) : d_G(x_u,y_u) < 2m^{-2} \land 2^j \leq S(u,\varepsilon') < 2^{j+1}\} \]

Let (see Definition 4.4 and (9), (10))

\[ C_{j,R,m}^1(x,y) := C_{j,R,m} \cap \{u \in A(x,y) : x_u \in \text{Kak}^{1,\delta}(2^j,\varepsilon',y_u)\} \quad (31) \]

and

\[ C_{j,R,m}^2(x,y) := C_{j,R,m} \cap \{u \in A(x,y) : x_u \in \text{Kak}^{2,\delta}(2^j,\varepsilon',y_u)\}. \quad (32) \]

By definition, \( C_{j,R,m}(x,y) = C_{j,R,m}^1(x,y) \cup C_{j,R,m}^2(x,y) \). Let \( d \) and be as in Lemma 5.9. The following proposition implies Theorem 6.1.

**Proposition 7.1.** There exists \( c(d) > 0 \) and \( \delta_0 > 0 \) such that for every \( \delta_0 > \delta > 0 \) there exists a set \( E_\delta \subset M, \mu(E_\delta) \geq \frac{99}{100} \) and \( m_\delta, R_\delta \in \mathbb{N} \) such that for every \( m \geq m_\delta, R \geq R_\delta \) and every \( x,y \in E_\delta \) there exist \( W_R(x,y) \subset A(x,y) \) satisfying:

(a). \([W_R(x,y)] \geq \frac{99}{100} R;\]

(b). for every \( j \in \mathbb{N} \) satisfying \( 2^j \leq R^{1-40\delta} \), any \((1/100, P_m)\)-matching of \( x \) and \( y \) (with matching function \( h \)), we have for \( w = 1,2 \)

\[ |C_{j,R,m}^w(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \leq c(d)R \frac{2^{j-1}\delta}{2^{j-1}\delta}. \]

We will prove Proposition 7.1 in a separate subsection. Let us now show how the above proposition implies Theorem 6.1.

**Proof of Theorem 6.1.** Fix \( \delta > 0, R \geq R_\delta, m \geq m_\delta, x,y \in E_\delta \) and a \((1/100, P_m)\)-good matching of \( x \) and \( y \) with the set \( A(x,y) \) and the matching function \( h \). Notice that by the definition of \( P_m \) and \( C_{j,R,m}(x,y) \), we have

\[ A(x,y) \cap W_R(x,y) \subset \bigcup_{j \in \mathbb{N}} C_{j,R,m}(x,y). \]

Moreover, by (12) and the definition of \( C_{j,R,m}(x,y) \), for \( j \leq \log_2 f(m) \),

\[ C_{j,R,m}(x,y) = \emptyset. \quad (33) \]

Hence, by (a) and the definition of \( C_{j,R,m}(x,y) \), we have (recall also that \( h \) is the matching function, hence it is absolutely continuous)

\[ \frac{1}{m} > \bar{f}_R^{P_m}(x,y) \]

\[ \geq 1 - \frac{|((W_R(x,y)) \cup h^{-1}(W_R(x,y))) \cap [0,R]|}{R} \]

\[ \geq \frac{1}{R} \sum_{j \geq 0} |C_{j,R,m}(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \quad (34) \]

\[ \geq \frac{9}{10} \frac{1}{R} \sum_{j \geq 0} |C_{j,R,m}(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \]
Let $j_R$ be such that,
\[ 2^{j_R} \leq R^{1-4\delta} < 2^{j_R+1}. \] (35)

By (33), (b) and $C_{j,R,m}(x,y) = C_{j,R,m}^1(x,y) \cup C_{j,R,m}^2(x,y)$, we have
\[ \frac{1}{R} \sum_{j<j_R} |C_{j,R,m}(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \leq \frac{1}{R} \sum_{j<j_R} \frac{2c(d)R}{2^{d-1}\delta j} \leq \frac{1}{1000}, \] (36)
by enlarging $m$ if necessary (since $f(m)$ goes to $\infty$). Therefore and by (33) there exists $j_1 \geq j_R$ such that
\[ C_{j_1,R,m}(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y)) \neq \emptyset. \] (37)

By definition of $C_{j_1,R,m}(x,y)$ and (35) it follows that there exists $u \in A(x,y)$ such that
\[ S(u, \varepsilon') \geq 2^{j_R} \geq \frac{1}{2} R^{1-4\delta} \geq R^{1-41\delta}. \] (38)

Since $\delta > 0$ is arbitrary (changing $\delta' = 41\delta$), this finishes the proof of Theorem 6.1. \[ \square \]

7.1 Proof of Proposition 7.1

We will formulate a proposition which will imply Proposition 7.1.

**Proposition 7.2.** There exists $c'(d) > 0$ and $\delta_0 > 0$ such that for every $\delta > 0$ there exists a set $E_\delta \subset M$, $\mu(E_\delta) \geq \frac{90}{100}$ and $m_\delta$, $R_\delta \in \mathbb{N}$ such that for every $m \geq m_\delta$, $R \geq R_\delta$ and every $x, y \in E_\delta$ there exist $W_R(x,y) \subset [0,R]$ such that (a) holds for every $(1/100, P_m)$-matching of $x, y$ (with matching function $h$) and for every $j \in \mathbb{N}$ satisfying $2^j \leq R^{1-4\delta}$, we have

(A) for every $u \in W_R(x,y)$ (see Definition 2.7)
\[ |C_{j,R,m}^1(x,y) \cap B(u, 2^{1+2\delta}) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \leq c'(d)2^{(1+2\delta-2\delta/d)j}; \]

(B) for every $u \in W_R(x,y)$
\[ |C_{j,R,m}^2(x,y) \cap B(u, 2^{1+4\delta}) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \leq 2^{(1+20\delta)d)j}; \]

Before we prove Proposition 7.2 let us show how it implies Proposition 7.1.

**Proof of Proposition 7.1.** Notice that by assumptions of Proposition 7.2 it follows that we only need to prove (b) in Theorem 6.1 (with $w = 1, 2$ in (b)). The proof for $w = 1$ uses (A) and the proof for $w = 2$ uses (B). Since the proofs in both cases follow the same lines, we will give the proof in case $w = 1$. Fix $j$ as in (b). Divide the interval $[0,R]$ into disjoint intervals $I_1, \ldots, I_k$ of length $2^{(1+2\delta)j}$ in the following way. Fix the smallest element
\[ u_1 \in C_{j,R,m}^1(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y)). \]
Let $I_1$ be an interval with right endpoint $u_1$ and length $l_1 := 2^{(1+2\delta)j}$. Now inductively for $u > 1$, we pick $u_w$ to be the smallest element in $C^1_{j,R,m}(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y)) \setminus (I_1 \cup \ldots \cup I_{u-1})$. As $u_w$ satisfies (A), we let $I_w$ be the interval with right endpoint $u_w$ and length $l_j$. We continue until we cover $C^1_{j,R,m}(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))$.

Since $2^j \leq R^{1-20\delta}$, we have $2^{(1+2\delta)j} \leq R^{1-\delta} \ll R$ and hence $k > 1$. Moreover by definition, we have

$$k \leq \left\lfloor \frac{R}{l_j} \right\rfloor + 2. \quad (39)$$

Notice that by Definition 2.2, the fact that $C^1_{j,R,m}(x,y) \subset A(x,y)$ and the definition of $(I_i)$ it follows that

$$C^1_{j,R,m}(x,y) \cap I_i \cap W_R(x,y) \cap h^{-1}(W_R(x,y)) \subset$$

$$\subset C^1_{j,R,m}(x,y) \cap B(u_i, 2^{(1+2\delta)j}) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))$$

Therefore and by (A), we have

$$|C^1_{j,R,m}(x,y) \cap I_i \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \leq c'(d)2^{(1+2\delta-2\delta/d)j}. \quad (40)$$

Summing over $i \in \{1, \ldots, k\}$ by (40) and (39), we get

$$|C^1_{j,R,m}(x,y) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \leq \left( \frac{R}{l_j} + 2 \right)c'(d)2^{(1+2\delta-2\delta/d)j} \leq \frac{c(d)R}{2^{d+\delta}},$$

the last inequality by the definition of $l_j$ and defining $c(d) = 2c'(d)$. This finishes the proof. \hfill \Box

7.2 Proof of Proposition 7.2

In this section we will prove Proposition 7.2. It is the most technical part of the paper. We will start by giving an outline of the proof:

Outline of the proof of Proposition 7.2. The arguments are very different in proving (A) and (B). All the difficulty in proving (B) is transferred to Lemma 7.3. Indeed, Lemma 7.3 directly implies Corollary 7.4 which implies (B). We will give the outline of proof of Lemma 7.3 in the next section. The method in proving (A) is based on the fact that if $x \in \text{Kak}^{1,\delta}(2^j, \varepsilon', y)$ then the first non-orbit divergence occurs in a direction other than $V$, $X$ or $U$ (the $V$ coordinate is too small by definition of $\text{Kak}^{1,\delta}$) and consequently the direction in which $x,y$ split belongs to the centralizer $C(U)$ but is different than the flow direction. Since $f$-metric allows one to “slide” along the orbits of points (ie, $U$ direction) only, there is no way to correct the splitting in the $C(U)$ direction if it is different than the flow (which is the case for (A)). Let us also stress out that the condition $\dim G - \dim(C(X)) - 3 > 0$ is only used in the proof of Lemma 7.3.

We divide the proof in several subsections. For $j \in \mathbb{N}$ let (see (9))

$$G(\delta, j, y) = \text{Kak}^{2,\delta}(2^j, \varepsilon', y) \cap \left( \bigcup_{p,q \in [2^j(1+20\delta), 2^j(1+40\delta)+1]} \phi_{-p}(\text{Kak}^{2,\delta}(2^j, \varepsilon', \phi_q y)) \right), \quad (41)$$

We have the following lemma:
Lemma 7.3. For $j \in \mathbb{N}$ let
\[ EG(\delta, j) = \{ y \in G/\Gamma : G(j, y) = \emptyset \}. \]
There exists $j_\delta > 0$ such that for $j \geq j_\delta$,
\[ \mu(EG(\delta, j)) \geq 1 - j^{-2}. \]

By the above lemma it follows that if we define
\[ E'_\delta := \bigcap_{j > j'_\delta} EG(\delta, j), \]
then $\mu(E'_\delta) \geq 1 - 10^6$ if $j'_\delta := \max(j_\delta, 10^{10})$.

We will prove Lemma 7.3 in the last section, let us first state the following immediate corollary:

Corollary 7.4. For $y \in E'_\delta$ and $x \in \text{Kak}^{2, \delta}(2^j, \varepsilon', y)$ with $j > j_\delta$, we have
\[ \phi_{p, x} \notin \text{Kak}^{2, \delta}(2^j, \varepsilon', \phi_{q, y}), \]
for any $p, q \in [2^{j(1+20\delta)}, 2^{j(1+40\delta)+1}]$.

Proof. This just follows by the definition of $E'_\delta$, since for $y \in E'_\delta$, we have $G(\delta, j, y) = \emptyset$ for $j \geq j_\delta$.

We can now prove Proposition 7.2.

Proof of Proposition 7.2 By ergodic theorem for $\chi_{E'_\delta}$ it follows that there exists a set $E_\delta \subset E'_\delta \subset F$, $\mu(E_\delta) > 999/1000$ and $N_\delta > 0$ such that for every $x, y \in E_\delta$ and every $R \geq N_\delta$, we have
\[ |W_R(x, y) := W_R(x) \cap W_R(y)| \geq 99/100, \]
where, for $z \in E_\delta$,
\[ W_R(z) := \{ t \in [0, R] : \phi_{t, z} \in E'_\delta \}. \]

Notice that by the definition of $W_R(x, y)$ and the set $E'_\delta$ it follows that (a) and (b) hold. Hence we only need to show that (A) and (B) in Proposition 7.2 hold. The methods of proof are different for (A) and (B), (B) being a simple consequence of Corollary 7.4.

Proof of (B) Notice that by the definition of $C^2_{j, R, m}(x, y)$ it follows that if $u \in C^2_{j, R, m}(x, y)$, then $y_u = \phi_{h(u)} y \in E'_\delta \subset E_\delta$ and
\[ x_u \in \text{Kak}^{2, \delta}(2^j, \varepsilon', y_u). \]

Notice that since $h$ is an $(1/100, \mathcal{P}_m)$-matching function it follows that for every $v - u \in [2^{j(1+20\delta)}, 2^{j(1+40\delta)}]$, we have
\[ |h(v) - h(u)| < 2|v - u|. \]

Therefore and by Corollary 7.4 for $y = y_u$ and $x = x_u$ it follows that for any $v \in [u + 2^{j(1+20\delta)}, u + 2^{j(1+40\delta)}]$ (notice that $|h(v) - h(u)| < 2^{j(1+40\delta)+1}$)
\[ \phi_{v-u}(x_u) \notin \text{Kak}^{2, \delta}(2^j, \varepsilon', \phi_{h(v)-h(u)}(y_u)). \]
Note that $\phi_{v-u}(x_u) = x_v$ and $\phi_{h(v) - h(u)}(y_u) = \phi_{h(v)}y = y_v$.

Therefore
\[
C^2_{j,R,m}(x,y) \cap \left(B(u, 2^{j(1+10\delta)}) \setminus B(u, 2^{j(1+20\delta)})\right) = \emptyset.
\]

So
\[
|C^2_{j,R,m}(x,y) \cap B(u, 2^{(1+40)\delta}) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))| \leq 2^{j(1+20\delta)}.
\]

and this finishes the proof of (B).

**Proof of (A):** Let $u$ be the smallest element in $C^1_{j,R,m}(x,y) \cap B(u, 2^{j(1+2\delta)}) \cap W_R(x,y) \cap h^{-1}(W_R(x,y))$ (if such $u$ doesn’t exist, then the intersection is empty and the proof is finished).

We will show that there exists a set $A'_u \subseteq [h(u), h(u) + 2^{j(1+2\delta)}]$, $|A'_u| \leq C'(d)2^{j(1+2\delta - 2\delta/d)}$, such that for every $h(v) \in [h(u), h(u) + 2^{j(1+2\delta)}] \setminus A'_u$,
\[
v \not\in C^1_{j,R,m}(x,y) \cap B(u, 2^{(1+2)\delta}) \cap W_R(x,y) \cap h^{-1}(W_R(x,y)).
\]

This will finish the proof of (A) since $|h' - 1| < \varepsilon$ on $A(x,y)$.

Assume for contradiction that (43) does not hold and let $v$ belong to the RHS of (43).

By the definition of $C^1_{j,R,m}(x,y)$ and Definition 4.4 and 4.3 it follows that for $w = u, v$, we have
\[
x_w = \exp(a_w V) \exp(b_w X)g_w y_w,
\]
where $|a_w| < 2^{-j(1+10\delta)}$, $|b_w| < m^{-2}$ and $g_w \in \text{Bow}(2^j, \varepsilon', e)$ and $a_w g_w = a_v g_w = 0$ with $d_G(g_w, e) < m^{-2}$. Moreover, since $|a_w| \leq 2^{-j(1+10\delta)}$, by Definition 4.3 Definition 4.4 and Definition 4.2, it follows that
\[
g_w \notin \text{Bow}(2^{j+1}, \varepsilon', e).
\]

Notice that $\varepsilon'$ is fixed, so we may choose $m$ sufficiently large so that the condition $d_G(g_w, e) < m^{-2}$ will imply that $d_G(g_w, e) < C(d)^{-1}\varepsilon'$, where $C(d)$ is as in Corollary 5.10. Applying this Corollary with $R = 2^{j+1}$ and $\varepsilon = \varepsilon'$ it follows that there exists $A_u \subseteq [0, 2^{j(1+2\delta)}]$, $|A_u| \leq C(d)2^{j(1+2\delta - 2\delta/d)}$ and such that for every $r \in [0, 2^{j(1+2\delta)}] \setminus A_u$, we have
\[
d_G(\exp(rU)g_u \exp(-rU), e) > 10\varepsilon'.
\]

Hence, by Lemma 5.11 with $R = 2^j$, and $\eta' = 2\delta$ we have that for every $r \in [0, 2^{j(1+2\delta)}] \setminus A_u$ there exists $s_r \in [0, 4(L_U + 1)\delta]$
\[
d_G(\exp(-s_rX) \exp(rU)g_u \exp(-rU) \exp(s_rX), \exp(C_r)) < 2^{-\delta j},
\]
where $C_r \in C(U)$ is as in the statement of Lemma 5.11.

Notice that $x_u = \exp((v - u)U)x_u$ and $y_u = \exp((h(v) - h(u))U)y_u$. Therefore and by (44) (setting $p = v - u$ and $q = h(v) - h(u)$) it follows that if $\tilde{x}$ and $\tilde{y}$ are lifts of $x$ and $y$, for some $\gamma \in \Gamma$,
\[
\begin{align*}
\exp(pU) \exp(a_u V) \exp(b_u X)g_u \exp(h(u)U)\tilde{y} &= \exp(a_v V) \exp(b_v X)g_v \exp(qU) \exp(h(u)U)\tilde{y} \gamma
\end{align*}
\]

$^4$The splitting time is defined through the V coefficient and dynamical control on $g_w$. Since we are in the set $C^1$ (see the definition of $\text{Kak}^{1,4}$) it follows that the V coefficient is of lower order, hence the splitting has to be produced by $g_w$. 

We will consider two cases:

**Case I.** $\gamma = e$ in (17). In this case we get

$$\exp(pU) \exp(a_u V) \exp(b_v X) g_u \exp(-qU) = \exp(a_v V) \exp(b_v X) g_v. \quad (48)$$

By Lemma 5.3 and since $|a_u| < 2^{-j(1+10\delta)}$, it follows that there exists $l = l(p)$ with $p/2 < |l| < 2p$, such that

$$\exp(pU) \exp(a_u V) \exp(b_v X) = \exp(b_v V) \exp(c_p X) \exp(lU),$$

and $|b_p| < 2|a_u|, |c_p| < 4m^{-2}$. Therefore, using (48), we have

$$\exp((l-q)U) \exp(qU) g_u \exp(-qU) = \exp(-c_p X) \exp(-b_v V) \exp(a_v V) \exp(b_v X) g_v. \quad (49)$$

If $q \in [0, 2^{j(1+2\delta)}] \setminus A_u$, then by (16), we have, that there exists $s \in [0, 4(L_U + 1)j\delta]$ such that

$$d_G(\exp(-sX) \exp(qU) g_u \exp(-qU) \exp(sX), \exp(C)) < 2^{-j\delta}. \quad (50)$$

Conjugate (49) by $\exp(sX)$. Since $s \leq 4(L_U + 1)j\delta \leq \frac{1}{2} \log 2^j$ (since $\delta$ is of lower order than $L_U$), by Lemma 5.7, we have

$$d_G(\exp(-sX) g_u \exp(sX), e) < m^{-1},$$

since $\|z_u\| \leq d_G(z_u, e) \leq m^{-2}$. Moreover since $|a_v|, |b_p| \leq 2^{-j(1+10\delta)}$ and $|b_v|, |c_p| \leq 4m^{-2}$ and since $s \leq 4(L_U + 1)j\delta$, the RHS of (49) after conjugating by $\exp(sX)$ (term by term) is $40m^{-2}$ close to $e$. Hence and by (50), we get

$$d_G(\exp(e^{-2s}(l-q)U), \exp(-C)) \leq 40m^{-2},$$

this however contradicts the properties of $C$ (see Lemma 5.11), if $m$ is chosen so that $\varepsilon' > m^{-1/100}$. Consequently, for $q \in [0, 2^{j(1+2\delta)}] \setminus A_u$ it follows that (50) does not hold. Recall that $q = h(v) - h(u)$. This shows that for $v \in h^{-1}([0, 2^{j(1+2\delta)}] \setminus A_u)$, (13) holds. It remains to define $A_u' = h^{-1}(A_u)$ and notice that $|A_u'| \leq (1 + \varepsilon)|A_u|$ (since $h$ is a matching function).

**Case II.** $\gamma \neq e$ in (17). In this case (17) is equivalent to

$$\exp(-qU) g_v^{-1} \exp(-b_v X) \exp(-a_v V) \exp(pU) \exp(a_v V) \exp(b_v X) g_u = \exp(h(u) U) \tilde{y}\gamma \tilde{y} \exp(-h(u)U). \quad (51)$$

Let $s := \frac{1}{2} \log 2^{j(1+3\delta)}$. We have

$$d_G(\exp(-sX) \exp(pU) \exp(a_v V) \exp(b_v X) g_u \exp(sX), e) < \frac{8}{m^{1/3}}. \quad (52)$$

Indeed, this follows from applying the conjugation term-wise and by $|p| < 2^{j(1+2\delta)}$, $|a_v| < 2^{-j(1+10\delta)}$, $|b_v| \leq m^{-2}$ and $d_G(\exp(-sX) g_u \exp(sX), e) < \frac{1}{m^{1/3}}$ (see Lemma 5.6 with $\delta' = 3\delta$). By an analogous reasoning, (52) holds also for $v$ (and $g_v^{-1}$ and $-q$ instead of $g_u$ and $p$). So by (51), for every, we have

$$d_G((g-sy_u)\gamma' (g-sy_u)^{-1}, e) < \frac{16}{m^{1/3}},$$

and this contradicts Lemma 4.11 since $m$ is chosen after inj$(G/T)$.  □
7.3 Proof of Lemma 7.3

Let us start by giving outline of the proof:

Outline of the proof: We start by \((53)\), which is the definition of \(G(j, y)\) not being empty. Next, we transform \((53)\) (using the bounds on coefficients) to \((59)\). Then using the properties \(P1\) and \(P2\) (which we prove in the appendix) we further transform it to \((63)\). We then have Lemma 7.5 which tells us that the set of solutions to \((63)\) is small. Then using theorems on the cardinality of lattice points in balls in semisimple Lie groups, it follows that Lemma 7.5 follows by Lemma 7.6 (in Lemma 7.5 we are summing over \(\gamma \in \Gamma\) and in Lemma 7.6 the element \(\gamma \in \Gamma\) is fixed). The crucial result here is Lemma 7.4. It allows to show that there are (sufficiently many) small translations of the set \(S_{i,\gamma}\) which are disjoint. Therefore the measure of the set cannot be too large (this is made precise in the proof of Lemma 7.6). The crucial condition \(\dim G = \dim(C(X)) - 3 > 0\) is used to show that the "sufficiently many" translates is enough to get the estimates since the cardinality of translates which are pairwise disjoint is strongly related to the number \(\dim G = \dim(\Omega_\Gamma) - 3\) (see the statement of Lemma 7.7). In fact this is the only place in the whole proof of Theorem 1.1 in which we need this assumption. The method of proving Lemma 7.4 goes by a straightforward calculation using the properties of the adjoint representation of the \(sl(2, \mathbb{R})\) triple.

Proof of Lemma 7.3: Fix \(\delta > 0\) and \(j \in \mathbb{N}\). If \(G(\delta, j, y) \neq \emptyset\), there exists \(x \in G/\Gamma\) such that \(x \in \text{Kak}^{2,\delta}(2^j, \varepsilon', y)\) and \(\phi_p x \in \text{Kak}^{2,\delta}(2^j, \varepsilon', \phi_q y)\). By \((10)\) this implies that

\[
x \in \exp(aV) \exp(bX)gy, \quad g \in \text{Bow}(2^j, \varepsilon', e),
\]

with \(|a| \in [2^{-j(1+10\delta)}, 2^{-j}]\), \(b \leq \varepsilon'\) and \(aV(g) = aX(g) = 0\). Analogously,

\[
\exp(pU)x \in \exp(cV) \exp(dX) \tau(g')\exp(qU)y, \quad g' \in \text{Bow}(2^j, \varepsilon', e)
\]

with \(c, d\) satisfy the same estimates as \(a, b\) and \(aV(g') = aX(g') = 0\). We may lift \(x\) and \(y\) to \(\tilde{x}\) and \(\tilde{y}\) so that the first equation holds for the lifts as well. We may without loss of generality assume that \(\tilde{y}\) is in a fixed fundamental domain \(F\). Combining the two above equations and denoting \(\tilde{y}_q = \phi_q \tilde{y}\) and \(\tilde{y}_q = \exp(qU)\tilde{y}_q\), yields for some \(\gamma \in \Gamma\):

\[
\left(\exp(pU) \exp(aV) \exp(bX) \text{Bow}(2^j, \varepsilon', e) \exp(-qU)\right) \cap \\
\left(\exp(cV) \exp(dX) \tau(g') \gamma \gamma_\gamma^{-1}\right) \neq \emptyset.
\]

In what follows below, we will conjugate equations by \(\exp(sX)\) and \(\exp(rU)\) and use the fact that they preserve measure, hence the measure of the set of \(y\) for which the above holds will be equal to that of the conjugated equation.

Conjugating the above equation by \(\exp(-sX)\), with \(s = \frac{1}{2} \log(2^j)\) and using \([X, U] = 2U\) and \([X, V] = -2V\), implies that for \(\tilde{g}' = \exp(-sX)\tilde{y}_q\) and some \(\gamma \in \Gamma\), we have

\[
\left(\exp(p'U) \exp(a'V) \exp(bX) \text{Bow}(2^{j/2}, \varepsilon', e) \exp(sX) \exp(-q'U)\right) \cap \\
\left(\exp(c'V) \exp(dX) \exp(-sX) \text{Bow}(2^{j/2}, \varepsilon', e) \exp(sX) \tilde{g}' \gamma \gamma_\gamma^{-1}\right) \neq \emptyset,
\]

It follows that Lemma 7.7 is one of the main new tools which allows to generalize Ratner’s results from [23].
where
\[ p', q' \in [2^{20}j^6, 2^{40}j^6] \mid a', |c'| \in [2^{-10}j^6, 1] \text{ and } |b|, |d| \leq \epsilon'. \] (55)

Moreover by Lemma 3.7, it follows that
\[
\exp(-sX) B_{GW}(2^{1/2}, \epsilon', e) \exp(sX) \subset \exp(C(U, X)_{\epsilon'}) B_G(e, 2^{-j/2}).
\]

where \( C(U, X)_{\epsilon'} = B_g(0, \epsilon) \cap C(U, X) \). Denote
\[
h = \exp(-dX) \exp(-c'V) \exp(p'U) \exp(p'V) \exp(bX) \quad \text{and} \quad \tilde{B} = h B_G(e, 2^{-j/2}) h^{-1}.
\]

Notice that for \( w \in \{d, -c', a', b\} \) and \( W \in \{U, V, X\} \) (since all the numbers are small), we have for every small enough \( r > 0 \), \( \exp(wW) B_G(e, r) \exp(-wW) \subset B_G(e, Cr) \) (with a global constant \( C \)). Moreover, by the bound on \( p' \) (see (55)) it follows that \( \exp(p'U) B_G(e, 2^{-j/2}) \exp(-p'U) \subset B_G(e, 2^{-j/2+K_U j^6}) \), for some global constant \( K_U \).

Therefore, (by enlarging \( K_U \) if necessary), we have
\[
\tilde{B} \subset B_G(e, 2^{-j/2+K_U j^6}).
\] (56)

Therefore (54), (56) and Lemma 3.3 imply that
\[
\tilde{y}' \gamma(\tilde{y}')^{-1} \subset B_G(e, 2^{-j(1/4-K_U j^6)}) m(p', q', a', b, c', d) \exp(C(U, X)_{\epsilon'1/3})
\] (57)

where \( m(p', q', a', b, c', d) = \exp(-dX) \exp(-c'V) \exp(p'U) \exp(p'V) \exp(bX) \exp(-q'U) \).

From now, instead of the previous equation, we consider the square of the previous equation. The reason is that in the Appendix we do the computations in \( SL(2, \mathbb{R}) \) (and then transfer to \( G \)) and this allows to deal with \(-id \in SL(2, \mathbb{R})\). Notice that considering the adjoint action of components of \( m(p', q', a', b, c', d) \) one by one (see e.g. (13) and (14) ) on the ball \( B_G(e, 2^{-j(1/2-K_U j^6)}) \) and since the elements in \( C(U, X)_{\epsilon'1/3} \) are small, we have if \( m = m(p', q', a', b, c', d) \), then
\[
(m \exp(C(U, X)_{\epsilon'1/3}))^{-1} B_G(e, 2^{-j(1/4-K_U j^6)}) m \exp(C(U, X)_{2\epsilon'1/3}) \subset B_G(e, 2^{-j(1/4-2K_U j^6)}).
\] (58)

Then our new equation is:
\[
\tilde{y}' \gamma^2(\tilde{y}')^{-1} \subset B_G(e, 2^{-j(1/4-K_U j^6)}) m^2(p', q', a', b, c', d) \exp(C(U, X)_{2\epsilon'1/3}).
\] (59)

In the Appendix we will show that for every \( p', q', a', b, c', d \) as above we have the following: there exists \( K' \) such that

**P1.** \( \delta G(m^2(p', q', a', b, c', d), e) < K' j^6 \),

and

**P2.** \( m^2(p', q', a', b, c', d) = h \exp(sX) h^{-1} \), where \( 40j^6 - 5 \leq |s| \leq 160j^6 + 6 \), \( h \) commutes with \( C(U, X) \) and
\[
h^{-1} B_G(e, 2^{-j(1/2-2K_U j^6)+1}) h \subset B_G(e, 2^{-j(1/4-K_U j^6)}).
\] (60)

Let \( (m_i)_{i=0}^{T(j)} \in G \), be a \( 2^{-j/2} \) dense set in \( m^2(p', q', a', b, c', d) \), i.e. for every \( m^2 \), there exists \( m_i \) such that \( d(m^2, m_i) \leq 2^{-j/2} \). Notice that by **P1** and Lemma 3.3 it can be done with
\[
T(j) \leq 2K_{ij} j^6 + 3j^2/2.
\] (61)

---

Here the reasoning follows from (13) with \( s = 0 \) and \( t = p' \).
for some $K''_0$ depending only on $U$, since $B_G(x, \varepsilon) \supset B_{SL(2,\mathbb{R})}(x, \varepsilon)$. Then (69) implies that for some $i \in \{1, \ldots T(j)\}$, we have

$$
\hat{y}' \gamma'^2(\hat{y}')^{-1} \in B_G(e, 2^{-j(\frac{1}{2} - 2K_U \delta) + 1})m_i \exp(C(U, X)_{2^{1/3}}). \tag{62}
$$

By P2 it follows that $m_i = g_i \exp(s_i X) g_i^{-1}$ and moreover that $g_i$ commutes with $C(U, X)$. Therefore and by (60), if we denote $\gamma' = \gamma'^2$ and $\tilde{z} = g_i \hat{y}'$, then (62) implies that

$$
\tilde{z} \gamma'(\tilde{z})^{-1} \in B_G(e, 2^{-j(\frac{1}{2} - K'_0 \delta)}) \exp(s_i X) \exp(C(U, X)_{2^{1/3}}) \tag{63}
$$

Since $T(j) \leq 2K''_0 j \delta(2^{3/2})$, by the above reasoning. Lemma 7.2 follows straightforward by the following lemma:

**Lemma 7.5.** For $j > j_\delta$ and for every $i \in \{1, \ldots T(j)\}$

$$
\mu(\{\tilde{z} \in F : \exists \gamma' \in \Gamma, \text{ such that } \tilde{z} \text{ satisfies } (63)\}) \leq \frac{1}{T(j)j^3}. \tag{64}
$$

We will now show Lemma 7.5.

Notice that if $\tilde{z} \in F$ is a solution of (63), then by triangle inequality and the bound on $s_i$ (see P2) it follows that

$$
d_G(\gamma', e) \leq 2d_C(\tilde{z}, e) + 2 + |s_i| \leq 81 j \delta.
$$

By [8] Theorem 1.7 it follows that for some constant $C_0 > 0$

$$
|\gamma \in \Gamma : d_G(\gamma, e) \leq 81 j \delta| \leq 2^{C_0 j \delta}. \tag{65}
$$

Define

$$
S = S_{i, \gamma} := \{\tilde{z} \in F : \text{ such that } \tilde{z} \text{ satisfies } (63) \text{ with } \gamma' = \gamma\}. \tag{66}
$$

By (65), Lemma 7.5 follows by showing:

**Lemma 7.6.** For every $j \geq J_\delta$ and $\gamma \in \Gamma$, with $d_G(\gamma, e) < 81 j \delta$ and every $i \in \{1, \ldots, T(j)\}$, we have

$$
\mu(S_{i, \gamma}) \leq \frac{1}{j^3 T(j)^{2^{C_0 j \delta}}}. \tag{67}
$$

For the remainder of the proof, we will let $U, X, V$ be denoted by $X^0_0, X^0_1$ and $X^0_2$ so as to simplify notation (our index on the superscript began at 1 before). In particular, \(\{X^k_l : k = 0, \ldots, n \text{ and } l = 0, \ldots, m_k\}\) is a basis of $\mathfrak{g}$. Observe that $C(X) = \langle X^k_{m_k/2} : k = 1, \ldots, n \text{ and } m_k \in 2\mathbb{Z}\rangle$. To prove Lemma 7.6, we need the following lemma:

**Lemma 7.7.** There exists a constant $D_U$ (depending on $U$ only) such that for $0 \leq t_{kl}, s_{kl} \leq 2^{-D_U j \delta}$ for $k = 1, \ldots, n$ and $l = 0, \ldots, m_k$, $l \neq m_k/2$, if there exists $(k_0, l_0)$ such that $|t_{k_0 l_0} - s_{k_0 l_0}| \geq 2^{-j(\frac{1}{2} - D_U \delta)}$, then

$$
\exp \left( \sum_k \sum_{l \neq m_k/2} t_{kl} X^k_l \right) S_{i, \gamma} \cap \exp \left( \sum_k \sum_{l \neq m_k/2} s_{kl} X^k_l \right) S_{i, \gamma} = \emptyset. \tag{68}
$$

\[\text{We have assumed (63) and transformed the equation to (63). Hence the measure of } y \text{ solving (63) (and hence also belonging to the set } EG(j) \text{ Lemma 7.3) is no larger than the measure of } y' \text{ solutions to (63).}\]
Before we prove Lemma 7.7, let us show how it implies Lemma 7.6 and hence also finishes the proof of Lemma 7.3.

**Proof of Lemma 7.6.** We take a maximal \(2^{-j \left(\frac{1}{2} - 2D_U \delta\right)}\)-separated set \(A\) inside the cube \([0, 2^{-D_U j \delta}]^{\dim G - \dim C(X)}\). Then by definition of maximal separated sets, we have

\[
|A| \geq 2^{j \left(\frac{1}{2} - 2D_U \delta\right) - 1)(\dim G - \dim C(X))
\]

(69)

Suppose \(s, t \in A\), then they differ in one direction by at least \(2^{-j \left(\frac{1}{2} - 2D_U \delta\right)}\). Notice that since \((t_{kl}) \in [0, 2^{-D_U j \delta}]^{\dim G - \dim C(X)}\), we have

\[
\exp \left( \sum t_{kl} X_k^k \right) S_{i, \gamma} \subset V_i(F),
\]

where \(V_i(F) = \{ x \in G : d_G(x, F) \leq 1 \}\).

Therefore, by (68), we have

\[
\mu(V_i(F)) \geq \mu \left( \bigcup_{(t_{kl}) \in A} \exp \left( \sum t_{kl} X_k^k \right) S_{i, \gamma} \right) = \sum_{(t_{kl}) \in A} \mu \left( \exp \left( \sum t_{kl} V_i^k \right) S_{i, \gamma} \right) = |A| \mu(S_{i, \gamma}),
\]

(70)

Letting \(C'_\Gamma = \mu(V_i(F))\), the above inequality and (69) implies:

\[
\mu(S_{i, \gamma}) \leq C'_\Gamma |A|^{-1} \leq C'_\Gamma 2^{-j \left(\frac{1}{2} - 2D_U \delta\right) - 1)(\dim G - \dim C(X))}.
\]

(71)

Since by assumption \(\dim G - \dim(C(X)) \geq 4\), for small enough \(\delta\) (smallness depending only on constants \(D_U\), \(K'_U\) and \(C_\Gamma\)), we have

\[
(j \left(\frac{1}{2} - 2D_U \delta\right) - 1)(\dim G - \dim C(X)) \geq 4(j \left(\frac{1}{2} - 2D_U \delta\right) - 1) \geq [C_\Gamma j \delta + K'_U j \delta + \frac{3}{2} j] + \frac{j}{10}.
\]

So by the bound on \(T(j)\) (see (61)), if \(2^{j/10} > C'_\Gamma^{-1} j^3\) (which is always true for large \(j\)), we have

\[
C_\Gamma 2^{-j \left(\frac{1}{2} - 2D_U \delta\right) - 1)(\dim G - \dim C(X))} \leq \frac{1}{j T(j) 2^{C_\Gamma j \delta}}.
\]

This and (71) finish the proof of Lemma 7.6. 

So it only remains to prove Lemma 7.7.

**Proof of Lemma 7.7.** Suppose that for \(t = (t_{kl}), s = (s_{kl})\) there is a \((k_0, l_0)\) such that \(|t_{k_0 l_0} - s_{k_0 l_0}| \geq 2^{-j \left(\frac{1}{2} - 2D_U \delta\right)}\) and suppose by contradiction that there is a \(\tilde{x}\) (different than the \(\tilde{x}\) from the start of this section) such that

\[
\tilde{x} \in \exp \left( \sum t_{kl} X_k^k \right) S_{i, \gamma} \cap \exp \left( \sum s_{kl} X_k^k \right) S_{i, \gamma}.
\]
By definition (see §60 and §63), that means there exists \( g_1, g_2 \in B(e, 2^{-j(1-K'_U \delta)}) \), \( 160j \delta + 6 > |\alpha| > 40j \delta - 5 \) (in fact \( \alpha = s_i \)) and \( h_1 = \exp(Y_1), h_2 = \exp(Y_2) \) with \( Y_1, Y_2 \in C(U, X)_{2^{\nu_1/3}} \), such that

\[
g_1 h_1 \exp(\alpha X) = \exp \left( - \sum t_{kl} X_i^k \right) \tilde{x} \gamma \tilde{x}^{-1} \exp \left( \sum t_{kl} X_i^k \right),
\]

\[
g_2 h_2 \exp(\alpha X) = \exp \left( - \sum s_{kl} X_i^k \right) \tilde{x} \gamma \tilde{x}^{-1} \exp \left( \sum s_{kl} X_i^k \right).
\]

(72)

Computing \( \tilde{x} \gamma \tilde{x}^{-1} \) from the second equation and plugging into the first one, we get

\[
g_1 h_1 \exp(\alpha X) = \exp \left( - \sum t_{kl} X_i^k \right) \exp \left( \sum s_{kl} X_i^k \right) g_2 h_2 \exp(\alpha X) \exp \left( - \sum s_{kl} X_i^k \right) \exp \left( \sum t_{kl} X_i^k \right).
\]

(73)

**Lemma 7.8.** Let

\[
M : (Y_1, \ldots, Y_n) \mapsto \log_{\mathfrak{sl}}(\exp(Y_1) \ldots \exp(Y_2)).
\]

Then there exists \( \kappa > 0, \sigma > 0 \) such that if \( ||(Y_1, \ldots, Y_n)|| < \sigma \), then

\[
||M(Y_1, \ldots, Y_n) - \sum Y_i|| < \kappa \max \{ ||Y_i|| \cdot ||Y_j|| \}.
\]

**Proof.** This is an easy corollary of Taylor’s theorem with the knowledge that \([Y_i, Y_j] = 0\) for every \( i \). \( \square \)

As \( 0 \leq s_{kl}, t_{kl} \leq 2^{-D_U j \delta} \), by Lemma 7.8 it follows that there exists \( K''_U > 0 \) and \( g'_2 \in B(e, 2^{-j(\frac{1}{2} - K''_U \delta)}) \) such that

\[
\exp \left( - \sum t_{kl} X_i^k \right) \exp \left( \sum s_{kl} X_i^k \right) g_2 = g'_2 \exp \left( - \sum t_{kl} X_i^k \right) \exp \left( \sum s_{kl} X_i^k \right).
\]

Using this and multiplying (73) by \( \exp(\alpha X)^{-1} \) from the right, we have

\[
g'_2^{-1} g_1 \exp(Y_1) = \exp \left( - \sum t_{kl} X_i^k \right) \exp \left( \sum s_{kl} X_i^k \right) \exp(Y_2) \exp(\alpha X) \cdot \exp \left( - \sum s_{kl} X_i^k \right) \exp \left( \sum t_{kl} X_i^k \right) \exp(-\alpha X). \tag{74}
\]

We apply Lemma 7.8 to write \( \exp(- \sum t_{kl} X_i^k) \exp(\sum s_{kl} X_i^k) \) as

\[
\exp \left( Y' + \sum (s_{kl} - t_{kl}) X_i^k \right),
\]

where \( ||Y'|| < \kappa \max \{ |s_{kl}|, |t_{kl}| \} \). Notice that since the chain basis elements are eigenvectors for \( \text{ad}_X \), we have by applying the conjugation to \( \exp(Y' + \sum (s_{kl} - t_{kl}) X_i^k) \):

\[
\exp(\alpha X) \exp \left( - \sum s_{kl} X_i^k \right) \exp \left( \sum t_{kl} X_i^k \right) \exp(-\alpha X) = \exp \left( Y'' + \sum (s_{kl} - t_{kl}) e^{(m_k-2)\alpha} X_i^k \right) \tag{75}
\]

\[
\exp \left( - \sum t_{kl} X_i^k \right) \exp \left( \sum s_{kl} X_i^k \right) \exp(Y_2) \exp(\alpha X) \cdot \exp \left( - \sum s_{kl} X_i^k \right) \exp \left( \sum t_{kl} X_i^k \right) \exp(-\alpha X).
\]
where if $Y' = \sum y_{kl} X^k_i$, $Y'' = \sum e^{(m_{k-2})\alpha} y_{kl} X^k_i$, so

$$||Y'' - Y'|| < \kappa \max \left\{ 1 - e^{(m_{k-2})\alpha} \right\} \max \left\{ \left| s_{k1l}, t_{k2l} \right| \right\} .$$

Therefore, since all terms are small

$$\exp \left( -\sum t_{kl} X^k_i \right) \exp \left( \sum s_{kl} X^k_i \right) \exp(Y_2) \exp(\alpha X) \cdot$$

$$\exp \left( -\sum s_{kl} X^k_i \right) \exp \left( \sum t_{kl} X^k_i \right) \exp(-\alpha X) =$$

$$\exp \left( Y'' + \sum (s_{kl} - t_{kl}) X^k_i \right) \exp(Y_2) \exp \left( -Y'' - \sum (s_{kl} - t_{kl}) e^{(m_{k-2})\alpha} X^k_i \right) =$$

$$\exp \left( (Y' - Y'') + Y'' + \sum (s_{kl} - t_{kl})(1 - e^{(m_{k-2})\alpha}) X^k_i + Y_2 \right) \quad (76)$$

where

$$||Y'|| \leq \kappa \max \left\{ \left| Y' + \sum (s_i - t_j) V_i \right| \cdot ||Y_2||, \left| Y'' + \sum (s_i - t_j) e^{\lambda_i V_i} \right| \cdot ||Y_2||, \left| Y' + \sum (s_i - t_j) V_i \right| \cdot \left| Y'' + \sum (s_i - t_j) e^{\lambda_i V_i} \right| \right\} .$$

Thus if we set $Y = Y'' + Y' - Y''$, we get the following very rough bound since all terms in $Y''$ and $Y' - Y''$ are products of terms bounded by this small number:

$$||Y|| \leq e^{j/100} \left\| \sum (s_{kl} - t_{kl})(1 - e^{(m_{k-2})\alpha}) X^k_i \right\|. \quad (77)$$

Finally notice that since $40j\delta - 5 < |\alpha| < 160j\delta + 6$ and by the choice of $k_0, l_0$ (recall $|s_{k0l0} - t_{k0l0}| \geq 2^{-j(\frac{1}{2} - D_U \delta)}$), we have

$$\max \left\{ \left| (s_{kl} - t_{kl})(1 - e^{(m_{k-2})\alpha}) \right| \geq \left| (s_{k0l0} - t_{k0l0})(1 - e^{(m_{k-2})\alpha}) \right| \right\} \geq$$

$$\frac{1}{2} |s_{k0l0} - t_{k0l0}| \geq \frac{1}{2} \cdot 2^{-j(\frac{1}{2} - D_U \delta)}. \quad (78)$$

Recall that $g_1 \in B(e, 2^{-j(\frac{1}{2} - K''_U \delta)})$, $g_2' \in B(e, 2^{-j(\frac{1}{2} - K''_U \delta)})$ and then there exists a constant $K'''_U > 0$ only depends on chain structure such that

$$g_2'^{-1} \cdot g_1 \in B(e, 2^{-j(\frac{1}{2} - K'''_U \delta)}).$$

Thus, by Lemma 7.8, we have

$$g_2'^{-1} \cdot g_1 \exp(Y_1) = \exp \left( Y_1 + \sum c_{kl} X^k_i \right)$$

where $|c_i| \leq 2^{-j(\frac{1}{2} - K'''_U \delta)}$. It follows that the $X^k_i$ coefficients of the LHS of (74) are less than $2^{-j(\frac{1}{2} - K'''_U \delta)}$ (since $Y_1 \in C(U, X) \subset C(X)$). However, the coefficient by $k_0, l_0$ of the RHS of (74) is by the definition of $Y$, (78), (77) and (76) bounded below by $\frac{1}{2} \cdot 2^{-j(\frac{1}{2} - D_U \delta)}$, hence if $D_U$ is large enough (in terms of $K'''_U$) we get a contradiction with (74). This finishes the proof. \qed
Appendix: Proof of P1 and P2.

Recall that we have the homomorphism $\phi : \mathfrak{sl}(2, \mathbb{R}) \rightarrow \mathfrak{g}$ taking the standard horocyclic generator of $\mathfrak{sl}(2, \mathbb{R})$ to $U$. By Lemma 3.4, $\phi$ extends to $\tilde{\phi}$. Therefore to prove P1 and P2 it is enough to make the computations in $SL(2, \mathbb{R})$.

Recall that

$$m(p, q, a, b, c, d) = \exp(-dX) \exp(-cV) \exp(pU) \exp(aV) \exp(bX) \exp(-qU), \quad (79)$$

where $U, V, X$ with

$$p, q \in [2^{20}j^2, 2^{40}j^2], \ |a|, |c| \in [2^{-10}j^2, 1], \ |b|, |d| \leq \varepsilon'.$$  

(80)

We can consider this matrix in $SL(2, \mathbb{R})$ or $G$ equivalently. We have

**Lemma 7.9.** For every $p, q, a, b, c, d$ as in (80), we have that $m^2 = h \exp(sX)h^{-1}$, where $|s| \in [40j^2 - 5, 160j^2 + 6]$ and $h = k \exp(\alpha U)$, with $k \in SO(2)$ and $|\alpha| \leq 2K_U \delta j$ for some fixed $K_U$.

Before we prove the above Lemma let us show how it implies P1 and P2.

**Proof of P1 and P2.** Notice that by right invariance of $d_{SL(2, \mathbb{R})}$ and triangle inequality, we have

$$\frac{1}{2} d_{SL(2, \mathbb{R})}(m^2, e) \leq d_{SL(2, \mathbb{R})}(m, e) \leq d_{SL(2, \mathbb{R})}(\exp(-dX), e) + d_{SL(2, \mathbb{R})}(\exp(-cV), e) + d_{SL(2, \mathbb{R})}(\exp(pU), e) + d_{SL(2, \mathbb{R})}(\exp(aV), e) + d_{SL(2, \mathbb{R})}(\exp(bX), e) + d_{SL(2, \mathbb{R})}(\exp(-qU), e)$$

Notice that by (80) it follows that the terms of the RHS with $d, c, a, b$ are bounded. Moreover, by Lemma 3.6 it follows that

$$d_{SL(2, \mathbb{R})}(\exp(pU), e), d_{SL(2, \mathbb{R})}(\exp(-qU), e) \leq C \max(\log p, \log q) \leq 40Cj\delta.$$ 

This finishes the proof of P1.

Notice also that the first part of P2 follows from Lemma 7.9 with $h = k \exp(\alpha U)$. Let us now show (60). Notice that

$$k \exp(\alpha U)B_G(e, 2^{-j(1/2 - 2K_U \delta) + 1}) \exp(-\alpha U)k^{-1} =$$

$$\exp(\alpha U)B_G(e, 2^{-j(1/2 - 2K_U \delta) + 1}) \exp(-\alpha U). \quad (81)$$

Futhermore, by the bound on $\alpha$ and (13), we get that the right hand side above is contained in $B_G(e, 2^{-j(1/2 - K^\delta_U \delta)})$ for sufficiently large $j$. This finishes the proof.

So we only need to prove Lemma 7.9.

**Proof of Lemma 7.9.** Then by direct computation, we have

$$m = \begin{pmatrix}
e b \left( ae^{-d}p + e^{-d} \right) & e^{-b-d}p - e^{-b} \left( ae^{-d}p + e^{-d} \right) q \\
e b \left( a \left( e^{-d}p - ce^d \right) \right) - e^{-b} \left( e^{-d} - ce^d \right) & e^{-b} \left( e^{-d} - ce^d \right) - e^{-b} \left( a \left( e^{-d} - ce^d \right) \right) q
\end{pmatrix}. \quad (82)$$

The trace of this matrix is

$$e^{-b-d} \left( qe^{2(b+d)}(a(cp - 1) + c) + e^{2b}(ap + 1) + e^{2d}(1 - cp) \right).$$
Notice that $|acpq|e^{b+d}$ is the dominating term above (see (80)), we have $\frac{9}{10}|acpq|e^{b+d} \leq \text{Tr}(m) \leq \frac{11}{10}|acpq|e^{b+d}$. Therefore,

$$2^{20j\delta-1} \leq |\text{Tr}(m)| \leq 2^{80j\delta+1}. \quad (83)$$

Since $|\text{Tr}(m)| > 2$, $m$ is diagonalizable. As a result we have $\text{Tr}(m) = \lambda^{-1} + \lambda$ (suppose $|\lambda| > 1$), with $2^{20j\delta-2} \leq |\lambda| \leq 2^{80j\delta+2}$. The estimate of $m^2$'s eigenvalue will follow from this.

So $m$ can be diagonalized as $m = h'(\begin{pmatrix} \lambda \\ \lambda^{-1} \end{pmatrix})h'^{-1}$. Write $h' = kn'$, with $k \in SO(2, \mathbb{R})$, $a$ a diagonal matrix, and $n' = \exp(\alpha'U)$ for some $\alpha' \in \mathbb{R}$. Then $h' = kn$, and notice that if $\pm h' \exp(sX)h'^{-1} = m$, then if $h = kn$, $\pm h \exp(sX)h^{-1} = m$.

Now, $d(\pm h \exp(sX)h^{-1}, e) = d(\pm kn \exp(sX)n^{-1}k^{-1}, e) = d(\pm n \exp(sX)n^{-1}, e)$. If $n = \exp(\alpha U)$, then $n \exp(sX)n^{-1} = \exp((1 - \lambda^{-2})\alpha U) \exp(sX)$, so

$$d(\pm \exp((1 - \lambda^{-2})\alpha U), e) = d(n \exp(sX)n^{-1} \exp(-sX), e) \leq d(n \exp(sX)n^{-1}, e) + d(\exp(-sX), e) = d(m, e) + |s| \leq 160Cj\delta + 160j\delta + 4. \quad (84)$$

On the other hand, $d(\pm \exp((1 - \lambda^{-2})\alpha U), e) \geq c \log |(1 - \lambda^{-2})\alpha| - \pi \geq c \log |\alpha| + c \log |1 - \lambda^{-2}| - \pi$ (we get a $\pi$ because we may need to multiply by $-\text{id}$ which has distance $\pi$ to $e$). Therefore, $\log |\alpha| \leq 2K'j\delta.$

References

[1] L. M. Abramov, *The entropy of a derived automorphism*. (Russian) Dokl. Akad. Nauk SSSR 128 1959 647–650.

[2] F. Beleznay, M. Foreman, *The complexity of the collection of measure-distal transformations*. Ergodic Theory Dynam. Systems 16 (1996), no. 5, 929–962.

[3] I. U. Bronšteĭn, *Extensions of minimal transformation groups*. Translated from the Russian. Martinus Nijhoff Publishers, The Hague, 1979. viii+319 pp. ISBN: 90-286-0368-9

[4] A. Brudnyi, *On local behavior of analytic functions*, (English summary) J. Funct. Anal. 169 (1999), no. 2, 481–493.

[5] T. de la Rue, *Systèmes dynamiques gaussiens d’entropie nulle, lâchement et non lâchemen t Bernoulli*. (French) [Zero-entropy Gaussian dynamical systems that are loosely and not loosely Bernoulli] Ergodic Theory Dynam. Systems 16 (1996), no. 2, 379–404.

[6] H. A. Dye, *On groups of measure preserving transformation. I*. Amer. J. Math. 81 1959 119–159.

[7] H. A. Dye, *On groups of measure preserving transformations. II*. Amer. J. Math. 85 1963 551–576.
[8] A. Gorodnik, A. Nevo, *Counting lattice points*. (English summary) J. Reine Angew. Math. 663 (2012), 127–176.

[9] J. Feldman, *New K-automorphisms and a problem of Kakutani*. Israel J. Math. 24.1 (1976): 16-38.

[10] S. Ferenczi, *Systems of finite rank*. Colloq. Math. 73 (1997), no. 1, 35–65.

[11] M. Foreman, D. Rudolph, B. Weiss. The conjugacy problem in ergodic theory. *Ann. of Math. (2)* 173 (2011), no. 3, 1529 – 1586.

[12] M. Foreman, B. Weiss, *An anti-classification theorem for ergodic measure preserving transformations*. (English summary) J. Eur. Math. Soc. (JEMS) 6 (2004), no. 3, 277–292.

[13] S. Kakutani, *Induced measure preserving transformations*. Proc. Imp. Acad. Tokyo 19, (1943). 635–641.

[14] A. B. Katok, *Time change, monotone equivalence, and standard dynamical systems*. (Russian) Dokl. Akad. Nauk SSSR 223 (1975), no. 4, 789–792.

[15] A. B. Katok, *Monotone equivalence in ergodic theory*. (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 41 (1977), no. 1, 104–157, 231.

[16] A. B. Katok, J. P. Thouvenot, *Slow entropy type invariants and smooth realization of commuting measure-preserving transformations*. Ann. Inst. H. Poincaré Probab. Statist. 33 (1997), no. 3, 323–338.

[17] A. Kanigowski, K. Vinhage, D. Wei, *Slow entropy of parabolic flows*. preprint arXiv:1711.09460

[18] D. Y. Kleinbock, N. Shah, A. Starkov, *Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory*. Handbook of dynamical systems. Vol. 1A, 813–930, North-Holland, Amsterdam, 2002.

[19] D. S. Ornstein, *Ergodic theory, randomness, and dynamical systems*. James K. Whittemore Lectures in Mathematics given at Yale University. Yale Mathematical Monographs, No. 5. Yale University Press, New Haven, Conn.-London, 1974. vii+141 pp.

[20] D. Ornstein, D. Rudolph, B. Weiss, *Equivalence of measure preserving transformations*, Mem. Amer. Math. Soc., 37(262), 1982.

[21] M. Ratner, *Horocycle flows are loosely Bernoulli*. Israel J. Math. (1978), 31: 122-132.

[22] M. Ratner, *The Cartesian square of the horocycle flow is not loosely Bernoulli*. Israel J. Math. 34 (1979), no. 1-2, 72–96 (1980).

[23] M. Ratner, *Some invariants of Kakutani equivalence*. Israel J. Math. 38 (1981), no. 3, 231–240.

[24] M. Ratner, *Horocycle flows, joinings and rigidity of products*. Ann. of Math. (2) 118 (1983), no. 2, 277–313.
[25] M. Ratner, \textit{Interactions between ergodic theory, Lie groups, and number theory}. Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), 157–182, Birkhäuser, Basel, 1995.

[26] M. Ratner, \textit{Time change invariants for measure preserving flows}. (English summary) Modern theory of dynamical systems, 263–273, Contemp. Math., 692, Amer. Math. Soc., Providence, RI, 2017.

[27] Ya. G. Sinai, \textit{A weak isomorphism of transformations with invariant measure}, (Russian) Dokl. Akad. Nauk SSSR 147 (1962), 797–800.

[28] D. Witte, \textit{Rigidity of some translations on homogeneous spaces}, Invent. Math. 81 (1985), no. 1, 1–27.