ON MEASURES INVARIANT UNDER DIAGONALIZABLE ACTIONS – THE RANK ONE CASE AND THE GENERAL LOW ENTROPY METHOD

MANFRED EINSIEDLER AND ELON LINDENSTRAUSS

This paper is dedicated to Gregory Margulis and Marina Ratner.

Abstract. We consider measures on locally homogeneous spaces $\Gamma\backslash G$ which are invariant and have positive entropy with respect to the action of a *single* diagonalizable element $a \in G$ by translations, and prove a rigidity statement regarding a certain type of measurable factors of this action.

This rigidity theorem, which is a generalized and more conceptual form of the low entropy method of [Lin2, EKL] is used to classify positive entropy measures invariant under a one parameter group with an additional recurrence condition for $G = G_1 \times G_2$ with G_1 a rank one algebraic group. Further applications of this rigidity statement will appear in forthcoming papers.

Contents

1. Introduction

A well known problem in modern dynamics is to classify measures invariant under natural partially hyperbolic algebraic $\mathbb{Z}^d$ and $\mathbb{R}^d$-actions for $d \geq 2$. The simplest example is Furstenberg's Conjecture regarding the classification of probability measures on $\mathbb{R}/\mathbb{Z}$ invariant under a nonlacunary multiplicative subgroup of $\mathbb{Z}^\times$ (e.g. the group generated by $\times 2$ and $\times 3$).

A slightly more sophisticated class of systems exhibiting the same type of phenomenology is the action of multidimensional $\mathbb{R}$-diagonalizable groups H on a locally compact space $\Gamma\backslash G$ where G is an algebraic group over $\mathbb{R}$. A prototypical

---

Date: November 8, 2007.

This work was supported by the Clay Mathematics Institute and by the NSF grants DMS-0554373, 0622397 (ME), 0500205 and 0554345 (EL).

example of such an action is the action of the subgroup of $n \times n$-diagonal matrices of $\mathrm{SL}(n,\mathbb{R})$ on $\mathrm{SL}(n,\mathbb{Z})\backslash\mathrm{SL}(n,\mathbb{R})$. This particular example is linked to several open problems in number theory, including — as pointed out by Margulis — a conjecture by Littlewood regarding simultaneous Diophantine approximations. It is natural to expand the class of systems we are considering by working over more general local fields instead of $\mathbb{R}$, as well as considering products of linear groups over several different fields. In particular, the case where G is an S-algebraic group, i.e. a product of linear groups over $\mathbb{R}$ or p-adic fields appears naturally in applications (see e.g. [Lin2, EKl]).

So far, all of the progress regarding classifying invariant measures in such systems has been under the additional assumption of positive entropy; a prototypical example are the theorems of Rudolph and Johnson which state that a probability measure μ on $\mathbb{R}/\mathbb{Z}$ invariant and ergodic under the action of a nonlacunary multiplicative semigroup either has zero entropy with respect to any single element of the acting semigroup or is Lebesgue.

In the locally homogeneous context, work by Katok, Spatzier, and Kalinin [KS, KaS] required additional assumptions regarding the mixing properties of the invariant measure, assumptions which are very hard to verify in many applications.

One method which overcomes this problem in the locally homogeneous context — the *high entropy method* — has been introduced by M. E. and A. Katok. This method is inherently based on having non-proportional Lyapunov exponents and hence indeed an action of a multidimensional group. It moreover requires the restriction of the measure to leaves of two different (and transverse) foliations of the space by orbits of two distinct noncommuting unipotent subgroups be nontrivial, and in particular requires more than just positive entropy for one element of the action. Initially it was developed only for actions of $\mathbb{R}$-split Cartan subgroups of $\mathbb{R}$-split simple algebraic groups [EK1], but has been generalized to essentially an action of an arbitrary two-dimensional diagonalizable subgroup of S-algebraic groups in [EK2].

A different method, the *low entropy method*, has been introduced by E.L. in [Lin2], and subsequently used in conjunction with the high entropy method in our paper with A. Katok in [EKL]. It is this method that we set to generalize in this paper. In essence, it is not about the action of a multiparameter diagonalizable group but about a single parameter group, and gives a subtle restriction on the rich and rather wild class of probability measures invariant under such a one parameter group.

We remark that in all currently known approaches to measure classification in the locally homogeneous context, including both the low entropy and the high entropy method as well as [KS] and implicitly even in [Rud2], the notion of leafwise measures (which are also known as restricted measures or conditional measures along the leaves of a foliation) play a central role. Very generally, whenever we have a reasonable action of a locally compact group U on a locally compact[(1)] space X, and for any locally finite measure μ on X, we can obtain a system of leafwise measures μ_x^U which can be viewed as a map from X to the space of locally finite measures on U satisfying certain compatibility relations (see Section 3 for details). In the context of classifying measures invariant under diagonalizable actions, one is

[(1)]We implicitly assume that all locally compact spaces are also σ-compact.

mostly concerned with leafwise measures on orbits of one or more unipotent groups normalized by the action.

In [Lin2] (as well as [EKL]) essential use was made of ideas used by Ratner in her study of unipotent flows, and particularly from her earlier work on the horocyclic flow [Ra2, Ra1, Ra3]. Superficially, the measure classification results of [Lin2, EKL] have little in common with these results of Ratner. But as shown below, our main theorem can be interpreted as a theorem about factors of non-measure preserving actions, reminiscent of Ratner's factor rigidity theorem [Ra1].

To present our main result, we need the following:

Definition 1.1. Let X be a locally compact metric space with a Borel probability measure μ, let H be a locally compact metric group acting continuously and locally free on X. We denote the action by $h.x$ for $h \in H$ and $x \in X$. Let $\phi : X \to Y$ be a measurable map to a Borel space Y. Then ϕ is *locally H-aligned modulo μ* (or simply *locally H-aligned* if μ is understood) if for every $\epsilon > 0$ and neighborhood $O \ni e$ in H there exists $X' \subset X$ with $\mu(X') > 1 - \epsilon$ and some $\delta > 0$ such that for every $x \in X'$

$$\{x' \in X' : \phi(x') = \phi(x)\} \cap B_\delta(x) \subset O.x. \tag{1.1}$$

In other (less precise) words, ϕ is locally H-aligned modulo μ if up to a set of negligible μ measure, the level set $\phi(x) = c$ are locally contained in a single H-orbit.

Another notion we will need is recurrence, and more specifically relative to a Borel map ϕ.

Definition 1.2. With the notations of Definition 1.1:

(i) μ is *H-recurrent* if for every set B of positive μ-measure and a.e. $x \in B$ the set $\{h \in H : h.x \in B\}$ is unbounded (i.e. has non-compact closure).

(ii) μ is *H-transient* if for every $\epsilon > 0$ there is a set $B \subset X$ with $\mu(B) > 1 - \epsilon$ so that for every $x \in B$ the set $\{h \in H : h.x \in B\}$ is bounded.

(iii) μ is *H-recurrent relative to ϕ* if for every set B of positive μ-measure and a.e. $x \in B$

$$\{h \in H : h.x \in B \textit{ and } \phi(x) = \phi(h.x)\} \tag{1.2}$$

is unbounded.

(iv) μ is *H-transient relative to ϕ* if for every $\epsilon > 0$ there is a set $B \subset X$ with $\mu(B) > 1-\epsilon$ so that for every $x \in B$ the set of return times (1.2) is bounded.

Finally, we will need the following standard definition[(2)]. For any Borel space X, Y, etc. we let $\mathcal{X}, \mathcal{Y}, \dots$ denote the corresponding Borel σ-algebra.

Definition 1.3. Let X and Y be Borel spaces, H a locally compact group acting (Borel measurably) on both X and Y, and μ a Borel probability measure on X. A Borel map $\phi : X \to Y$ is a *factor map modulo μ* for H if there is a set $X' \subset X$ of full μ-measure so that for every pair $x, h.x \in X'$ $(h \in H)$

$$\phi(h.x) = h.\phi(x).$$

The space X we consider will be of the from $\Gamma\backslash G$ where G is either a Lie group or an S-algebraic linear group, and $\Gamma < G$ a discrete subgroup. By an S-algebraic linear group we mean a finite product of linear algebraic groups G_σ defined over

[(2)]The notion of factor is standard, but perhaps less so in the context of actions on measure spaces which do not even preserve the measure class.

various local fields $\mathbb{K}_\sigma$ for $\sigma \in S$ (where we allow repetitions). Unless otherwise stated, $\mathbb{K}_\sigma$ may have any characteristic (which may even depend on σ). An S-algebraic group will be said to be of characteristic zero if all of the fields $\mathbb{K}_\sigma$ are such.

For a Lie group G we will use $\mathfrak{g}$ to denote its Lie algebra, and similarly for algebraic groups over a field with zero characteristic. We extend the conventional definition of $\mathfrak{g}$, log, Ad more generally as follows:

- if $G = \prod_{\sigma \in S} G_\sigma$ is an S-algebraic group of fields of zero characteristic, we let $\mathfrak{g}_\sigma$ denote the Lie algebra of G_σ and $\mathfrak{g} = \prod_{\sigma \in S} \mathfrak{g}_\sigma$, and extend the definitions of log and Ad in the obvious way (component by component).
- in the case where at least one of the fields $\mathbb{K}_\sigma$ has positive characteristic, it will be useful to use a more generalized notion[(3)]: for every $\sigma \in S$ we take $\mathfrak{g}_\sigma$ to be some (more or less arbitrary) vector space over $\mathbb{K}_\sigma$, equipped with a homeomorphism log from a neighborhood of $e \in G_\sigma$ into a neighborhood Ω_σ of 0 in some algebraic subvariety $\bar{\Omega}_\sigma^Z$ of $\mathfrak{g}_\sigma$ containing zero, with $\log(e) = 0$. We take Ad to be a linear representation of G on $\mathfrak{g}$ so that
$$\mathrm{Ad}(g)[\log(h)] = \log(ghg^{-1}) \qquad \text{whenever both sides are defined.}$$
We can now again define $\mathfrak{g} = \prod_{\sigma \in S} \mathfrak{g}_\sigma$ and extend log and Ad to $\mathfrak{g}$ in the obvious way; also let $\Omega = \prod_\sigma \Omega_\sigma$.

We will say that an element $a = (a_\sigma)$ in an S-algebraic group G is *diagonalizable* if $\mathrm{Ad}(a_\sigma)$ is diagonalizable over the respective field of definition $\mathbb{K}_\sigma$ at every $\sigma \in S$. We say a is *of class $\mathcal{A}$* if additionally[(4)] (i) for every σ, if λ, λ' are eigenvalues of $\mathrm{Ad}(a_\sigma)$ with $|\lambda|_\sigma = |\lambda'|_\sigma$ then $\lambda = \lambda'$, (ii) for every σ, if λ is an eigenvalue of $\mathrm{Ad}(a_\sigma)$ with $|\lambda|_\sigma = 1$ then $\lambda = 1$, (iii) for some σ the map $\mathrm{Ad}(a_\sigma)$ has at least one eigenvalue of absolute value > 1 and at least one with absolute value < 1. If G is a Lie group we say that $a \in G$ is diagonalizable if $\mathrm{Ad}(a)$ is diagonalizable over $\mathbb{R}$ and that it is class $\mathcal{A}$ if all the eigenvalues are positive, at least one is > 1 and at least one is < 1.

We fix some class $\mathcal{A}$ element $a \in G$, and let $\mathfrak{g}_-$ (and similarly $\mathfrak{g}_+$) denote the subspace of $\mathfrak{g}$ (in the S-algebraic case, the $\bigoplus_{\sigma \in S} \mathbb{K}_\sigma$-submodule) generated by all eigenspaces of $\mathrm{Ad}(a)$ corresponding to eigenvalues of absolute value < 1 (respectively, > 1). We let $\mathfrak{g}_0$ denote the eigenspace(s) of $\mathrm{Ad}(a)$ for the eigenvalue 1. We will also need the related groups
$$\begin{aligned} G^- &= \{g \in G : a^n g a^{-n} \to e \text{ for } n \to \infty\}, \\ G^+ &= \{g \in G : a^n g a^{-n} \to e \text{ for } n \to -\infty\}, \\ G^0 &= C_G(a). \end{aligned} \tag{1.3}$$
Note that log is a continuous injective map from a neighborhood of $e \in G^-$ to a neighborhood of $0 \in \mathfrak{g}_-$, and similarly for G^+, G^0.

Let $U < G$ be a closed group which is (i) normalized by a, (ii) $U \le G^-$. If G is an S-algebraic group, we assume further that (iii) U is a direct product of Zariski

[(3)]While Lie algebras make perfect sense over any field, there is no reasonable log function defined on a neighborhood of the identity $e \in G_\sigma$ to the Lie algebra. Since we do want to use a map to a linear space — still called the log, we need to loosen the definition of Lie algebra accordingly (see Section 6.1).

[(4)]This is somewhat more general than the notion of class $\mathcal{A}$ elements used by Tomanov and Margulis in [MT2].

closed unipotent subgroups U_σ of G_σ for $\sigma \in S$ and (iv) Ad restricted to U_σ is an algebraic representation (this is automatic in the zero characteristic case, but needs to be assumed explicitly in the more general context). *When we talk about Zariski closed subgroups of U we always mean direct products of Zariski closed subgroups of U_σ for $\sigma \in S$* (even if two algebraic groups for different elements of S are over the same field). If G is a (not necessarily algebraic) Lie group, we can still consider U as an algebraic group over $\mathbb{R}$ since $\log(\cdot)$ and $\exp(\cdot)$ can be extended to give bijections between U and its Lie algebra which we can consider as an affine space. Moreover, Zariski closed subgroups of U are precisely the subgroups closed and connected with respect to the Hausdorff topology.

Theorem 1.4 (Main Theorem). *Suppose $X = \Gamma\backslash G$, $a \in G$, and $U < G$ are as above, with G a Lie group or an S-algebraic group, $\mathfrak{g}$ its "Lie algebra" in the generalized sense considered above, and a of class $\mathcal{A}$. Let μ be an a-invariant U-recurrent measure on X, and let $\phi : X \to Y$ be a factor map modulo μ for the action of the group $a^{\mathbb{Z}}U$. Assume that*

(Φ) *the map $x \mapsto \mu_x^U$ is $\phi^{-1}(\mathcal{Y})$-measurable,*

(U-1) *for almost every x there is no a-normalized Zariski closed proper subgroup of U supporting μ_x^U,*

(U-2) *for all nonzero $w \in \mathfrak{g}_+$, it holds that $\operatorname{Ad}(U)[w] \not\subset \mathfrak{g}_+ \oplus \mathfrak{g}_0$.*

Then μ is the convex combination of two a-invariant measures μ_1 and μ_2 such that:

(LE-1) *ϕ is locally $C_G(U) \cap G^0$-aligned modulo μ_1*

(LE-2) *μ_2 is $C_G(U) \cap G^-$-recurrent relative to ϕ.*

Note that in the case of an ergodic measure, the above establishes one of the properties (LE-1) or (LE-2) must hold for μ. A particularly interesting case for the factor map ϕ is the map $x \mapsto \mu_x^U$ (or more precisely the map taking x to the equivalence class of μ_x^U under proportionality, which determines μ_x^U — see Section 3 for more details). This is, for instance, the choice of ϕ used in the proof of Theorem 1.5 below.

The above theorem generalizes the low entropy method developed by E.L. in [Lin2] and extended in a joint paper of the authors with A. Katok [EKL]. While the main outline of the proof remains the same, the general case considered here requires several new ideas. The main new difficulty is that unlike the case in [Lin2, EKL] the group U need not be one dimensional — indeed, it may be a pretty general nilpotent group. This requires careful analysis of the structure of the leafwise measure of μ_x^U. Another novelty is the extraction of an abstract form of the low entropy method — which is essentially a rigidity statement regarding a *one* parameter diagonalizable flow[(5)]. We note that in the terminology of [EKL] the possibility (LE-1) corresponds to exceptional returns. Moreover, the condition (LE-2) satisfied by μ_2 implies in many cases, just as in [Lin2] and [EKL], that μ_2 has further invariance properties under groups generated by unipotents, which allows one to use the powerful tools from the theory of unipotent flows such as Ratner's measure classification theorem [Ra4] to study the ergodic components of μ (or more precisely μ_2).

[(5)]Note that even the additional recurrence condition which was present in [Lin2] as a weak substitute for additional invariance is not explicitly used, though implicitly additional invariance or recurrence is needed to find a suitable factor map for which both (LE-1) and (LE-2) convey meaningful information.

Note that by decomposing μ and by passing to an algebraic subgroup if necessary, it is always possible to reduce to the case that U satisfies (U-1). Assumption (U-2) however is a substantial assumption which seems fairly difficult to remove. As the applications below show it is mild enough to allow us to apply our main theorem in many contexts, e.g. it is always satisfied if G is semisimple and $U = G^-$. However there certainly are interesting cases where Theorem 1.4 cannot be applied because of it. A simple example of a situation where Theorem 1.4 is *not* applicable is when $G = \mathrm{SL}(3,\mathbb{R})$, $a = \begin{pmatrix} t & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & t^{-1} \end{pmatrix}$ and U the one-dimensional unipotent subgroup $\begin{pmatrix} 1 & s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ of G since in this case $w = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ is an element of $\mathfrak{g}_+$ which is invariant under $\mathrm{Ad}(U)$. On the other hand, for the same a and G the group $U' = \begin{pmatrix} 1 & 0 & s \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ does satisfy (U-2) of Theorem 1.4.

The low entropy method was originally developed in [Lin2] to study measures on $\Gamma\backslash\mathrm{SL}(2,\mathbb{R})\times\mathrm{SL}(2,\mathbb{Q}_p)$, Γ an irreducible lattice such as $\mathrm{SL}(2,\mathbb{Z}[1/p])$, which are invariant under the diagonal subgroup of $\mathrm{SL}(2,\mathbb{R})$ and recurrent under the action of the group $\mathrm{SL}(2,\mathbb{Q}_p)$. Such measures arise naturally when one studies how Hecke Maass forms on $\Gamma\backslash\mathbb{H}$ are distributed. As a relatively straightforward application of our main theorem we prove the following generalization of [Lin2, Theorem 1.1] to products $G = G_1 \times G_2$ where G_1 is a general (zero characteristic) rank one semisimple algebraic group. Theorem 1.4 can also be used to study this question in the positive characteristic case, but as this involves some complications we defer the positive characteristic case to a later paper.

Recall that if G_1 is a $\mathbb{K}_\sigma$-rank one semisimple algebraic group defined over a local field $\mathbb{K}_\sigma$, if $A_1 < G_1$ is a maximal $\mathbb{K}_\sigma$-split torus then the centralizer $C_{G_1}(A_1)$ is a reductive group with a nontrival $\mathbb{K}_\sigma$-character χ; if we set $M_1 = \ker\chi$ then M_1 is compact (cf. [PR, Thm. 3.1, pg. 108]) and A_1M_1 is of finite index in $C_{G_1}(A_1)$.

We recall that an orbit of a group H on $\Gamma\backslash G$ is called *periodic* if it supports a finite H-invariant measure; a measure on $\Gamma\backslash G$ is called *homogeneous* if it is the unique H-invariant probability measure on a periodic H-orbit.

Theorem 1.5. *Let $G = G_1 \times G_2$ where G_1 is a semisimple linear algebraic group over a characteristic zero local field $\mathbb{K}_\sigma$ with $\mathbb{K}_\sigma$-rank 1 and G_2 is a zero characteristic S-algebraic group. Let $\Gamma \subset G$ be a discrete subgroup. Let A_1 be a $\mathbb{K}_\sigma$-split torus of G_1 and let χ be a nontrivial $\mathbb{K}_\sigma$-character of A_1 that can be extended to $C_G(A_1)$. Let $M_1 = \{h \in C_G(A_1) : \chi(h) = 1\}$. Let μ be an A_1-invariant, G_2-recurrent probability measure on $\Gamma\backslash G$ such that*

(1) *almost every A_1-ergodic component of μ has positive ergodic theoretic entropy with respect to some $a \in A_1$ with $|\chi(a)|_\sigma \neq 1$ and*
(2) *for μ-a.e. x the group*
$$\{h \in M_1 \times G_2 : h.x = x\}$$
is finite.

Then μ is a convex combination of homogeneous measures. Each of these homogeneous measures is supported on an orbit of a subgroup H which, after restriction *of scalars*[(6)] *to a local subfield* $\mathbb{F}_\sigma$*, contains a finite index subgroup of a semisimple algebraic subgroup of* G_1 *of* $\mathbb{F}_\sigma$*-rank one.*

In case of characteristic zero considered here the subfield $\mathbb{F}_\sigma$ of $\mathbb{K}_\sigma$ can be taken to be $\mathbb{R}$ resp. $\mathbb{Q}_p$.

This theorem for the characteristic zero case has been announced[(7)] in [EL2, Thm. 2.9]. Condition (1) regarding entropy is due to limitations of our techniques (as well as any other technique known in this context) to deal with zero entropy measures. We expect a similar theorem to hold without it, but the proof of such a theorem is well beyond the reach of current technology. Condition (2) on the other hand is essential: as explained in Section 8, if it fails very little information is conveyed by the fact that μ is G_2-recurrent, and without this recurrence condition it is well-known that there is an abundance of a-invariant measures on X.

There is also a straightforward analog for Lie groups:

Theorem 1.6. *Let* $G = G_1 \times G_2$ *where* G_1 *is a rank 1 semisimple Lie group with finite center and* G_2 *is any Lie group. Let* $\Gamma \subset G$ *be a discrete subgroup. Let* $a \in G$ *be of class* $\mathcal{A}$*, and let* M_1 *be a maximal compact subgroup of* G_1^0*. Let* μ *be an* a*-invariant,* G_2*-recurrent probability measure on* $\Gamma\backslash G$ *such that*

(1) *almost every* a*-ergodic component of* μ *has positive ergodic theoretic entropy and*
(2) *for* μ*-a.e.* x *the group*
$$\{h \in M_1 \times G_2 : h.x = x\}$$
is finite.

Then μ *is a convex combination of* a*-ergodic and invariant homogeneous measures.*

The proof is very similar to that of Theorem 1.5 and is left to the reader.

Further applications of Theorem 1.4 will appear in subsequent papers, in particular:

(1) A partial classification of A-invariant and ergodic probability measures on quotients of S-algebraic groups G where A is a maximal S-split torus in G, which generalizes the results from [EKL] (see [EL2, §2] for more details). Here we have to assume positive entropy and have to allow the possibility that the measure is supported on an orbit of a subgroup which allows a rank one factor (which is the situation occuring in certain examples constructed by M. Rees [Ree] and their generalizations).
(2) A full classification of joinings for the actions of higher rank S-split tori on quotients of S-algebraic groups by lattices which do not have (global) rank one factors. This generalizes the result of [EL1] (where it was assumed that the action had no local rank one factors, i.e. all simple factors of the groups involved had rank ≥ 2) to the general case of S-algebraic groups

[(6)] More formally: If $\mathbb{G}_1$ is the algebraic group defined over $\mathbb{K}_\sigma$ whose group of $\mathbb{K}_\sigma$-points equals $G_1 = \mathbb{G}_1(\mathbb{K}_\sigma)$, then H as in the theorem contains a finite index subgroup $\mathbb{L}(\mathbb{F}_\sigma)^\circ$ of some algebraic semisimple subgroup $\mathbb{L} < \mathrm{Res}_{\mathbb{K}_\sigma/\mathbb{F}_\sigma} \mathbb{G}_1$ defined over $\mathbb{F}_\sigma$. The groups $\mathbb{L}(\mathbb{F}_\sigma)^\circ$ are in an appropriate sense the connected component of $\mathbb{L}(\mathbb{F}_\sigma)$ containing e (this is literally true in the Archimedean case). See Section 8.1–8.2 for more details.

[(7)] The statement given in [EL2] is inaccurate: the assumption $|\Gamma \cap \{e\} \times G_2| < \infty$ given there is insufficient for a general rank one group G_1 and needs to be replaced with the assumption (2) of Theorem 1.5.

where the quotients do not allow rank one factors. An example of a case which has local rank one factors but no (global) rank one factors and whose analysis requires the use of Theorem 1.4 is the classification of self-joinings of the action of a maximal $\mathbb{R}$-split torus A on $\Gamma\backslash\mathrm{SL}(\mathbb{R}) \times \mathrm{SL}(\mathbb{R})$ with Γ an irreducible lattice.

Acknowledgements. We thank Shahar Mozes for helpful comments; we also thank the anonymous referee for his timely and helpful referee report. Much of results presented in this paper have been obtained while first E.L. and later M.E. were employed by Clay Mathematics Institute (CMI) as Research Fellow and Research Scholar respectively. This support was extremely helpful and is much appreciated. We also thank the NSF for its support.

We dedicate this paper, which is a culmination of a long period of intensive work, to two giants in the field — Gregory Margulis and Marina Ratner — whose deep work inspired us and continues to inspire us in our research.

2. Outline of the proof of the main theorem

In this section we give a sketch of the proof of the following slightly weaker version of Theorem 1.4:

Theorem 2.1. *Under the assumptions of Theorem 1.4, at least one of the following two possibilities holds:*

(LE-1′) *ϕ is locally $C_G(U)$-aligned modulo μ*

(LE-2′) *μ is not $C_G(U) \cap G^-$-transient relative to ϕ.*

For convenience we let

$$H = C_G(U) \qquad H^- = H \cap G^- \qquad H^0 = H \cap G^0$$

here and throught the proof of Theorems 1.4 and 2.1.

To see that Theorem 2.1 is indeed a weaker form of Theorem 1.4, suppose $\mu = t\mu_1 + (1-t)\mu_2$ for $t \in [0,1]$, μ_1 satisfying (LE-1) and μ_2 satisfying (LE-2). If $t = 1$, i.e. $\mu = \mu_1$, by (LE-1) we have that ϕ is H^0-aligned modulo μ, and so in particular H-aligned and μ satisfies (LE-1′). If $t < 1$ the H^--recurrence of μ_2 relative to ϕ implies that μ is not H^--transient relative to ϕ, and μ satisfies (LE-2′). In the Section 4 we will show how to deduce the formally stronger Theorem 1.4 from Theorem 2.1.

We also note that condition (LE-2′) above is equivalent in this context to the following statement (see Corollary 4.4):

(LE-2″) Every subset $X_0 \subset X$ of full measure contains two distinct points x, y with $H^-.x = H^-.y$ and $\phi(x) = \phi(y)$.

As we have mentioned in the introduction, our proof borrows heavily ideas from Ratner's papers [Ra2, Ra1, Ra3]. A fundamental observations there is the following[(8)]: for a unipotent subgroup $U < G$ and two nearby points x, y the fastest divergence (shearing) of $u.x, u.y$ for $u \in U$ is along $H = C_G(U)$. Furthermore, this divergence is a polynomial function of u. This (together with the pointwise ergodic

[(8)]The papers [Ra2, Ra1, Ra3] are written for $G = SL(2,\mathbb{R})$; these results where generalized by D. Witte-Morris in [Mor]. We also note that the polynomial nature of unipotent actions is central in G. A. Margulis' proof of nondivergence of unipotent trajectories [Mar].

theorem, and using Lusin theorem) eventually leads to the statement that any U-invariant and ergodic probability measure μ is stabilized by the fastest divergence direction of two nearby (sufficiently typical) points. In [Ra3] Ratner extracts an axiomatic version of the properties of U-action used in her proof and calls a system satisfying these properties an $\mathcal{H}$-flow. This property of unipotent actions is also referred to as Ratner's H-property.

The low entropy method, also uses the above mentioned polynomial divergence. However, now the measure μ is *not* assumed to be invariant under a unipotent group. Instead we will assume invariance under $a \in G$, and use in a crucial way the leafwise measures μ_x^U. A basic property of these leafwise measures is that μ_x^U is the Haar measure on U a.s. if and only if μ is U-invariant — the case considered by Ratner. The assumption (U-1) of Theorem 1.4 that there is a.s. no a-normalized Zariski closed proper subgroup of U supporting μ_x^U can be viewed as a week substitute to U-invariance, which when coupled with invariance under a, is still sufficient for us to make effective use the shearing properties of the U-flow (cf. D. Rudolph's paper [Rud1] for another use of a similar idea).

We will have to study these leafwise measures and their properties quite carefully in Section 3 and Section 6 below; but for the purposes of this section the main property we shall use is that the map that sends $x \in X$ to μ_x^U is essentially a factor map modulo μ for the action of the group $a^{\mathbb{Z}}U$ in the sense of Definition 1.3 — more precisely: (i) the map that takes x to the equivalence class $[\mu_x^U]$ of the measure μ_x^U under proportionality is a measurable map from X to a compact space $PM_\infty^*(U)$ of equivalence classes of locally finite measures on U satisfying a certain growth condition, (ii) this map is a factor map modulo μ and moreover (iii) the equivalence class $[\mu_x^U]$ determines μ_x^U a.s.

Let $\phi : X \to Y$ be a factor map modulo μ as in Theorem 1.4; since Y is a (standard) Borel space we can (and will) endow Y with the structure of a locally compact metric space. Assume in contradiction to Theorem 2.1 that the factor map ϕ is not locally H-aligned modulo μ but the measure μ is H^--transient relative to ϕ.

By Lusin's theorem we can find a large compact set $K \subset X$ on which ϕ is continuous. Then for any n_0 and for most x we have that μ_x^U-most $u \in a^{-n_0} B_1^U a^{n_0}$ satisfy $u.x \in K$, where B_1^U denote the unit ball around $e \in U$. If ϕ is not locally H-aligned, we can find two distinct nearby points x, y not on the same local H-orbit with $\phi(x) = \phi(y)$ (and hence by condition (Φ) of Theorem 1.4, $\mu_x^U = \mu_y^U$) so that for μ_x^U-most $u \in a^{-n_0} B_1^U a^{n_0}$ both $u.x$ and $u.y$ are in K. If we chose K to also be a subset of the conull set on which ϕ behaves nicely, then $\phi(u.x) = \phi(u.y)$ for any such u. If u is additionally of the right size and sufficiently generic — because of the polynomial divergence property an element not too close to certain varieties will do — $u.x$ and $u.y$ differ approximately by some bounded nontrivial element of H^- (i.e. an element belonging to some fixed compact set in $H^- \setminus \{e\}$). We will choose n_0 such that sufficiently generic $u \in a^{-n_0} B_1^U a^{n_0}$ satisfy this.

However, the two conditions on $u \in a^{-n_0} B_1^U a^{n_0}$ above

- u is μ_x^U-typical inside $a^{-n_0} B_1^U a^{n_0}$ and
- u is algebraically sufficiently generic

might not be compatible since *a priori* we do not know how μ_x^U is distributed on $a^{-n_0} B_1^U a^{n_0}$. In other words, the leafwise measure μ_x^U could give almost all of the mass of $a^{-n_0} B_1^U a^{n_0}$ to a small neighborhood of some subvariety of U. Since $\mu_x^U = a\mu_{a.x}^U a^{-1}$ (this follows form $x \mapsto [\mu_x^U]$ being a factor map for a; cf. (LM-5) in

Section 3.1) this property of μ_x^U is equivalent to $\mu_{a^{n_0}.x}^U$ assigning almost all of the mass of B_1^U to a small neighborhood of some subvariety of U.

Using the assumption (U-1) of Theorem 1.4, namely that U is the smallest Zariski closed a-normalized subgroup of U supporting μ_x^U a.s., we will show in Section 6 that for most points z this does not happen for μ_z^U. But $a^{n_0}.x$ might be an exception, since we have very little control on its position (n_0 was dictated to us by algebraic considerations)!

To avoid that problem we will replace x and y by $x_k = a^k.x$ and $y_k = a^k.y$ for some k. This changes the relative position of the two points and so the size of u for which $u.x_k$ and $u.y_k$ have the right distance from each other. Let n_k be such that sufficiently generic $u \in a^{-n_k} B_1^U a^{n_k}$ have that property. Since ϕ is a factor map for $a^{\mathbb{Z}}U$, we still have $\phi(x_k) = \phi(y_k)$. Again we need to ask how the leafwise measure $\mu_{x_k}^U$ restricted to $a^{-n_k} B_1^U a^{n_k}$, or equivalently, $\mu_{a^{n_k}.x_k}^U$ restricted to B_1^U looks like. Here $a^{n_k}.x_k = a^{n_k+k}.x$, and it will be crucial to study n_k as a function of k (as we will do in Section 7). For if $n_k + k = n_0$ we are still asking about the properties of the point $a^{n_0}.x$. However, if $n_k + k$ is changing, we have a chance for choosing k such that $a^{n_k}.x_k$ is generic enough so that $\mu_{a^{n_k}.x_k}^U$ has good properties. This will happen precisely when the direction of maximal shear between $u.x$ and $u.y$ is not along G^0, and our technical assumption (U-2) on $a \in G$ and $U < G$ is used to ensure this does not happen. When choosing k, we also want to make sure that x_k and y_k are still close together (since we want to use the polynomial divergence of U to separate them). It turns out that in order to ensure x_k and y_k are still close together we need to be able to control the $\mathfrak{g}^+$ component of the difference between x and y in terms of the shearing "time" n_0 - which will again follow from our technical assumption (U-2).

Using now the polynomial divergence of the action of U we will find $u \in a^{-n_k} B_1^U a^{n_k}$ such that two things happen at once: $u.x_k, u.y_k \in K$ and these two points differ approximately by some nontrivial element of H^-. The construction makes sure that $\phi(u.x_k) = \phi(u.y_k)$. Choosing the original two points x, y ever closer together, going to the limit, and using continuity of ϕ inside K, the argument above shows that every set of sufficiently high measure contains two points of the same H-orbit, establishing (LE-2$''$), and hence showing that μ is not transient relative to ϕ.

3. Leafwise measures along orbits and their properties

3.1. **Basic properties of leafwise measures.** Let $\mathcal{A}$ be a countably generated σ-algebra of Borel sets in a locally compact metric space X. For any $x \in X$ the *atom* $[x]_{\mathcal{A}}$ of x is defined to be the intersection of all elements A of $\mathcal{A}$ that contain x.

Any countably generated σ-algebra $\mathcal{A}$ gives a system of *conditional measures* $\mu_x^{\mathcal{A}}$ with each such measure supported by the respective atom $[x]_{\mathcal{A}}$ and the map $x \mapsto \mu_x^{\mathcal{A}}$ is $\mathcal{A}$-measurable. For any $f \in L^1_\mu$,

$$E_\mu(f|\mathcal{A})(x) = \int f \, \mathrm{d}\mu_x^{\mathcal{A}} \qquad \text{for a.e. } x \tag{3.1}$$

and this equation determines $\mu_x^{\mathcal{A}}$ up to a set of measure zero.

Suppose now that U is a locally compact group acting continuously on X, and let μ be a probability measure on X so that the action of U on X is free outside a μ-null set.

We recall that a countably generated σ-algebra $\mathcal{A}$ is said to be *subordinate*[(9)] *to* U if for μ-a.e. x, there is some $\delta > 0$ so that

$$B_\delta^U.x \subset [x]_{\mathcal{A}} \subset B_{\delta^{-1}}^U.x. \tag{3.2}$$

A countably generated σ-algebra $\mathcal{A}$ is *subordinate to* U *on* Y if $Y \in \mathcal{A}$ and (3.2) holds for μ-a.e. $x \in Y$.

The foliation of X into orbits of U also allows us to define a system of leafwise measures for this foliation. Despite the similar notation, this construction is quite different from the conditional measures for a countably generated σ-algebra discussed above. It would be most convenient to us, following [Lin2] to view the leafwise measures as a Borel measurable map $x \mapsto \mu_x^U$ from X to the space $M_\infty(U)$ of locally finite Borel measures on U. We use the weak* topology on $M_\infty(U)$, i.e. the coarsest topology for which $\rho \mapsto \int_U f(y)\,\mathrm{d}\rho(y)$ is continuous for every compactly supported continuous f. For any two measures ν_1 and ν_2 on U we write $\nu_1 \propto \nu_2$ if ν_2 and ν_2 are proportional, i.e. $\nu_1 = C\nu_2$ for some $C > 0$.

While the equivalence class with respect to proportionality of μ_x^U is defined in a pretty canonical way, the exact representative is chosen in a fairly arbitrary way: indeed, in [Lin2] by demanding that $\mu_x^U(B_1^U) = 1$ (so, in particular, the $\propto$ equivalence class $[\mu_x^U]$ determines the measure μ_x^U).

We recall the basic properties of leafwise measures for leaves of a foliation (see [Lin2, Sect. 3] for details); in the following X' is an appropriately chosen subset of X of full μ-measure:

(LM-1) The map $x \mapsto \mu_x^U \in \mathcal{M}_\infty(U)$ is measurable.

(LM-2) For every $x \in X'$ and $u \in U$ with $u.x \in X'$, we have that $\mu_x^U \propto (\mu_{u.x}^U)u$, where $(\mu_{u.x}^U)u$ denotes the push forward of the measure $\mu_{u.x}^U$ under the map $v \mapsto vu$.

The following two properties uniquely determine the μ_x^U:

(LM-3) For every $x \in X'$ we have $\mu_x^U(B_1^U) = 1$, and $\mu_x^U(B_\epsilon^U) > 0$ for every $\epsilon > 0$.

(LM-4) For any $Y \subset X$, for every σ-algebra $\mathcal{A}$ subordinate to U on Y and a.e. $x \in Y$ the probability measure $\mu_x^{\mathcal{A}}$ is proportional to $\mu_x^U.x|_{[x]_{\mathcal{A}}}$.[(10)]

If, like in our case we have a larger group $G > U$ acting on X, and the measure μ is preserved under the action of an element a in the normalizer $N_G(U)$ of U in G, we furthermore have that:[(11)]

(LM-5) For $x \in X'$ we have $a.x \in X'$ and $\mu_{a.x}^U \propto a(\mu_x^U)a^{-1}$.

The space $M_\infty(U)$ is not locally compact. However, in most cases, and in particular in the case at hand here where $X = \Gamma\backslash G$, $U < G$ a unipotent subgroup, and μ a measure on X invariant under the action of an element $a \in G$ contracting U, one can impose an *a priori* growth condition on the measures μ_x^U of the form

$$\int_U \rho(u)\, d\mu_x^U(u) < \infty \tag{3.3}$$

for some strictly positive function ρ on U (see Proposition 3.9 below). The space $M_\infty^*(U)$ of locally finite measures on U which satisfy (3.3) (for an implicitly fixed

---

(9) Note that implicitly this notion also depends on the measure μ.

(10) More formally: define $V_x \subset U$ by $V_x.x = [x]_{\mathcal{A}}$ (this can be done for a.e. x by definition of subordinate σ-algebra). Then $\mu_x^{\mathcal{A}}$ is proportional to the push forward of $\mu_x^U|_{V_x}$ under the map $u \mapsto u.x$.

(11) We think of a as fixed, and in particular the choice of X' is allowed to be dependent on a.

function ρ) is a nice locally compact space (and so in particular is a standard Borel space). Another advantage with working with $M^*_\infty(U)$ is that the space $PM^*_\infty(U)$ of equivalence classes of measures in $M^*_\infty(U)$ under proportionality can also be given the structure of a locally compact metric space by choosing from each equivalence class $[\nu]$ the representative according to which the integral of ρ is 1.

With Proposition 3.9 in mind, (LM-1)–(LM-2) can be summarized as saying that the map $x \mapsto [\mu^U_x]$ is a factor map for U in the sense of Definition 1.3, where U acts on $PM^*_\infty(U)$ by right translations. Property (LM-5) shows that this map is in fact a factor map for the action of a bigger group — the solvable group generated by U and a (where a act on $PM^*_\infty(U)$ by taking an equivalence class of measures $[\nu]$ to its push forward under the map $u \mapsto aua^{-1}$).

Leafwise measures along U-orbits convey much information about how the measure interacts with the action of the group U. In particular, we have the following:

Lemma 3.1. *Let U be a locally compact group acting continuously on a locally compact metric space X. Suppose that μ is a locally finite measure on X and that the action of U is free outside a μ-null set. Then:*

(1) *μ is U-recurrent iff for μ-a.e. x the leafwise measure μ^U_x is an infinite measure;*
(2) *μ is U-transient if and only if for μ-a.e. x the leafwise measure μ^U_x is a finite measure;*
(3) *μ is U-invariant iff for μ-a.e. x the leafwise measure μ^U_x is left Haar measure;*
(4) *more generally, μ is L-invariant for $L < U$ iff for μ-a.e. x the leafwise measure μ^U_x is left L-invariant.*

For proof of (1) and (2) see [Lin2, Prop. 4.1]; (3) is proved in [Lin2, Prop. 4.3]; the same proof also gives (4) (we leave the details to the reader).

The following useful property is not explicitly stated in [Lin2, Sect. 3], but is an immediate corollary of the construction there (again, we leave the proof to the reader):

Lemma 3.2. *Let U be a closed subgroup of a locally compact group V. Assume that V acts continuously on a locally compact metric space X, that μ is a locally finite measure on X and that the action of V is free outside a μ-null set. Assume moreover that for μ-a.e. x, the measure μ^V_x is supported on U. Then, identifying $M_\infty(U)$ as a subspace of $M_\infty(V)$ in the obvious way*[12]*,*

$$\mu^V_x \propto \mu^U_x \qquad \text{for } \mu\text{-a.e. } x.$$

3.2. **An almost subordinate σ-algebra and some consequences.** We will be mostly considering the leafwise measures μ^U_x for U a unipotent group as in Theorem 1.4 normalized and contacted by a, where μ is an a-invariant probability measure on $X = \Gamma\backslash G$. This setting simplifies some of the aspects of the study of these leafwise measures.

The results we present in this vain here are fairly standard, and related results can be found e.g. in [LS, LY1, LY2], and were adapted to the locally homogeneous setting in [MT1, Sect. 9]. In particular, we quote the following:

[12] I.e. via the push forward under the identity map $U \hookrightarrow V$.

Proposition 3.3 ([MT1, Prop 9.2]). *Let G be a Lie group or a product of linear groups over various local fields, let $a \in G$ be diagonalizable, and let $\Gamma < G$ be a discrete subgroup. Let G^- be the contracting horospheric subgroup of a as above. Let μ be an a-invariant and ergodic probability measure on $X = \Gamma\backslash G$. Then there is a countably generated σ-algebra $\mathcal{A}$ such that*

(1) *$\mathcal{A}$ is a-decreasing (i.e. $a^{-1}\mathcal{A} \subset \mathcal{A}$),*
(2) *$\mathcal{A}$ is subordinate to G^- (i.e. for a.e. x (3.2) holds).*
(3) *$h_\mu(a) = H_\mu(\mathcal{A} \mid a^{-1}\mathcal{A})$.*

We remark that Margulis and Tomanov considered only fields of zero characteristic; however, the proof carries through also to the positive characteristic context. For an explicit treatment of both the zero and the positive characteristic case see [EL3].

It would be convenient in Section 5.2 to be able to work with a given (non-ergodic) a-invariant measure μ without passing first to ergodic components. For this will give the following (simpler) variant of Proposition 3.3:

Proposition 3.4. *Let G be a Lie group or a product of linear groups over various local fields, let $a \in G$ be diagonalizable, and let $\Gamma < G$ be a discrete subgroup. Let μ be an a-invariant probability measure on $X = \Gamma\backslash G$. Let $U < G^-$ be an a-normalized and contracted subgroup. For every $\epsilon > 0$ there exists some $R > 1$, a measurable set $Q \subset X$, and a countably generated σ-algebra $\mathcal{A}_U$ such that*

(1) *$\mu(Q) > 1 - \epsilon$,*
(2) *$\mathcal{A}_U$ is a-decreasing,*
(3) *for a.e. $x \in Q$ the $\mathcal{A}_U$-atom $[x]_{\mathcal{A}_U}$ of x satisfies $B_1^U.x \subset [x]_{\mathcal{A}_U} \subset B_R^U.x$.*

In particular, if μ is ergodic, this σ-algebra will be U-subordinate, since for almost every x there will be some $n \geq 0$ for which $a^n.x \in Q$ and hence

$$[x]_{\mathcal{A}} \subset [x]_{a^{-n}\mathcal{A}} = a^{-n}[a^n x]_{\mathcal{A}} \subset a^{-n} B_R^U a^n.x.$$

Similarly since a.s. $a^{-n'}.x \in Q$ for some $n' \geq 0$

$$[x]_{\mathcal{A}} \supset [x]_{a^{n'}\mathcal{A}} = a^{n'}[a^{-n'}x]_{\mathcal{A}} \supset a^{n'} B_1^U a^{-n'}.x.$$

Proposition 3.3.(3) implies the following regarding $\mathcal{A}_U$:

Lemma 3.5. *Let $G, a, \Gamma, U < G^-$ be as above, and let μ be a-invariant and ergodic. Let $\mathcal{A}_U$ be a countably generated U-subordinate σ-algebra. Then $H_\mu(\mathcal{A}_U \mid a^{-1}\mathcal{A}_U) \leq h_\mu(a)$.*

Proof. Let

$$I_\mu(a^n\mathcal{A}_U \mid \mathcal{A}_U)(x) = -\log(\mu_x^{\mathcal{A}_U}([x]_{a^n\mathcal{A}_U}));$$

note that for any x, n this value is nonnegative. Moreover,

$$I_\mu(a^{n+m}\mathcal{A}_U \mid \mathcal{A}_U)(x) = I_\mu(a^m\mathcal{A}_U \mid \mathcal{A}_U)(a^n x) + I_\mu(a^n\mathcal{A}_U \mid \mathcal{A}_U)(x). \tag{3.4}$$

Let $I_1 = I_\mu(a\mathcal{A}_U \mid \mathcal{A}_U)$; then by (3.4)

$$I_\mu(a^n\mathcal{A}_U \mid \mathcal{A}_U)(x) = \sum_{k=0}^{n-1} I_1(a^k x).$$

It follows (using the pointwise ergodic theorem) that

$$H_\mu(\mathcal{A}_U \mid a^{-1}\mathcal{A}_U) = \int I_1(x)\, d\mu(x) = \lim_{n\to\infty} \frac{1}{n} I_\mu(a^n\mathcal{A}_U \mid \mathcal{A}_U).$$

Let $\mathcal{A}$ be a G^--subordinate σ-algebra as in Proposition 3.3, $\epsilon > 0$ arbitrarily. By (3) of that proposition (and the discussion above) a.s. $\frac{1}{n} I_\mu(a^n\mathcal{A} \,|\, \mathcal{A}) \to h_\mu(a)$. The random variable

$$Z_n = \frac{\mu_x^{\mathcal{A}}([x]_{a^n\mathcal{A}})}{\mu_x^{\mathcal{A}_U \vee \mathcal{A}}([x]_{a^n\mathcal{A}})}$$

has expected value 1 for every n. Fix some $\epsilon > 0$. By Borel-Cantelli and Chebyshev we have for a.e. x and all n large enough that $Z_n \leq \exp(\epsilon n)$. Therefore,

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} I_\mu(a^n\mathcal{A} \mid \mathcal{A} \vee \mathcal{A}_U)(x) &= \limsup_{n\to\infty} \frac{1}{n}(\log Z_n + I_\mu(a^n\mathcal{A} \mid \mathcal{A}))(x) \\ &\leq \epsilon + \lim_{n\to\infty} \frac{1}{n} I_\mu(a^n\mathcal{A} \mid \mathcal{A})(x) \\ &= \epsilon + h_\mu(a) \qquad \text{for } \mu\text{-a.e. } x \end{aligned}$$

Since $\mathcal{A}_U$ is U-subordinate and $\mathcal{A}$ is G^--subordinate we can find $\delta > 0$ small enough so that on a set $Q \subset X$ of measure $\mu(Q) > 0.99$

$$[x]_{\mathcal{A}} \subset B^{G^-}_{\delta^{-1}}(x) \qquad [x]_{\mathcal{A}_U} \supset B^U_\delta(x)$$

and let k_0 be such that $a^{k_0}(B^{G^-}_{\delta^{-1}} \cap U)a^{-k_0} \subset B^U_\delta(x)$. Since $\mu(Q) > 0.99$, $\mu(Q \cap a^{k_0}Q) > 0.98$. Suppose for some x, $a^n x \in Q \cap a^{-k_0}Q$. Then

$$[x]_{a^{n+k_0}\mathcal{A}} \cap [x]_{\mathcal{A}_U} \subset [x]_{a^n\mathcal{A}_U},$$

and

$$I_\mu(a^{n+k_0}\mathcal{A} \vee \mathcal{A}_U \mid \mathcal{A}_U)(x) = I_\mu(a^{n+k_0}\mathcal{A} \mid \mathcal{A}_U)(x) \geq I_\mu(a^n\mathcal{A}_U \mid \mathcal{A}_U)(x).$$

Since μ is ergodic, for μ-a.e. x this will happen infinitely often, hence if $\mu_{\mathcal{A}_U}([x]_{\mathcal{A}_U \vee \mathcal{A}}) > 0$ (which again happens a.e.)

$$\begin{aligned} H_\mu(\mathcal{A}_U \mid a^{-1}\mathcal{A}_U) &= \lim_{n\to\infty} \frac{1}{n} I_\mu(a^n\mathcal{A}_U \mid \mathcal{A}_U)(x) \\ &\leq \limsup_{n\to\infty} \frac{1}{n} I_\mu(a^n\mathcal{A} \mid \mathcal{A}_U)(x) \\ &= \limsup_{n\to\infty} \frac{1}{n} I_\mu(a^n\mathcal{A} \mid \mathcal{A}_U \vee \mathcal{A})(x) \leq \epsilon + h_\mu(a). \end{aligned}$$

□

The proof of Proposition 3.4 relies on the following two elementary lemmata:

Lemma 3.6. *Let X be some metric space and let μ be a probability measure on X. Then for any $x \in X$ and (Lebesgue-)a.e. $r > 0$ the set $P = B_r(x)$ satisfies $\mu(B_\rho(P) \cap B_\rho(X \setminus P)) < C\rho$ for some constant $C = C(x,r) > 0$ and every $0 < \rho < r$. Here $B_\rho(A) = \bigcup_{x\in A} B_\rho(x)$ are all points with distance less than ρ to some point in a set $A \subset X$. In particular, the boundary ∂P must be a null set.*

Proof. Define $f(r) = \mu(B_r(x))$. Then clearly f is an increasing function and so for a.e. r differentiable. Fix one such r and let $P = B_r(x)$. Since $B_\rho(P) \cap B_\rho(X \setminus P) \subset B_{r+\rho}(x) \setminus B_{r-\rho}(x)$ this implies for small enough ρ that

$$\mu(B_\rho(P) \cap B_\rho(X \setminus P)) \leq f(r+\rho) - f(r-\rho) \leq 2(f'(r)+1)\rho.$$

□

Lemma 3.7. *Let G_σ be a Lie group or a linear algebraic group over a local field $\mathbb{K}_\sigma$, let $a \in G$ be semisimple, and let $U_\sigma < G_\sigma^-$ be a closed a-normalized and contracted subgroup. There exists a left invariant metric $d_\sigma(\cdot,\cdot)$ on G, a monotone function $f:(0,1)\to\mathbb{R}^+$ with $\lim_{t\to 0} f(t)=0$, and a constant $\chi\in(0,1)$ such that*

$$d_\sigma(a^n u a^{-n}, e) \le \chi^n f(d_\sigma(u,e)) \tag{3.5}$$

for all $n\ge 0$ and $u\in B_1^{U_\sigma}$.

Proof. If G is a Lie group or $\mathbb{K}_\sigma$ is $\mathbb{R}$ or $\mathbb{C}$, then we fix some inner product on the Lie algebra of G and using it define a Riemannian metric $d_\sigma(\cdot,\cdot)$ on G. Taking the restriction of the inner product to the Lie algebra of U_σ we get similarly a Riemannian metric $d_{U_\sigma}(\cdot,\cdot)$ on U_σ with $d_\sigma(\cdot,\cdot)\le d_{U_\sigma}(\cdot,\cdot)$ on U_σ. Now define $f(t)=\operatorname{diam}_{U_\sigma}(B_t^{G_\sigma}\cap U_\sigma)$ to be the diameter of $B_t^{G_\sigma}\cap U_\sigma$ with respect to $d_{U_\sigma}(\cdot,\cdot)$. Since U_σ is closed we have $\lim_{t\to0} f(t)=0$. Finally one checks easily that there exists some $\chi<1$ with $d_{U_\sigma}(aua^{-1},e)\le \chi d_{U_\sigma}(u,e)$. This implies the lemma in the Archimedean case.

For p-adic and positive characteristic fields there always exists a compact and open subgroup $K<G_\sigma$. We then define the left invariant metric on K by averaging any matrix norm $d_K(k_1,k_2)=\int_K \|k(k_1-k_2)\|\,\mathrm{d}k$ where $\mathrm{d}k$ denotes the Haar measure on K. Without loss of generality we may assume that $d_K(\cdot,\cdot)<1$. Now define $d_\sigma(g_1,g_2)=1$ if $g_1K\ne g_2K$ and $d_\sigma(g_1,g_2)=d(e,k)$ if $g_2=g_1k$ for some $k\in K$. That this defines a left invariant metric follows from the left invariance of $d_K(\cdot,\cdot)$ under K. It is easy to check that for $u\in U_\sigma\cap K$ there exists $\chi<1$ and $C>0$ with $d_\sigma(a^nua^{-n},e)<C\chi^n d_\sigma(u,e)$. The function f is used to absorb C. $\square$

Proof of Proposition 3.4. By assumption G is a Lie group or a product of linear groups. Using Lemma 3.7 we get a metric on G also satisfying the conclusion of Lemma 3.7 for all $u\in U$.

Let $K\subset X$ be compact with $\mu(K)>1-\epsilon$. Choose some $r\in(0,1)$ such that for every $x\in K$ the set $\overline{B_{2r}(x)}$ is the injective image $\overline{B_{2r}^G}.x$ of $\overline{B_{2r}^G}\subset G$. Now choose for every $x\in K$ a ball $B_{r_x}(x)$ with $r_x<r$ such that the conclusion of Lemma 3.6 holds. Choose finitely many points $x_1,\dots,x_\ell$ such that the corresponding balls form a subcover of K and let $\mathcal{P}$ be the partition of X generated by these balls. We define $\mathcal{P}_U$ to be the σ-algebra generated by $\mathcal{P}$ and the images $(A\cap B_r^G).x_i$ for all left U-invariant Borel sets $A=UA\subset G$ and $i=1,\dots,\ell$. Then $\mathcal{P}_U$ is countably generated and for almost every $x\in K$

$$B_\rho^U.x\subset[x]_{\mathcal{P}_U}\subset B_R^U.x \tag{3.6}$$

for some $\rho>0$ (that is allowed to depend on x) and some fixed $R>0$. (The only points for which this fails are the boundary points of one of the balls in the construction.) In fact, the construction is such that

$$[x]_{\mathcal{P}_U}=((B_{2r}^G\cap U).x)\cap P \tag{3.7}$$

for any x that belongs to one of the balls $B_{r_{x_i}}(x_i)$ and $x\in P\in\mathcal{P}$.

We claim that $\mathcal{A}=\bigvee_{n=0}^\infty a^{-n}\mathcal{P}_U$ also satisfies (3.6) instead of $\mathcal{P}_U$ for almost every $x\in K$ (but possibly a different ρ). First note that the upper bound carries over trivially since $\mathcal{A}\supset\mathcal{P}_U$.

Now define for $\rho\in(0,1)$

$$\partial_\rho\mathcal{P}=\{x\in X: B_\rho(x)\not\subset P \text{ for all } P\in\mathcal{P}\}.$$

Then by Lemma 3.6 and the construction of $\mathcal{P}$ we know

$$\mu(\partial_\rho \mathcal{P}) \leq C\rho \tag{3.8}$$

for some fixed constant $C > 0$. Suppose $B_\rho^U.x \not\subset [x]_{\mathcal{A}}$, then there exists $n \geq 0$ such that $a^n B_\rho^U.x \not\subset [a^n x]_{\mathcal{P}_U}$. By the properties of the chosen metric

$$a^n B_\rho^U.x = (a^n B_\rho^U a^{-n}).(a^n.x) \subset B_{\chi^n f(\rho)}^U.(a^n.x).$$

If $a^n.x$ belongs to $P \in \mathcal{P}$, then (3.7) shows (if $f(\rho) < 2r$) that $B_{\chi^n f(\rho)}^U.(a^n.x) \not\subset P$ for $a^n.x \in P \in \mathcal{P}$. This again implies $a^n.x \in \partial_{\chi^n f(\rho)}\mathcal{P}$ and so

$$x \in E_\rho = \bigcup_{n=0}^{\infty} a^{-n} \partial_{\chi^n f(\rho)}(\mathcal{P}).$$

Since a preserves μ we have $\mu(E_\rho) \leq \sum_{n=0}^\infty C\chi^n f(\rho) = \frac{Cf(\rho)}{1-\chi}$ by (3.8). Since the upper bound goes to zero with $\rho \to 0$, we have that almost every $x \in K$ satisfies the claim. More precisely we have shown that $B_\rho^U.x \subset [x]_{\mathcal{A}} \subset B_R^U.x$ for a.e. x where ρ depends on x and is such that $x \in X \setminus E_\rho$.

Now choose ρ such that $\mu(K \setminus E_\rho) > 1 - \epsilon$. For large enough m we have $B_1^U \subset a^{-m} B_\rho^U a^m$ and so the set $Q = a^{-m}(K \setminus E_\rho)$ and the σ-algebra $a^{-m}\mathcal{A}$ satisfy the claims of the proposition. □

We also quote the following two facts:

Lemma 3.8 ([Lin2, Cor. 5.4])**.** *Let a, $U < G^-$, G, Γ, and $X = \Gamma\backslash G$ be as in Proposition 3.4. Let μ be an a-invariant probability measure on $X = \Gamma\backslash G$ and $\mathcal{E}_0$ a countably generated σ-algebra of a-invariant sets. Then for μ-a.e. x, for $\mu_x^{\mathcal{E}_0}$ a.e. y*

$$(\mu_x^{\mathcal{E}_0})_y^U = \mu_y^U.$$

The proof of this lemma follows easily from the well-known fact that for any a-invariant measurable subset $E \subset X$ one can find a G^--invariant set E' so that $\mu(E \setminus E') = 0$ and the properties of leafwise measures discussed in Section 3.1. We refer the reader to [Lin2] for details.

Proposition 3.9. *Again let a, $U < G^-$, G, Γ, and $X = \Gamma\backslash G$ be as in Proposition 3.4. Let μ be a a-invariant probability measure on X, and $\mu = \int_X \mu_x^{\mathcal{E}}\, d\mu(x)$ its ergodic decomposition. Then for μ-a.e. $x \in X$*

(1)
$$D_\mu(U,a)[x] = \lim_{|n|\to\infty} \frac{\left|\log \mu_x^U(a^n B_1^U a^{-n})\right|}{|n|}$$
exists;

(2) $D_\mu(U,a)[x] \leq h_{\mu_x^{\mathcal{E}}}(a)$, *with equality holding if $U = G^-$ (where $h_{\mu_x^{\mathcal{E}}}(a)$ denote the ergodic theoretic entropy of $\mu_x^{\mathcal{E}}$ with respect to the action of a);*

(3) $D_\mu(U,a)[x] = 0$ *iff μ_x^U is equal to the δ-measure at $e \in U$.*

Proof. All of these facts are well-known, we give a proof for the convenience of the reader. Note that by Lemma 3.8 we may assume that μ is ergodic. Let $\epsilon > 0$ be arbitrary, and let $\mathcal{A}_U$ be as in Proposition 3.4, or in the case $U = G^-$ as in Proposition 3.3; note that in both cases we obtain a σ-algebra which is subordinate to U.

Let, for every x and σ-algebra $\tilde{\mathcal{A}}$ subordinate to U, $B^U(x,\tilde{\mathcal{A}})$ be a subset of U satisfying[13]

$$B^U(x,\tilde{\mathcal{A}}).x = [x]_{\tilde{\mathcal{A}}}.$$

By (LM-4) it follows that for a.e. x

$$\log\left(\frac{\mu_x^U(B^U(x,\mathcal{A}_U))}{\mu_x^U(B^U(x,a^{-n}\mathcal{A}_U))}\right) = \log \mu_x^{a^{-n}\mathcal{A}_U}([x]_{\mathcal{A}_U}) = -I_\mu(\mathcal{A}_U \mid a^{-n}\mathcal{A}_U)(x) \tag{3.9}$$

and it follows that

(3.10)

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\log\mu_x^U(B^U(x,a^{-n}\mathcal{A}_U)) &= \lim_{n\to\infty}\frac{1}{n}I_\mu(\mathcal{A}_U \mid a^{-n}\mathcal{A}_U)(x)\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} I_\mu(\mathcal{A}_U \mid a^{-1}\mathcal{A}_U)(a^i x)\\ &= \int_X I_\mu(\mathcal{A}_U \mid a^{-1}\mathcal{A}_U)(x)\,d\mu(x) = H_\mu(\mathcal{A}_U \mid a^{-1}\mathcal{A}_U),\end{aligned}$$

where we have used (3.9) and the pointwise ergodic theorem to pass from the second to the third line of (3.10).

Let

$$\tilde{Q} = \left\{x\in X : \sup_n \frac{1}{n}\sum_{i=0}^{n-1} 1_{Q^\complement}(a^i x) \le \sqrt{\epsilon}\right\}.$$

By the maximal inequality, $\mu(\tilde{Q}) \ge 1-\sqrt{\epsilon}$. From the definition of $\tilde{Q}$, it follows that for any $x\in\tilde{Q}$, for any $n\in\mathbf{N}$ there are n_1,n_2 with $n(1-2\sqrt{\epsilon}) \le n_1 \le n \le n_2 \le n(1+2\sqrt{\epsilon})$ with $a^{n_1}.x, a^{n_2}.x \in Q$, hence

$$a^{-n_1}B_1^U a^{n_1} \subset a^{-n_1}B^U(a^{n_1}x,\mathcal{A})a^{n_1} = B^U(x,a^{-n_1}\mathcal{A}) \subset B^U(x,a^{-n}\mathcal{A}) \tag{3.11}$$

and similarly

$$B^U(x,a^{-n}\mathcal{A}) \subset B^U(x,a^{-n_2}\mathcal{A}) \subset a^{-n_2}B_R^U a^{n_2}. \tag{3.12}$$

It follows from (3.11) and (3.10) that for every $x\in\tilde{Q}$

$$\limsup_{n\to\infty}\frac{1}{n}\log\mu_x^U(a^{-n}B_1^U a^n) \le (1-2\sqrt{\epsilon})^{-1}\lim_{n\to\infty}\frac{1}{n}\log\mu_x^U(B^U(x,a^{-n}\mathcal{A})) \tag{3.13}$$

$$= (1-2\sqrt{\epsilon})^{-1}H_\mu(\mathcal{A}\mid a^{-1}\mathcal{A}). \tag{3.14}$$

Similarly, since there is some fixed n_0 for which $B_R^U \subset a^{-n_0}B_1^U a^{n_0}$, we have by (3.12) and (3.10) that

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\mu_x^U(a^{-n}B_1^U a^n) &= \liminf_{n\to\infty}\frac{1}{n}\log\mu_x^U(a^{-n}B_R^U a^n)\\ &\ge (1+2\sqrt{\epsilon})^{-1}\lim_{n\to\infty}\frac{1}{n}\log\mu_x^U(B^U(x,a^{-n}\mathcal{A}))\\ &= (1+2\sqrt{\epsilon})^{-1}H_\mu(\mathcal{A}\mid a^{-1}\mathcal{A}).\end{aligned}$$

Since ϵ was arbitrary, we see that a.s.

$$\limsup_{n\to\infty}\frac{1}{n}\log\mu_x^U(a^{-n}B_1^U a^n) = \liminf_{n\to\infty}\frac{1}{n}\log\mu_x^U(a^{-n}B_1^U a^n) = H_\mu(\mathcal{A}_U\mid a^{-1}\mathcal{A}_U).$$

---

[13] Since a contracts U by Poincare recurrence if μ is preserved by a for μ-a.e. x, the map $u\mapsto u.x$ is injective on U and hence $B^U(x,\tilde{\mathcal{A}})$ is well-defined.

By Lemma 3.5 and Proposition 3.3 $H_\mu(\mathcal{A}_U \mid a^{-1}\mathcal{A}_U) \leq h_\mu(a)$, with equality if $U = G^-$. This establishes statements (1) and (2) of the proposition.

Part (3) of this proposition also follows as (3.13) implies that

$$\lim_{n\to\infty} \frac{1}{n} \log \mu_x^U(a^{-n}B_1^U a^n) = 0 \quad \text{only if} \quad H_\mu(\mathcal{A} \mid a^{-1}\mathcal{A}) = 0.$$

In this case, for a.e. x we have that for every $n, m \in \mathbb{Z}$ with $m < n'$

$$\mu_x^{a^{-n}\mathcal{A}}([x]_{a_{-m}\mathcal{A}}) = 1$$

or in other words that

$$\mu_x^U(B^U(x, a^{-n}\mathcal{A}))$$

is independent of n. In view of (LM-3) and equations (3.11) and (3.12) (and the corresponding similar inclusions for negative n) the only way this could happen is if μ_x^U is the δ-measure $e \in U$. □

We also quote the following related phenomenon (the reference given is by no means the first place this lemma is proved; indeed, this fact has been used already in [KS] and many other papers):

Lemma 3.10 ([EK2, Lem 5.10])**.** *Under the assumptions of Proposition 3.9, for μ-a.e. $x \in X$ the groups*

$$\mathrm{Stab}_U(\mu_x^U) = \left\{u \in U : u\mu_x^U = \mu_x^U\right\}$$

and

$$\mathrm{Stab}_U([\mu_x^U]) = \left\{u \in U : u\mu_x^U \textit{ is proportional to } \mu_x^U\right\}$$

coincide. The same statement holds for right multiplication $\mu_x^U \mapsto \mu_x^U u$.

We remark that Proposition 3.9 can be used to give an alternative proof to Lemma 3.10 to the one given in [EK2]. Indeed, if $\mathrm{Stab}_U(\mu_x^U) \subsetneq \mathrm{Stab}_U([\mu_x^U])$ for a set of x of positive measure, then for those x it would follow that the measure $\mu_x^U(a^{-n}B_1^U a^n)$ grows super exponentially in n contradicting Proposition 3.9.

Another useful corollary of Proposition 3.9 is the following:

Proposition 3.11. *Notations as in Proposition 3.9. Let μ be a-invariant probability measure on $X = \Gamma\backslash G$. Then μ is G^--recurrent if and only if for a.e. ergodic component $\mu_x^{\mathcal{E}}$ of μ with respect to the action of a satisfies that $h_{\mu_x^{\mathcal{E}}}(a) > 0$.*

Proof. This is an immediate corollary of Proposition 3.9 and Lemma 3.1.(1). □

4. Reduction of main theorem from Theorem 2.1

In this section we will deduce Theorem 1.4 from the formally weaker Theorem 2.1. Recall that the assumptions to these two theorems were identical, but the conclusions were different: in Theorem 2.1 we conclude that one of the following two possibilities holds:

(LE-1′) ϕ is locally $C_G(U)$-aligned modulo μ

(LE-2′) μ is not $C_G(U) \cap G^-$-transient relative to ϕ

whereas Theorem 1.4 gives that we can express μ as a convex combination of two measures μ_1, μ_2 satisfying the stronger restrictions that:

(LE-1) ϕ is locally $C_G(U) \cap G^0$-aligned modulo μ_1

(LE-2) μ_2 is $C_G(U) \cap G^-$-recurrent relative to ϕ.

Recall the notation

$$H = C_G(U) \qquad H^- = H \cap G^- \qquad H^0 = H \cap G^0$$

from Section 2.

We deduce Theorem 1.4 from Theorem 2.1 byt showing that we can decompose μ into $t\mu_1 + (1-t)\mu_2$ with both μ_1 and μ_2 satisfying the conditions of Theorem 1.4 and where μ_1 is H^--transient relative to ϕ and μ_2 is H^--recurrent relative to ϕ. Along the way we prove that (LE-2′) is equivalent to (LE-2″), which is the form of this condition we have used in the sketch of proof of Theorem 2.1.

Applying Theorem 2.1 to μ_1 we may conclude that μ_1 satisfies (LE-1′) i.e. ϕ is locally H-aligned modulo μ_1, and using the a-invariance of μ_1 and Poincare recurrence we will conclude that in fact μ_1 satisfies the formally stronger (LE-1) (see Proposition 4.1).

4.1. **(LE-1′) implies (LE-1).** This will follow from the following more general statement which generalizes [EKL, Lemma 4.2]:

Proposition 4.1. *Let G be a Lie group or an S-algebraic group and $a \in G$ a class $\mathcal{A}$ element as in Theorem 1.4. Let μ be an a-invariant probability measure on $X = \Gamma\backslash G$, and ϕ a factor map with respect to the action of the group generated by a. Suppose that $H < G$ is a closed subgroup so that for sufficiently small δ_0 we have that $\log(B^H_{\delta_0}) \subset \mathfrak{g}_0 \oplus \mathfrak{g}_-$. Then ϕ is locally H-aligned modulo μ if and only if ϕ is locally $H \cap G^0$-aligned modulo μ.*

Proof. We need to show that if $\phi : X \to Y$ is H-aligned modulo μ it is also $H \cap G^0$-aligned (the other direction is tautological). Let $\tilde{\mathcal{Y}}$ be the image under ϕ^{-1} of the Borel σ-algebra on Y, let $\epsilon > 0$, and let $X' \subset X$ be a subset as in Definition 1.1 with $\mu(X') > 1 - \epsilon$.

For every measurable subset $B \subset X$, we have that $\mu(B \setminus \tilde{B}) = 0$ where

$$\tilde{B} = \left\{ x : \mu_x^{\tilde{\mathcal{Y}}}(B) > 0 \right\} \in \tilde{\mathcal{Y}}$$

since

$$\mu(B \setminus \tilde{B}) = \int_{X \setminus \tilde{B}} \mu_x^{\tilde{\mathcal{Y}}}(B \setminus \tilde{B}) \, d\mu(x) \leq \int_{X \setminus \tilde{B}} \mu_x^{\tilde{\mathcal{Y}}}(B) \, d\mu(x) = 0.$$

By applying this to $X' \cap U_i$ with U_i a countable basis for the topology of X it follows that for almost every $x \in X'$ and every $\delta > 0$

$$\mu_x^{\tilde{\mathcal{Y}}}(X' \cap B_\delta(x)) > 0. \tag{4.1}$$

Fix norms on $\mathfrak{g}_0$ and $\mathfrak{g}_-$ satisfying that

(1) the norm on $\mathfrak{g}_0$ is $\mathrm{Ad}(a)$-invariant
(2) there is some $c < 1$ so that for every $w \in \mathfrak{g}_-$, $\|\mathrm{Ad}(a)[w]\| < c \|w\|$.

Let δ be much smaller than δ_0 — sufficiently small so that

$$\left(B_\delta^{\mathfrak{g}_0}(0) \times B_\delta^{\mathfrak{g}_-}(0)\right) \cap \log(B^H_{\delta_0}(e)) = \left(B_\delta^{\mathfrak{g}_0}(0) \times B_\delta^{\mathfrak{g}_-}(0)\right) \cap \log(B^H_{\delta_0/10}(e)).$$

For any $\eta \leq \delta$ we let

$$B(\delta, \eta) = \left(B_\delta^{\mathfrak{g}_0}(0) \times B_\eta^{\mathfrak{g}_-}(0)\right) \cap \log(B^H_{\delta_0}(e)).$$

Applying (4.1) and the alignment condition (1.1) we see that for a.e. $x \in X'$ and every $\eta < \delta$ it holds that

$$f_\delta(x, \eta) = \mu_x^{\tilde{\mathcal{Y}}}(\exp B(\delta, \eta).x) > 0.$$

For a fixed x the function $f_\delta(x,\eta)$ is monotone non-decreasing in η. Furthermore, since ϕ is a factor for a and μ is a-invariant, it follows that $a.\mu_x^{\tilde{\mathcal{Y}}} = \mu_{a.x}^{\tilde{\mathcal{Y}}}$ and hence

$$\begin{aligned} f_\delta(x,\eta) &= \mu_x^{\tilde{\mathcal{Y}}}(\exp B(\delta,\eta).x) \\ &\mu_{a.x}^{\tilde{\mathcal{Y}}}(\exp\left(Ad(a)[B(\delta,\eta)]\right).(a.x)) \\ &\leq \mu_{a.x}^{\tilde{\mathcal{Y}}}(\exp B(\delta,c\eta).(a.x)) = f(a.x,c\eta) \leq f_\delta(a.x,\eta). \end{aligned} \tag{4.2}$$

Using Poincare recurrence we conclude that equality holds throughout (4.2) and consequently almost surely $f_\delta(x,\eta) = f_\delta(a.x,\eta)$. It follows that $f_\delta(x,\eta)$ is almost surely constant on the interval $\eta \in (0,\delta)$.

Let

$$\begin{aligned} X'' = \bigl\{x \in X' : f_\delta(x,\tau) = f_\delta(x,\tau') \quad \forall \tau,\tau' \in (0,\delta) \\ \text{and } \mu_x^{\tilde{\mathcal{Y}}}(\exp B(\tau,\tau).x) > 0 \text{ for every } \tau > 0\bigr\}. \end{aligned}$$

It follows from the discussion above that $\mu(X' \setminus X'') = 0$ and in particular $\mu(X'') > 1-\epsilon$.

If $x, x' \in X''$ are sufficiently close with $\phi(x) = \phi(x')$ then $x' \in \exp B(\delta/2,\delta/2).x$ (since $X'' \subset X'$ and the definition of X'). Also, by definition of leafwise measures, $\mu_x^{\tilde{\mathcal{Y}}} = \mu_{x'}^{\tilde{\mathcal{Y}}}$, and since $x' \in X''$ we also know that $\mu_{x'}^{\tilde{\mathcal{Y}}}(\exp B(\tau,\tau).x') > 0$ for every $\tau > 0$. Suppose $x' = \exp(w_0+w_-).x$ with $w_0 \in \mathfrak{g}_0, w_- \in \mathfrak{g}_-,\ w_0+w_- \in B(\delta/2,\delta/2)$ and $w_- \neq 0$. Then for any $\epsilon > 0$

$$\mu_x^{\tilde{\mathcal{Y}}}(\exp B(\delta,\|w_-\|-\epsilon).x) < \mu_x^{\tilde{\mathcal{Y}}}(\exp B(\delta,\|w_-\|+\epsilon).x)$$

since if τ is sufficiently small

$$\exp B(\tau,\tau).x' \subset (\exp B(\delta,\|w_-\|+\epsilon).x) \setminus (\exp B(\delta,\|w_-\|-\epsilon).x)$$

— and hence $f_\delta(x,\cdot)$ is not constant.

Thus we conclude that $w_- = 0$, i.e. that $x' \in \exp B(\delta/2,0).x$. Since $\exp B(\delta/2,0)$ gives arbitrarily small neighborhood of $e \in H \cap G^0$ for sufficiently small δ we conclude that ϕ is locally $H \cap G^0$-aligned modulo μ. $\square$

4.2. **Recurrence and transience relative to ϕ.** Let X be a locally compact space, U a locally compact group acting on X and $\phi : X \to Y$ a Borel map. Let $\Phi : X \to X \times Y$ be the map $\Phi(x) = (x,\phi(x))$. The group U acts on $X \times Y$ by $u.(x,y) = (u.x,y)$.

We start by considering the notions of recurrence and transience relative to ϕ in more detail, showing that it can be rephrased as ordinary recurrence (transience) on $X \times Y$.

Proposition 4.2. *A probability measure μ on X is U-recurrent (transient) relative to ϕ if and only if the measure $\tilde{\mu} = \Phi_*(\mu)$ on $X \times Y$ is U-recurrent (transient, respectively).*

Proof. Suppose X is U-recurrent relative to ϕ. Let $\tilde{B} \subset X \times Y$ be a Borel sets with $\tilde{\mu}(\tilde{B}) > 0$. Let $B = \Phi^{-1}(\tilde{B})$; by definition of $\tilde{\mu}$, we have that $\mu(B) = \tilde{\mu}(\tilde{B}) > 0$. It follows that for every $x \in B$ in a subset $B' \subset B$ of full measure

$$\{u \in U : u.x \in B \text{ and } \phi(u.x) = \phi(x)\} = \{u \in U : u.\Phi(x) \in \Phi(B)\} \tag{4.3}$$

is unbounded. Since $\tilde{\mu}$-a.e. $\tilde{x} \in \tilde{B}$ is of the form $\Phi(x)$ for $x \in B'$ it follows that $\tilde{\mu}$ is U-recurrent.

Conversely, suppose $\tilde{\mu}$ is U-recurrent. Let $B \subset X$ be arbitrary with $\mu(B) > 0$. Let $\tilde{B} = \Phi(B)$. Since Φ is an injective Borel map $\tilde{B}$ is a Borel subsets of $X \times Y$ and by definition $\tilde{\mu}(\tilde{B}) > 0$. Since $\tilde{\mu}$ is U-recurrent, it follows that for $\tilde{\mu}$-a.e. $\tilde{x} = (x, \phi(x)) \in X \times Y$

$$\{u \in U : u.\Phi(x) \in \Phi(B)\}$$

is unbounded, where again invoking (4.3) we may conclude that μ is U-recurrent relative to ϕ.

We leave the proof of the equivalence of U-transience of the measure μ on X relative to ϕ and U-transience of the corresponding measure $\tilde{\mu}$ on $X \times Y$ to the reader. □

Let U be a locally compact group acting on a locally compact space X. Any locally finite measure μ on X can be written as $\mu = \mu_1 + \mu_2$ with μ_1 transient under U and μ_2 recurrent under U. This can be shown directly (see [Lin1, Prop. 2.4] or via the connection with leafwise measures given in Lemma 3.1: simply take $B = \{x : \mu_x^U \text{ is a finite measure}\}$. Then one can easily check that $\mu_1 = \mu|_B$ is U-transient, $\mu_2 = \mu|_{B^\complement}$ is U-recurrent, and of course $\mu = \mu_1 + \mu_2$. Using Proposition 4.2 we can deduce:

Corollary 4.3. *Under the conditions of Theorem 1.4 we may decompose μ as a convex combination of two probability measures μ_1, μ_2 satisfying the conditions of Theorem 1.4, with μ_1 H^--transient relative to ϕ, and μ_2 H^--recurrent relative to ϕ.*

Proof. Let $\Phi : X \to X \times Y$ and $\tilde{\mu}$ be as in Proposition 4.2. As we have noted above, we can represent $\tilde{\mu}$ as a convex combination of two probability measures $\tilde{\mu}_1$ and $\tilde{\mu}_2$ with $\tilde{\mu}_1$ a H^--recurrent measure and $\tilde{\mu}_2$ a H_--transient measure on $X \times Y$. Let π denote the natural projection $X \times Y \to X$. Then by definition $\mu = \pi_*(\tilde{\mu})$, and we can represent it as a convex combination of $\mu_1 = \pi_*(\tilde{\mu}_1)$ and $\mu_2 = \pi_*(\tilde{\mu}_2)$. As both $\tilde{\mu}_i$ are supported on the graph of Φ we have that for both $i = 1, 2$ $\tilde{\mu}_i = \Phi_*(\mu_i)$. It now follows from Proposition 4.2 that μ_1 is H^--transient relative to ϕ and μ_2 is H^--recurrent relative to ϕ. □

We conclude our discussion of relative recurrence and transience by showing that (LE-2′) and (LE-2″) of Section 2 are equivalent:

Corollary 4.4. *Let a, ϕ, $X = \Gamma\backslash G$ and μ be as in Theorem 1.4. Let H^- be a subgroup of G^- normalized by a. Then μ is H^--transient modulo ϕ iff there is a subset $X_0 \subset X$ of full measure so that for any $y \in Y$ we have that $\phi^{-1}(y) \cap X_0$ intersects every H^--orbit in at most one point.*

Proof. The existence of such an X_0 clearly implies that μ is H^--transient modulo ϕ. Suppose μ is H^--transient relative to ϕ, and let Φ and $\tilde{\mu} = \Phi_*\mu$ be as in Proposition 4.2. Recall that we have assumed that a acts also on Y and that ϕ is a factor map for this action. Denote the diagonal action of a on $X \times Y$ by $\Delta(a).(x, y) = (a.x, a.y)$; then $\Phi(a.x) = \Delta(a)\Phi(x)$ and hence $\tilde{\mu} = \Phi_*\mu$ is $\Delta(a)$-invariant. Since μ is H^- transient relative to ϕ, the measure $\tilde{\mu}$ on $X \times Y$ is H^--transient (with respect to the action $h.(x, y) = (h.x, y)$), and so by Lemma 3.1 the leafwise measure $\tilde{\mu}_x^{H^-}$ is finite almost surely. But since $\tilde{\mu}$ is $\Delta(a)$-invariant, and since $\Delta(a).(h.\tilde{x}) = (aha^{-1}).(\Delta(a).\tilde{x})$ we have that $\tilde{\mu}$-a.e.,

$$[\tilde{\mu}_{\Delta(a).\tilde{x}}^{H^-}] = [a\tilde{\mu}_{\tilde{x}}^{H^-} a^{-1}].$$

Since conjugation by a contracts H^- it follows that $\tilde{\mu}_{\tilde{x}}^{H^-}$ is a.e. proportional to the δ-measure supported on $e \in H^-$.

Let as before π be the projection $X \times Y \to X$, let $\tilde{X}' \subset X \times Y$ be a set on which (LM-1)–(LM-5) are satisfied for the action of H^- on $X \times Y$. Set

$$\tilde{X}_0 = \left\{ \tilde{x} \in X \times Y : [\tilde{\mu}_{\tilde{x}}^{H^-}] = [\delta_e] \right\} \cap \tilde{X}'.$$

This is a subset of full measure of $X \times Y$ and by (LM-2) and (LM-3) there can be no pair of distinct points $\tilde{x}, \tilde{y} \in \tilde{X}_0$ on the same H^--orbit.

Since $\Phi(X)$ also a $\tilde{\mu}$ co-null subset of $X \times Y$, we have that $\tilde{X}_0 \cap \Phi(X)$ is co-null and since π is a Borel isomorphism on $\Phi(X)$ we have that $X_0 = \pi(\tilde{X}_0 \cap \Phi(X_0)) = \Phi^{-1}(\tilde{X}_0)$ is μ-co null. But the fact that $\tilde{X}_0$ contains no two points in the same H^--orbit implies that X_0 does not contain any two distinct points on the same H^--orbit with the same ϕ-value. $\square$

5. Proof of the main theorem (modulo some technical steps)

In this section we will give an almost complete proof of Theorem 1.4 (or more precisely of Theorem 2.1 which implies the formally stronger Theorem 1.4 as already shown in Section 4). To make the outline of the proof given in Section 2 work, we need to start by defining various sets which are needed to justify the arguments.

5.1. The assumptions, our goal, and the set X'. Let $X = \Gamma\backslash G$, $a \in G$, $U < G$, ϕ and μ be as in Theorems 1.4 and 2.1, $H = C_G(U)$ and $H^- = C_G(U) \cap G^-$. Assume in contradiction to Theorem 2.1 that neither (LE-1′) nor (LE-2′) hold, i.e. that ϕ is not locally H-aligned modulo μ but that μ is H^--transient relative ϕ.

The first set we are going to use is an a-invariant subset $X' \subset X$ of full measure on which various typical properties occure. In particular, we assume that

(A-1) the properties (LM-2), (LM-3), and (LM-5) (cf. Section 3.1) hold on X',
(A-2) for every $y \in X'$ a leafwise measure μ_x^U is not supported by any a-normalized Zariski closed proper subgroup of U (cf. (U-1) of Theorem 1.4),
(A-3) X' does not contain any two distinct points x, y on the same H^--orbit with $\phi(x) = \phi(y)$ (cf. Corollary 4.4).

Consider now a neighborhood $O_H \ni e$ in H, sufficiently small to be in the domain of the definition of log. Since log is injective on a small neighborhood of $e \in G$ if O_H is sufficiently small

$$\log O_H \subset C_{\mathfrak{g}}(U) := \{ w \in \mathfrak{g} : \mathrm{Ad}_u(w) = w \quad \forall u \in U \}.$$

Let O_H be some fixed relatively compact neighborhood of $e \in H$ of this form.

Our assumption that ϕ is not locally H-aligned modulo μ clearly implies the following:

(*) There is some ϵ_0 so that for every measurable $B \subset X$ with $\mu(B) > 1 - \epsilon_0$ and every $\delta > 0$ there exists $x, y \in B$ with distance $d(x, y) < \delta$ and with $\phi(x) = \phi(y)$ but $y \notin O_H.x$

Throughout the proof we assume $\epsilon < \epsilon_0$ and also that ϵ is smaller than some constant which we will derive from a, U and G in Section 5.6.

5.2. **$\mathcal{A}, \mu_x^U$, and the set X_1.** We now apply Proposition 3.4 to find an a-decreasing σ-algebra $\mathcal{A}$, a measurable subset $X_1 \subset X$, and some (probably big) $R > 1$ so that

$$\mu(X_1) > 1 - \frac{1}{2}\epsilon^2, \text{ and} \tag{5.1}$$

$$B_1^U.x \subset [x]_{\mathcal{A}} \subset B_R^U.x \text{ for } x \in X_1. \tag{5.2}$$

Recall that a σ-algebra $\mathcal{A}$ is a-decreasing if $a^{-1}.\mathcal{A} \subset \mathcal{A}$.

In addition to $\mathcal{A}$ we will also need some regularity properties of μ_x^U on X_1. The first of which is easy; modifying X_1 there exists some (probably big) $M > 0$ such that

$$\mu_x^U(B_R^U) < M \text{ for } x \in X_1. \tag{5.3}$$

To see this recall that μ_x^U is a Radon measure, i.e. it is locally finite and so $\mu_x^U(B_R^U) < \infty$ a.s. By choosing M accordingly we can ensure that both (5.1) and (5.3) hold.

The second regularity property we require for $x \in X_1$ is that μ_x^U restricted to B_1^U should not have most of its mass too close to proper Zariski closed subsets of U. We will show in Section 6 that our assumption (U-1) regarding the support of μ_x^U will imply this type of regularity on a set of large measure. More precisely by Corollary 6.4 we have the following: Let D be the maximal degree of $\mathrm{Ad}_u(w)$ as a polynomial in u for any $w \in \mathfrak{g}$. There exists $\eta > 0$ such that (after shrinking X_1 once more, but by a sufficiently small amount so that (5.1) is not violated) we have

$$\mu_x^U\big(\{u \in B_1^U : \|f(u)\| \geq \eta\}\big) \geq \eta \tag{5.4}$$

for every $x \in X_1$ and every polynomial f of degree less than D (defined on any of the factors of U) that has a coefficient of norm one. Here $\|w\|$ is some norm on $\mathfrak{g}$ (more precisely, $\|w\| = \max_\sigma \|w_\sigma\|_\sigma$, with $\|\cdot\|_\sigma$ some norm on $\mathfrak{g}_\sigma$), and we may assume without loss of generality (since we can rescale $\|\cdot\|$ at our pleasure) that for any polynomial f as above with coefficients of norm at most one, if the image of $f(B_1^U) \subset \bar{\Omega}^Z = \prod_\sigma \bar{\Omega}_\sigma^Z$ then in fact $f(B_1^U)$ is contained in some fixed compact subset of Ω (where $\Omega, \bar{\Omega}^Z$ are as on p. 4).

5.3. **The compact set K.** As we have already done several times, we may assume that ϕ is a map from X to some locally compact space Y. Having found ϵ, M, and η we let $K \subset X'$ be a compact Lusin set for ϕ, i.e. ϕ depends continuously on $x \in K$, with measure

$$\mu(K) > 1 - \frac{\epsilon^2\eta}{4M}. \tag{5.5}$$

5.4. **The set X_2.** Since $\mathcal{A}$ is a-decreasing, the sequence $a^{-n}.\mathcal{A}$ is a decreasing sequence of σ-algebras. By the martingale maximal inequality we have

$$\mu\left(\left\{x : \sup_{n>0} \mathrm{E}_\mu\big(1 - \chi_K \big| a^{-n}.\mathcal{A}\big) \geq \frac{\eta}{2M}\right\}\right) \leq \frac{\|1 - \chi_K\|_1}{\eta/(2M)} < \frac{\epsilon^2}{2}.$$

Thus there is a set $X_2 \subset X'$ such that

$$\mathrm{E}_\mu\big(\chi_K | a^{-n}.\mathcal{A}\big)(x) \geq 1 - \frac{\eta}{2M} \text{ for } x \in X_2 \text{ and } n > 0 \tag{5.6}$$

satisfying that

$$\mu(X_2) > 1 - \frac{1}{2}\epsilon^2. \tag{5.7}$$

5.5. **The set X_3, and the points $x_\delta, y_\delta \in X_3$.** Since a preserves μ, we can apply the maximal inequality for ergodic averages, and obtain that for any integrable $f : X \to \mathbb{R}_{\geq 0}$

$$\mu\left(\left\{x : \sup_{N\geq 1} \frac{1}{N}\sum_{n=0}^{N-1} f(a^n.x) \geq \epsilon\right\}\right) < \frac{\|f\|_1}{\epsilon}.$$

We define $f = \chi_{X\setminus(X_1\cap X_2)}$, then $\|f\|_1 < \epsilon^2$ by (5.1) and (5.7). Therefore, we find the set X_3 of measure $\mu(X_3) > 1-\epsilon$ such that

$$\left|\left\{n \in [0, N-1] : a^n.x \in X_1 \cap X_2\right\}\right| > (1-\epsilon)N$$

for $N \geq 1$ and $x \in X_3$. Without loss of generality we may require that X_3 is compact.

Since $\mu(X_3) > 1-\epsilon$, there exists by assumption (*) in Section 5.1 for every $\delta > 0$ two points $x_\delta, y_\delta = g_\delta.x_\delta \in X_3$ with (i) g_δ is inside the domain of definition of log and $\log(g_\delta) \notin C_{\mathfrak{g}}(U)$ (ii) $\|\log(g_\delta)\| < \delta$, (iii) $\phi(x_\delta) = \phi(y_\delta)$. Note that (iii) implies that $\mu^U_{x_\delta} = \mu^U_{y_\delta}$, since by assumption ($\Phi$) of Theorem 1.4 the Borel map $x \mapsto \mu^U_x$ is measurable with respect to the σ-algebra corresponding to ϕ.

5.6. **The points $x'_\delta, y'_\delta \in X_2$ with regularity of the leafwise measures at the correct scale.** Our aim in this section is to find starting from the pair of points x_δ, y_δ constructed in the previous subsection a new pair of points which would satisfy some additional regularity properties. These points would have the form $x'_\delta = a^k.x$, $y'_\delta = a^k.y_\delta$ and we restrict the range of permissible k so as to ensure that x'_δ and y'_δ are still fairly close together (though possibly not as close as x_δ and y_δ). In order to find these new points we shall make use of some auxiliary results we defer to Section 7. Note that since $\phi(x_\delta) = \phi(y_\delta)$, and since ϕ respects the action of a, we will automatically have that $\phi(x'_\delta) = \phi(y'_\delta)$.

Recalling the assumptions made in the introduction, $\mathrm{Ad}_u(\log(g_\delta))$ is a $\mathfrak{g}$-valued polynomial in u, and since $\log(g_\delta) \notin C_{\mathfrak{g}}(U)$ this polynomial is not constant. Let $w = \log(g_\delta)$.

Since a contracts U, the coefficients of the polynomial $\mathrm{Ad}_{a^{-n_0}ua^{n_0}}(w)$, which we recall are elements in the Lie algebra $\mathfrak{g}$, grow (exponentially) in n_0. In particular, we can choose n_0 such that some coefficient has norm in $(c, 1]$ and all others have norm in $[0, 1]$ — where $c > 0$ is some fixed constant that only depends on $a \in G$ and $U < G$. Here n_0 will grow indefinitely when δ becomes smaller. More generally, for any k we define in Section 7 a function $n_k = n_k(w)$ so that the norms of the coefficients of $\mathrm{Ad}_{a^{-n_k}ua^{n_k}}(\mathrm{Ad}_{a^k}(w))$ satisfy the above stated conditions. As we will show in Proposition 7.5 below for $\|w\| < \delta_0$ and when k is restricted to be in $[0, \kappa n_0]$ (with both δ_0, κ positive constants depending only on a, U, G) the function n_k can be written as

$$n_k = \min_{1\leq j\leq J} \lfloor s_j k + \tau_j(w)\rfloor . \tag{5.8}$$

Here the slopes s_j depend only on a, U, G (the dependence of n_k on w is via the constant coefficients $\tau_j = \tau_j(w)$), and what is crucial for us is that all the $s_j \neq -1$. In particular, this implies that for any j, the function $k + \lfloor s_j k + \tau_j\rfloor$ attains any value at most $1 + |s_j + 1|^{-1}$ times.

As we shall see in Proposition 7.4, assumption (U-2) implies that for an appropriate choice of κ, for every $k \in [0, \kappa n_0]$

$$\|\mathrm{Ad}_{a^k}(w)\| \leq \|w\|^{1/2} . \tag{5.9}$$

In particular (assuming δ_0 was chosen to be sufficiently small), this guarantees that $n_k > 0$ when k is in the above range. On the other hand, (5.8) applied to $k = 0$ implies that for some j we have $\tau_j < n_0+1$. Therefore $n_k \le k \max_j s_j + n_0$ for all k, and if k is in the above range (i.e. $k \in [0, \kappa n_0]$) in fact $n_k \le N := (1+\kappa \max_j s_j)n_0$.

We want to find $k \in [0, \kappa n_0]$ so that both

$$a^k.x_\delta, a^k.y_\delta \in X_2 \tag{5.10}$$

$$a^{k+n_k}.x_\delta, a^{k+n_k}.y_\delta \in X_1. \tag{5.11}$$

By definition of $X_3 \ni x_\delta, y_\delta$,

$$\left|\left\{0 \le k \le \kappa n_0 : a^k.x_\delta \notin X_2\right\}\right| \le \epsilon\kappa n_0,$$

and similarly for y_δ. By (5.8), the definition of X_3, and the bounds $0 \le n_k \le N$ for any $k \in [0, \kappa n_0]$ show that

$$\begin{aligned}\left|\left\{0 \le k \le \kappa n_0 : a^{k+n_k}.x_\delta \notin X_1\right\}\right| &\le J(1 + \max_j |s_j+1|^{-1})\times \\ &\quad \times \left|\left\{0 \le \ell \le N : a^\ell.x_\delta \notin X_1\right\}\right| \\ &\le C_0\epsilon n_0,\end{aligned}$$

with C_0 depending on a, U, G. Again the same holds for y_δ. Assuming we have chosen $\epsilon < \min(0.01, 0.01\kappa C_0^{-1})$ we are guaranteed by these bounds to have some $k \in [0, \kappa n_0]$ so that both (5.10) and (5.11) are satisfied.

5.7. **Applying the regularity to find $x''_\delta, y''_\delta \in K$.** We claim the properties of $x'_\delta, y'_\delta = g'.x'_\delta$ with $g' = a^k g_\delta a^{-k}$ that are implied by (5.10) and (5.11) are enough to apply the shearing properties of the U-flow successfully. This will give at the end of this step points $x''_\delta = u.x'_\delta, y''_\delta = u.y'_\delta \in K$ that differ approximately by some element of $G^- \cap H$. Since ϕ is a factor map for $a^{\mathbb{Z}}U$, the equality of ϕ-values $\phi(x''_\delta) = \phi(y''_\delta)$ is preserved. Let $n = n_k$ and $w' = \log(g')$; recall that by (5.9) $\|w'\| \le \delta^{1/2}$.

Since $x'_\delta \in X_2$ we have

$$\mu_{x'_\delta}^{a^{-n}.\mathcal{A}}(K) > 1 - \frac{\eta}{2M} \tag{5.12}$$

by (3.1) and (5.6). Since $a^n.x'_\delta \in X_1$ and $\mathcal{A}$ satisfies (5.2) we have

$$B_1^U.(a^n.x'_\delta) \subset [a^n.x'_\delta]_{\mathcal{A}} \subset B_R^U.(a^n.x'_\delta)$$

and so by going back to x'_δ we get

$$(a^{-n}B_1^U a^n).x'_\delta \subset [x'_\delta]_{a^{-n}.\mathcal{A}} \subset (a^{-n}B_R^U a^n).x'_\delta. \tag{5.13}$$

The above equation that allow us to work with $(a^{-n}B_1^U a^n).x'_\delta$ and $(a^{-n}B_1^U a^n).y'_\delta$ instead of the somewhat random pieces of U-leaves $[x'_\delta]_{a^{-n}.\mathcal{A}}$ and $[y'_\delta]_{a^{-n}.\mathcal{A}}$ respectively.

Moreover, $\mu^U_{a^n.x'_\delta}(B_R^U) < M$ and so $\mu^U_{x'_\delta}(a^{-n}B_R^U a^n) < M\mu^U_{x'_\delta}(a^{-n}B_1^U a^n)$ by (LM-5). By (LM-1) and the upper bound in (5.13) this implies $\mu_{x'_\delta}^{a^{-n}\mathcal{A}}(a^{-n}B_1^U a^n.x'_\delta) > \frac{1}{M}$, i.e. when we restrict ourselves to $(a^{-n}B_1^U a^n).x'_\delta$ we do, in a quantitative way, not lose all of the mass. From this and (5.12) we see that

$$\frac{\mu_{x'_\delta}^{a^{-n}.\mathcal{A}}\big((a^{-n}B_1^U a^n).x'_\delta \cap K\big)}{\mu_{x'_\delta}^{a^{-n}.\mathcal{A}}\big((a^{-n}B_1^U a^n).x'_\delta\big)} > 1 - \frac{\eta}{2}$$

or equivalently by (LM-1)

$$\mu^U_{x'_\delta}\big(\{u\in a^{-n}B^U_1a^n : u.x'_\delta\in K\}\big) > \Big(1-\frac{\eta}{2}\Big)\,\mu^U_{x'_\delta}(a^{-n}B^U_1a^n).$$

The same holds for y'_δ, and since $\mu^U_{x'_\delta}=\mu^U_{y'_\delta}$ we get

$$\mu^U_{x'_\delta}\big(\{u\in a^{-n}B^U_1a^n : u.x'_\delta, u.y'_\delta\in K\}\big) > (1-\eta)\mu^U_{x'_\delta}(a^{-n}B^U_1a^n). \tag{5.14}$$

Let $p(u)$ denote the $\mathfrak{g}$-valued polynomial $p(u)=\mathrm{Ad}_{a^{-n}ua^n}(w')$. This polynomial has degree less than D and by definition of $n_k\equiv n$ has coefficients of norm at most one and at least one of norm bigger than some fixed constant $c>0$. Therefore, since $a^n.x'_\delta\in X_1$ by (5.4)

$$\mu^U_{a^n.x'_\delta}\big(\{u\in B^U_1 : \|p(u)\|\geq \eta c\}\big)\geq\eta.$$

Using (LM-5) this shows for x'_δ that

$$\mu^U_{x'_\delta}\big(\{u\in a^{-n}B^U_1a^n : \|\mathrm{Ad}_u(w')\|\geq\eta c\}\big)\geq\eta\mu^U_{x'_\delta}(a^{-n}B^U_1a^n). \tag{5.15}$$

Combining (5.14)–(5.15) we find that there exists some $u_\delta\in a^{-n}B^U_1a^n$ such that

$$x''_\delta=u_\delta.x'_\delta,\qquad y''_\delta=u_\delta.y'_\delta\in K\qquad \|\mathrm{Ad}_{u_\delta}(w')\|\geq\eta c. \tag{5.16}$$

We will show in Section 7.2 that $u_\delta g'u_\delta^{-1}$ is very close to H^- so that x''_δ and y''_δ are almost on the same H^--orbit. More precisely: by Lemma 7.3 and Lemma 7.2 for any $\rho>0$ if δ is small enough there is some $v\in C_{\mathfrak{g}}(U)\cap\mathfrak{g}^-$ so that

$$\|\mathrm{Ad}_{u_\delta}(w')-v\|<\rho.$$

Let $\Omega\subset\mathfrak{g}$ be as in p. 4. Since for small u we have that the polynomial $p(u)$ defined above takes values in Ω — indeed, by its very definition $p(u)=\log(a^{-n}ua^nga^{-n}u^{-1}a^n)$ — and since all the coefficients of $p(u)$ have norm at most one, our choice of $\|\cdot\|$ in Section 5.2 implies that $\mathrm{Ad}_{u_\delta}(w')\in p(B^U_1)$ is in some fixed compact subset of Ω; using equation (5.16) we can in fact conclude that $\mathrm{Ad}_{u_\delta}(w')\in p(B^U_1)$ is in some fixed compact subset of $\Omega\setminus\{0\}$. In particular $\mathrm{Ad}_{u_\delta}(w')=\log(g''_\delta)$ for some $g''_\delta\in G$.

5.8. **Letting $\delta\to 0$.** We apply the above procedure to find $x''_\delta, y''_\delta\in K$ for a sequence of δ that converge to zero, with $y''_\delta=g''_\delta.y''_\delta$, and recall that there is a sequence of v_δ in a fixed compact subset of $C_{\mathfrak{g}}(U)\cap\mathfrak{g}^-\setminus\{0\}$ so that $\|\log(g''_\delta)-v_\delta\|\to 0$ as $\delta\to 0$. By compactness of K and continuity of ϕ on K there exist some limit points $\overline{x}$ and $\overline{y}$ with $\phi(\overline{x})=\phi(\overline{y})$ and $\overline{y}=\overline{g}.\overline{x}$ for some $\overline{g}$ with $\log(\overline{g})\in(C_{\mathfrak{g}}(U)\cap\mathfrak{g}^-)\setminus\{0\}$. But $\log(\overline{g})\in C_{\mathfrak{g}}(U)$ implies that $\overline{g}$ commutes with a neighborhood of $e\in U$: hence since this is an algebraic condition $\overline{g}\in C_G(U)$. Similarly, $\log(\overline{g})\in\mathfrak{g}^-$ implies that $\overline{g}\in G^-$. And obviously $\log(\overline{g})\neq 0$ implies that $\overline{g}\neq e$.

Since $K\subset X'$ we have found two distinct points in X' with the same ϕ-value on the same H^--orbit contradicting assumption (A-3) of Section 5.1. This contradiction completes the proof of Theorem 2.1, and hence in view of the results of Section 4 establishes our main theorem, Theorem 1.4.

6. Regularity of leafwise measures

In this section we consider the leafwise measures μ_x^U under the assumption (U-1) of Theorem 1.4 and will show that typically μ_x^U does not have most of its mass close to subvarieties. Recall our assumption that G is either a Lie group or a direct product of linear groups and that in the latter case U is isomorphic to a direct product of Zariski closed unipotent groups. We start by analyzing polynomials on U.

6.1. **The** log**-map in positive characteristic.** In the case where $G_\sigma \subset \mathrm{GL}_m(\mathbb{K}_\sigma)$ is an algebraic group over a local field $\mathbb{K}_\sigma$ of positive characteristic the standard power series for the logarithm does not define a map (due to all primes appearing in the denominator of some term in the series). Here we may use the following crude replacement[14] of the Lie algebra and the logarithm map: We set $\mathfrak{g} = \mathrm{Mat}_m(\mathbb{K}_\sigma)$ to be the matrix algebra containing G and let $\log(g) = g - I$ where I is the identity matrix. Moreover, if we define the adjoint action by conjugation $\mathrm{Ad}_g(w) = gwg^{-1}$ for $g \in G$ and $w \in \mathfrak{g}$, then clearly $\mathrm{Ad}_g(\log h) = \log(ghg^{-1})$ for all $g, h \in G$. With this notation in mind, we do not have to distinuish between zero and positive characteristic below.

6.2. a**-homogeneous polynomials.** The basic assumption we had on a in Theorem 1.4 is that Ad_a is diagonalizable. We denote the eigenvalues of Ad_a acting on $\mathfrak{g}$ by letters $\xi, \lambda, \ldots$ and the eigenspaces by $\mathfrak{g}^\xi, \mathfrak{g}^\lambda, \ldots$. Note that in the S-algebraic case, the eigenvalues are elements in any one of the fields $\mathbb{K}_\sigma$ over which G is defined. Moreover, even if $\mathbb{K}_\sigma \cong \mathbb{K}_{\sigma'}$ for $\sigma \neq \sigma'$ we view them as distinct fields, and so in particular each $\mathfrak{g}^\lambda$ is by definition a subspace of some $\mathfrak{g}_\sigma$. We study the conjugation map $u \mapsto aua^{-1}$ for $u \in U$ in the following lemma.

Lemma 6.1. *Under the assumptions of Theorem 1.4 each of the subgroups U_σ is isomorphic as a variety to a product of affine spaces $\prod_\lambda \mathbb{A}_\lambda$ such that conjugation by a corresponds to multiplication by λ on $\mathbb{A}_\lambda \cong \mathbb{K}_\sigma^{d(\lambda)}$.*

For characteristic zero (or large enough positive characteristic) the above is an easy consequence of the polynomial maps log and exp defined between U and its Lie algebra.

A $\mathbb{K}_\sigma$-valued or $\mathfrak{g}$-valued polynomial $p(u)$ defined on U_σ is *a-homogeneous of weight* λ if $\lambda \in \mathbb{K}_\sigma$ and $p(aua^{-1}) = \lambda p(u)$ for all $u \in U_\sigma$. Here we assume that multiplication by $\lambda \in \mathbb{K}_\sigma$ annihilates $\mathfrak{g}_{\sigma'}$ for any $\sigma' \neq \sigma$, so that an a-homogeneous $\mathfrak{g}$-valued polynomial automatically only depends on the coordinates in U_σ and has its values in $\mathfrak{g}_\sigma$. Since a is Ad-diagonalizable, every polynomial is a unique sum of a-homogeneous polynomials.

The above lemma provides us with the choice of independent a-homogeneous variables generating the coordinate ring $\mathbb{K}_\sigma(U_\sigma)$. When we consider coefficients of a polynomial on U we will implicitly assume that we expressed the polynomial in terms of these variables.

Proof of Lemma 6.1 in positive characteristic. We only sketch the proof by indicating how to use the Lie algebra of U_σ to find independent a-homogeneous variables in $\mathbb{K}_\sigma(U_\sigma)$. For this recall the following definition of the Lie algebra: If $\mathfrak{m}_e < \mathbb{K}_\sigma(U_\sigma)$

[14] Using these definitions, strictly speaking, (U-2) becomes an assumption on the concrete realization of G as a linear group.

is the ideal of polynomials vanishing at $e \in U$, then the Lie algebra is given by the dual space to $\mathfrak{m}_e/\mathfrak{m}_e^2$. The semisimple a acts also on $\mathfrak{m}_e/\mathfrak{m}_e^2$ which also decomposes into eigenspaces. Now we can choose for every eigenvalue λ some a-homogeneous polynomials $x_{\lambda,1}, \ldots, x_{\lambda,d(\lambda)}$ of weight λ such that they form modulo $\mathfrak{m}_e^2$ a basis of the corresponding eigenspaces in $\mathfrak{m}_e/\mathfrak{m}_e^2$. By induction on the weights one can show that these variables generate the coordinate ring $\mathbb{K}_\sigma(U_\sigma)$. However, since the dimension of $\mathfrak{m}_e/\mathfrak{m}_e^2$ equals the dimension of U it follows that these variables must be algebraically independent. $\square$

6.3. **The mass distribution for the leafwise measures.** We will first show that the Zariski closure of the support of μ_x^U is almost surely a subgroup.

Proposition 6.2. *Let $X = \Gamma\backslash G$, let $a \in G$ be class $\mathcal{A}$, and let $U < G^-$ be an a-normalized direct product of closed unipotent subgroups U_σ of G_σ for $\sigma \in S$. Let μ be an a-invariant probability measure on X. Then for almost every x the Zariski closure P_x of $\operatorname{supp}(\mu_x^U|_{B_1^U})$ is a Zariski closed a-normalized subgroup of U with $\operatorname{supp}(\mu_x^U) \subset P_x$.*

The above proposition is general in the sense that we have not assumed that μ satisfies the assumption (U-1) in Theorem 1.4, under this assumption we get the following immediate corollary.

Corollary 6.3. *Suppose in addition to the assumptions of Proposition 6.2 that μ satisfies (U-1) in Theorem 1.4. Then for almost every x and any nonzero polynomial (defined on any of the factors of U) we have*

$$\operatorname{supp}(\mu_x^U|_{B_1^U}) \not\subset Z(f) = \{u \in U : f(u) = 0\}.$$

However, what we need for the low entropy argument is not only that the support of the leafwise measure is not contained in a variety, but in fact the stronger statement saying that the mass is not too closely concentrated near varieties. As we shall see the above and a simple compactness argument implies this. We will prove the next corollary after we finish the proof of Proposition 6.2.

Corollary 6.4. *Let $D \geq 1$ and $\epsilon > 0$. Under the same assumption as in Corollary 6.3, there exists $\eta > 0$ and a set $Q \subset X$ of measure $\mu(Q) > 1 - \epsilon$ such that*

$$\mu_x^U(\{u \in B_1^U : \|f(u)\| \geq \eta\}) \geq \eta$$

for every $x \in Q$ and every polynomial f of degree less than D (defined on any of the factors of U) that has a coefficient of norm one.

This justifies the assertion (5.4) made in Section 5.2. The main step towards the proposition is contained in the following lemma.

Lemma 6.5. *Let P_x be the Zariski closure of $\operatorname{supp}(\mu_x^U|_{B_1^U})$. Then for almost every x the ideal defining P_x is a-homogeneous, i.e. it is generated by finitely many a-homogeneous polynomials and so P_x is normalized under the action of a. Moreover, $\operatorname{supp}(\mu_x^U) \subset P_x$ so that P_x is also the Zariski closure of $\operatorname{supp}(\mu_x^U)$.*

Proof. Let K be a Lusin set for μ_x^U of almost full measure such that (LM-5) holds for points in K. Then by Poincaré recurrence for almost every $x \in K$ there exists a sequence $n_k \to -\infty$ such that $a^{n_k}.x \in K$ and $a^{n_k}.x \to x$ for $k \to \infty$. For any such x we claim that the lemma holds. Let $m \geq 1$ be fixed. Since $\mu_{a^{n_k}.x}^U \propto$

$a^{n_k}\mu_x^U a^{-n_k}$ by (LM-5) and $a^{n_k}B_1^U a^{-n_k}$ (for large negative n_k) contains B_m^U we see that $\mathrm{supp}(\mu^U_{a^{n_k}.x}|_{B_m^U}) \subset a^{n_k}(\mathrm{supp}\mu_x^U|_{B_1^U})a^{-n_k}$.

Let $f(u)$ be a polynomial that vanishes on $\mathrm{supp}(\mu_x^U|_{B_1^U})$. Then

$$f(a^{-n_k}ua^{n_k}) \text{ vanishes on } \mathrm{supp}(\mu^U_{a^{n_k}.x}|_{B_m^U}). \tag{6.1}$$

If f is not already homogeneous, then we can write $f = f_{\mathrm{lt}} + f_{\mathrm{rem}}$. Here the leading term f_{lt} is a-homogeneous of some weight λ and the remaining polynomial f_{rem} is a sum of a-homogeneous polynomials of weights that all have norm bigger than $\|\lambda\|$. For this notice that we cannot have two weights of equal norm since a is of class $\mathcal{A}$. Also recall that all nontrivial weights are in norm less than 1 since U is contracted by a. Then

$$\lim_{k\to\infty} \lambda^{n_k} f(a^{-n_k}ua^{n_k}) = f_{\mathrm{lt}}(u)$$

is a-homogeneous. Using (6.1) and that $\mu^U_{a^{n_k}.x}|_{B_m^U} \to \mu_x^U|_{B_m^U}$ we see that f_{lt} also vanishes on $\mathrm{supp}(\mu_x^U|_{B_m^U})$. Replacing f by $f - f_{\mathrm{lt}}$ and continuing by induction we can write f as a sum of a-homogeneous polynomials that all vanish on $\mathrm{supp}(\mu_x^U|_{B_m^U})$. This is true for all generators of the ideal defining P_x, and — since m was arbitrary — $\mathrm{supp}(\mu_x^U) \subset P_x$. □

Lemma 6.6. *For μ-a.e. x and μ_x^U-a.e. $u_0 \in U$ we have $P_{u_0.x}u_0 = P_x$.*

Proof. First recall that $\mu^U_{u_0.x}u_0 \propto \mu_x^U$ whenever $x, u_0.x \in X'$ by (LM-4). Therefore, the support satisfies $\mathrm{supp}(\mu^U_{u_0.x})u_0 = (\mathrm{supp}(\mu_x^U))$ and so $P_{u_0.x}u_0 = P_x$ by Lemma 6.5. □

Lemma 6.7. *Let $f(u)$ be an a-homogeneous polynomial of weight λ. Let $u_0 \in U$ be fixed. Then $f(uu_0^{-1}) = f(u) + f_0(u)$ where any a-homogeneous term in $f_0(u)$ has a weight of absolute value bigger than λ.*

Proof. By definition a polynomial is obtained by taking sums of products of matrix coefficients — where we can assume that a is a diagonal matrix with its eigenvalues arranged in increasing order. Therefore, it is enough to check the statement for matrix coefficients. However, for these the statement follows from the structure of multiplication for upper triangular matrices: Any matrix coefficient of uu_0^{-1} is the sum of the corresponding matrix coefficients of u and u_0^{-1} and of products of matrix coefficients of u and u_0^{-1}. Here the weights of the matrix coefficients of u respectively u_0^{-1} must give as a product the weights of the matrix coefficient of uu_0^{-1} currently considered. However, we leave u_0 fixed and so the lemma follows once we recall that all weights considered are in absolute value less than one. □

Proof of Proposition 6.2. Suppose x satisfies Lemma 6.5 and 6.6 and, moreover, that $u_0.x$ also satisfies Lemma 6.5 for μ_x^U-a.e. $u_0 \in U$. Take such an u_0, and let f be an a-homogeneous polynomials of weight λ that vanishes on $P_{u_0.x}$. Then the polynomial $f(uu_0^{-1})$ in u vanishes on P_x by Lemma 6.6. By Lemma 6.7 we know $f(uu_0^{-1}) = f(u) + f_0(u)$ where $f_0(u)$ has only terms of weight that are in absolute value bigger than λ. By Lemma 6.5 we conclude that $f(u)$ vanishes on P_x. Since the ideal for $P_{u_0.x}$ is generated by such polynomials, $P_x \subset P_{u_0.x}$. The proof of the opposite inclusion is the same.

Therefore, $P_x = P_{u_0.x} = P_{u_0.x}u_0 = P_xu_0$ for μ_x^U-a.e. $u_0 \in U$ by Lemma 6.6 again. Since $V = \{u_0 \in U : P_x = P_xu_0\}$ is itself a Zariski closed subset of U, it follows that $P_x \subset V$ by definition. Therefore, P_x is a subgroup of U. □

Corollary 6.3 follows immediately from Proposition 6.2, to see Corollary 6.4 we will combine Corollary 6.3 with the following lemma. The norm of a polynomial defined on U is the maximum of the norms of the coefficients when writing the polynomial in terms of the variables given in Lemma 6.1.

Lemma 6.8. *Fix some $D \geq 1$. Then for every probability measure ν on $\overline{B_1^U}$ either*

(1) *there is a nonzero polynomial f on U of degree less than D such that* $\mathrm{supp}(\nu) \subset Z(f) = \{u \in U : f(u) = 0\}$, *or*

(2) *there exists an $\eta > 0$ depending only on ν such that for all polynomials of degree less than D with norm one we have*

$$\nu\big(\{u \in \overline{B_1^U} : \|f(u)\| \geq \eta\}\big) \geq \eta. \tag{6.2}$$

We define $\eta_\nu = 0$ in the first case and let η_ν be the maximal η satisfying (6.2) *in the second case. Then the set of probability measures ν with $\eta_\nu \geq \epsilon$ is weak* closed.*

Proof. Suppose (2) fails for ν. Then for every $\eta = \frac{1}{n}$ there exists some polynomial f_n as in (2) with $\nu(\{u \in \overline{B_1^U} : \|f_n(u)\| \geq \eta\}) \leq \eta$. Choosing a subsequence we may assume that $f_{n_k} \to f$ where f is a polynomial of degree less than D and norm one. Let $\epsilon > 0$, then for large enough k we have $\{u \in \overline{B_1^U} : \|f(u)\| < \epsilon\} \supset \{u \in \overline{B_1^U} : \|f_{n_k}(u)\| < \frac{1}{n_k}\}$. This implies that $\nu(\{u \in \overline{B_1^U} : \|f(u)\| < \epsilon\}) \geq 1 - \epsilon$ and so ν is supported on $Z(f)$ as claimed in (1).

For the second claim let $\nu_n \to \nu$ and let $\eta_{\nu_n} \geq \epsilon$. Let f be a polynomial of degree less than D and norm one, then by definition $\nu_n(\{u \in \overline{B_1^U} : \|f(u)\| \geq \epsilon\}) \geq \epsilon$ for all n. Since $\{u \in \overline{B_1^U} : \|f(u)\| \geq \epsilon\}$ is compact, the same holds as well for the weak* limit ν. □

Proof of Corollary 6.4. Define ν_x to be the renormalized restriction of μ_x^U to $\overline{B_1^U}$. By Corollary 6.3 case (1) in Lemma 6.8 happens μ-almost never, i.e. $\eta_{\nu_x} > 0$ a.e. and depends measurably on x. This implies that for small enough η

$$\mu_x^U(\{u \in \overline{B_1^U} : \|f(u)\| \geq \eta\}) \geq \eta\mu_x^U(\overline{B_1^U}) \geq \eta$$

on a set Q of measure bigger than $1-\epsilon$. To get rid of the closure we can use (LM-5) and take the image of Q under a to get

$$\mu_x^U(\{u \in a\overline{B_1^U}a^{-1} : \|f(u)\| \geq \eta\}) \geq \eta\mu_x^U(a\overline{B_1^U}a^{-1}).$$

for $x \in a.Q$. Here $a\overline{B_1^U}a^{-1} \subset B_1^U$ is some fixed neighborhood of $e \in U$ and so by (LM-3) we can give a uniform lower bound $c \in (0,1)$ on its measure for $x \in a.Q$ after possibly making Q smaller. Now the corollary follows for $c\eta$ instead of η. □

7. Quantative shearing estimates and the function n_k

In this section we define n_k and provide the missing details of the step described in Section 5.6. Let us note that this section is purely algebraic, the considerations do not use μ or properties of a particular point in X. All the results hinge upon the assumption (U-2) of Theorem 1.4:

(U-2) for all nonzero $w \in \mathfrak{g}_+$, it holds that $\mathrm{Ad}(U)[w] \not\subset \mathfrak{g}_+ \oplus \mathfrak{g}_0$.

we isolate explicitly the following corollary of this assumption:

(U-$*$) $C_{\mathfrak{g}}(U) = \{w \in \mathfrak{g} : \mathrm{Ad}_u(w) = w \text{ for all } u \in U\} \subset \mathfrak{g}^0 + \mathfrak{g}^-$.

7.1. **The definition of n_k.** Given $w \in \mathfrak{g}$ we have assumed that $\mathrm{Ad}_u(w)$ depends polynomially on $u \in U$. We decompose $\mathrm{Ad}_u(w)$ in two different ways: We first decompose w into its components in the eigenspaces $w^\lambda \in \mathfrak{g}^\lambda$ and consider $\mathrm{Ad}_u(w^\lambda)$. We can further decompose $\mathrm{Ad}_u(w^\lambda)$ into components in the eigenspaces $\mathrm{Ad}_u^\xi(w^\lambda) \in \mathfrak{g}^\xi$.

Lemma 7.1. *The $\mathfrak{g}^\xi$-valued polynomial map $\mathrm{Ad}_u^\xi(w^\lambda)$ in $u \in U$ is a-homogeneous of weight $\frac{\xi}{\lambda}$.*

Proof. By definition
$$\mathrm{Ad}_{aua^{-1}}(w^\lambda) = \lambda^{-1}\,\mathrm{Ad}_{au}(w^\lambda).$$
By taking components according to the weight ξ we get
$$\mathrm{Ad}^\xi_{aua^{-1}}(w^\lambda) = \lambda^{-1}\,\mathrm{Ad}_a \circ \mathrm{Ad}^\xi_u(w^\lambda) = \frac{\xi}{\lambda}\,\mathrm{Ad}^\xi_u(w^\lambda)$$
which proves the lemma. □

Applying the lemma we get
$$\mathrm{Ad}_{a^{-n}ua^n} \circ \mathrm{Ad}_{a^k}(w) = \sum_{\lambda,\xi} \mathrm{Ad}^\xi_{a^{-n}ua^n} \circ \mathrm{Ad}_{a^k}(w^\lambda) = \sum_{\lambda,\xi} \lambda^k \left(\frac{\lambda}{\xi}\right)^n \mathrm{Ad}^\xi_u(w^\lambda).$$
Since U is uniformly contracted, we must have $\|\xi/\lambda\| < 1$ for all nontrivial terms in the sum unless the corresponding term $\mathrm{Ad}^\xi_u(v^\lambda) = v^\lambda$ is a constant polynomial and $\xi = \lambda$. In particular, this shows that the coefficients of all nonconstant $\mathrm{Ad}_{a^{-n}ua^n}(a^k w a^{-k})$ grow exponentially with growing n.

We define the norm of a $\mathfrak{g}$-valued polynomial on U as the maximum of the norms of the coefficients, where we use a fixed set of independent homogeneous variables in U as in Lemma 6.1 and a fixed basis of $\mathfrak{g}$ consisting of eigenvectors. Then the integer valued linear function
$$n_k(\lambda,\xi) = \left\lfloor \frac{\log(\|\lambda\|^{-k}\|\,\mathrm{Ad}^\xi_u(w^\lambda)\|^{-1})}{\log\|\lambda\| - \log\|\xi\|} \right\rfloor = \lfloor s_{\lambda,\xi} k + \tau_{\lambda,\xi}(w) \rfloor$$
for
$$s_{\lambda,\xi} = -\frac{\log\|\lambda\|}{\log\|\lambda\| - \log\|\xi\|} \qquad \tau_{\lambda,\xi}(w) = -\frac{\log\|\,\mathrm{Ad}^\xi_u(w^\lambda)\|}{\log\|\lambda\| - \log\|\xi\|} \tag{7.1}$$
satisfies
$$\|\,\mathrm{Ad}^\xi_{a^{-n}ua^n}(a^k w^\lambda a^{-k})\| \in (c, 1]$$
for $n = n_k(\lambda,\xi)$ — unless $\mathrm{Ad}^\xi_u(w^\lambda)$ is constant and $n_k(\lambda,\xi) = \infty$. Here $c > 0$ is some absolute constant that only depends on the action of a on U. Clearly for $0 \le n < n_k(\lambda,\xi)$ the norm is also less than 1. Therefore, if
$$n_k = n_k(w) = \min_{\lambda,\xi} n_k(\lambda,\xi)$$
is finite, it satisfies
$$\|\,\mathrm{Ad}_{a^{-n_k}ua^{n_k}}(a^k w a^{-k})\| \in (c, 1]. \tag{7.2}$$
Assume now $w \notin C_{\mathfrak{g}}(U)$, then by construction $n_k : \mathbb{Z} \to \mathbb{Z}$ is (the integer part of) a piecewise linear function and the slopes only depend on the weights of a.

7.2. **The directions $\mathfrak{v}$ of fastest divergence along U.** Just as in the last subsection, we treat $\mathrm{Ad}_u(w)$ as a polynomial (i.e. regular function) in u with coefficients in $\mathfrak{g}$.

For any $w \in \mathfrak{g}$ let $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ be the sum of the smallest weight terms in $\mathrm{Ad}_u(w)$ for the a-action on U (smallest a-homogeneous weight), i.e. for $w \in C_{\mathfrak{g}}(U)$ we take $\mathrm{Ad}_u^{\mathrm{lt}}(w) = w$ and for $w \notin C_{\mathfrak{g}}(U)$ we write

$$\mathrm{Ad}_u(w) = \mathrm{Ad}_u^{\mathrm{lt}}(w) + \sum_{\xi,\lambda:\|\xi/\lambda\|>\tilde{\zeta}} \mathrm{Ad}_u^{\xi}(w^{\lambda}); \qquad \mathrm{Ad}_u^{\mathrm{lt}}(w) = \sum_{\xi,\lambda:\|\xi/\lambda\|=\tilde{\zeta}} \mathrm{Ad}_u^{\xi}(w^{\lambda})$$

with $\tilde{\zeta}$ chosen so that $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ is not identically equal to zero.

Note that $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ is the part of $\mathrm{Ad}_u(w)$ with fastest expansion when a^{-1} acts on U. It is not necessarily true that $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ has values in one of the eigenspaces of $\mathfrak{g}$. However, since by assumption all eigenvalues of $\mathrm{Ad}(a)$ have distinct absolute values, if $w \in \mathfrak{g}^{\lambda}$ then $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ is a-homogeneous with some weight ζ, and its values lie in $\mathfrak{g}^{\xi}$ with $\zeta = \frac{\xi}{\lambda}$ by Lemma 7.1.

We define the subspace $\mathfrak{v} \subset \mathfrak{g}$ of *directions of fastest divergence along U* to be the linear hull of the coefficients of $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ for all $w \in \mathfrak{g} \setminus C_{\mathfrak{g}}(U)$.

Lemma 7.2. *We have*

(1) *that $\mathfrak{v}$ is contained in $C_{\mathfrak{g}}(U)$,*
(2) *that $\mathfrak{v}$ is a direct sum of subspaces of weight spaces $\mathfrak{g}^{\xi}$ for various weights ξ (and hence is normalized by Ad_a),*
(3) *and under the assumption (U-$*$), that $\mathfrak{v}$ is contained in $\mathfrak{g}^{-}$.*

Proof. We first remark that for any $w = \sum_{\lambda} w^{\lambda} \in \mathfrak{g}$ we have that the polynomial $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ is in the linear span of the polynomials $\mathrm{Ad}_u^{\mathrm{lt}}(w^{\lambda})$. This holds since there is no cancellation between terms arising from different weights — indeed, since we have assumed that a is class $\mathcal{A}$ by Lemma 7.1 for any ξ the polynomials $\mathrm{Ad}_u^{\xi}(w^{\lambda})$ for all possible values of λ have different a-homogeneous weight, and hence do not cancel. Moreover, if $w \notin C_{\mathfrak{g}}(U)$ the a-homogeneous weight of the low order term $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ has absolute value < 1, and hence $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ is in the linear span of the polynomials $\mathrm{Ad}_u^{\mathrm{lt}}(w^{\lambda})$ for $w_{\lambda} \notin C_{\mathfrak{g}}(U)$. This fact implies (2) above, and shows that in proving (1) and (3) we need only to consider eigenvectors $w^{\lambda} \in \mathfrak{g}^{\lambda} \setminus C_{\mathfrak{g}}(U)$.

Consider now for $w^{\lambda} \in \mathfrak{g}^{\lambda}$ (not in $C_{\mathfrak{g}}(U)$) the expansion

$$\mathrm{Ad}_{u_0}(w^{\lambda}) = \mathrm{Ad}_{u_0}^{\mathrm{lt}}(w^{\lambda}) + \dots \tag{7.3}$$

and assume that ξ is such that $\mathrm{Ad}_{u_0}^{\mathrm{lt}}(w^{\lambda}) = \mathrm{Ad}_{u_0}^{\xi}(w^{\lambda})$ — it is uniquely determined by w^{λ}. Here and below the dots indicate terms in weight spaces of bigger weight. Now fix u_0 and apply Ad_u to get

$$\mathrm{Ad}_{uu_0}(w^{\lambda}) = \mathrm{Ad}_u\big(\mathrm{Ad}_{u_0}^{\xi}(w^{\lambda}) + \dots\big) = \mathrm{Ad}_{uu_0}^{\xi}(w^{\lambda}) + \dots \tag{7.4}$$

Here we do not have any terms of weight in absolute value smaller than ξ on the far right of (7.4) since $uu_0 \in U$ and ξ was choosen minimally in (7.3).

Recall that $\mathrm{Ad}_u\big(\mathrm{Ad}_{u_0}^{\xi}(w^{\lambda})\big) - \mathrm{Ad}_{u_0}^{\xi}(w^{\lambda})$ can only have nontrivial components in $\mathfrak{g}^{\eta}$ where η has absolute value less than $\|\xi\|$ (again using Lemma 7.1 and that U is contracted by a). However, if this polynomial would have a nonzero component $f(u)$ in the weight space $\mathfrak{g}^{\eta}$ for some η as above then some eigenvector component v_{u_0} of $\mathrm{Ad}_{u_0}(w^{\lambda})$ that appears on the right of (7.3) as part of the dots would also have to give a component $\mathrm{Ad}_u^{\eta}(v_{u_0}) \in \mathfrak{g}^{\eta}$ in order for (7.4) to hold. However,

$\mathrm{Ad}_u^\eta(v_{u_0})$ and $f(u)$ considered as a polynomial in u have different a-homogeneous degree by Lemma 7.1. Therefore, $\mathrm{Ad}_u(\mathrm{Ad}_{u_0}^\xi(w^\lambda)) = \mathrm{Ad}_{u_0}^\xi(w^\lambda)$ cannot contain any nonzero terms in a weight space of lower weight. This shows that the leading terms coming from eigenvectors w^λ belong to $C_{\mathfrak{g}}(U)$, establishing (1) of the lemma.

To prove (3), we distinguish between two cases: If $\|\lambda\| \le 1$, then since $U \in G^-$ for any $u \in U$

$$\mathrm{Ad}_u(w^\lambda) - w^\lambda \in \mathfrak{g}^- \tag{7.5}$$

and hence all the coefficients of this polynomial (which in particular include the coefficients of the polynomial $\mathrm{Ad}_u^{\mathrm{lt}}(w^\lambda)$) are in $\mathfrak{g}^-$. Note that (7.5) is not identically zero since $w^\lambda \notin C_{\mathfrak{g}}(U)$. If $\|\lambda\| > 1$, then for any $w^\lambda \in \mathfrak{g}^\lambda$ we know that there exists some $u \in U$ with $\mathrm{Ad}_u(w^\lambda) \notin \mathfrak{g}^+ + \mathfrak{g}^0$ by assumption (U-$*$). In particular, this shows that in this case $\mathrm{Ad}_u^{\mathrm{lt}}(w^\lambda)$ also has its coefficients in $\mathfrak{g}^-$, completing the proof of statement (3) of the lemma. □

The next lemma establishes that indeed $\mathfrak{v}$ can be interpreted as the space of directions of fastest divergence along U:

Lemma 7.3. *For every $\rho > 0$ there exists $\delta > 0$ such that $\|w\| < \delta$ for $w \in \mathfrak{g}$ implies that for any $n \le n_0 = n_0(w)$ and any $u \in a^{-n}B_1^U a^n$ there exists an $v \in \mathfrak{v}$ such that $\|\mathrm{Ad}_u(w) - v\| < \rho$.*

Proof. Without loss of generality we may assume that $w \in \mathfrak{g}^\lambda$, for otherwise we can decompose w accordingly and the result for the components easily implies the same for w.

Fix inside every weight space $\mathfrak{g}^\xi$ a linear complement to $\mathfrak{v} \cap \mathfrak{g}^\xi$ and take the sum of these to obtain a linear complement $\mathfrak{v}^\perp \subset \mathfrak{g}$ to $\mathfrak{v}$. For every $w \in \mathfrak{g}^\lambda$ and a given weight ξ of smaller norm than λ we claim that the norm of the component of $\mathrm{Ad}_u^\xi(w)$ in $\mathfrak{v}^\perp$ is bounded by a multiple of the maximum of the norms of $\mathrm{Ad}_u^\zeta(w)$ for weights ζ smaller than ξ. This follows by considering the kernel of the corresponding maps; if the sum of $\mathrm{Ad}_u^\zeta(w)$ for all weights ζ that are smaller than ξ would have a bigger kernel than the projection of $\mathrm{Ad}_u^\xi(w)$ to $\mathfrak{v}^\perp$ we would have an element $w \in \mathfrak{g}^\lambda$ whose leading term $\mathrm{Ad}_u^{\mathrm{lt}}(w)$ has weight ζ and a nontrivial component in $\mathfrak{v}^\perp$ which contradicts the definition of $\mathfrak{v}$. Therefore, the norm of the component of $\mathrm{Ad}_u^\xi(w)$ in $\mathfrak{v}^\perp$ is bounded by a constant multiple of the norm of $\mathrm{Ad}_u^\zeta(w)$ for some ζ of smaller norm.

By Lemma 7.1 we know that $\mathrm{Ad}_u^\xi(w)$ has bigger a-homogeneous weight than $\mathrm{Ad}_u^\zeta(w)$, in other words when applying conjugation by a^{-1} to u the projection of $\mathrm{Ad}_u^\xi(w)$ to $\mathfrak{v}^\perp$ will grow at a slower exponential rate than $\mathrm{Ad}_u^\zeta(w)$. Recall that $n = n_0(w)$ was chosen so that at least one coefficient of $\mathrm{Ad}_{a^{-n}ua^n}(w)$ has size one. Together with the above the lemma follows. □

7.3. **(U-2) implies control on the a-divergence.** In our main argument we need to show that the points x'_δ and y'_δ are still close together. These points are produced out of two nearby points x_δ and y_δ via the action of a^k which makes it important to restrict the parameter k.

Proposition 7.4. *Assuming (U-2) there exists some $\kappa > 0$ such that for sufficiently small $w \in \mathfrak{g}$ and $0 \le k \le k_0 = \kappa n_0(w)$ implies $\|\mathrm{Ad}_{a^k}(w)\| \le \|w\|^{1/2}$.*

Proof. By assumption (U-2) the expression $\mathrm{Ad}_u(w) - w$ (considered as a $\mathfrak{g}$-valued polynomial in u) depends injectively on w when restricted to $\mathfrak{g}^+$. In particular, $\|w\|$ is bounded by a multiple of $\|\mathrm{Ad}_u(w) - w\|$. Therefore, for every weight λ of norm bigger than one and every $w^\lambda \in \mathfrak{g}^\lambda$ there exists a weight ξ such that the constant term $\tau_{\lambda,\xi} = -\frac{\log\|\mathrm{Ad}_u^\xi(w^\lambda)\|}{\log\|\lambda\| - \log\|\xi\|}$ in the linear function $n_k(\lambda, \xi)$ (cf. (7.1)) is bounded by $-C_1 \log\|w^\lambda\| + C_2$ for two absolute constants $C_1, C_2 > 0$. This shows that $n_0(w^\lambda) \le -C_1 \log\|w^\lambda\| + C_2$ whenever λ has norm bigger than one. We choose $\kappa > 0$ such that $\kappa < (2 \log C_1 \|\lambda\|)^{-1}$ for all such weights. In particular, this shows that $\kappa n_0(w^\lambda) < -\frac{\log\|w^\lambda\|}{2\log\|\lambda\|}$ for all small enough w^λ.

Let $w \in \mathfrak{g}$ and $w = \sum_\lambda w^\lambda$. Then $\mathrm{Ad}_{a^k}(w) = \sum_\lambda \lambda^k w^\lambda$. The lemma follows since by the above $\|\lambda^k w^\lambda\| < \|w^\lambda\|^{1/2}$ for all λ of norm greater than one and all $k \le \kappa n_0(w) \le \kappa n_0(w^\lambda)$. □

7.4. **(U-$*$) implies that there is no sub interval of slope -1.** To be able to use a density argument to find k in Section 5.6 we need to know that n_k has no sub intervals of slope -1.

Proposition 7.5. *Assumption (U-$*$) implies that there are some $\delta_0, \kappa > 0$, an integer J and real numbers $s_1, \ldots, s_J$ all $\neq -1$ (all of these parameters depend only on a, U, G), so that for $w \in \mathfrak{g}$ with $\|w\| < \delta_0$ and $k \in [0, \kappa n_0(w)]$ we may write*

$$n_k(w) = \min_{1 \le j \le J} \lfloor s_j k + \tau_j(w) \rfloor.$$

Proof. Recall that n_k is defined as the minimum of $n_k(\lambda, \xi) = \lfloor s_{\lambda,\xi} k + \tau_{\lambda,\xi}(w) \rfloor$ for all weights λ, ξ and that $s_{\lambda,\xi} = -1$ iff $\xi = 1$ (see (7.1)).

To prove the proposition, we need to show that if w and k satisfy the conditions of the proposition n_k cannot equal to $n_k(\lambda, \xi)$ if $\xi = 1$. Suppose this were false. Without loss of generality we may suppose $w = w^\lambda \in \mathfrak{g}^\lambda$ and let $w' = \mathrm{Ad}_{a^k}(w)$. By Proposition 7.4 $\|w'\| < \|w\|^{1/2}$. By definition of $n_k(\lambda, \xi)$ the polynomial $\mathrm{Ad}^\xi_{a^{-n_k} u a^{n_k}}(w')$ has a coefficient of almost norm one – in particular there exist some $u \in B_1^U$ such that $\|\mathrm{Ad}^1_{a^{-n_k} u a^{n_k}}(w')\|$ has a fixed lower bound. However, Lemma 7.3 shows that $\mathrm{Ad}_{a^{-n_k} u a^{n_k}}(w')$ is close to $\mathfrak{v}$ if w is small enough. This gives a contradiction to $\mathfrak{v} < \mathfrak{g}^-$, therefore $n_k < n_k(\lambda, 1)$ for all weights λ if only δ_0 is small enough (smaller than some constant depending only on a, U, G). □

8. Application: Invariant measures on $\Gamma\backslash G_1 \times G_2$

In this section, we prove Theorem 1.5 regarding A_1-invariant, G_2-recurrent measures on $\Gamma\backslash G_1 \times G_2$ where G_1 is a rank one group over a local field $\mathbb{K}_\sigma$, $A_1 \subset G_1$ is a $\mathbb{K}_\sigma$-split torus and G_2 is any zero characteristic S-algebraic group. We let $a \in A_1$ be an element for which $\chi(a) \in \mathbb{Q}_\sigma$ and $|\chi(a)|_\sigma \neq 1$ for all roots χ. (Clearly, it is enough to ask this for the unique positive indivisible root of G_1 since all other roots are powers of this one.)

8.1. **Some algebraic facts for semisimple groups of rank one.** The following two lemmata will be helpful for the discussion of all possible homogeneous measures as in Theorem 1.5 as well as in the justification of the assumptions to Theorem 1.4. To simplify the notation in these algebraic facts, we will write G for the semisimple group of $\mathbb{K}_\sigma$-rank one instead of G_1.

Lemma 8.1. *Let G be a semisimple algebraic group over a characteristic zero local field $\mathbb{K}_\sigma$ of $\mathbb{K}_\sigma$-rank one. Let $a \in A$ be a diagonalizable element of the maximal split torus A in G with $|\chi(a)|_\sigma \neq 1$ as above, let $U \leq G^-$ be a nontrivial algebraic subgroup normalized by a. Then for any $g \in G^+$ there exists some $u \in U$ such that $ugu^{-1} \notin G^+G^0$.*

Note that this is quite easy from the structure of linear representations in the case where G is $\mathrm{SL}(2, \mathbb{K}_\sigma)$, and as shown below the general case can be reduced to this one.

Proof. Let $u = \exp \mathbf{u}$ for some eigenvector $\mathbf{u}$ of $\mathrm{Ad}(a)$ restricted to the Lie algebra of U. By the Jacobson-Morozov theorem, we can find $\mathbf{h}, \mathbf{n}$ completing $\mathbf{u}$ to an $\mathrm{sl}(2, \mathbb{K}_\sigma)$-triplet, i.e. $[\mathbf{h}, \mathbf{u}] = -2\mathbf{u}$, $[\mathbf{h}, \mathbf{n}] = 2\mathbf{n}$ and $[\mathbf{u}, \mathbf{n}] = 2\mathbf{h}$.

Since $\mathrm{Ad}(a)$ is diagonalizable over $\mathbb{K}_\sigma$, we can decompose the Lie algebra $\mathfrak{g}$ of G into weight spaces $\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_\lambda$ which are by definition eigenspaces for $\mathrm{Ad}(a)$. We have assumed that $\mathbf{u}$ is an eigenvector, i.e. it belongs to $\mathfrak{g}_{-\lambda_0}$ for some positive root λ_0 (where we use additive notation for the roots). Let $\mathbf{h} = \sum_\lambda \mathbf{h}_\lambda$ and $\mathbf{n} = \sum_\lambda \mathbf{n}_\lambda$ with $\mathbf{h}_\lambda, \mathbf{n}_\lambda \in \mathfrak{g}_\lambda$. Then as $[\mathfrak{g}_\lambda, \mathfrak{g}_\eta] \subset \mathfrak{g}_{\lambda+\eta}$ we have that $\mathbf{u}$, $\mathbf{h}_0$, and $\mathbf{n}_{\lambda_0}$ are also an $\mathrm{sl}(2, \mathbb{K}_\sigma)$ -triplet, hence without loss of generality we can assume that $\mathbf{h} = \mathbf{h}_0$ and $\mathbf{n} = \mathbf{n}_{\lambda_0}$.

Let $\mathfrak{h}$ be the Lie algebra generated by $\mathbf{u}, \mathbf{h}$ and $\mathbf{n}$. Clearly, both $\mathbf{u}$ and $\mathbf{n}$ generate Lie algebras to unipotent algebraic subgroups of G. The algebraic subgroup H generated by these must have $\mathfrak{h}$ as its Lie algebra(15). Therefore, H is isomorphic to either $\mathrm{SL}(2, \mathbb{K}_\sigma)$ or to $\mathrm{PGL}(2, \mathbb{K}_\sigma)$. Since G is of $\mathbb{K}_\sigma$-rank one and since the torus A_H in H whose Lie algebra contains $\mathbf{h}$ commutes with a, it follows that $A = A_H$ is this torus in H.

Now let $x = \exp \mathbf{x} \in G^+$. Write $\mathbf{x} = \sum_\lambda \mathbf{x}_\lambda$ where $\mathbf{x}_\lambda \in \mathfrak{g}_\lambda$. For each λ, for sufficiently large k, $\mathrm{ad}(\mathbf{u})^k \mathbf{x}_\lambda \in \mathfrak{g}^-$. Moreover, from the representation theory of $\mathrm{sl}(2, \mathbb{K}_\sigma)$ it follows also that if $\mathbf{x}_\lambda \in \mathfrak{g}_\lambda$ is nonzero for some positive λ, then $\mathrm{ad}(\mathbf{u})^{2\lambda} \mathbf{x}_\lambda \in \mathfrak{g}^{-\lambda}$ is nonzero. Therefore,

$$\mathrm{Ad}(\exp(t\mathbf{u}))(\mathbf{x}_\lambda) = \sum_n \frac{1}{n!} \mathrm{ad}(t\mathbf{u})^n \mathbf{x}_\lambda \notin \mathfrak{g}_+,$$

and the result follows by summing over all λ and choosing $t \in \mathbb{K}_\sigma$ to avoid cancellation. $\square$

Lemma 8.2. *Let G be a semisimple algebraic group over a characteristic zero local field $\mathbb{K}_\sigma$ of $\mathbb{K}_\sigma$-rank one. Let $a \in A$ be a diagonalizable element of the maximal split torus A in G with $|\chi(a)|_\sigma \neq 1$, let $U \leq G^-$ and $U_+ \leq G^+$ be nontrivial algebraic subgroups normalized by a. Then U and U_+ generate a finite index normal subgroup L° of a simple algebraic subgroup L of G containing a. In fact, L° is the subgroup generated by all one dimensional unipotent subgroups(16) of G.*

(15)To see this note first there is no proper Lie subalgebra of $\mathfrak{h}$ containing $\mathbf{u}$ and $\mathbf{n}$, so that the Lie algebra of H contains $\mathfrak{h}$. Next notice that the unipotent groups leave $\mathfrak{h}$ invariant and so the same must be true for H. This shows that $\mathfrak{h}$ is a Lie ideal in the Lie algebra of H. However, for algebraic groups semisimple ideals in the Lie algebra correspond to normal Zariski closed subgroups which shows the claim.

(16)As before, we suppress here a few of the algebraic phrases, i.e. more precisely L° is the subgroup generated by all $\mathbb{K}_\sigma$-points of one dimensional unipotent algebraic subgroups of G defined over $\mathbb{K}_\sigma$.

Proof. Let L be the algebraic subgroup generated by the algebraic groups U and U_+. We claim first that L is a simple algebraic subgroup of G which contains A. For this we proceed similar to the proof of Lemma 8.1. Let $\mathfrak{l}$ be the Lie algebra of L. Let $u = \exp \mathbf{u}$ for some eigenvector $\mathbf{u} \in \mathfrak{l}_{-\lambda_0}$ of $\mathrm{Ad}(a)$ for some positive λ_0. Note that $\mathfrak{l}$ contains the Lie algebra of U, so there is such an element. As before we know by the Jacobson-Morozov theorem that there exists an $\mathrm{sl}(2, \mathbb{K}_\sigma)$-triplet $\mathbf{u}$, $\mathbf{h}$, $\mathbf{n}$ consisting of eigenvectors for $\mathrm{Ad}(a)$. Then, $\mathrm{ad}(\mathbf{u})$ maps the eigenspace $\mathfrak{l}_\lambda$ into $\mathfrak{l}_{\lambda-\lambda_0}$ and the correct iterate gives for positive λ an injective map from $\mathfrak{l}_\lambda$ to $\mathfrak{l}_{-\lambda}$. By the symmetry of positive and negative weights in our assumptions, this shows that the map $\mathrm{ad}(\mathbf{u})^2$ restricted to $\mathfrak{l}_{\lambda_0}$ is surjective onto $\mathfrak{l}_{-\lambda_0}$. This shows that $\mathfrak{l}$ contains an $\mathrm{sl}(2, \mathbb{K}_\sigma)$-triplet consisting of eigenvectors for $\mathrm{Ad}(a)$. In fact, the above shows that there is some $\mathbf{n}' \in \mathfrak{l}_{\lambda_0}$ for which $\mathbf{u} = \mathrm{ad}(\mathbf{u})^2(\mathbf{n}')$, which implies that $\mathbf{n}'$, $\mathbf{h}' = \mathrm{ad}(\mathbf{u})(\mathbf{n}')$, and $\mathbf{u}$ is $\mathrm{sl}(2, \mathbb{K}_\sigma)$-triplet. (In fact, this triplet is equal to the orginal triplet considered above.) Just as in the proof of Lemma 8.1 this implies that $A < L$.

For the above claim it remains to show that L is simple. So suppose first that L has a unipotent radical, which by the above is normalized by A. So let $\mathbf{u}$ be an eigenvector for $\mathrm{Ad}(a)$ restricted to the Lie algebra of the radical. Clearly, the weight cannot be zero since the centralizer of A does not contain any unipotent subgroups in a semisimple group. So we may assume the weight is positive, but then the above shows that $\mathbf{u}$ is contained in an $\mathrm{sl}(2, \mathbb{K}_\sigma)$-triplet which is contained in $\mathfrak{l}$. This is a contradiction to $\mathbf{u}$ being in the Lie algebra of the unipotent radical of L. Finally note that since L is unipotently generated its radical equals the unipotent radical, therefore L is semisimple. If L would have several simple factors, then only one of them could be isotropic since G has $\mathbb{K}_\sigma$-rank one. However, this is impossible since L is generated by unipotent subgroups and these can only be contained in the isotropic factor.

Let now L° be as in the lemma. We first claim that L° is open, i.e. it contains a neighborhood of the identity. Let $U' = (u(t) : t \in \mathbb{K}_\sigma) < U$ be an A-normalized one dimensional unipotent subgroup. Note that a neighborhood of the identity in L is homeomorphic to a neighborhood of zero in $\mathbb{K}_\sigma^{\dim L}$ (via log and exp). We will show that it is possible to conjugate U' by elements of L° to obtain $\dim L$ many conjugates $U_j = (u_j(t) : t \in \mathbb{K}_\sigma)$ whose Lie algebras together span $\mathfrak{l}$. Assuming this, it follows that the map from $\mathbb{K}_\sigma^{\dim L}$ to L which is by $(t_1, \ldots, t_{\dim L}) \mapsto u_1(t_1) \cdots u_{\dim L}(t_{\dim L})$ has an invertible derivative at 0. In this case, the claim follows by the inverse function theorem.

Let $U_1 = U'$. We now construct the conjugates U_j for $j > 1$. Suppose we have already constructed $k < \dim L$ many such conjugates that have transverse Lie algebras. Then the span $\mathfrak{l}'$ of these Lie algebras is a subspace of $\mathfrak{l}$. If $\mathfrak{l}'$ is not invariant under both U and U^+, then we may conjugate some U_j by an element of U or U^+ to obtain an new conjugate U_{k+1} whose Lie algebra is transversal to the ones constructed earlier. Otherwise, $\mathfrak{l}'$ is normalized by L since L is the algebraic subgroup generated by U and U^+. Since the Lie bracket $[\cdot, \cdot]$ is obtained by taking the derivative of the adjoint representation, it follows that $\mathfrak{l}'$ is a normal Lie subalgebra of $\mathfrak{l}$. This is impossible unless $k = \dim L$ since L and, therefore also, $\mathfrak{l}$ are simple.

Note that in the case of $\mathbb{K}_\sigma = \mathbb{R}$ it follows that L° is precisely the connected component of L (in the Hausdorff topology). This proves the lemma in the real case. In the case of p-adic groups we have to use the structure theory of algebraic groups once more. Our first goal here is to show that $L^\circ \cap A$ has finite index in A. For this, note first that as before there exists an algebraic subgroup $H < L$ which is isomorphic to $\mathrm{SL}(2, \mathbb{K}_\sigma)$ or $\mathrm{PGL}(2, \mathbb{K}_\sigma)$ and contains both A and U'. It is easy to check for $\mathrm{SL}(2, \mathbb{K}_\sigma)$, and so also for H, that some $a' \in A$ with $|\chi(a')|_\sigma \neq 1$ can be written as $a' = u_1 u_1' u_2 u_2'$ with $u_1', u_2' \in U' = H^-$ and $u_1, u_2 \in H^+$. Conjugating this with some power of a' we may assume that $u_1, u_2 \in L^\circ$. However, this shows that $a' \in L^\circ$ since $U' \subset L^\circ$ by definition. Since A is a group extension of the compact group $\{a \in A : |\chi(a)| = 1\}$ by the integers and $L^\circ \cap A$ is open in A, it follows that $L^\circ \cap A$ has finite index in A.

Moreover, L° contains the stable horospherical subgroup L^-, since $L^\circ \cap L^-$ is open and since a' can be used to expand this to the whole of L^-. Let P be the parabolic subgroup of L containing L^-. Then P/L^- is again an extension of a compact group by the integers, and as with A we get that $L^\circ \cap P$ has finite index in P. Now let K be a good maximal compact open subgroup of L so that we have the Iwasawa decomposition $L = PK$. Then we have

$$L = PK \subset L^\circ p_1 K \cup \cdots \cup L^\circ p_n K$$

for some $p_1, \ldots, p_n \in P$. Here $L^\circ p_i k$ is open in $p_i K$, and so by compactness we deduce that L° has finite index in L.

Finally, we claim that L° is containing all one dimensional unipotent subgroups of L which implies also normality. For this, consider any one dimensional unipotent subgroup U''. Since L° is open, $L^\circ \cap U''$ is an open neighborhood of the identity in U''. Moreover, U'' is contained in an algebraic subgroup of L isomorphic to $\mathrm{SL}(2, \mathbb{K}_\sigma)$ or $\mathrm{PGL}(2, \mathbb{K}_\sigma)$. By finite index, there is some $a'' \in L^\circ$ which expands U'' which implies that $U'' \subset L^\circ$. □

8.2. **A_1-ergodic and invariant homogeneous measures on $X = \Gamma\backslash G_1 \times G_2$.** Assuming that G_1 has $\mathbb{K}_\sigma$-rank one, we will discuss here the structure of A_1-ergodic and invariant homogeneous measures. The following is a combination of Lemma 8.2 and of the classification of probability measures invariant under unipotent subgroups in the S-algebraic setting, see [Ra5] and [MT1].

Lemma 8.3. *Let G_1 be a $\mathbb{K}_\sigma$-rank one semisimple algebraic group defined over $\mathbb{K}_\sigma$ with split torus $A_1 < G_1$, let G_2 be any S-algebraic group defined over local fields of characteristic zero, and let $\Gamma < G_1 \times G_2$ be a discrete subgroup. Let ν be an A_1-ergodic and invariant probability measures on $X = \Gamma\backslash G_1 \times G_2$. Suppose ν is invariant under some nontrivial element of G_1^- and some nontrivial element of G_1^+. Then ν is a convex combination of homogeneous measures as in Theorem 1.5. If in addition $\mathbb{K}_\sigma$ equals either $\mathbb{R}$ or $\mathbb{Q}_p$ for some prime p, then this is a finite average.*

Proof. We may assume $\mathbb{K}_\sigma$ equals either $\mathbb{Q}_\infty = \mathbb{R}$ or $\mathbb{Q}_p$ for some prime p, for otherwise we may apply restriction of scalar to the group G_1. Let s be ∞ or p according to the case we are in. Note that in this restriction of scalar the $\mathbb{K}_\sigma$-split rank one torus A_1 becomes higher rank with $\mathbb{Q}_s$-rank one. Therefore, if we want to keep working with measures invariant and ergodic under the maximal split torus, we have to do an ergodic decomposition.

So suppose now $\mathbb{K}_\sigma = \mathbb{Q}_s$. Since μ is invariant under A_1 and some nontrivial element of G_1^- it follows easily that μ is also invariant under some A_1-normalized unipotent algebraic subgroup of G_1^- and similarly of G_1^+. We let U be the maximal such subgroup of G_1^- and similarly for $U^+ \subset G_1^+$.

Let L be the algebraic subgroup generated by U and U^+, and let $L^\circ < L$ be the actual subgroup generated by U and U^+ as in Lemma 8.2. By this lemma L contains A_1 and $L^\circ \cap A$ has finite index in A. Therefore, the A-ergodic L°-invariant probability measure ν can be decomposed into finitely many L°-ergodic and invariant measures ν_i. Since L° is generated by unipotents, we may apply the classification of measures invariant under such groups as in [Ra5] and [MT1]. Therefore, ν_i is homogeneous, i.e. there exists a closed subgroup $H < G_1 \times G_2$ such that ν_i is the unique H-invariant probability measure on a closed H-orbit. □

8.3. **Proof of Theorem 1.5 — first part.** In this subsection we study the leafwise measure for $G^- = G_1^-$. As a step towards the classification of the measure μ as in Theorem 1.5 we establish the invariance of the leafwise measures $\mu_x^{G^-}$ under left translations by a nontrivial subgroup of G^- depending on x. At this point the main obstacle for applying Theorem 1.4 is that we have to decompose the measure μ in order to have (U-1) satisfied for one and the same subgroup $U < G^-$. Condition (U-2) will follow from Lemma 8.1 for any nontrivial U.

Let $\phi : X \to Y = PM_\infty^*(G^-)$ denote the factor map $x \mapsto [\mu_x^{G^-}]$ as in Section 3, and let $\mathcal{Y}$ be the Borel σ-algebra on Y. Set $\nu = \phi_*\mu$ and let $\mathcal{E}_Y \subset \mathcal{Y}$ be a countably generated sub σ-algebra which is equivalent modulo ν to the σ-algebra of a-invariant subsets of Y. Finally, we let

$$\mathcal{E}_\phi = \phi^{-1}(\mathcal{E}_Y).$$

In these notations, $\nu = \int_Y \nu_y^{\mathcal{E}_Y}\, d\nu(y)$ gives the ergodic decomposition of ν (in particular, for a.e. y, the measure $\nu_y^{\mathcal{E}_Y}$ is ergodic under the action of a), and we have correspondingly a decomposition $\mu = \int_X \mu_x^{\mathcal{E}_\phi}\, d\mu(x)$; furthermore, the basic properties of decomposition of measures imply that

$$\phi_* \mu_x^{\mathcal{E}_\phi} = \nu_{\phi(x)}^{\mathcal{E}_Y} \qquad \text{for } \mu\text{-a.e. } x \in X.$$

For notational convenience, set $\mu^x = \mu_x^{\mathcal{E}_\phi}$.

We need to understand the relation between the leafwise measures along G^--orbits of μ and that of the probability measures μ^x. Concretely, we have for μ-a.e. x, for μ^x-a.e. y that

$$\mu_y^{G^-} = (\mu^x)_y^{G^-} \tag{8.1}$$

by Lemma 3.8.

By construction and the way the leafwise measures are defined (specifically, by the normalization in (LM-3)) $\mu_x^{G^-}$ is determined by $[\mu_x^{G^-}]$ and therefore the map $x \mapsto \mu_x^{G^-}$ is $\phi^{-1}(\mathcal{Y})$-measurable. (8.1) shows that $(\mu^x)_y^{G^-} = \mu_y^{G^-}$ for μ^x-a.e. y. Since $(\mu^x)_y^{G^-}$ is only defined up to a set of μ_x measure zero we may choose it so that μ-a.s. $(\mu^x)_y^{G^-} = \mu_y^{G^-}$ for *all* y. Therefore, for a.e. x, the assumption (Φ) of Theorem 1.4 is satisfied by ϕ and μ^x.

We let P_y be the subgroup in Proposition 6.2 applied with $U = G^-$. Since P_y is normalized by a, it is clearly constant on the ergodic components for a. In fact, it depends (measurably) only on the ergodic component of $\Phi(y)$ and so it is constant a.e. (and equal to P_x) with respect to μ^x for a.e. x.

We wish to apply Theorem 1.4 for μ^x and $U = P_x$ for μ-a.e. x. Then (U-1) is satisfied by definition of P_x and (U-2) follows from Lemma 8.1 whenever P_x or equivalently $\mu_x^{G^-}$ is nontrival. By Proposition 3.9 and our assumption (1) in Theorem 1.5 this holds μ-a.e.

It follows that μ^x and ϕ satisfy the conditions of Theorem 1.4 for a.e. x, and hence for a.e. x we can decompose μ^x into two measures μ_1^x and μ_2^x such that

(LE-1_x) ϕ is locally $C_G(U_x) \cap G^0$-aligned modulo μ_1^x, and
(LE-2_x) μ_2^x is $C_G(U_x) \cap G^-$-recurrent relative to ϕ.

We prove next that actually more is true.

Lemma 8.4. *In the above notation, we have for a.e. x that either $\mu^x = \mu_1^x$ or μ^x satisfies (LE-2_x). Moreover, in the latter case the measure μ^x is invariant under a nontrivial unipotent element of G^-.*

Proof. Let for simplicity of notation $\nu = \mu^x$ be a measure for which $\phi_*\nu$ is a-ergodic and ν allows a decomposition into parts ν_1 and ν_2 for which (LE-1) resp. (LE-2) holds. Moreover, we may also assume that $[\nu_y^{G^-}] = \phi(y)$. Recall that by construction of μ^x and the discussion immediately after (8.1) we know this happens a.e.

Suppose now that $\nu \neq \nu_1$, i.e. ν_2 is appearing nontrivially in the decomposition. By (LE-2) and (LM-2) this shows that there is a set X_2 of positive ν measure such that for $y \in X_2$ there exists some $u_y \in G^-$ with $\nu_y^{G^-}$ and $\nu_y^{G^-} u_y$ are proportional. By Lemma 3.10 this shows that $\nu_y^{G^-} = \nu_y^{G^-} u_y$ for a.e. $y \in X_2$. By [EK2, Prop. 6.2] this shows, again for a.e. $y \in X_2$, that there is an a-normalized subgroup H_y containing u_y (or any other elements with the same property) which preserves $\nu_y^{G^-}$ under the right action. Since this is an a-invariant property of $\phi_*\nu$ only, we conclude by ergodicity that this property holds a.e. with the same H. Moreover, we may assume that H is for a.e. y the maximal subgroup preserving $\nu_y^{G^-}$ on the right. We claim this implies that ν is H-invariant.

Assume (LM-2) for y and $u.y$ and that $\nu_y^{G^-}$ and $\nu_{u.y}^{G^-}$ are invariant on the right under H for some $u \in G^-$ and $y \in X$, which happens for a.e. y and $\nu_y^{G^-}$-a.e. u. Then by (LM-2) we have that $\nu_y^{G^-}$ is invariant under the right action of $u^{-1}Hu$. By maximality of H this shows that for a.e. y the subgroup H is normalized by any $u \in \operatorname{supp}(\nu_y^{G^-}) \subset G^-$. This can be used to show that $\nu_y^{G^-}$ is actually invariant under left multiplication by elements of H too, c.f. [EK2, Lemma 8.11]. The claim follows now from the proof of Lemma 3.1, i.e. [Lin2, Prop. 4.3]. In that lemma it was assumed that the leafwise measure for a subgroup U is the left Haar measure of U and obtained invariance under the full group U, if instead one knows only left invariance of the leafwise measures for G^- under one and the same subgroup $H < G^-$ one obtains invariance under H. □

We will show in the next section that μ^x satisfies (LE-2_x) a.e. This implies Theorem 1.5. More precisely, in this case we get that μ^x is invariant under a nontrivial element of G^- by Lemma 8.4. This implies that the ergodic components of μ^x are invariant under a nontrivial element of G^-. Since the ergodic components of μ can be obtained from the ergodic components of μ^x for a.e. x, it follows that a.e. ergodic component is invariant under a nontrivial element of G^-. By symmetry the same holds for G^+. Therefore, a.e. ergodic component of μ satisfies the assumptions of Lemma 8.3 and the theorem follows.

8.4. **Proof of Theorem 1.5 — (LE-2_x) holds for μ^x a.e.** To complete the proof of Theorem 1.5 we need to show that a.e. (LE-1_x) does not hold for μ^x. For this we will need the product lemma, which is one of the main ingredients of the high entropy method of [EK1]; but we need the more general [Lin2, Prop 6.4] or [EK2, Thm 7.5]. Since conjugation by a contracts G^- while acting trivially on G_2 it implies that there is a set $X' \subset X$ of full measure w.r.t. μ on which

$$\mu_x^{G^- \times G_2} = \mu_x^{G^-} \times \mu_x^{G_2},$$

hence using the property (LM-2) of the construction of leafwise measures there is a set of full measure (which we also denote by X') on which

$$\mu_x^{G^-} = \mu_y^{G^-} \qquad \text{for every } x, y \in X' \text{ with } y \in G_2.x; \tag{8.2}$$

since $\phi(x) = [\mu_x^{G^-}]$ this implies that ϕ is constant on the intersection of X' with G_2-orbits in X. Therefore, elements of $\mathcal{E}_\phi$ consist (modulo μ) of unions of complete G_2-orbits, and so the leafwise measures for G_2-orbits of μ and of μ^x are equal in the sense that for μ-a.e. x and μ^x-a.e. y we have $\mu_y^{G_2} = (\mu^x)_y^{G_2}$. In view of Lemma 3.1.(1), and since we have assumed that μ is G_2-recurrent, this equality implies that for a.e. x the measure μ^x is G_2-recurrent.

To analyze the set where (LE-1_x) holds and its connection to the assumption (2) in Theorem 1.5 we need another lemma regarding algebraic groups. Recall that $M_1 = \{h \in G_1^0 : \chi(h) = 1\}$.

Lemma 8.5. *For any nontrivial a-normalized $U \le G^-$, the group $M_1 \cap C_G(U)$ has finite index in $G_1^0 \cap C_G(U)$.*

Proof. Let $\mathbf{u}_1, \dots, \mathbf{u}_\ell$ be a basis for the Lie algebra of U consisting of eigenvalues for the action of A_1. Then $C_{G_1}(U) = C_{G_1}(\mathbf{u}_1) \cap \dots \cap C_{G_1}(\mathbf{u}_\ell)$.

Let $L = G_1^0 \cap N_G(\mathbb{K}_\sigma \mathbf{u}_1) \cap \dots \cap N_G(\mathbb{K}_\sigma \mathbf{u}_\ell)$. This is a reductive group (G_1^0 and so also L have no unipotents, hence no unipotent radical) containing A_1. For every $g \in L$, $\mathbf{u}_i$ is an eigenvector of $\operatorname{Ad} g$, and the corresponding eigenvalue gives a $\mathbb{K}_\sigma$-character χ_i on L. The group M_1 was defined as the kernel of a $\mathbb{K}_\sigma$-character χ on $C_{G_1}(A_1)$. Since L has $\mathbb{K}_\sigma$-rank one, the characters $\chi|_L$ and χ_i for $i = 1, \dots, \ell$ are rationally related, hence the intersection $M_1 \cap C_G(U)$ of the kernels of all of these characters and the intersection $G_1^0 \cap C_{G_1}(U)$ of the kernels of χ_i for $i = 1, \dots, \ell$ are commensurable. □

Claim: For μ-a.e. x, (LE-1_x) does not hold for μ^x.

Proof of this claim. Indeed, suppose in contradiction that there is some set Y with $\mu(Y) > 0$ on which (LE-1_x) holds for μ^x. By assumption (2) of Theorem 1.5 we have that for a.e. x the group

$$\operatorname{Stab}_{M_1 \times G_2}(x) := \{h \in M_1 \times G_2 : h.x = x\}$$

is finite. Another assumption is that μ is G_2-recurrent, which as discussed implies that for μ-a.e. x, the measure μ^x is G_2-recurrent. Therefore, if indeed $\mu(Y) > 0$, there will be some x for which:

(1) μ^x is G_2-recurrent;
(2) ϕ is locally $C_G(U_x) \cap G^0$-aligned modulo μ^x;

(3) there is a set $X' \subset X$ of full μ^x-measure so that $\phi(y') = \phi(y)$ for any $y', y \in X'$ on the same orbit of G_2;

(4) for μ^x-a.e. y, say again for all $y \in X'$, we have that $|\mathrm{Stab}_{M_1 \times G_2}(y)| < \infty$.

Note that (3) is a consequence of the product lemma; cf. (8.2).

These four properties, however, are not compatible. Indeed, let $X_1 \subset X'$, $\delta > 0$ and $O \subset C_G(U_x) \cap G^0$ be as in (1.1) of Definition 1.1 applied to the measure μ^x and the group $C_G(U_x) \cap G^0$ with $\mu^x(X_1) > 1/2$. By recurrence of μ^x there will be some $y \in X_1$ and an unbounded sequence $h_i \in G_2$ such that for every i

$$h_i.y \in X_1 \cap B_\delta(y).$$

By (3) we have also that $\phi(h_i.y) = \phi(y)$ for every i. Applying (1.1), we conclude that $h_i.y \in O.y$. Since $O \subset C_G(U_x) \cap G^0$ has compact closure, since $h_i \in G_2 \subset C_G(U_x) \cap G^0$, and since $h_i \to \infty$ we have

$$\left|\mathrm{Stab}_{C_G(U_x) \cap G^0}(y)\right| = \infty. \tag{8.3}$$

Observe that $C_G(U_x) \cap G^0 = \left(C_{G_1}(U_x) \cap G_1^0\right) \times G_2$, so by Lemma 8.5 we have that $M_1 \times G_2$ contains a finite index subgroup of $C_G(U_x) \cap G^0$. The infinitude of the stabilizer group in (8.3) thus implies that $|\mathrm{Stab}_{M_1 \times G_2}(y)| = \infty$ — in contradiction to (4). Thus our claim has been proved. □

As discussed in the last subsection, the above was the remaining step of the proof of Theorem 1.5.

8.5. **Short discussion about assumption (2) in Theorem 1.5.** We end this paper by noting that assumption (2) is absolute necessary in order to give a meaningful description of invariant measures.

To see this, note first that $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{SO}(3)$ embeds into $G_1 = \mathrm{SO}(5, 1)$ and that the latter has real rank one. If we let $G_2 = \mathrm{SO}(2, 1)$ we see that $G = G_1 \times G_2$ contains the subgroup $\mathrm{PSL}(2, \mathbb{R}) \times \mathrm{SO}(3) \times \mathrm{SO}(2, 1)$. We define Γ to be the discrete subgroup which is contained in the latter and equals $\mathrm{PSL}(2, \mathbb{Z}) \times \Gamma'$ where Γ' is an irreducible lattice in $\mathrm{SO}(3) \times \mathrm{SO}(2, 1)$.

If now μ equals any product of an A_1-ergodic and invariant probability measure on $\mathrm{SL}(2, \mathbb{Z})\backslash\mathrm{SL}(2, \mathbb{R})$ and e.g. the Haar measure on $\Gamma'\backslash\mathrm{SO}(3) \times \mathrm{SO}(2, 1)$, then μ is only a linear combination of A_1-invariant homogeneous measure if the original measure on $\mathrm{SL}(2, \mathbb{Z})\backslash\mathrm{SL}(2, \mathbb{R})$ was homogeneous. However, as is well known there is an abundance of non-homogeneous probability measure that are invariant and ergodic with respect to the geodesic flow for the modular surface, even with positive entropy. Here we have G_2-recurrence but in a way in which the assumption in Theorem 1.5 (2) is not satisfied.

References

[EK1] M. Einsiedler and A. Katok. Invariant measures on G/Γ for split simple Lie-groups G. *Comm. Pure Appl. Math.* **56**(2003), 1184–1221.

[EK2] M. Einsiedler and A. Katok. Rigidity of measures—the high entropy case and non-commuting foliations. *Israel J. Math.* **148**(2005), 169–238. Probability in mathematics.

[EKL] M. Einsiedler, A. Katok, and E. Lindenstrauss. Invariant measures and the set of exceptions to Littlewood's conjecture. to appear in Ann. of Math.

[EKl] Manfred Einsiedler and Dmitry Kleinbock. Measure rigidity and p-adic Littlewood-type problems. preprint (17 pages), 2005.

[EL1] Manfred Einsiedler and Elon Lindenstrauss. Joining of higher rank diagonalizable actions on locally homogeneous spaces. to appear in Duke Math. J. (29 pages), 2005.

[EL2] Manfred Einsiedler and Elon Lindenstrauss. Diagonal flows on locally homogeneous spaces and number theory. to appear in the Proceedings of the International Congress of Mathematicians 2006 (29 pages), 2006.

[EL3] Manfred Einsiedler and Elon Lindenstrauss. Diagonal actions on locally homogeneous spaces. lecture notes for CMI summer school Homogeneous Flows, Moduli Spaces and Arithmetic, in preparation, 2007.

[KaS] Boris Kalinin and Ralf Spatzier. Rigidity of the measurable structure for algebraic actions of higher-rank Abelian groups. *Ergodic Theory Dynam. Systems* **25**(2005), 175–200.

[KS] A. Katok and R. J. Spatzier. Invariant measures for higher-rank hyperbolic abelian actions. *Ergodic Theory Dynam. Systems* **16**(1996), 751–778.

[LY1] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula. *Ann. of Math. (2)* **122**(1985), 509–539.

[LY2] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension. *Ann. of Math. (2)* **122**(1985), 540–574.

[LS] François Ledrappier and Jean-Marie Strelcyn. A proof of the estimation from below in Pesin's entropy formula. *Ergodic Theory Dynam. Systems* **2**(1982), 203–219 (1983).

[Lin1] Elon Lindenstrauss. Arithmetic quantum unique ergodicity and Adelic dynamics. to appear in the proceedings of Current Developments in Mathematics conference, Harvard 2004 (30 pages), 2005.

[Lin2] Elon Lindenstrauss. Invariant measures and arithmetic quantum unique ergodicity. *Ann. of Math. (2)* **163**(2006), 165–219.

[Mar] G. A. Margulis. The action of unipotent groups in a lattice space. *Mat. Sb. (N.S.)* **86(128)**(1971), 552–556.

[MT1] G. A. Margulis and G. M. Tomanov. Invariant measures for actions of unipotent groups over local fields on homogeneous spaces. *Invent. Math.* **116**(1994), 347–392.

[MT2] G. A. Margulis and G. M. Tomanov. Measure rigidity for almost linear groups and its applications. *J. Anal. Math.* **69**(1996), 25–54.

[PR] Vladimir Platonov and Andrei Rapinchuk. *Algebraic groups and number theory*, volume 139 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1994. Translated from the 1991 Russian original by Rachel Rowen.

[Ra1] Marina Ratner. Factors of horocycle flows. *Ergodic Theory Dynam. Systems* **2**(1982), 465–489.

[Ra2] Marina Ratner. Rigidity of horocycle flows. *Ann. of Math. (2)* **115**(1982), 597–614.

[Ra3] Marina Ratner. Horocycle flows, joinings and rigidity of products. *Ann. of Math. (2)* **118**(1983), 277–313.

[Ra4] Marina Ratner. On Raghunathan's measure conjecture. *Ann. of Math. (2)* **134**(1991), 545–607.

[Ra5] Marina Ratner. Raghunathan's conjectures for Cartesian products of real and p-adic Lie groups. *Duke Math. J.* **77**(1995), 275–382.

[Ree] M. Rees. Some R^2-anosov flows. 1982.

[Rud1] Daniel J. Rudolph. Ergodic behaviour of Sullivan's geometric measure on a geometrically finite hyperbolic manifold. *Ergodic Theory Dynam. Systems* **2**(1982), 491–512 (1983).

[Rud2] Daniel J. Rudolph. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergodic Theory Dynam. Systems* **10**(1990), 395–406.

[Mor] Dave Witte. Rigidity of some translations on homogeneous spaces. *Invent. Math.* **81**(1985), 1–27.