Clay Mathematics Proceedings
Volume 8, 2008

Diagonal actions on locally homogeneous spaces

M. Einsiedler and E. Lindenstrauss

Contents

1. Introduction

1.1. In these notes we present some aspects of work we have conducted, in parts jointly with Anatole Katok, regarding dynamics of higher rank diagonalizable groups on (locally) homogeneous spaces[(1)] $\Gamma\backslash G$. A prototypical example of such an action is the action of the group of determinant one diagonal matrices A on the space of lattices in $\mathbb{R}^n$ with covolume one for $n \geq 3$ which can be identified with the quotient space $\mathrm{SL}(n,\mathbb{Z})\backslash\mathrm{SL}(n,\mathbb{R})$. More specifically, we consider the problem of classifying measures invariant under such an action, and present two of the applications of this measure classification.

[(1)]The space $X = \Gamma\backslash G$ we define is in fact a homogeneous space for the group G in the abstract sense of algebra but if we also consider the metric structure, see §7.1, the phrase "locally homogeneous" seems more appropriate.

There have been several surveys on this topic, including some that we have written (specifically, [**Lin05**] and [**EL06**]). For this reason we will be brief in our historical discussions and the discussion of the important work of the pioneers of the subject.

1.2. For the more general setup let $G = \mathbb{G}(\mathbb{R})$ be the group of $\mathbb{R}$-points of a linear algebraic group over $\mathbb{R}$, and let $\Gamma < G$ be a lattice (i.e., a discrete, finite covolume subgroup). In this setup it is natural to consider for any subgroup $H < G$, in particular for any algebraic subgroup, the action of H on the symmetric space $\Gamma\backslash G$. Ratner's landmark measure classification theorem (which is somewhat more general as it considers the case of G a general Lie group) states the following:

1.3. Theorem (M. Ratner [**Rt91**])**.** *Let G, Γ be as above, and let $H < G$ be an algebraic subgroup generated by one parameter unipotent subgroups. Then any H-invariant and ergodic probability measure μ is the natural (i.e., L-invariant) probability measure on a single orbit of some closed subgroup $L < G$ ($L = G$ is allowed).*

We shall call a probability measure of the type above (i.e., supported on a single orbit of its stabilizer group) *homogeneous*.

1.4. For one parameter diagonalizable flows the (partial) hyperbolicity of the flow guarantees the existence of many invariant measures. It is, however, not unreasonable to hope that for multiparameter diagonalizable flows the situation is better. For example one has the following conjecture attributed to Furstenberg, Katok-Spatzier and Margulis:

1.5. Conjecture. *Let A be the group of diagonal matrices in $\mathrm{SL}(n,\mathbb{R})$, $n \geq 3$. Then any A-invariant and ergodic probability measure on $\mathrm{SL}(n,\mathbb{Z})\backslash \mathrm{SL}(n,\mathbb{R})$ is homogeneous.*

The reader may note that we have phrased Conjecture 1.5 in much more specialized way than Theorem 1.3. While the basic phenomena behind the conjecture is expected to be quite general, care must be exercised when stating it more generally (even for the groups A and G given above).

1.6. Conjecture 1.5 is quite open. But progress has been made. Specifically, in our joint paper with Katok [**EKL06**], Conjecture 1.5 is proved under the condition that there is some $a \in A$ with positive entropy (see Theorem 11.5 below for a more formal statement).

1.7. These lecture notes are based on our joint course given in the CMI Pisa summer school as well as a graduate course given by the second named author in Princeton the previous semester. Notes for both were carefully taken by Shimon Brooks and thoroughly edited by us. The material presented here has almost entirely been published in several research papers, in particular [**EK03, Lin06, EK05, EKL06, EL08**].

1.8. The treatment here differs from the original treatment in places, hopefully for the better. In particular, we use this opportunity to give an alternative simplified treatment to the high entropy method developed by M.E. and Katok in [**EK03, EK05**]. For this reason our treatment of the high entropy method in §9 is much more careful and thorough than our treatment of the low entropy method in the following section (the reader who wishes to learn this technique in greater detail is advised to look at our recent paper [**EL08**]).

It is interesting to note that what we call the low entropy method for studying measures invariant under diagonalizable groups uses heavily unipotent dynamics, and, in particular, ideas of Ratner developed in her study of isomorphism and joining rigidity in [**Rt82b, Rt82a, Rt83**] which was a precursor to her more general results on unipotent flows in [**Rt90, Rt91**].

1.9**.** More generally, the amount of detail given on the various topics is not uniform. Our treatment of the basic machinery of leafwise measures as well as entropy in §3-7 is very thorough as are the next two sections §8-9. This has some correlation to the material given in the Princeton graduate course, though the presentation of the high entropy method given here is more elaborate.

The last two sections of these notes give a sample of some of the applications of the measure classification results given in earlier chapters. We have chosen to present only two: our result with Katok on the set of exceptions to Littlewood's Conjecture from [**EKL06**] and the result of E.L. on Arithmetic Quantum Unique Ergodicity from [**Lin06**]. The measure classification results presented here also have other applications; in particular we mention our joint work with P. Michel and A. Venkatesh on the distribution properties of periodic torus orbits [**ELMV06, ELMV07**].

1.10**.** One day a more definitive and complete treatment of these measure rigidity results would be written, perhaps by us. Until that day we hope that these notes, despite their obvious shortcomings, might be useful.

Acknowledgements. This work owes a debt to the Clay Mathematical Institute in more than one way. We thank CMI for its support of both of us (E.L. was supported by CMI during the years 2003-2005, and M.E. was supported by CMI in the second half of 2005). Many of the ideas we present here were developed during this period. We also thank CMI for the opportunity it provided us to present our work to a wide and stimulating audience in the Pisa summer school. We also thank Shimon Brooks for his careful notetaking. Finally we thank Shirali Kadyrov, Beverly Lytle, Fabrizio Polo, Alex Ustian, and in particular Uri Shapira for comments on the manuscript. The work presented here has been obtained over several years and supported by several NSF grants, in particular DMS grants 0554373, 0622397 (ME), 0500205 and 0554345 (EL).

2. Ergodic theory: some background

We start by summarizing a few basic notions of ergodic theory, and refer the reader with the desire to see more details to any book on ergodic theory, e.g. [**Wal82**], [**Gla03**], or [**EW09a**].

2.1. Definition. *Let X be a locally compact space, equipped with an action of a noncompact (but locally compact) group*[(2)] *H which we denote by $(h,x) \mapsto h.x$ for $h \in H$ and $x \in X$. An H-invariant probability measure μ on X is said to be* ergodic *if one of the following equivalent conditions holds:*

[(2)]All groups will be assumed to be second countable locally compact, all measures Borel probability measures unless otherwise specified.

(i) *Suppose $Y \subset X$ is an H-invariant set, i.e., $h.Y = Y$ for every $h \in H$. Then $\mu(Y) = 0$ or $\mu(X \setminus Y) = 0$.*
(ii) *Suppose f is a measurable function on X with the property that for every $h \in H$, for μ-a.e. x, $f(h.x) = f(x)$. Then f is constant a.e.*
(iii) *μ is an extreme point of the convex set of all H-invariant Borel probability measures on X.*

2.2. A stronger condition which implies ergodicity is mixing:

2.3. Definition. *Let X, H and μ be as in Definition 2.1. The action of H is said to be mixing if for any sequence $h_i \to \infty$ in H[3] and any measurable subsets $B, C \subset X$,*

$$\mu(B \cap h_i.C) \to \mu(B)\mu(C) \qquad \textit{as } i \to \infty.$$

Recall that two sets B, C in a probability space are called independent if $\mu(B \cap C) = \mu(B)\mu(C)$. So mixing is asking for two sets to be asymptotically independent (when one of the sets is moved by bigger and bigger elements of H).

2.4. A basic fact about H-invariant measures is that any H-invariant measure is an average of ergodic measures, i.e., there is some auxiliary probability space (Ξ, ν) and a (measurable) map attaching to each $\xi \in \Xi$ an H-invariant and ergodic probability measure μ_ξ on X so that

$$\mu = \int_\Xi \mu_\xi d\nu(\xi).$$

This is a special case of Choquet's theorem on representing points in a compact convex set as generalized convex combinations of extremal points.

2.6. Definition. *An action of a group H on a locally compact topological space X is said to be uniquely ergodic if there is only one H-invariant probability measure on X.*

2.7. The simplest example of a uniquely ergodic transformation is the map $T_\alpha : x \mapsto x + \alpha$ on the one dimensional torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ where α is irrational. Clearly Lebesgue measure m on $\mathbb{T}$ is T_α-invariant; we need to show it is the only such probability measure.

To prove this, let μ be an arbitrary T_α-invariant probability measure. Since μ is T_α-invariant,

$$\hat{\mu}(n) = \int_{\mathbb{T}} e(nx) d\mu(x) = \int_{\mathbb{T}} e(n(x + \alpha)) d\mu(x) = e(n\alpha)\hat{\mu}(n),$$

where as usual $e(x) = \exp(2\pi i x)$. Since α is irrational, $e(n\alpha) \neq 1$ for all $n \neq 0$, hence $\hat{\mu}(n) = 0$ for all $n \neq 0$ and clearly $\hat{\mu}(0) = 1$. Since the functions $e(nx)$ span a dense subalgebra of the space of continuous functions on $\mathbb{T}$ we have $\mu = m$.

2.9. Definition. *Let X be a locally compact space, and suppose that $H = \{h_t\} \cong \mathbb{R}$ acts continuously on X. Let μ be an H-invariant measure on X. We say that $x \in X$*

[3] I.e., a sequence so that for any compact $K \subset H$ only finitely many of the h_i are in K.

is generic for μ *if for every* $f \in C_0(X)$ *we have*[(4)]*:*

$$\frac{1}{T}\int_0^T f(h_t.x)\,dt \to \int_X f(y)\,d\mu(y) \qquad \text{as } T \to \infty.$$

Equidistribution is another closely related notion:

2.11. Definition. *A sequence of probability measures* μ_n *on a locally compact space* X *is said to be* equidistributed *with respect to a (usually implicit) measure* m *if they converge to* m *in the weak* topology, i.e., if* $\int f\,d\mu_n \to \int f\,dm$ *for every* $f \in C_0(X)$*.*

A sequence of points $\{x_n\}$ *in* X *is said to be equidistributed if the sequence of probability measures* $\mu_N = N^{-1}\sum_{n=1}^{N}\delta_{x_n}$ *is equidistributed, i.e., if for every* $f \in C_0(X)$

$$\frac{1}{N}\sum_{n=1}^{N} f(x_n) \to \int_X f(y)\,dm(y) \qquad \text{as } N \to \infty.$$

Clearly there is a lot of overlap between the two definitions, and in many situations" equidistributed" and "generic" can be used interchangeably.

2.12. For an arbitrary $H \cong \mathbb{R}$-invariant measure μ on X, the Birkhoff pointwise ergodic theorem shows that μ-almost every point $x \in X$ is generic with respect to some H-invariant and ergodic probability measure on X. If μ is ergodic, μ-a.e. $x \in X$ is generic for μ.

If X is compact, and if the action of $H \cong \mathbb{R}$ on X is uniquely ergodic with μ being the unique H-invariant measure, then something much stronger is true: *every* $x \in X$ is generic for μ!

Indeed, let μ_T denote the probability measure

$$\mu_T = \frac{1}{T}\int_0^T \delta_{h_t.x}\,dt$$

for any $T > 0$. Then any weak* limit of μ_T as $T \to \infty$ will be H-invariant. However, there is only one H-invariant probability measure[(5)] on X, namely μ, so $\mu_T \to \mu$, i.e., x is generic for μ.

E.g. for the irrational rotation considered in §2.7 it follows that orbits are equidistributed. A more interesting example is provided by the horocycle flow on *compact* quotients $\Gamma\backslash \mathrm{SL}(2,\mathbb{R})$. The unique ergodicity of this system is a theorem due to Furstenberg [**Fur73**] and is covered in the lecture notes [**Esk**] by Eskin.

3. Entropy of dynamical systems: some more background

3.1. A very basic and important invariant in ergodic theory is entropy. It can be defined for any action of a (not too pathological) unimodular amenable group H preserving a probability measure [**OW87**], but for our purposes we will only need (and only consider) the case $H \cong \mathbb{R}$ or $H \cong \mathbb{Z}$. For more details we again refer to [**Wal82**], [**Gla03**], or [**EW09b**].

---

(4)Where $C_0(X)$ denotes the space of continuous functions on X which decay at infinity, i.e., so that for any $\epsilon > 0$ the set $\{x : |f(x)| \geq \epsilon\}$ is compact.

(5)This uses that X is compact. If X is non-compact, one would have to address the possibility of the limit not being a probability measure. This possibility is often described as *escape of mass*.

Entropy is an important tool also in the study of unipotent flows[(6)], but plays a much more prominent role in the study of diagonalizable actions which we will consider in these notes.

3.2. Let (X, μ) be a probability space. The static entropy $H_\mu(\mathcal{P})$ of a finite or countable partition $\mathcal{P}$ of X is defined to be

$$H_\mu(\mathcal{P}) = -\sum_{P \in \mathcal{P}} \mu(P) \log \mu(P).$$

which in the case where $\mathcal{P}$ is countable may be finite or infinite.

One basic property of entropy is sub-additivity; the entropy of the refinement $\mathcal{P} \vee \mathcal{Q} = \{P \cap Q : P \in \mathcal{P}, Q \in \mathcal{Q}\}$ satisfies

$$H_\mu(\mathcal{P} \vee \mathcal{Q}) \leq H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q}). \tag{3.2a}$$

However, this is just a starting point for many more natural identities and properties of entropy, e.g. equality holds in (3.2a) if and only if $\mathcal{P}$ and $\mathcal{Q}$ are independent, the latter means that any element of $\mathcal{P}$ is independent of any element of $\mathcal{Q}$. All these natural properties find good explanations if one interprets $H_\mu(\mathcal{P})$ as the average of the information function

$$I_\mu(\mathcal{P})(x) = -\log \mu(P) \text{ for } x \in P \in \mathcal{P}$$

which measures the amount of information revealed about x if one is given the partition element $P \in \mathcal{P}$ that contains $x \in P$.

3.3. The ergodic theoretic entropy $h_\mu(T)$ associated to a measure preserving map $T : X \to X$ can be defined using the entropy function H_μ as follows:

3.4. Definition. *Let μ be a probability measure on X and $T : X \to X$ a measurable map preserving μ. Let $\mathcal{P}$ be either a finite or a countable[(7)] partition of X with $H_\mu(\mathcal{P}) < \infty$. The entropy of the four-tuple $(X, \mu, T, \mathcal{P})$ is defined to be[(8)]*

$$h_\mu(T, \mathcal{P}) = \lim_{N \to \infty} \frac{1}{N} H_\mu \left(\bigvee_{n=0}^{N-1} T^{-n} \mathcal{P} \right). \tag{3.4a}$$

The ergodic theoretic entropy of (X, μ, T) is defined to be

$$h_\mu(T) = \sup_{\mathcal{P} : H_\mu(\mathcal{P}) < \infty} h_\mu(T, \mathcal{P}).$$

The ergodic theoretic entropy was introduced by A. Kolmogorov and Ya. Sinai and is often called the Kolmogorov-Sinai entropy.[(9)] We may interpret the entropy $h_\mu(T)$ as a measure of the complexity of the transformation with respect to the measure μ. We will discuss this in greater detail later, but the geodesic flow has positive entropy with respect to the Haar measure on $\Gamma \backslash \mathrm{SL}(2, \mathbb{R})$ while the horocycle flow has zero entropy. However, vanishing entropy does not mean that the

---

(6) In particular, in [**MT94**] Margulis and Tomanov give a shorter proof of Ratner's measure classification theorem using entropy theory.

(7) One may also restrict oneself to finite partitions without changing the outcome, but we will see situations where it will be convenient to allow countable partitions.

(8) Note that by the subadditivity of the entropy function H_μ the limit in (3.4a) exists and is equal to $\inf_N \frac{1}{N} H_\mu(\bigvee_{n=0}^{N-1} T^{-n} \mathcal{P})$.

(9) Ergodic theoretic entropy is also somewhat confusingly called the metric entropy (even though it has nothing to do with any metric that might be defined on X!).

dynamics of the transformation or the flow is simple, e.g. the horocycle flow is mixing with respect to the Haar measure on $\Gamma\backslash\operatorname{SL}(2,\mathbb{R})$. Also, one can find quite complicated measures μ on $\Gamma\backslash\operatorname{SL}(2,\mathbb{R})$ that are invariant under the geodesic flow and with respect to which the geodesic flow has zero entropy.

3.5. If μ is a T-invariant but not necessarily ergodic measure, it can be shown that the entropy of μ is the average of the entropies of its ergodic components: i.e., if μ has the ergodic decomposition $\mu = \int \mu_\xi d\nu(\xi)$, then

$$h_\mu(T) = \int h_{\mu_\xi}(T) d\nu(\xi). \tag{3.5a}$$

Therefore, it follows that an invariant measure with positive entropy has in its ergodic decomposition a *positive fraction* of ergodic measures with positive entropy.

3.6. We will see in §7 concrete formulas and estimates for the entropy of flows on locally homogeneous spaces $\Gamma\backslash G$. To obtain these the main tool is the following notion: A partition $\mathcal{P}$ is said to be a *generating partition* for T and μ if the σ-algebra $\bigvee_{n=-\infty}^{\infty} T^{-n}\mathcal{P}$ (i.e., the σ-algebra generated by the sets $\{T^n P : n \in \mathbb{Z}, P \in \mathcal{P}\}$) separates points; that is, for μ-almost every x, the atom of x with respect to this σ-algebra is $\{x\}$.[10] The Kolmogorov-Sinai theorem asserts the non-obvious fact that $h_\mu(T) = h_\mu(T, \mathcal{P})$ whenever $\mathcal{P}$ is a generating partition.

3.7. We have already indicated that we will be interested in the entropy of flows. So we need to define the ergodic theoretic entropy for flows (i.e., for actions of groups $H \cong \mathbb{R}$). Suppose $H = \{a_t\}$ is a one parameter group acting on X. Then it can be shown that for $s \neq 0$, $\frac{1}{|s|} h_\mu(x \mapsto a_s.x)$ is independent of s. We define the entropy of μ with respect to $\{a_t\}$, denoted $h_\mu(a_\bullet)$, to be this common value of $\frac{1}{|s|} h_\mu(x \mapsto a_s.x)$.[11]

3.8. Suppose now that (X, d) is a compact metric space, and that $T : X \to X$ is a homeomorphism (the pair (X, T) is often implicitly identified with the generated $\mathbb{Z}$-action and is called a topological dynamical system).

3.9. Definition. *The $\mathbb{Z}$-action on X generated by T is said to be* expansive *if there is some $\delta > 0$ so that for every $x \neq y \in X$ there is some $n \in \mathbb{Z}$ so that $d(T^n x, T^n y) > \delta$.*

If X is expansive then any measurable partition $\mathcal{P}$ of X for which the diameter of every element of the partition is $< \delta$ is generating (with respect to any measure μ) in the sense of §3.6.

3.10. Problem. *Let A be a $d \times d$ integer matrix with determinant* 1 *or* -1. *Then A defines a dynamical system on $X = \mathbb{R}^n/\mathbb{Z}^n$. Characterize when A is expansive with respect to the metric derived from the Euclidean metric on $\mathbb{R}^n$. Also determine whether an element of the geodesic flow on a compact quotient* $\Gamma\backslash\operatorname{SL}(2,\mathbb{R})$ *is expansive.*

---

(10) Recall that the atom of x with respect to a countably generated σ-algebra $\mathcal{A}$ is the intersection of all $B \in \mathcal{A}$ containing x and is denoted by $[x]_{\mathcal{A}}$. We will discuss that and related notions in greater detail in §5.

(11) Note that $h_\mu(a_\bullet)$ depends not only on H as a group but on the particular parametrization a_t.

3.11. For some applications presented later, an important fact is that for many dynamical systems (X,T) the map $\mu \mapsto h_\mu(T)$ defined on the space of T-invariant probability measures on X is semicontinuous. This phenomenon is easiest to see when (X,T) is expansive.

3.12. Proposition. *Suppose (X,T) is expansive, and that μ_i, μ are T-invariant probability measures on X with $\mu_i \to \mu$ in the weak* topology. Then*

$$h_\mu(T) \geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T).$$

In less technical terms, for expansive dynamical systems, a "complicated" invariant measure might be approximated by a sequence of "simple" ones, but not vice versa.

3.13. **Proof.** Let $\mathcal{P}$ be a partition of X such that for each $P \in \mathcal{P}$

(i) $\mu(\partial P) = 0$
(ii) P has diameter $< \delta$ (δ as in the definition of expansiveness).

As X is compact, such a partition can easily obtained from a (finite sub-cover of a) cover of X consisting of small enough balls satisfying (i).

Since $\mu(\partial P) = 0$ and $\mu_i \to \mu$ weak*, for every $P \in \mathcal{P}$ we have that $\mu_i(P) \to \mu(P)$. Then for a fixed N we have (using footnote (8) for the measure μ_i) that

$$\begin{aligned}\frac{1}{N} H_\mu\left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P}\right) &= \lim_{i\to\infty} \frac{1}{N} H_{\mu_i}\left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P}\right) \\ &\geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T,\mathcal{P}) \overset{\text{(by (ii))}}{=} \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T).\end{aligned}$$

Taking the limit as $N\to\infty$ we get

$$h_\mu(T) = h_\mu(T,\mathcal{P}) = \lim_{N\to\infty} \frac{1}{N} H_\mu\left(\bigvee_{n=0}^{N-1} T^{-n}\mathcal{P}\right) \geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(T)$$

as required. □

Note that we have used both (ii) and expansiveness only to establish

(ii′) $h_\nu(T) = h_\nu(T,\mathcal{P})$ for $\nu = \mu_1, \mu_2, \ldots$.

We could have used the following weaker condition: for every ϵ, there is a partition $\mathcal{P}$ satisfying (i) and

(ii″) $h_\nu(T) \leq h_\nu(T,\mathcal{P}) + \epsilon$ for $\nu = \mu_1, \mu_2, \ldots$.

3.14. We are interested in dynamical systems of the form $X = \Gamma\backslash G$ (G a connected Lie group and $\Gamma < G$ a lattice) and

$$T : x \mapsto g.x = xg^{-1}.$$

Many such systems[12] will not be expansive, and furthermore in the most interesting case of $X_n = \mathrm{SL}(n,\mathbb{Z})\backslash \mathrm{SL}(n,\mathbb{R})$ the space X is not compact (which we assumed throughout the above discussion of expansiveness).

Even worse, on $X_2 = \mathrm{SL}(2,\mathbb{Z})\backslash \mathrm{SL}(2,\mathbb{R})$ one may have a sequence of probability measures μ_i ergodic and invariant under the one parameter group $\left\{ a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2}\end{pmatrix}\right\}$

[12] For example, the geodesic flow defined on quotients of $G = \mathrm{SL}(2,\mathbb{R})$.

with $\underline{\lim}_{i\to\infty} h_{\mu_i}(a_\bullet) > 0$ converging weak* to a measure μ which is not a probability measure and furthermore has zero entropy[(13)].

However, one has the following "folklore theorem"[(14)] :

3.15. Proposition. *Let G be a connected Lie group, $\Gamma < G$ a lattice, and $H = \{a_t\}$ a one parameter subgroup of G. Suppose that μ_i, μ are H-invariant* probability[(15)] *measures on X with $\mu_i \to \mu$ in the weak* topology. Then*

$$h_\mu(a_\bullet) \geq \overline{\lim_{i\to\infty}}\, h_{\mu_i}(a_\bullet).$$

For X compact (and possibly by some clever compactification also for general X), this follows from deep (and complicated) work of Yomdin, Newhouse and Buzzi (see e.g. [**Buz97**] for more details); however Proposition 3.15 can be established quite elementarily. In order to prove this proposition, one shows that any sufficiently fine finite partition of X satisfies §3.11(ii′′).

3.16**.** The following example shows that this semicontinuity does not hold for a general dynamical system:

3.17. Example. *Let $S = \left\{1, \frac{1}{2}, \frac{1}{3}, \dots, 0\right\}$, and $X = S^{\mathbb{Z}}$ (equipped with the usual Tychonoff topology). Let $\sigma : X \to X$ be the shift map defined by $\sigma(x)_n = x_{n+1}$ for $x = (x_n)_{n\in\mathbb{Z}} \in X$.*

Let μ_n be the probability measure on X obtained by taking the product of the probability measures on S giving equal probability to 0 and $\frac{1}{n}$, and $\delta_{\mathbf{0}}$ the probability measure supported on the fixed point $\mathbf{0} = (\dots, 0, 0, \dots)$ of σ. Then $\mu_n \to \delta_{\mathbf{0}}$ weak, $h_{\mu_n}(\sigma) = \log 2$ but $h_{\delta_{\mathbf{0}}}(\sigma) = 0$.*

3.18**.** Let (X, d) be a compact metric space and let $T : X \to X$ be continuous. Two points $x, x' \in X$ are said to be *k, ϵ-separated* if for some $0 \leq \ell < k$ we have that $d(T^\ell x, T^\ell x') \geq \epsilon$. Let $N(X, T, k, \epsilon)$ denote the maximal cardinality of a k, ϵ-separated subset of X.

3.19. Definition. *The* topological entropy[(16)] *of (X, T) is defined by*

$$H(X, T, \epsilon) = \underline{\lim_{k\to\infty}} \frac{\log N(X, T, k, \epsilon)}{k}$$
$$h_{\mathrm{top}}(X, T) = \lim_{\epsilon\to 0} H(X, T, \epsilon).$$

The topological entropy of a flow $\{a_t\}$ is defined as in §3.7 and denoted by $h_{\mathrm{top}}(X, a_\bullet)$.

---

(13)Strictly speaking, we define entropy only for probability measures, so one needs to rescale μ first.

(14)Which means in particular that there seems to be no good reference for it. A special case of this proposition is proved in [**EKL06**, Section 9]. The proof of this proposition is left as an exercise to the energetic reader.

(15) Here we assume that the weak* limit is a probability measure as, unlike the case of unipotent flows, there is no general fact that rules out various weird situations. E.g., for the geodesic flow on a noncompact quotient X of $\mathrm{SL}(2, \mathbb{R})$ it is possible to construct a sequence of invariant probability measures whose limit μ satisfies $\mu(X) = 1/2$.

(16)For X which is only locally compact, one can extend T to a map $\tilde{T}$ on its one-point compactification $\tilde{X} = X \cup \{\infty\}$ fixing ∞ and define $h_{\mathrm{top}}(X, T) = h_{\mathrm{top}}(\tilde{X}, \tilde{T})$.

3.20**.** Topological entropy and the ergodic theoretic entropy are related by the *variational principle* (see e.g. [**Gla03**, Theorem 17.6] or [**KH95**, Theorem 4.5.3])

3.21. Proposition. *Let X be a compact metric space and $T : X \to X$ a homeomorphism.*(17) *Then*

$$h_{\mathrm{top}}(X,T) = \sup_{\mu} h_\mu(T)$$

where the sup *runs over all T-invariant probability measures on X.*

Note that when $\mu \mapsto h_\mu(T)$ is upper semicontinuous (see §3.11) the supremum is actually attained by some T-invariant measure on X. These *measures of maximal entropy* are often quite natural measures, e.g. in many cases they are Haar measures on $\Gamma\backslash G$.

3.22**.** To further develop the theory of entropy we need to recall in the next few sections some more notions from measure theory.

4. Conditional Expectation and Martingale theorems

The material of this and the following section can be found in greater detail e.g. in [**EW09a**].

4.1. Proposition. *Let $(X, \mathcal{B}, \mu)$ be a probability space, and $\mathcal{A} \subset \mathcal{B}$ a sub-σ-algebra. Then there exists a continuous linear functional*

$$E_\mu(\cdot|\mathcal{A}) : L^1(X, \mathcal{B}, \mu) \to L^1(X, \mathcal{A}, \mu)$$

called the conditional expectation of f given $\mathcal{A}$*, such that*

$$E_\mu(f|\mathcal{A}) \text{ is } \mathcal{A}\text{-measurable} \tag{4.1a}$$

for any $f \in L^1(X, \mathcal{B}, \mu)$, and we have

$$\int_A E_\mu(f|\mathcal{A})d\mu = \int_A f d\mu \text{ for all } A \in \mathcal{A}. \tag{4.1b}$$

Moreover, together Equations (4.1a)–(4.1b) *characterizes the function $E_\mu(f|\mathcal{A}) \in L^1(X, \mathcal{B}, \mu)$.*

On $L^2(X, \mathcal{B}, \mu)$ the operator $E_\mu(\cdot|\mathcal{A})$ is simply the orthogonal projection to the closed subspace $L^2(X, \mathcal{A}, \mu)$. From there one can extend the definition by continuity to $L^1(X, \mathcal{B}, \mu)$. Often, when we only consider one measure we will drop the measure in the subscript.

Below we will rely our arguments on the dynamical behavior of points. Because of that we prefer to work with functions instead of equivalence classes of functions and hence the above uniqueness has be understood accordingly. We will need the following useful properties of the conditional expectation $E(f|\mathcal{A})$, which we already phrase in terms of functions rather than equivalence classes of functions:

4.2. Proposition.
(i) *$E(\cdot|\mathcal{A})$ is a positive operator of norm 1, and moreover, $|E(f|\mathcal{A})| \le E(|f||\mathcal{A})$ almost everywhere.*
(ii) *For $f \in L^1(X, \mathcal{B}, \mu)$ and $g \in L^\infty(X, \mathcal{A}, \mu)$, we have $E(gf|\mathcal{A}) = gE(f|\mathcal{A})$ almost everywhere.*

(17) This proposition also easily implies the analogous statement for flows $\{a_t\}$.

(iii) *If $\mathcal{A}' \subset \mathcal{A}$ is a sub-σ-algebra, then*
$$E(E(f|\mathcal{A})|\mathcal{A}') = E(f|\mathcal{A}')$$
almost everywhere. Moreover, if $f \in L^1(X, \mathcal{A}, \mu)$, then $E(f|\mathcal{A}) = f$ almost everywhere.

(iv) *If $T : X \to Y$ sends the probability measure μ on X to $T_*\mu = \mu \circ T^{-1} = \nu$ on Y, and if $\mathcal{C}$ is a sub-σ-algebra of the σ-algebra $\mathcal{B}_Y$ of measurable sets on Y, then $E_\mu(f \circ T|T^{-1}\mathcal{C}) = E_\nu(f|\mathcal{C}) \circ T$ for any $f \in L^1(Y, \mathcal{B}_Y, \nu)$.*

We only prove the last two claims. Take any $A \in \mathcal{A}' \subset \mathcal{A}$. By the characterizing property of conditional expectation, we have
$$\int_A E(E(f|\mathcal{A})|\mathcal{A}') = \int_A E(f|\mathcal{A}) = \int_A f$$
Therefore by uniqueness, we have $E(E(f|\mathcal{A})|\mathcal{A}') = E(f|\mathcal{A}')$ almost everywhere. If $f \in L^1(X, \mathcal{A}, \mu)$, then f satisfies the first characterizing property of $E(f|\mathcal{A})$, while trivially satisfying the second. Again invoking uniqueness, we have $E(f|\mathcal{A}) = f$ almost everywhere.

We consider now the situation of the pushforward $T_*\mu = \nu$ of the measure and the pullback $T^{-1}\mathcal{C}$ of the σ-algebra. By the definitions we have for any $C \in \mathcal{C}$ that
$$\int_{T^{-1}C} E_\nu(f|\mathcal{C}) \circ T d\mu = \int_C E_\nu(f|\mathcal{C}) d\nu = \int_C f d\nu = \int_{T^{-1}C} f \circ T d\mu,$$
which implies the claim by the uniqueness properties of conditional expectation.

4.3**.** The next two theorems describes how the conditional expectation behaves with respect to a sequence of sub-σ-algebras, and can be thought of as continuity properties.

4.4. Theorem (Increasing Martingale Convergence Theorem)**.** *Let $\mathcal{A}_1$, $\mathcal{A}_2, \ldots$ be a sequence of σ-algebras, such that $\mathcal{A}_i \subset \mathcal{A}_j$ for all $i < j$. Let $\mathcal{A}$ be the smallest σ-algebras containing all of the $\mathcal{A}_n$ (in this case, we write $\mathcal{A}_n \nearrow \mathcal{A}$). Then*
$$E(f|\mathcal{A}_n) \to E(f|\mathcal{A})$$
almost everywhere and in L^1.

4.5. Theorem (Decreasing Martingale Convergence Theorem)**.** *Suppose that we have a sequence of σ-algebras $\mathcal{A}_i \searrow \mathcal{A}$, i.e., such that $\mathcal{A}_i \supset \mathcal{A}_j$ for $i < j$, and $\mathcal{A} = \bigcap \mathcal{A}_i$. Then $E(f|\mathcal{A}_n)(x) \to E(f|\mathcal{A})(x)$ almost everywhere and in L^1.*

4.6**. Remark:** In many ways, the Decreasing Martingale Convergence Theorem is similar to the pointwise ergodic theorem. Both theorems have many similarities in their proof with the pointwise ergodic theorem and other theorems; the proofs consists of two steps, convergence in L^1, and a maximum inequality to deduce pointwise convergence.

5. Countably generated σ-algebras and Conditional measures

Note that the algebra generated by a countable set of subsets of X is countable, but that in general the same is not true for the σ-algebra generated by a countable set of subsets of X. E.g. the Borel σ-algebra of any space we consider is countably generated in the following sense.

5.1. Definition. *A σ-algebra $\mathcal{A}$ in a space X is* countably generated *if there is a countable set (or equivalently algebra) $\mathcal{A}_0$ of subsets of X such that the smallest σ-algebra $\sigma(\mathcal{A}_0)$ that contains $\mathcal{A}_0$ is precisely $\mathcal{A}$.*

5.2**.** One nice feature of countably generated σ-algebras is that we can study the atoms of the algebra. If $\mathcal{A}$ is generated by a countable algebra $\mathcal{A}_0$, then we define the $\mathcal{A}$-*atom* of a point x to be

$$[x]_{\mathcal{A}} := \bigcap_{x \in A \in \mathcal{A}_0} A = \bigcap_{x \in A \in \mathcal{A}} A.$$

The equality follows since $\mathcal{A}_0$ is a generating algebra for the σ-algebra $\mathcal{A}$. In particular, it shows that the atom $[x]_{\mathcal{A}}$ does not depend on a choice of the generating algebra. Notice that by countability of $\mathcal{A}_0$ we have $[x]_{\mathcal{A}} \in \mathcal{A}$. In other words, $[x]_{\mathcal{A}}$ is the smallest set of $\mathcal{A}$ containing x. Hence the terminology — the atom of x cannot be broken up into smaller sets within the σ-algebra $\mathcal{A}$.

Note, in particular, that $[x]_{\mathcal{A}}$ could consist of the singleton x; in fact, this is the case for all atoms of the Borel σ-algebra on, say, $\mathbb{R}$. The notion of atoms is convenient when we want to consider conditional measures for smaller σ-algebras.

5.3**. Caution:** A sub σ-algebra of a countably generated σ-algebra need not be countably generated!

5.4. Lemma. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability preserving system such that individual points have zero measure. Then the σ-algebra $\mathcal{E}$ of T–invariant sets (i.e., sets $B \in \mathcal{B}$ such that $B = T^{-1}B = TB$) is* not *countably generated.*

5.5**. Proof:** Since T is ergodic, any set in $\mathcal{E}$ has measure 0 or 1, and in particular, this holds for any generating set. Suppose that $\mathcal{E}$ is generated by a countable collection $\{E_1, E_2, \ldots\}$, each E_i having measure 0 or 1. Taking the intersection of all generators E_i of measure one and the complement $X \setminus E_i$ of those of measure zero, we obtain an $\mathcal{E}$-atom $[x]_{\mathcal{E}}$ of measure 1. Since the orbit of x is invariant under T, we have that $[x]_{\mathcal{E}}$ must be the orbit of x. Since the orbit is at most countable, this is a contradiction. □

5.6**.** We will now restrict ourselves to the case of X a locally compact, second-countable metric space, $\mathcal{B}$ will be the Borel σ-algebra on X. A space and σ-algebra of this form will be referred as a *standard Borel space*, and we will always take μ to be a Borel measure. We note that for such X, the Borel σ-algebra is countably generated by open neighborhoods of points in a countable dense subset of X. When working with a Borel measure on X, we may replace X by the one-point-compactification of X, extend the measure trivially to the compactification, and assume without loss of generality that X is compact.

5.7. Definition. *Let $\mathcal{A}$, $\mathcal{A}'$ be sub-σ-algebras of the σ-algebra $\mathcal{B}$ of a probability space $(X, \mathcal{B}, \mu)$. We say that $\mathcal{A}$ is* equivalent *to $\mathcal{A}'$* modulo μ *(denoted $\mathcal{A} \doteq_\mu \mathcal{A}'$) if for every $A \in \mathcal{A}$ there exists $A' \in \mathcal{A}'$ such that $\mu(A \triangle A') = 0$, and vice versa.*

5.8. Proposition. *Let $(X, \mathcal{B})$ be a standard Borel space, and let μ be a Borel probability measure on X. Then for every sub-σ-algebra $\mathcal{A} \subset \mathcal{B}$, there exists $\tilde{\mathcal{A}} \subset \mathcal{A}$ such that $\tilde{\mathcal{A}}$ is countably generated, and $\tilde{\mathcal{A}} \doteq_\mu \mathcal{A}$.*

Roughly speaking the proposition follows since the space $L^1(X, \mathcal{A}, \mu)$ is separable, which in turn is true because it is as a subspace of $L^1(X, \mathcal{B}, \mu)$. One can define

$\tilde{\mathcal{A}}$ by a countable collection of sets $A_i \in \mathcal{A}$ for which the characteristic functions χ_{A_i} are dense in the set of all characteristic functions χ_A with $A \in \mathcal{A}$.

This Proposition conveniently allows us to ignore issues of countable generation, as long as we do so with respect to a measure (i.e., up to null sets) on a nice space.

We now wish to prove the existence and fundamental properties of conditional measures:

5.9. Theorem. *Let $(X, \mathcal{B}, \mu)$ be a probability space with $(X, \mathcal{B})$ being a standard Borel space, and let $\mathcal{A} \subset \mathcal{B}$ a sub-σ-algebra. Then there exists a subset $X' \subset X$ of full measure (i.e., $\mu(X \backslash X') = 0$), belonging to $\mathcal{A}$, and Borel probability measures $\mu_x^{\mathcal{A}}$ for $x \in X'$ such that:*

(i) *For every $f \in L^1(X, \mathcal{B}, \mu)$ we have $E(f|\mathcal{A})(x) = \int f(y) d\mu_x^{\mathcal{A}}(y)$ for almost every x. In particular, the right-hand side is $\mathcal{A}$-measurable as a function of x.*

(ii) *If $\mathcal{A} \doteq_\mu \mathcal{A}'$ are equivalent σ-algebras modulo μ, then we have $\mu_x^{\mathcal{A}} = \mu_x^{\mathcal{A}'}$ for almost every x.*

(iii) *If $\mathcal{A}$ is countably generated, then $\mu_x^{\mathcal{A}}([x]_{\mathcal{A}}) = 1$ for every $x \in X'$, and for $x, y \in X'$ we have that $[x]_{\mathcal{A}} = [y]_{\mathcal{A}}$ implies $\mu_x^{\mathcal{A}} = \mu_y^{\mathcal{A}}$.*

(iv) *The set X' and the map $x \overset{\tau}{\mapsto} \mu_x^{\mathcal{A}}$ are $\mathcal{A}$-measurable on X'; i.e., if U is open in $\mathcal{P}(X)$, the space of probability measures on X equipped with the weak* topology, then $\tau^{-1}(U) \in \mathcal{A}|_{X'}$.*

Moreover, the family of conditional measures $\mu_x^{\mathcal{A}}$ is almost everywhere uniquely determined by its relationship to the conditional expectation described above.

If $\mathcal{A}$ is countably generated, then $x, y \in X$ are called equivalent w.r.t. $\mathcal{A}$ if $[x]_{\mathcal{A}} = [y]_{\mathcal{A}}$. Hence (iii) also says that equivalent points have identical conditional measures.

5.10**. Caution:** In general we will only prove facts concerning the conditional measures $\mu_x^{\mathcal{A}}$ for almost every $x \in X$. In fact, we even restricted ourselves to a set X' of full measure in the existence of $\mu_x^{\mathcal{A}}$. However, even the set X' is by no means canonical. We also must understand the last claim regarding the uniqueness in that way; if we have two families of conditional measure defined on sets of full measure X' and X'', then one can find a *subset* of $X' \cap X''$ of full measure where they agree.

5.11**. Comments:** If $N \subset X$ is a null set, it is clear that $\mu_x^{\mathcal{A}}(N) = 0$ for a.e. x. (Use Theorem 5.9.(i) and Proposition 4.1 to check this.) However, we cannot expect more as, for a given x, the set $[x]_{\mathcal{A}}$ is often a null set.

If $B \subset X$ is measurable, then

$$\mu\big(\{x \in B : \mu_x^{\mathcal{A}}(B) = 0\}\big) = 0. \tag{5.11a}$$

To see this define $A = \{x : \mu_x^{\mathcal{A}}(B) = 0\} \in \mathcal{A}$ and use again Theorem 5.9.(i) and Proposition 4.1 to get

$$\mu(A \cap B) = \int_A \chi_B d\mu = \int_A \mu_x^{\mathcal{A}}(B) d\mu(x) = 0.$$

5.12**. Proof:** Since we are working in a standard Borel space, we may assume that X is a compact, metric space. Hence, we may choose a countable set of continuous functions which give a dense $\mathbb{Q}$-vector space $\{f_0 \equiv 1, f_1, \ldots\} \subset C(X)$. Set $g_0 = f_0 \equiv 1$, and for each f_i with $i \geq 1$, pick[(18)] $g_i = E(f_i|\mathcal{A}) \in L^1(X, \mathcal{A}, \mu)$. Taking the union of countably many null sets there exists a null set N for the measure μ such that for all $\alpha, \beta \in \mathbb{Q}$ and all i, j, k:

- If $\alpha \leq f_i \leq \beta$ (on all of X), then $\alpha \leq g_i(x) \leq \beta$ for all $x \notin N$
- If $\alpha f_i + \beta f_j = f_k$, then $\alpha g_i(x) + \beta g_j(x) = g_k(x)$ for $x \notin N$

Now for all $x \notin N$, we have a continuous linear functional $\mathcal{L}_x : f_i \mapsto g_i(x)$ from $C(X) \to \mathbb{R}$ of norm $||\mathcal{L}_x|| \leq 1$. By the Riesz Representation Theorem, this yields a measure $\mu_x^{\mathcal{A}}$ on $C(X)$. This measure is characterized by $E(f|\mathcal{A})(x) = \mathcal{L}_x(f) = \int f(y) d\mu_x^{\mathcal{A}}(y)$ for all $f \in C(X)$. Using monotone convergence this can be extended to other class of functions: first to characteristic functions of compact and of open sets, then to characteristic functions of all Borel sets and finally to integrable functions, i.e., we have part (i) of the Theorem. As already remarked, this implies that $x \mapsto \int f(y) d\mu_x^{\mathcal{A}}(y)$ is an $\mathcal{A}$-measurable function for $x \notin N$. This implies part (iv).

Now suppose we have two equivalent σ-algebras $\mathcal{A}$ and $\mathcal{A}'$ modulo μ, and take their common refinement $\tilde{\mathcal{A}}$. Then for any $f \in C(X)$, we see that both $g = E(f|\mathcal{A})$ and $g' = E(f|\mathcal{A}')$ satisfy the characterizing properties of $E(f|\tilde{\mathcal{A}})$, and so they are equal almost everywhere. Again taking a countable union of null sets, corresponding to a countable dense subset of $C(X)$, we see that $\mu_x^{\mathcal{A}} = \mu_x^{\mathcal{A}'}$ almost everywhere, giving part (ii).

For part (iii), suppose that $\mathcal{A} = \sigma(\{A_1, \ldots\})$ is countably generated. For every i, we have that $1_{A_i}(x) = E(1_{A_i}|\mathcal{A})(x) = \mu_x^{\mathcal{A}}(A_i)$ almost everywhere. Hence there exists a set N of μ-measure 0, given by the union of the these null sets for each i, such that $\mu_x^{\mathcal{A}}(A_i) = 1$ for all i and every $x \in A_i \backslash N$. Therefore, since $[x]_{\mathcal{A}}$ is the countable intersection of A_i's containing x, we have $\mu_x^{\mathcal{A}}([x]_{\mathcal{A}}) = 1$ for all $x \notin N$. Finally, since $x \to \mu_x^{\mathcal{A}}$ is $\mathcal{A}$-measurable, we have that $[x]_{\mathcal{A}} = [y]_{\mathcal{A}} \Rightarrow \mu_x^{\mathcal{A}} = \mu_y^{\mathcal{A}}$ whenever both are defined (i.e., $x, y \in X'$). $\square$

5.13**. Another construction** An alternate construction for the conditional measure for a countably generated σ-algebra is to start by finding a sequence of finite partitions $\mathcal{A}_n \nearrow \mathcal{A}$. For finite partitions, the conditional measures are particularly simple; we have

$$\mu_x^{\mathcal{A}_n} = \frac{\mu|_{[x]_{\mathcal{A}_n}}}{\mu([x]_{\mathcal{A}_n})}$$

Now, for any $f \in C(X)$, the Increasing Martingale Convergence Theorem tells us that for any continuous f and for almost every x, we have $\int f d\mu_x^{\mathcal{A}_n} = E(f|\mathcal{A}_n)(x) \to E(f|\mathcal{A})(x)$. Again by choosing a countable dense subset of $C(X)$ we show a.e. that $\mu_x^{\mathcal{A}_n}$ converge in the weak* topology to a measure $\mu_x^{\mathcal{A}}$ as in (i) of the theorem.

5.14**. The ergodic decomposition revisited.** One application for the notion of conditional measures is that it can be used to prove the existence of the ergodic decomposition. In fact, for any H-invariant measure μ, we have the ergodic

(18) Here the word "pick" refers to the choice of a representative of the equivalence class of integrable measurable functions.

decomposition

$$\mu = \int \mu_x^{\mathcal{E}} d\mu(x),$$

where $\mathcal{E}$ is (alternatively a countably generated σ-algebra equivalent to) the σ-algebra of all H-invariant sets, and $\mu_x^{\mathcal{E}}$ is the conditional measure (on the $\mathcal{E}$-atom of x). This is a somewhat more intrinsic way to write the ergodic decomposition as one does not have to introduce an auxiliary probability space.

5.15. Definition. *Two countably generated σ-algebras $\mathcal{A}$ and $\mathcal{C}$ on a space X are* countably equivalent *if any atom of $\mathcal{A}$ can be covered by at most countably many atoms of $\mathcal{C}$, and vice versa.*

5.16**. Remark:** This *is* an equivalence relation. Symmetry is part of the definition, reflexivity is obvious, and transitivity can be readily checked.

5.17. Proposition. *Suppose $\mathcal{A}$ and $\mathcal{A}'$ are countably equivalent sub-σ-algebras. Then for μ-a.e. x, we have*

$$\mu_x^{\mathcal{A}}|_{[x]_{\mathcal{A}\vee\mathcal{A}'}} \propto \mu_x^{\mathcal{A}'}|_{[x]_{\mathcal{A}\vee\mathcal{A}'}}$$

Or, put another way,

$$\mu_x^{\mathcal{A}\vee\mathcal{A}'} = \frac{\mu_x^{\mathcal{A}}|_{[x]_{\mathcal{A}\vee\mathcal{A}'}}}{\mu_x^{\mathcal{A}}([x]_{\mathcal{A}\vee\mathcal{A}'})} = \frac{\mu_x^{\mathcal{A}'}|_{[x]_{\mathcal{A}\vee\mathcal{A}'}}}{\mu_x^{\mathcal{A}'}([x]_{\mathcal{A}\vee\mathcal{A}'})}$$

Here and in the following the notation $\mu \propto \nu$ for two measures on a space X denotes proportionality, i.e. that there exists some $c > 0$ with $\mu = c\nu$.

5.18**. Proof:** As a first step, we observe that $\mathcal{A}$ is countably equivalent to $\mathcal{A}'$ if and only if $\mathcal{A}$ is countably equivalent to the σ-algebra generated by $\mathcal{A}$ and $\mathcal{A}'$. Hence we may assume that $\mathcal{A} \subset \mathcal{A}'$, and the statement of the Proposition reduces to

$$\mu_x^{\mathcal{A}'} = \frac{\mu_x^{\mathcal{A}}|_{[x]_{\mathcal{A}'}}}{\mu_x^{\mathcal{A}}([x]_{\mathcal{A}'})}$$

The next step is to verify that the denominator on the right-hand side is actually $\mathcal{A}'$-measurable (as a function of x). As $\mathcal{A}'$ is countably generated, we may take a sequence $\mathcal{A}'_n \nearrow \mathcal{A}'$ of finite algebras, and consider the decreasing chain of sets $[x]_{\mathcal{A}'_n}$. Notice that $E(1_{[x]_{\mathcal{A}'_n}}|\mathcal{A})(x) = \mu_x^{\mathcal{A}}([x]_{\mathcal{A}'_n})$ is a perfectly good $\mathcal{A} \vee \mathcal{A}'_n$-measurable function. In the limit as $n \to \infty$, the set $[x]_{\mathcal{A}'_n} \searrow [x]_{\mathcal{A}'} = \bigcap_n [x]_{\mathcal{A}'_n}$ as $(\mathcal{A}'_n \vee \mathcal{A}) \nearrow \mathcal{A}'$, and so $x \mapsto \mu_x^{\mathcal{A}}([x]_{\mathcal{A}'})$ is $\mathcal{A}'$-measurable.

We still also have to verify that this denominator is non-zero (almost everywhere). Consider the set $Y = \{x : \mu_x^{\mathcal{A}}([x]_{\mathcal{A}'}) = 0\}$. We must show that $\mu(Y) = 0$ when $\mathcal{A}$ and $\mathcal{A}'$ are countably equivalent. The previous step guarantees that Y is measurable, and we can integrate fibre by fibre: $\mu(Y) = \int \mu_x^{\mathcal{A}}(Y) d\mu(x)$. But $[x]_{\mathcal{A}}$ is a finite or countable union $\bigcup_{i\in I}[x_i]_{\mathcal{A}'}$ of $\mathcal{A}'$-atoms, and so

$$\mu_x^{\mathcal{A}}(Y) = \sum_{i\in I} \mu_x^{\mathcal{A}}([x_i]_{\mathcal{A}'} \cap Y)$$

and so it suffices to show that each term on the right-hand side is 0. If $[x_i]_{\mathcal{A}'} \cap Y = \emptyset$, then there is nothing to show. On the other hand, if there exists some $y \in [x_i]_{\mathcal{A}'} \cap Y$, then by definition of Y we have $\mu_y^{\mathcal{A}}([x_i]_{\mathcal{A}'}) = 0$. But $[x_i]_{\mathcal{A}'} \subset [x]_{\mathcal{A}}$, and so $y \in [x]_{\mathcal{A}}$, which by Theorem 5.9 (and the subsequent Remark) implies that $\mu_x^{\mathcal{A}}([x_i]_{\mathcal{A}'}) = \mu_y^{\mathcal{A}}([x_i]_{\mathcal{A}'}) = 0$.

We now know that $\frac{\mu_x^{\mathcal{A}}|_{[x]_{\mathcal{A}'}}}{\mu_x^{\mathcal{A}}([x]_{\mathcal{A}'})}$ makes sense. We easily verify that it satisfies the characterizing properties of $\mu_x^{\mathcal{A}'}$, and we are done. □

6. Leaf-wise Measures, the construction

We will need later (e.g. in the discussion of entropy) another generalization of conditional measures that allows us to discuss "the restrictions of the measure" to the orbits of a group action just like the conditional measures describe "the restriction of the measure" to the atoms. However, as we have seen in Lemma 5.4, one cannot expect to have a σ-algebra whose atoms are precisely the orbits.

As we will see these restricted measures for orbits, which we will call *leaf-wise measures*, can be constructed by patching together conditional measures for various σ-algebras whose atoms are pieces of orbits. Such a construction (with little detail provided) is used by Katok and Spatzier in [**KS96**]; we follow here the general framework outlined in [**Lin06**], with some simplifications and improvements (e.g. Theorem 6.30 which in this generality seems to be new).

6.1. **A few assumptions.** Let T be a locally compact, second countable group. We assume that T is equipped with a right-invariant metric such that any ball of finite radius has compact closure. We write $B_r^T(t_0) = \{t \in T : d(t, t_0) < r\}$ for the open ball of radius r around $t_0 \in T$, and write $B_r^T = B_r^T(e)$ for the ball around the identity $e \in T$. Also let X be a locally compact, second countable metric space. We assume that T acts continuously on X, i.e., that there is a continuous map $(t, x) \mapsto t.x \in X$ defined on $T \times X \to X$ satisfying $s.(t.x) = (st).x$ and $e.x = x$ for all $s, t \in T$ and $x \in X$. We also assume the T-action to be *locally free* in the following uniform way: for every compact $K \subset X$ there is some $\eta > 0$ such that $t \in B_\eta^T$, $x \in K$, and $t.x = x$ imply $t = e$. In particular, the identity element $e \in T$ is isolated in $\mathrm{Stab}_T(x) = \{t \in T : t.x = x\}$, so that the latter becomes a discrete group, for every $x \in X$— this property allows a nice foliation of X into T-orbits. Finally we assume that μ is a *Radon* (or *locally finite*) measure on X, meaning that $\mu(K) < \infty$ for any compact $K \subset X$.

6.2. Definition. *Let*[19] $x \in X$*. A set* $A \subset T.x$ *is an* open T-plaque *if for every* $a \in A$ *the set* $\{t : t.a \in A\}$ *is open and bounded.*

Note that by the above assumptions on T a set is bounded only if its closure is compact. We recall that $\mu \propto \nu$ for two measures on a space X denotes proportionality, i.e., that there exists some $c > 0$ with $\mu = c\nu$.

6.3. Theorem (Provisional[20]!)**.** *In addition to the above assume also that* $\mathrm{Stab}_T(x) = \{e\}$ *for* μ*-a.e.* $x \in X$*, i.e.,* $t \mapsto t.x$ *is injective for a.e.* x*. Then there is a system* $\{\mu_x^T\}_{x \in X'}$ *of Radon measures on* T *which we will call the* leaf-wise measures *which are determined uniquely, up to proportionality and outside a set of measure zero, by the following properties:*

(i) *The domain* $X' \subset X$ *of the function* $x \mapsto \mu_x^T$ *is a full measure subset in the sense that* $\mu(X \setminus X') = 0$.

(ii) *For every* $f \in C_c(T)$*, the map* $x \mapsto \int f d\mu_x^T$ *is Borel measurable.*

---

[19]Below we will work mostly with points x for which $t \in T \mapsto t.x$ is injective.

[20]Ideally, we would like to "normalize" by looking at equivalence classes of proportional Radon measures, but this will require further work. See Theorem 6.30.

(iii) *For every $x \in X'$ and $s \in T$ with $s.x \in X'$, we have $\mu_x^T \propto (\mu_{s.x}^T)s$, where the right-hand side is the push-forward of $\mu_{s.x}^T$ by the right translation (on T) $t \mapsto ts$, see Figure 1.*

(iv) *Suppose $Z \subset X$ and that there exists a countably generated σ-algebra $\mathcal{A}$ of subsets of Z such that for any $x \in Z$, the set $[x]_{\mathcal{A}}$ is an open T-plaque; i.e., $U_{x,\mathcal{A}} := \{t : t.x \in [x]_{\mathcal{A}}\}$ is open and bounded satisfying $[x]_{\mathcal{A}} = U_{x,\mathcal{A}}.x$. Then for μ-a.e. $x \in Z$,*

$$(\mu|_Z)_x^{\mathcal{A}} \propto \left(\mu_x^T|_{U_{x,\mathcal{A}}}\right).x$$

where the latter is the push-forward under the map $t \in U_{x,\mathcal{A}} \mapsto t.x \in [x]_{\mathcal{A}}$.

(v) *The identity element $e \in T$ is in the support of μ_x^T for μ-a.e. x.*

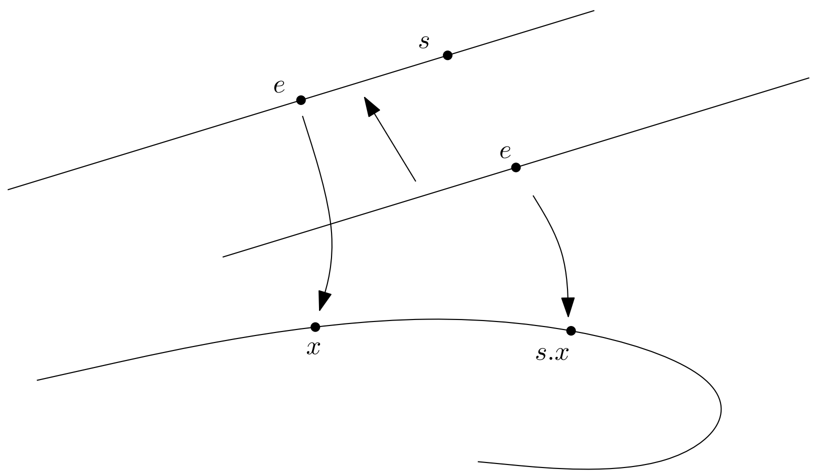

FIGURE 1. The two straight lines represent two copies of the group T and the curved line represents the orbit $T.x = T.(s.x)$. The arrows from the groups to the orbit represent the orbit maps $t \to t.x$ and $t \to t.(s.x)$. Right translation by s from T to T makes the diagram commutative. In other words, Thm. 6.3(iii) only says that the infinite measures $\mu_x^T.x$ and $\mu_{s.x}^T.(s.x)$ on X are proportional.

6.4. **Remarks:**

(i) The properties of leaf-wise measures are analogous to those of the conditional measures described in Theorem 5.9. With leaf-wise measures, we demand that the "atoms" correspond to entire (non-compact!) T-orbits, and herein lie most of the complications. On the other hand, these orbits inherit the group structure from T, and so the conditional measures μ_x^T are actually measures on the *group* T, which has structure that we can exploit.

(ii) Property 6.3.(iii) is the analogue of Property 5.9.(iii). Ideally, we would like to say that, since x and $g.x$ are in the same T-orbit, their leaf-wise measures should be the same. However, we prefer to work with measures on T so we move the measures from $T.x$ to T via $t.x \mapsto t$

(which implicitly makes use of the initial point x). Therefore, *points* on the orbit correspond to different *group* elements depending on the base point; hence we need to employ the right translation in order to have our measures (defined as measures on the group) agree at points of the orbit. Another difficulty is that the μ_x^T need not be probability measures, or even finite measures. There being no good way to "normalize" them, we must make do with proportionality instead of equality.

(iii) Property 6.3.(iv) is the most restrictive; this is the heart of the definition. It essentially says that one can restrict μ_x^T to $U_{x,\mathcal{A}}$ and get a finite measure, which looks just like (up to normalization) a good old conditional measure $\mu_x^{\mathcal{A}}$ derived from $\mathcal{A}$. So μ_x^T is in essence a global "patching" together of local conditional measures (up to proportionality issues).

6.5. **Examples:**

6.5.1. Let $X = \mathbb{T}^2$, on which $T = \mathbb{R}$ acts by $t.x = x + t\vec{v} \mod \mathbb{Z}^2$, for some irrational vector $\vec{v}$. If $\mu = \lambda$ is the Lebesgue measure on $\mathbb{T}^2$, then we can take $\mu_x^T = \lambda_{\mathbb{R}}$ to be Lebesgue measure on $\mathbb{R}$. Note that, even though the space X is quite nice (eg., compact), none of the leaf-wise measures are finite. Also, notice that the naive approach to constructing these measures would be to look at conditional measures for the sub-σ-algebra $\mathcal{A}$ of T-invariant Borel sets. Unfortunately, this σ-algebra is not countably generated, and is equivalent (see Lemma 5.4 and Proposition 5.8) to the trivial σ-algebra! This is a situation where passing to an equivalent σ-algebra to avoid uncountable generation actually destroys the information we want (T-orbits have measure 0). Instead, we define the leaf-wise measures on small pieces of T-orbits and then glue them together.

6.5.2. We now give an example of a p-adic group action. Let $X = (\mathbb{Q}_p \times \mathbb{R})/\mathbb{Z}[\frac{1}{p}] \cong (\mathbb{Z}_p \times \mathbb{R})/\mathbb{Z}$ where both $\mathbb{Z}[\frac{1}{p}]$ and $\mathbb{Z}$ are considered as subgroups via the canonical diagonal embedding. We let $T = \mathbb{Q}_p$ act on X by translations (where our group law is given by addition). To describe an interesting example of leaf-wise measures, we (measurably) identify X with the space of 2-sided sequences $\{x(i)\}_{i=-\infty}^{\infty}$ in base p (up to countably many nuisances) as follows: Note that $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the quotient of X by the subgroup $\mathbb{Z}_p$ and that we may use p-nary digit expansion in $[0, 1) \cong \mathbb{T}$. This way $x \in X$ determines a one-sided sequence of digits $x(i)$ for $i = 1, 2, \ldots$. Since multiplication by p is invertible on X, we may recover all digits $x(i)$ for $i = \ldots, -1, 0, 1, \ldots$ by applying the above to the points $p^{-n}x$. (The reader should verify that this procedure is well-defined at all but countably many points and that the assigned sequence of digits uniquely defines the initial point $x \in X$.)

Under this isomorphism of X with the space of sequences the action of translation by $\mathbb{Z}_p$ corresponds to changing (in a particular manner) the coordinates of the sequence corresponding to $i \leq 0$ such that the orbit under $\mathbb{Z}_p$ consists of all sequences that agree with the original sequence on all positive coordinates. For this recall that $\mathbb{Z}_p$ is isomorphic to $\{0, \ldots, p-1\}^{\mathbb{N}_0}$. More generally, the orbit of a point under $p^{-n}\mathbb{Z}_p$ corresponds to all sequences that have the same coordinates as the original sequence for $i > n$. Hence the $\mathbb{Q}_p$-orbit corresponds to all sequences that have the same digits as the original sequence for all $i > n$ for some n.

We now define a measure and discuss the leaf-wise measures for the action by $\mathbb{Q}_p$. Let μ be an identically independently distributed but biased Bernoulli measure – in other words we identify X again with the space of all 2-sided sequences, i.e., with $\{0,1,\dots,p-1\}^{\mathbb{Z}}$, and define μ as the infinite product measure using some fixed probability vector $v=(v_0,\dots,v_{p-1})\neq(\frac{1}{p},\dots,\frac{1}{p})$. We note that the map $\alpha: x\mapsto px$ defined by multiplication with p (which corresponds to shifting the sequences) preserves the measure μ and acts ergodically w.r.t. μ (in fact as one can check directly it is mixing w.r.t. μ which as mentioned before implies ergodicity). Note also that α preserves the foliation into $\mathbb{Q}_p$-orbits and in fact contracts them, i.e., $\alpha(x+\mathbb{Q}_p)=\alpha(x)+\mathbb{Q}_p$ and $\alpha(x+t)=\alpha(x)+pt$ for $t\in\mathbb{Q}_p$ and pt is p-adically smaller than t. Finally note that the $\mathbb{Q}_p$-action does not preserve the measure μ unless $v=(\frac{1}{p},\dots,\frac{1}{p})$. In this case there is very little difference to the above example on $\mathbb{T}^2$ – the leaf-wise measures end up being Haar measures on $\mathbb{Q}_p$. So let us assume the almost opposite extreme: suppose $v_0,\dots,v_{p-1}\in(0,1)$ and no two components of v are equal.

Let $\mathcal{A}$ be the countably generated σ-algebra (contained in the Borel σ-algebra of X) whose atoms are the $\mathbb{Z}_p$-orbits; it is generated by the *cylinder sets* of the form $\{x : x(i)=\epsilon_i \text{ for } 1\leq i\leq N\}$ for any $N>0$ and all possible finite sequences $(\epsilon_1,\dots,\epsilon_N)\in\{0,\dots,p-1\}^N$. Equivalently, the $\mathcal{A}$-atoms are all sequences that agree with a given one on all coordinates for $i\geq 1$ so that the atom has the structure of a one-sided shift space. By independence of the coordinates (w.r.t. μ) the conditional measures $\mu_x^{\mathcal{A}}$ are all Bernoulli i.i.d. measures according to the original probability vector v of μ; in other words, a random element of $[x]_{\mathcal{A}}$ according to $\mu_x^{\mathcal{A}}$ is a sequence $\{y(i)\}$ such that $y(i)=x(i)$ for $i\geq 1$, and the digits $y(i)$ for $-\infty<i\leq 0$ are picked independently at random according to the probability vector defining μ.

What does μ_x^T look like (where $T=\mathbb{Q}_p$)? For this notice that $\mathbb{Z}_p$ is open in $\mathbb{Q}_p$, so that the atoms for $\mathcal{A}$ are open T-plaques. Therefore, if we restrict μ_x^T to the subgroup $U=\mathbb{Z}_p$ of $T=\mathbb{Q}_p$, we should get by Theorem 6.3 (iv) that

$$x+\mu_x^T|_U\propto\mu_x^{\mathcal{A}}.$$

To understand this better, let's examine what a random point of $\frac{1}{\mu_x^T(U)}\mu_x^T|_U$ looks like. Of course, an element belonging to $\mathbb{Z}_p$ corresponds to a sequence $\{t(i)\}_{i=-\infty}^0$; how are the digits $t(i)$ distributed? Recall that if we translate by x, the resulting digits $(t+x)(i)$ (with addition formed in $\mathbb{Z}_p$ where the carry goes to the left) should be randomly selected according to the original probability vector. Hence the probability of $t(0)=\epsilon$ with respect to the normalized $\mu_x^T|_U$ becomes the original probability $v_{\epsilon+x(0)}$ of selecting the digit $\epsilon+x(0)$. By our assumption on the vector v this shift in the distribution determines $x(0)$. However, by using σ-algebras whose atoms are orbits of $p^n\mathbb{Z}^p$ for all $n\in\mathbb{Z}$ we conclude that μ_x^T determines all coordinates of x and hence x! (Of course had we used the theorem to construct the leaf-wise measures instead of directly finding it by using the structure of the given measure then the leaf-wise measure would only be defined on a set of full measure and the above conclusion would only hold on a set of full measure.)

This example shows that the seemingly mild assumption (which we will see satisfied frequently later) that there are different points with the same leaf-wise measures (after moving the measures to T as we did) is a rather special property of the underlying measure μ.

6.5.3. The final example is really more than an example – it is the reason we are developing the theory of leaf-wise measures and we will return to it in great detail (and greater generality) in the following sections. Let G be a Lie group, let T

be a closed subgroup, and let Γ be a discrete subgroup of G. Then T acts by right translation on $X = \Gamma\backslash G$, i.e., for $t \in T$ and $x = \Gamma g \in X$ we may define $t.x = xt^{-1}$. For a probability measure μ on X we have therefore a system of leaf-wise measures μ_x^T defined for a.e. $x \in X$ (provided the injectivity requirement is satisfied a.e.) which as we will see describes the properties of the measure along the direction of T. Moreover, if right translation by some $a \in G$ preserves μ, then with the correctly chosen subgroup T (namely the horospherical subgroups defined later) the leaf-wise measures for T will allow us to describe entropy of a w.r.t. μ.

The following definition and the existence established in Proposition 6.7 established afterwards will be a crucial tool for proving Theorem 6.3.

6.6. Definition. *Let $E \subset X$ be measurable and let $r > 0$. We say $C \subset X$ is an* r-cross-section for E *if*

(i) *C is Borel measurable,*
(ii) $|B_{r+1}^T.x \cap C| = |B_1^T.x \cap C| = 1$ *for all $x \in E \cup C$,*
(iii) *$t \in B_{r+1}^T \mapsto t.x$ is injective for all $x \in C$,*
(iv) *$B_{r+1}^T.x \cap B_{r+1}^T.x' = \emptyset$ if $x \neq x' \in C$, and*
(v) *the restriction of the action map $(t, x) \mapsto t.x$ to $B_{r+1}^T \times C \to B_{r+1}^T.C \supseteq B_r^T.E$ is a Borel isomorphism.*

The second property describes the heart of the definition; the piece $B_{r+1}^T.x$ of the T-orbit through $x \in E$ intersects C exactly once which justifies the term cross-section, see Figure 2. Also note that by the second property there is for

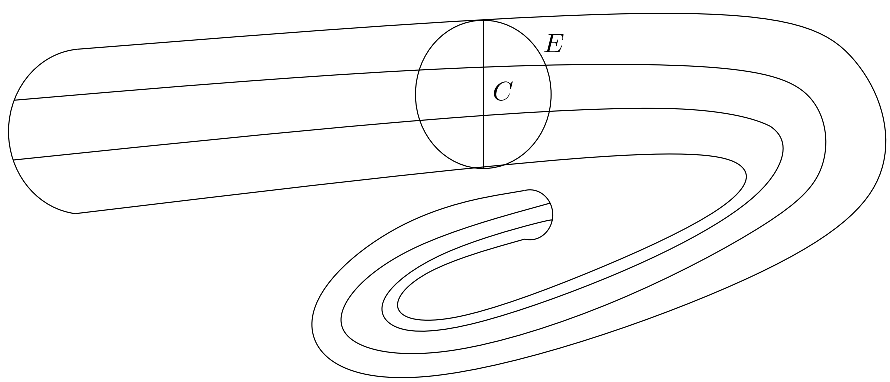

FIGURE 2. E (the circle) needs to be "small enough" in order for an r-cross section C (the vertical line through the circle) to exist. Otherwise, there may be large returns of points in E to E (in the picture if the circle is just a bit bigger) along the action of T (indicated by the curved lines).

every $x \in E$ some $t \in B_1^T$ with $t.x = x' \in C$. Hence, by right invariance of the metric on T we have $B_r^T t^{-1} \subset B_{r+1}^T$ and so the inclusion $B_{r+1}^T.C \supseteq B_r^T.E$ stated in the final property follows from the second property. Moreover, it is clear that the restriction of the continuous action is measurable, so the only requirement in the final property is injectivity of the map and the Borel measurability of the inverse. However, injectivity of this map is precisely the assertion in property (iii) and (iv). Finally, the measurability of the image and the inverse map are guaranteed by a general fact, see [**Sri98**, §4.5], saying that the image and the inverse of an injective Borel map are again Borel measurable. The reader who is unfamiliar with this theorem may construct (replacing the following general proposition) concrete cross-sections of sufficiently small balls in the important example in §6.5.3 using a transverse subspace to the Lie algebra of T inside the Lie algebra of G. This way one may obtain a compact cross-section and this implies measurability of the inverse map rather directly as the restriction of a continuous map to a compact set has compact image and a continuous inverse.

6.7. Proposition. *Let T act continuously on X satisfying the assumptions discussed in the beginning of this section. Assume $x_0 \in X$ is such that $t \in \overline{B_{r+1}^T} \mapsto t.x_0$ is injective for some $r > 1$. Then there exists some $\delta > 0$ such that for all $x \in E = \overline{B_\delta(x_0)}$ the map $t \in B_{r+1}^T \mapsto t.x$ is also injective and such that $t.x = t'.x'$ for some $x, x' \in E$ and $t, t' \in B_{r+1}^T$ implies $t't^{-1}, t^{-1}t' \in B_1^T$ and so $x' \in B_1^T.x$. Moreover, there exists some $C \subset E$ which is an r-cross-section for E.*

6.8**. Problem:** Prove the proposition in the case where $X = \Gamma\backslash G$ for a Lie group G (or a p-adic Lie group) and a closed subgroup $T < G$ by using a transverse to the Lie algebra of T as suggested above. The reader interested in only these cases may continue with §6.14.

6.9**. Proof, Construction of** E**:** If for every δ there exists some $x_\delta \in B_\delta(x_0)$ for which the restricted action $t \in B_{r+1}^T \mapsto t.x_\delta$ fails to be injective then there are $t_\delta \neq t'_\delta \in B_{r+1}^T$ with $t_\delta.x_\delta = t'_\delta.x_\delta$. Choosing converging subsequences of t_δ, t'_δ we get $t, t' \in \overline{B_{r+1}^T}$ with $t.x_0 = t'.x_0$. Moreover, we would have $t \neq t'$ as otherwise we would get a contradiction to the uniform local freeness of the action in §6.1 for the compact set $\overline{B_{r+1}^T}.\overline{B_\epsilon(x_0)}$ (where ϵ is small enough so that $\overline{B_\epsilon(x_0)}$ is compact).

Similarly, if for every $\delta > 0$ there are $x_\delta, x'_\delta \in B_\delta(x_0)$ and $t_\delta, t'_\delta \in B_{r+1}^T$ so that $t_\delta.x_\delta = t'_\delta.x'_\delta$ then in the limit we would have $t, t' \in \overline{B_{r+1}^T}$ with $t.x_0 = t'.x_0$. By assumption this implies $t = t'$, which shows that for sufficiently small δ, we must have $t'_\delta t_\delta^{-1}, t_\delta^{-1}t'_\delta \in B_1^T$ as claimed. Also notice that $\left(B_1^T\right)^{-1} = B_1^T$ by right invariance of the metric.

We now fix some $\delta > 0$ with the above properties and let $E = \overline{B_{\delta/2}(x_0)}$. Below we will construct a Borel subset $C \subset E$ such that $|B_1^T.x \cap C| = 1$ for all $x \in E$. This implies that C is an r-cross-section by the above properties: $t \in B_{r+1}^T$ and $x \in E$ with $t.x \in C \subset E$ implies $t \in B_1^T$ and so property (ii) of the definition holds. Injectivity of $t \in B_{r+1}^T \mapsto t.x$ for all $x \in E$ we have already checked. For the property (iv), note that $x, x' \in C$ and $t, t' \in B_{r+1}^T$ with $t.x = t'.x$ implies $x = t^{-1}t'.x' \in B_1^T.x'$ by the construction of E and so $x = x'$ by the assumed property of C. As explained after the definition the last property follows from the first four. Hence it remains to find a Borel subset $C \subset E$ with $|B_1^T.x \cap C| = 1$ for all $x \in E$.

6.10**. Outline of construction of** C**:** We will construct C by an inductive procedure where at every stage we define a set $C_{n+1} \subset C_n$ such that for every $x \in E$ the set $\{t \in B_1^T : t.x \in C_n\}$ is nonempty, compact, and the diameter of this set decreases to 0 as $n \to \infty$.

6.11. **Construction of P_w:** For the construction of C_n we first define for every n a partition of E which refines all prior partitions: For $n = 1$ we choose a finite cover of E by closed balls of radius[(21)] 1, choose some order of these balls, and define P_1 to be the first ball in this cover intersected with E, P_2 the second ball intersected with E minus P_1, and more generally if $P_1, \dots, P_i$ have been already defined then P_{i+1} is the $(i+1)$-th ball intersected with E and with $P_1 \cup \cdots \cup P_i$ removed from it.

For $n = 2$ we cover P_1 by finitely many closed balls of radius $1/2$ and construct with the same algorithm as above a finite partition of P_1 into sets $P_{1,1}, \dots, P_{1,i_1}$ of diameter less than $1/2$. We repeat this also for $P_2, \dots$

Continuing the construction we assume that we already defined the sets P_w where w is a word of length $|w| \le n$ (i.e., w is a list of m natural numbers and m is called the length $|w|$) with the obvious compatibilities arising from the construction: for any w of length $|w| = m \le n-1$ the sets $P_{w,1}, P_{w,2}, \dots$ (there are only finitely many) all have diameter less than $1/m$ and form a partition of P_w.

Roughly speaking, we will use these partitions to make decisions in a selection process: Given some $x \in E$ we want to make sure that there is one and only one element of the desired set C that belongs to $B_1^T.x$. Assuming this is not the case for $C = E$ (which can only happen for discrete groups T) we wish to remove, in some inductive manner obtaining the sets C_n along the way, some parts of E so as to make this true for the limiting object $C = \bigcap_n C_n$. Removing too much at once may be fatal as we may come to the situation where $B_1^T.x \cap C_n$ is empty for some $x \in E$. The partition elements P_w give us a way of ordering the elements of the space which we will use below.

6.12. **Definition of Q_w and C_n:** From the sequence of partitions defined by $\{P_w : w \text{ is a word of length } n\}$ we now define subsets $Q_w \subset P_w$ to define the C_n: We let $Q_1 = P_1$, and let $Q_2 = P_2 \setminus B_1^T.Q_1$, i.e. we remove from P_2 all points that already have on their B_1^T-orbit a point in Q_1. More generally, we define $Q_i = P_i \setminus \left(B_1^T.(Q_1 \cup \cdots \cup Q_{i-1})\right)$ for all i and define $C_1 = \bigcup_i Q_i$ (which as before is just a finite union). We now prove the claim from §6.10 for $n = 1$ that for every $x \in E$ the set $\{t \in B_1^T : t.x \in C_1\}$ is nonempty and compact. Here we will use without explicitly mentioning, as we will also do below, the already established fact that $t \in B_2^T$ and $x, t.x \in E$ implies $t \in B_1^T$ (note that by assumption $r > 1$). If i is chosen minimally with $B_1^T.x \cap P_i$ nonempty, then

$$\{t \in B_1^T : t.x \in C_1\} = \{t \in B_1^T : t.x \in Q_i\} = \\ \{t \in B_1^T : t.x \in P_i\} = \{t \in \overline{B_1^T} : t.x \in P_1 \cup \cdots \cup P_i\}.$$

Now note that $P_1 \cup \cdots \cup P_i$ is closed by the above construction (we used closed balls to cover E and $P_1 \cup \cdots \cup P_i$ equals the union of the first i balls intersected with E, a closed ball itself), and so the claim follows for $n = 1$ and any $x \in E$.

Proceeding to the general case for n, we assume $Q_w \subset P_w$ has been defined for $|w| = m < n$ with the following properties: we have $Q_{w,i} \subset Q_w$ for $i = 1, 2, \dots$ and for all $|w| < n-1$, for $|w| = |w'| < n$ and $w \ne w'$ the sets $B_1^T.Q_w$ and $B_1^T.Q_{w'}$ are disjoint, and the claim holds for $C_m = \bigcup\{Q_w : |w| = m\}$ and all $m < n$. Now fix some word w of length $n-1$, we define $Q_{w,1} = Q_w \cap P_{w,1}$,

[(21)] We ignore, for simplicity of notation, the likely possibility that $\delta < 1$.

$Q_{w,2} = Q_w \cap P_{w,2} \setminus (B_1^T.Q_{w,1})$, and for a general i we define inductively

$$Q_{w,i} = Q_w \cap P_{w,i} \setminus \big(B_1^T.(Q_{w,1} \cup \cdots \cup Q_{w,i-1})\big).$$

By the inductive assumption we know that for a given $x \in E$ there is some w of length $n-1$ such that the set

$$\{t \in B_1^T : t.x \in C_{n-1}\} = \{t \in B_1^T : t.x \in Q_w\} \tag{6.12a}$$

is closed and nonempty. Choose i minimally such that $B_1^T.x \cap Q_{w,i}$ (or equivalently $B_1^T.x \cap Q_w \cap P_{w,i}$) is nonempty, then as before

$$\{t \in B_1^T : t.x \in C_n\} = \{t \in B_1^T : t.x \in Q_{w,i}\} = \big\{t \in B_1^T : t.x \in Q_w \cap (P_{w,1} \cup \cdots \cup P_{w,i})\big\} \tag{6.12b}$$

is nonempty. Now recall that by construction $P_{w,1} \cup \cdots \cup P_{w,i}$ is relatively closed in P_w, so that the set in (6.12b) is relatively closed in the set in (6.12a). The latter is closed by assumption which concludes the induction that indeed for every n the set $\{t \in B_1^T : t.x \in C_n\}$ is closed and nonempty.

6.13. **Conclusion:** The above shows that $C_n = \bigcup_w Q_w$ (where the union is over all words w of length n) satisfies the claim that $\{t \in B_1^T : t.x \in C_n\}$ is compact and non-empty for every $x \in E$. Therefore, $C = \bigcap_n C_n \subset E$ satisfies that $C \cap B_1^T.x \neq \emptyset$ for every $x \in E$. Suppose now $t_1.x, t_2.x \in C$ for some $x \in E$ and $t_1, t_2 \in B_1^T$. Fix some $n \geq 1$. Recall that $\{t \in B_1^T : t.x \in C_n\} = \{t \in B_1^T : t.x \in Q_w\}$ for some Q_w corresponding to a word w of length n. As the diameter of $Q_w \subset P_w$ is less than $1/n$ we have $d(t_1.x, t_2.x) < 1/n$. This holds for every n, so that $t_1.x = t_2.x$ and so $t_1 = t_2$ as required. □

6.14. **σ-algebras.** Proposition 6.7 allows us to construct σ-algebras as they appear in Theorem 6.3(iv) in abundance. In fact we have found closed balls E and r-cross-sections $C \subset E$ such that $B_{r+1}^T \times C$ is measurably isomorphic to $Y = B_{r+1}^T.C$ (with respect to the natural map) so that we may take the countably generated σ-algebra on $B_{r+1}^T \times C$ whose atoms are of the form $B_{r+1}^T \times \{z\}$ for $z \in C$ and transport it to Y via the isomorphism. As we will work very frequently with σ-algebras of that type we introduce a name for them.

6.15. Definition. *Let $r > 1$. Given two measurable subsets $E \subset Y$ of X and a countably generated σ-algebra $\mathcal{A}$ of subsets of Y, we say that $(Y, \mathcal{A})$ is an (r,T)-*flower with base E*, if and only if:*

(i) *For every $x \in E$, we have that $[x]_{\mathcal{A}} = U_x.x$ is an open T-plaque such that $B_r^T \subset U_x \subset B_{r+2}^T$.*
(ii) *Every $y \in Y$ is equivalent to some $x \in E$, i.e., the atom $[y]_{\mathcal{A}} = [x]_{\mathcal{A}}$ is always an open T-plaque intersecting E nontrivially.*

We note that often the cross-section C will be a nullset (for the measure μ on X), but that the base E will not be a null set, hence it is important to introduce it — it may be thought of as a slightly thickened version of the cross-section so that we still know the rough shape of the atoms as required in (i). We may visualize the flower and the base using Figure 2. The base is the circle and the flower is the σ-algebra on the tube-like set whose atoms are the curved lines.

6.16. Corollary. *Assume as in Theorem 6.3 that $t \mapsto t.x$ for $t \in T$ is injective for μ-a.e. $x \in X$. Then for every n there exists a countable list of (n,T)-flowers such that the union of their bases is a set of full measure. In other words, there exists a countable collection of σ-algebras $\mathcal{A}_k$ of Borel subsets of Borel sets Y_k for $k = 1, 2, \ldots$ such that all of the $\mathcal{A}_k$-atoms are open T-plaques for all k, and such that for a.e. $x \in X$ and all $n \geq 1$ there exists k such that the $\mathcal{A}_k$-atom $[x]_{\mathcal{A}_k}$ contains $B_n^T.x$.*

6.17**. Proof.** By our assumption there exists a set X_0 of full measure such that $t \in T \mapsto t.x_0$ is injective for $x \in X_0$. Fix some n. By Prop. 6.7 applied to $r = n$ there exists an uncountable collection of closures E_x of balls for $x \in X_0$ such that x is contained in the interior E_x° and there is an n-cross-section $C_x \subset E_x$ for $x \in X_0$. Since X is second countable, there is a countable collection of these sets $C_m \subset E_m$ for which the union of the interiors is the same as the union of interiors of all of them.

As C_m is an n-cross-section for E_m, we have that $B_{n+1}^T.C_m \supset B_n^T.E_m$ and that $B_{n+1}^T \times C_m$ is measurably isomorphic to $Y_m = B_{n+1}^T.C_m$. We now define $\mathcal{A}_m$ to be the σ-algebra of subsets of Y_m which corresponds under the isomorphism to $\{B_{n+1}^T, \emptyset\} \otimes \mathcal{B}_{C_m}$ — here $\mathcal{B}_{C_m}$ is the Borel σ-algebra of the set C_m. It is clear that $\mathcal{A}_m$ is an (n,T)-flower with base E_m. Using this construction for all n, we get the countable list of (n,T)-flowers as required. □

It is natural to ask how the various σ-algebras in the above corollary fit together, where the next lemma gives the crucial property.

6.18. Lemma. *Let Y_1, Y_2 be Borel subsets of X, and $\mathcal{A}_1, \mathcal{A}_2$ be countably generated σ-algebras of Y_1, Y_2 respectively, such that atoms of each $\mathcal{A}_i$ are open T-plaques. Then the σ-algebras $\mathcal{C}_1 := \mathcal{A}_1|_{Y_1 \cap Y_2}$ and $\mathcal{C}_2 := \mathcal{A}_2|_{Y_1 \cap Y_2}$ are countably equivalent.*

6.19**. Proof:** Let $x \in Y_1 \cap Y_2$, and consider $[x]_{\mathcal{C}_1} = [x]_{\mathcal{A}_1} \cap Y_2$. By this and the assumption on $\mathcal{A}_1$ there exists a bounded set $U \subset T$ such that $[x]_{\mathcal{C}_1} = U.x$. Now, for each $t \in U$, we have the open T-plaque $[t.x]_{\mathcal{A}_2}$, which must be of the form $U_t.x$ for some open, bounded $U_t \subset T$. Now the collection $\{U_t\}_{t \in U}$ covers U, and since T is locally compact second countable, there exists a countable subcollection of the $\{U_t\}$ covering U. But this means that a countable collection of atoms of $\mathcal{A}_2$ covers $[x]_{\mathcal{C}_1}$; we then intersect each atom with Y_1 to get atoms of $\mathcal{C}_2$. Switch $\mathcal{C}_1$ and $\mathcal{C}_2$ and repeat the argument to get the converse. □

6.20**. Proof of Theorem 6.3, beginning.** We now combine Corollary 6.16, Lemma 6.18, and Proposition 5.17: Let $\mathcal{A}_k$ be the sequence of σ-algebras of subsets of Y_k as in Corollary 6.16. We define $Y_{k,\ell} = Y_k \cap Y_\ell$ and get that $(\mathcal{A}_k)|_{Y_{k,\ell}}$ and $(\mathcal{A}_\ell)|_{Y_{k,\ell}}$ are countably equivalent by Lemma 6.18. By Proposition 5.17 we get that

$$\mu_x^{\mathcal{A}_k}|_{[x]_{\mathcal{A}_\ell}} \text{ and } \mu_x^{\mathcal{A}_\ell}|_{[x]_{\mathcal{A}_k}} \tag{6.20a}$$

are proportional for a.e. $x \in Y_{k,\ell}$ (where we used additionally that the conditional measure for $\mu|_{Y_{k,\ell}}$ with respect to the σ-algebra $\mathcal{A}_k|_{Y_{k,\ell}}$ is just the normalized restriction of $\mu_x^{\mathcal{A}_k}$ to $Y_{k,\ell}$). Also recall that by Theorem 5.9(iii) for every k there is a null set in Y_k such that for $x, y \in Y_k$ not belonging to this null set and $[x]_{\mathcal{A}_k} = [y]_{\mathcal{A}_k}$ we have $\mu_x^{\mathcal{A}_k} = \mu_y^{\mathcal{A}_k}$. We collect all of these null sets to one null set $N \subset X$ and let X'' be the set of all points $x \in X \setminus N$ for which $t \mapsto t.x$ is injective. By construction of $\mathcal{A}_k$ we have $[x]_{\mathcal{A}_k} = U_{x,k}.x$ for some open and bounded $U_{x,k} \subset T$. For a bounded measurable set $D \subset T$ and $x \in X''$ we define

$$\mu_x^T(D) = \frac{1}{\mu_x^{\mathcal{A}_k}(B_1^T.x)} \mu_x^{\mathcal{A}_k}(D.x) \tag{6.20b}$$

where we choose k such that $D.x \subset [x]_{\mathcal{A}_k}$ which by the construction of the sequence of σ-algebras, i.e., by Corollary 6.16, is possible. Notice this definition is independent of k by the proportionality of the conditional measures in (6.20a).

However, we need to justify this definition by showing that the denominator does not vanish, at least for a.e. $x \in X''$. We prove this in the following lemma which will also prove Theorem 6.3(v).

6.21. Lemma. *Suppose $\mathcal{A}$ is a countably generated sub-σ-algebra of Borel subsets of a Borel set $Y \subset X$. Suppose further that the $\mathcal{A}$-atoms are open T-plaques. Let $U \subset T$ be an open neighborhood of the identity. Then for μ-a.e. $x \in Y$, we have $\mu_x^{\mathcal{A}}(U.x) > 0$.*

6.22**. Proof:** Set $B = \{x \in Y' : \mu_x^{\mathcal{A}}(U.x) = 0\}$, where $Y' \subset Y$ is a subset of full measure on which the conclusion of Theorem 5.9(iii) holds. We wish to show that $\mu(B) = 0$, and since we can integrate first over the atoms and then over the space (Theorem 5.9(i) and Proposition 4.1), it is sufficient to show for each $x \in Y'$ that $\mu_x^{\mathcal{A}}(B) = \mu_x^{\mathcal{A}}([x]_{\mathcal{A}} \cap B) = 0$. Now since atoms of $\mathcal{A}$ are open T-plaques, we can write $[x]_{\mathcal{A}} = (U_x).x$. Set $V_x \subset U_x$ to be the set of those t such that $t.x \in [x]_{\mathcal{A}} \cap B$.

Now clearly the collection $\{Ut\}_{t \in V_x}$ covers V_x, and we can find a countable subcollection $\{Ut_i\}_{i=1}^{\infty}$ that also covers V_x. This implies that $\{(Ut_i).x\}_{i=1}^{\infty}$ covers $[x]_{\mathcal{A}} \cap B$ by definition of V_x, so we have

$$\mu_x^{\mathcal{A}}([x]_{\mathcal{A}} \cap B) \leq \mu_x^{\mathcal{A}}(\bigcup_{i=1}^{\infty}(Ut_i).x) \leq \sum_{i=1}^{\infty} \mu_x^{\mathcal{A}}((Ut_i).x)$$

On the other hand, $t_i.x \in B$, so by definition of B we have that each term $\mu_x^{\mathcal{A}}((Ut_i).x) = \mu_x^{\mathcal{A}}(U.(t_i.x))$ on the right-hand side is 0. □

6.23**. Proof of Theorem 6.3, summary.** We let $X' \subset X''$ be a subset of full measure such that the conclusion of Lemma 6.21 holds for the σ-algebra $\mathcal{A}_k$, all $x \in Y_k \cap X'$, all k, and every ball $U = B_{1/n}^T$ for all n. This shows that for $x \in X'$ the expression on the right of (6.20b) is well defined. By the earlier established property it is also independent of k (as long as $D.x \subset [x]_{\mathcal{A}_k}$ as required before). Therefore, (6.20b) defines a Radon measure on T satisfying Theorem 6.3 (v). Property (iii) follows directly from the definition and the requirement that for $x, g.x \in X'' \cap Y_k$ with $[x]_{\mathcal{A}_k} = [g.x]_{\mathcal{A}_k}$ (which will be the case for many k) we have $\mu_x^{\mathcal{A}_k} = \mu_{g.x}^{\mathcal{A}_k}$, where we may have a proportionality factor appearing as μ_x^T is normalized via the set $B_1^T.x$ and $\mu_{g.x}^T$ is normalized via the set $B_1^T g.x$. Property (iv) follows from Lemma 6.18 and Proposition 5.17 similar to the discussion in 6.20. We leave property (ii) to the reader. □

We claimed before that the leaf-wise measure describes properties of the measure μ along the direction of the T-leaves, we now give three examples of this.

6.24**. Problem:** The most basic question one can ask is the following: What does it mean to have $\mu_x^T \propto \delta_e$ a.e.? Here δ_e is the Dirac measure at the identity of T, and this case is often described as *the leaf-wise measures are trivial a.e.* Show this happens if and only if there is a global cross-section of full measure, i.e., if there is a measurable set $B \subset X$ with $\mu(X \setminus B) = 0$ such that $x, t.x \in B$ for some $t \in T$ implies $t = e$.

6.25. Definition. *Suppose we have a measure space X, a group T acting on X, and μ a locally finite measure on X. Then μ is* T-recurrent *if for every measurable $B \subset X$ of positive measure, and for a.e. $x \in B$, the set $\{t : t.x \in B\}$ is unbounded (i.e., does not have compact closure in T).*

6.26. Theorem. *Let X, T, μ be as before, and suppose additionally that μ is a probability measure. Then μ is T-recurrent if and only if μ_x^T is infinite for almost every x.*

6.27**. Proof:** Assume T-recurrence. Let $Y = \{x : \mu_x^T(T) < \infty\}$, and suppose that $\mu(Y) > 0$. We may find a sufficiently large n such that the set $Y' = \{x \in Y : \mu_x^T(B_n^T) > 0.9\mu_x^T(T)\}$ also has positive measure. We will show that, for any $y \in Y'$, the set of return times $\{t : t.y \in Y'\}$ is bounded; in fact, that $\{t : t.y \in Y'\} \subset B_{2n}^T$ for any $y \in Y'$. Since $\mu(Y') > 0$, this then shows that μ is not T-recurrent.

Pick any return time t. By definition of Y', we know that $\mu_y^T(B_n^T) > 0.9\mu_y^T(T)$ and $\mu_{t.y}^T(B_n^T) > 0.9\mu_{t.y}^T(T)$. On the other hand, from Theorem 6.3.(iii) we know that $\mu_{t.y}^T \propto (\mu_y^T)t$, so that we have $\mu_y^T(B_n^T t) > 0.9\mu_y^T(Tt) = 0.9\mu_y^T(T)$. But now we have two sets B_n^T and $B_n^T t$ of very large μ_y^T measure, and so we must have $B_n^T \cap B_n^T t \neq \emptyset$. This means $t \in (B_n^T)^{-1}B_n^T$, as required.

Assume now that the leaf-wise measures satisfy $\mu_x^T(T) = \infty$ for a.e. x, but μ is not T-recurrent. This means there exists a set B of positive measure, and some compact $K \subset T$ such that $\{t : t.x \in B\} \subset K$ for every $x \in B$.

We may replace B by a subset of B of positive measure and assume that $B \subset E$ for a measurable $E \subset X$ for which there is an r-cross-section $C \subset E$ as in Proposition 6.7, where we chose r sufficiently big so that $B_r^T \supset B_1^T K B_1^T$. Let $(B_{r+1}^T.C, \mathcal{A})$ be the (r,T)-flower for which the atoms are of the form $B_{r+1}^T.z$ for $z \in C$. As C is a cross-section, the atoms of $\mathcal{A}$ are in one-to-one correspondence with elements of C. We define $D = \{z \in C : \mu_z^{\mathcal{A}}(B) > 0\}$, where we may require that $\mu_x^{\mathcal{A}}$ is defined on a set $X' \in \mathcal{A}$ and is strictly $\mathcal{A}$-measurable by removing possibly a null set from B. Therefore, the definition of D as a subset of the likely nullset C makes sense. Note that $B \setminus (B_{r+1}^T.D)$ is a null set, and so we may furthermore assume $B \subset B_1^T.D$ by the properties of C and E in Proposition 6.7.

Suppose now $t.z = t'.z'$ for some $t, t' \in T$ and $z, z' \in D$. By construction of D and by Proposition 6.7 we may write $z = t_x.x$ and $z' = t_{x'}.x'$ for some $t_x, t_{x'} \in B_1^T$ and $x, x' \in B$. Therefore, $tt_x.x = t't_{x'}.x'$ which implies that $t_{x'}^{-1}(t')^{-1}tt_x \in K$ by the assumed property of B. Thus $(t')^{-1}t \in B_1^T K B_1^T \subset B_r^T$, which implies $t = t'$ and $z = z'$ since $C \supset D$ is an r-cross-section. This shows that for every n we have that $B_{n+1}^T \times D \to B_{n+1}^T.D$ is injective and just as in Corollary 6.16 this gives rise to the (n,T)-flower $(B_{n+1}^T.D, \mathcal{A}_n)$ with center $B_1^T.D$ such that the atoms are of the form $B_{n+1}^T.z$ for $z \in D$.

By Theorem 6.3.(iv), we know that

$$\mu_x^{\mathcal{A}_n}(B) = \frac{\mu_x^T\big(\{t \in U_{x,n} : t.x \in B\}\big)}{\mu_x^T(U_{x,n})}$$

for a.e. $x \in B_n^T.D$. Here $U_{x,n} \subset T$ is the shape of the atom, i.e., is such that $[x]_{\mathcal{A}_n} = U_{x,n}.x$. Clearly, for $z \in D$ we have $U_{z,n} = B_{n+1}^T$ by construction. Therefore, we have for $y \in B \subset E \subset B_1^T.C$ that $U_{y,n} \supset B_n^T$. Also recall that by assumption $y \in B$, $t \in T$, and $t.y \in B$ implies $t \in K$. Together we get for a.e. $y \in B$ that

$$\mu_y^{\mathcal{A}_n}(B) \leq \frac{\mu_y^T(K)}{\mu_y^T(B_n^T)},$$

which approaches zero for a.e. $y \in B$ as $n \to \infty$ by assumption on the leaf-wise measures.

We define

$$B' = \{y \in B : \mu_y^{\mathcal{A}_n}(B) \to 0\},$$

which by the above is a subset of B of full measure. We also define the function f_n by the rule $f_n(x) = 0$ if $x \notin B_n^T.D$ and $f_n(x) = \mu_x^{\mathcal{A}_n}(B')$ if $x \in B_n^T.D$. Clearly, if $y \notin T.D$ then $f_n(y) = 0$ for all n. While if $y \in B_{n_0}^T.D$ and $f_{n_0}(y) = \mu_x^{\mathcal{A}_{n_0}}(B') > 0$ for some n_0 then we may find some $x \in B'$ equivalent to y with respect to all $\mathcal{A}_n$ for $n \geq n_0$, so that $f_n(y) = f_n(x)$ for $n \geq n_0$ by the properties of conditional measures. Therefore, $f_n(y) \to 0$ for a.e. $y \in X$. By dominated convergence (μ is a finite measure by assumption and $f_n \leq 1$) we have

$$\mu(B) = \int_{B_n^T.D} \mu_x^{\mathcal{A}_n}(B')d\mu = \int f_n d\mu \to 0,$$

i.e., $\mu(B) = 0$ contrary to the assumptions.□

6.28**. Problem:** With triviality of leaf-wise measures as one possible extreme for the behavior of μ along the T-leaves and T-recurrence in between, on the opposite extreme we have the following fact: μ is T-invariant if and only if the leaf-wise measures μ_x^T are a.e. left Haar measures on T. Show this using the flowers constructed in Corollary 6.16.

6.29**. Normalization.** One possible normalization of the leaf-wise measure μ_x^T, which is uniquely characterized by its properties up to a proportionality factor, is to normalize by a scalar (depending on x measurably) so that $\mu_x^T(B_1^T) = 1$. However, under this normalization we have no idea how big $\mu_x^T(B_n^T)$ can be for $n > 1$.

It would be convenient if the leaf-wise measures μ_x^T would belong to a fixed compact metric space in a natural way — then we could ask (and answer in a positive manner) the question whether the leaf-wise measures depend measurably on x where we consider the natural Borel σ-algebra on the compact metric space. Compare this with the case of conditional measures $\mu_x^{\mathcal{A}}$ for a σ-algebra $\mathcal{A}$ and a finite measure μ on a compact metric space X, here the conditional measures belong to the compact metric space of probability measures on X (where we use the weak* topology on the space of measures). Unfortunately, the lack of a bound of $\mu_x^T(B_2^T)$ shows, with μ_x^T normalized using the unit ball, that the leaf-wise measures do not belong to a compact subset in the space of Radon measures (using the weak* topology induced by compactly supported continuous functions on T). For that reason we are interested[22] in the possibly growth rate of $\mu_x^T(B_n^T)$, so that we can introduce a different normalization with respect to which we get values in a compact metric space.

---

[22]While convenient, this theorem is not completely necessary for the material presented in the following sections. The reader who is interested in those could skip the proof of this theorem and return to it later.

6.30. Theorem. *Assume in addition to the assumptions of Theorem 6.3 that μ is a probability measure on X and that T is unimodular. Denote the bi-invariant Haar measure on T by λ. Fix weights b_n such that $\sum_{n=1}^{\infty} b_n^{-1} < \infty$ (eg., think of $b_n = n^2$) and a sequence $r_n \nearrow \infty$. Then for μ-a.e. x we have*

$$\lim_{n\to\infty} \frac{\mu_x^T(B_{r_n}^T)}{b_n \lambda(B_{r_n+5}^T)} = 0$$

where B_r^T is the ball of radius r around $e \in T$.

In other words, the leaf-wise measure of big balls $B_{r_n}^T$ can't grow much faster than the Haar measure of a slightly bigger ball $B_{r_n+5}^T$. This is useful as it gives us a function $f : T \to \mathbb{R}^+$ which is integrable w.r.t. μ_x^T for a.e. $x \in X$, e.g. $f(x) = \frac{1}{b_n^2 \lambda(B_{r_n+5}^T)}$ for $x \in B_{r_n}^T \setminus B_{r_{n-1}}^T$. Hence we may normalize μ_x^T such that $\int_T f d\mu_x^T = 1$ and we get that μ_x^T belongs to the compact metric space of measures ν on T for which $\int_T f d\nu \leq 1$, where the latter space is equipped with the weak* topology induced by continuous functions with compact support. Hence it makes sense, and this is essentially Theorem 6.3.(ii), to ask for measurable dependence of μ_x^T as a function of x.

Before proving this theorem, we will need the following refinement regarding the existence of (r, T)-flowers.

6.31. Lemma. *For any measurable set $B \subset X$, $R > 0$, we can find a countable collection of (R, T)-flowers $(Y_k, \mathcal{A}_k)$ with base E_k so that*

(i) *any $x \in X$ is contained in only finitely many bases E_k, in fact the multiplicity is bounded with the bound depending only on T,*
(ii) *$\mu(B \setminus \bigcup_k E_k) = 0$,*
(iii) *for every $x \in E_k$ there is some $y \in [x]_{\mathcal{A}_k} \cap E_k \cap B$ so that*

$$B_1^T.y \subset [x]_{\mathcal{A}_k} \cap E_k,$$

for any two equivalent[23] *$x, y \in E_k$ we have $[x]_{\mathcal{A}_k} \cap E_k \subset B_4^T.y$, and*
(iv) *for every $x \in Y_k$ there is some $y \in [x]_{\mathcal{A}_k} \cap E_k \cap B$.*

The third property may, loosely speaking, be described as saying that for points x in the base E_k we require that there is some $y \in B \cap E_k$ equivalent to x such that y is deep inside the base E_k (has distance one to the complement) in the direction of T.

6.32**. Proof:** By Corollary 6.16 we already know that we can cover a subset of full measure by a countable collection of bases $\tilde{E}_k$ of $(R+1, T)$-flowers $(\tilde{Y}_k, \tilde{\mathcal{A}}_k)$ such that additionally there is some $(R+2)$-cross-section $\tilde{C}_k \subset \tilde{E}_k$, $\tilde{Y}_k = B_{R+2}^T.\tilde{C}_k$, and $\tilde{E}_k \subset B_1^T.\tilde{C}_k$. We will construct Y_k by an inductive procedure as subsets of $\tilde{Y}_k$ and will use the restriction $\mathcal{A}_k$ of $\tilde{\mathcal{A}}_k$ to Y_k as the σ-algebra.

For $k = 1$ we define

$$Y_1 = \left\{ x \in \tilde{Y}_1 : \mu_x^{\tilde{\mathcal{A}}_1}(B \cap \tilde{E}_1) > 0 \right\}, \tag{6.32a}$$

and $\mathcal{A}_1 = \tilde{\mathcal{A}}_1|_{Y_1}$. By definition we remove from $\tilde{Y}_1$ complete atoms to obtain Y_1, so that the shape of the remaining atoms is unchanged. From this it follows that

[23] Recall that x and y are equivalent w.r.t. $\mathcal{A}_k$ if $[x]_{\mathcal{A}_k} = [y]_{\mathcal{A}_k}$.

$(Y_1, \mathcal{A})$ is an $(R+1,T)$-flower with base $\tilde{E}_1 \cap Y_1$. Also note that $B \cap \tilde{E}_1 \cap Y_1$ is a subset of full measure of $B \cap \tilde{E}_1$ (cf. (5.11a) and (6.32a)). We define

$$E_1 = B_2^T.(\tilde{C}_1 \cap Y_1) \supset \tilde{E}_1 \cap Y_1,$$

where the inclusion follows because $\tilde{E}_1 \subset B_1^T.\tilde{C}_1$ holds by construction of the original flowers. Since we constructed Y_1 by removing whole atoms from $\tilde{Y}_1$, we obtain $E_1 \subset Y_1$.

Finally, by definition of Y_1 we have $\mu_x^{\mathcal{A}_1}(B \cap \tilde{E}_1) > 0$ for every $x \in E_1 \subset Y_1$, so there must indeed be some $y \in B \cap \tilde{E}_1$ which is equivalent to x. Again because Y_1 was obtained from $\tilde{Y}_1$ by removing entire atoms, we have $y \in \tilde{E}_1 \cap Y_1$. Moreover, $y \in B_1^T.C_1$ so that $B_1^T.y \subset (B_2^T.C_1) \cap Y_1 = E_1$. The conclusions in (iii) follow now easily for the case $k = 1$. At last notice that $(Y_1, \mathcal{A}_1)$ is an (R,T)-flower with base E_1.

For a general k we assume that we have already defined for any $\ell < k$ an (R,T)-flower $(Y_\ell, \mathcal{A}_\ell)$ with bases E_ℓ satisfying: $Y_\ell \subset \tilde{Y}_\ell$ is obtained by removing entire $\tilde{\mathcal{A}}_\ell$-atoms, $\mathcal{A}_\ell = \tilde{\mathcal{A}}_\ell|_{Y_\ell}$, properties (iii) and (iv) hold, and that $B \cap \bigcup_{\ell<k} E_\ell$ contains $B \cap \bigcup_{\ell<k} \tilde{E}_\ell$ except possibly for a nullset. The latter is the inductive assumption regarding (ii) as at the end of the construction it will imply (ii) by the assumption that the bases $\tilde{E}_j$ for $j = 1, 2, \ldots$ cover a set of full measure.

We now define

$$Y_k = \bigl\{x \in \tilde{Y}_k : \mu_x^{\tilde{\mathcal{A}}_k}\bigl(B \cap \tilde{E}_k \setminus \bigcup_{\ell<k} E_\ell\bigr) > 0\bigr\},$$

which as before is $\tilde{Y}_k$ minus a union of complete $\tilde{\mathcal{A}}_k$-atoms. In particular, we again get that $(Y_k, \mathcal{A}_k)$ (with $\mathcal{A}_k = \tilde{\mathcal{A}}_k|_{Y_k}$) is an $(R+1,T)$-flower with base $\tilde{E}_k \cap Y_k$ and that $B \cap \tilde{E}_k \setminus \bigcup_{\ell<k} E_\ell$ is contained in $\tilde{E}_k \cap Y_k$ except possibly for a null set. The latter ensures the inductive assumption regarding (ii) if we define E_k as a superset of $\tilde{E}_k \cap Y_k$. We define $E_k = B_2^T.(Y_k \cap C_k)$ which implies $\tilde{E}_k \cap Y_k \subset E_k$ and also property (iii) similar to the case $k = 1$. Indeed, if $x \in E_k$, then $x = t.z$ for some $t \in B_2^T$ and $z \in Y_k \cap C_k$ which implies

$$\mu_z^{\mathcal{A}_k}\bigl(B \cap \tilde{E}_k \setminus \bigcup_{\ell<k} E_\ell\bigr) > 0$$

by definition of Y_k. Hence there is some $y \in B \cap \tilde{E}_k$ equivalent to z (and to x) with $y = t_y.z$ for some $t_y \in B_1^T$ by the properties of $\tilde{E}_k$. This implies $B_1^T.y \subset E_k$ as required.

Suppose now we have completed the above construction defining Y_k and E_k and assume that x belongs to $E_{k_1}, E_{k_2}, \ldots, E_{k_m}$ for some $k_1 < k_2 < \cdots < k_m$. We wish to bound m in order to proof (i). By property (iii) we know for $j = 1, \ldots, m$ that $x = t_j.y_j$ for some $t_j \in B_1^T$ and $y_j \in E_{k_j} \cap B$. In fact, by the construction we know that $y_j \in B \cap \tilde{E}_{k_j} \setminus \bigcup_{\ell<k_j} E_\ell$. Also notice that

$$t_j^{-1} t_i.y_i = t_j^{-1}.x = y_j \text{ for any pair } i, j.$$

However, since $B_1^T.y_i \subset E_{k_i}$ for $i < j$ we must have $t_j^{-1} t_i \notin B_1^T$. As the metric on T is assumed to be right invariant we conclude that the elements $t_1^{-1}, \ldots, t_{k_m}^{-1}$ have all distance ≥ 1, and so m is bounded by the maximal number of 1-separated elements of B_1^T which has compact closure. This proves (i). $\square$

6.33. **Proof of Theorem 6.30.** We fix some $\delta > 0$, and some integer M. We define

$$B_m = \left\{ y : \frac{\mu_y^T(B_{r_n}^T)}{\mu_y^T(B_4^T)} \geq b_n \delta \frac{\lambda(B_{r_n+5}^T)}{\lambda(B_4^T)} \text{ for at least } m \text{ different } n \leq M \right\}.$$

We want to give a bound on $\mu(B_m)$ which will be independent of M and tends to 0 as $m \to \infty$. Let $R = r_M$, and let E_i and $\mathcal{A}_i$ be as in Lemma 6.31. (Note that by the choice of R the sequence of σ-algebras depends crucially on M.)

Consider the function

$$G = \sum_{n=1}^{M} \sum_{i=1}^{\infty} w_n \chi_{B_{r_n}^T . E_i}$$

with $w_n = \frac{1}{b_n \lambda(B_{r_n+5}^T)}$ and where χ_A denotes the characteristic function of a set A. We claim that G is bounded, with the bound independent of M.

Fixing n and x, let $I = \left\{ i : x \in B_{r_n}^T . E_i \right\}$. For each $i \in I$, let $h_i' \in B_{r_n}^T$ be such that $h_i'.x \in E_i$, and by Lemma 6.31.(iii), we can modify h_i' to some $h_i \in B_{r_n+4}^T$ so that $B_1^T h_i . x \subset [x]_{\mathcal{A}_i} \cap E_i$.

As the multiplicity of the sets $E_1, E_2, \ldots$ is bounded by some constant c_1 (that only depends on T) and since $B_1^T h_i . x \subset E_i$ we get that

$$\sum_{i \in I} \chi_{B_1^T h_i} \leq c_1 \chi_{B_{r_n+5}}.$$

This implies that $|I| \, \lambda(B_1^T) \leq c_1 \lambda(B_{r_n+5}^T)$. We conclude that

$$\sum_{i=1}^{\infty} w_n \chi_{B_{r_n}^T . E_i}(x) \leq w_n \, |I| \leq \frac{c_1 \lambda(B_{r_n+5}^T)}{b_n \lambda(B_1^T) \lambda(B_{r_n+5}^T)} \leq \frac{c_2}{b_n},$$

where c_2 again only depends on T. Therefore, $G(x) \leq c_3 = c_2 \sum_{n=1}^{\infty} b_n^{-1}$ for all M as claimed.

On the other hand, consider the (R, T)-flower $(Y_i, \mathcal{A}_i)$ with base E_i. By the properties of leaf-wise measures (Theorem 6.3.(iv)) and Lemma 6.31.(iii), we know that for every $y \in E_i \cap B_m$ and every n,

$$\frac{\mu_y^{\mathcal{A}_i}(E_i)}{\mu_y^{\mathcal{A}_i}(B_{r_n}^T . y)} \leq \frac{\mu_y^T(B_4^T)}{\mu_y^T(B_{r_n}^T)}.$$

So if $z \in Y_i$ and $y \in [z]_{\mathcal{A}_i} \cap B_m \cap E_i$ (the existence of such a y is guaranteed by Lemma 6.31.(iv)), then $\chi_{B_{r_n}^T . E_i} \geq \chi_{B_{r_n}^T . y}$ and so

$$\int_{Y_i} \chi_{B_{r_n}^T . E_i} \, d\mu_z^{\mathcal{A}_i} \geq \mu_y^{\mathcal{A}_i}(B_{r_n}^T . y) \geq \frac{\mu_y^T(B_{r_n}^T)}{\mu_y^T(B_4^T)} \mu_z^{\mathcal{A}_i}(E_i).$$

Multiplying with w_n and summing over $n = 1, \ldots, M$ we get

$$\begin{aligned} \int_{Y_i} \sum_{n=1}^{M} w_n \chi_{B_{r_n}^T . E_i} \, d\mu_z^{\mathcal{A}_i} &\geq \sum_{n=1}^{M} \frac{1}{b_n \lambda(B_{r_n+5}^T)} \frac{\mu_y^T(B_{r_n}^T)}{\mu_y^T(B_4^T)} \mu_z^{\mathcal{A}_i}(E_i) \\ &\geq m\delta \frac{1}{\lambda(B_4^T)} \mu_z^{\mathcal{A}_i}(E_i) \end{aligned}$$

where the latter follows from the definition of B_m. Integrating over $z \in Y_i$ we get

$$\int_{Y_i} \sum_{n=1}^{M} w_n \chi_{B_{r_n}^T . E_i} \, d\mu \geq m\delta c_4 \mu(E_i)$$

for a constant $c_4 > 0$ only depending on T. Summing the latter inequality over i, we get that

$$c_3 \mu(X) \geq \int_X G \, d\mu \geq c_4 m\delta \sum_i \mu(E_i) \geq c_4 m\delta \mu(B_m)$$

by Lemma 6.31.(ii). This implies $\mu(B_m) \leq \frac{c_3\mu(X)}{c_4 m\delta}$, independent of M. Hence we may lift the requirement that $n \leq M$ in the definition of B_m without effecting the above estimate and then let $m \to \infty$ and $\delta \to 0$ to obtain the theorem. □

7. Leaf-wise Measures and entropy

We return now to the study of entropy in the context of locally homogeneous spaces.

7.1. **General setup, real case.** Let $G \subset \mathrm{SL}(n, \mathbb{R})$ be a closed real linear group. (One may also take G to be a connected, simply connected real Lie group if so desired.) Let $\Gamma \subset G$ be a discrete subgroup and define $X = \Gamma \backslash G$. We may endow G with a left-invariant Riemannian metric which then induces a Riemannian metric on X too. With respect to this metric X is locally isometric to G, i.e., for every $x \in X$ there exists some $r > 0$ such that $g \mapsto xg$ is an isometry from the open r-ball B_r^G around the identity in G onto the open r-ball $B_r^X(x)$ around $x \in X$. Within compact subsets of X one may choose r uniformly, and we may refer to r as an *injectivity radius* at x (or on the compact subset).

Clearly any $g \in G$ acts on X simply by right translation $g.x = xg^{-1} = \Gamma(hg^{-1})$ for $x = \Gamma h \in X$, and one may check that this action is by Lipschitz automorphisms of X. For this recall that the metric on X is defined using a left-invariant metric on G, which in general is not right-invariant. By definition of X the G-action is transitive.

Recall that Γ is called a *lattice* if X carries a G-invariant probability measure m_X, which is called the *Haar measure* on X. This is the case if the quotient is compact, and in this case Γ is called a *uniform lattice*. From transitivity of the G-action it follows that the G-action is ergodic with respect to the Haar measure m_X. Although this is not clear a priori it is often true (in the non-commutative setting we are most interested in) that unbounded subgroups of G also act ergodically with respect to m_X.

If Γ is a lattice, then we may fix some $a \in G$ or a one-parameter subgroup $A = \{a_t = \exp(tw) : t \in \mathbb{R}\}$ and obtain a measure-preserving transformation $a.x = xa^{-1}$ or flow $a_t.x = xa_t^{-1}$ with respect to $\mu = m_X$. Our discussion of entropy below may be understood in that context. However, we will not assume that the measure μ on X, which we will be discussing, equals the Haar measure or that Γ is a lattice. Rather we will use the results here to obtain information about an unknown measure μ and in the best possible situations deduce from that μ equals the Haar measure.

7.2. **Arithmetic setup.** Fix a prime number p and let G be the group of $\mathbb{Q}_p$-points of an *algebraic subgroup* $\mathbb{G} \subset \mathrm{SL}(n)$, i.e., G would consists of all $\mathbb{Q}_p$-points of a variety $\mathbb{G}$ which is contained in the affine space of all n-by-n-matrices and whose points happen to form a group. Here a $\mathbb{Q}_p$-point of $\mathbb{G}$ is an element of the variety whose matrix entries are elements of $\mathbb{Q}_p$, as a shorthand we will write $G = \mathbb{G}(\mathbb{Q}_p)$ for the group of all $\mathbb{Q}_p$-points. In this setting (more precisely in the zero characteristic case) one may say $\mathbb{G}$ is *defined over* a field F if $\mathbb{G} = \{g \in \mathrm{SL}(n) : \phi(g)v \propto v\}$ where ϕ is an algebraic representation over F, i.e., an action of $\mathrm{SL}(n)$ by linear automorphisms of a finite dimensional vector space with a given basis such that the matrix entries corresponding to $\phi(g)$ are polynomials in the matrix entries of g with coefficients in F, and v equals an F-linear combination of the basis vectors. Again we will let $\Gamma \subset G$ be a discrete subgroup and study dynamics of subgroups of G on $X = \Gamma\backslash G$.

E.g. if G is the group of $\mathbb{Q}_p$-points of $SO(3)$ (defined in the usual way as the group of matrices of determinant one preserving $x_1^2+x_2^2+x_3^2$), which is an algebraic subgroup defined over $\mathbb{Q}$, then one may take Γ to be the group of $\mathbb{Z}[\frac{1}{p}]$-points of $SO(3)$. In this case G is noncompact if $p > 2$ but $X = \Gamma\backslash G$ is compact for any p.

A more general setup would be to allow products

$$G = G_\infty \times G_{p_1} \times \cdots \times G_{p_\ell}$$

over the real and finite places[24] of the group of $\mathbb{R}$-points $G_\infty = \mathbb{G}(\mathbb{R})$, resp., the group of $\mathbb{Q}_p$-points $G_p = \mathbb{G}(\mathbb{Q}_p)$ for some finite list of primes $p \in S_{\mathrm{fin}} = \{p_1, \ldots, p_\ell\}$, of an algebraic group $\mathbb{G}$ defined over $\mathbb{Q}$. In this case one may take $\Gamma = \mathbb{G}(\mathbb{Z}[\frac{1}{p} : p \in S_{\mathrm{fin}}])$ to be the $\mathbb{Z}[\frac{1}{p} : p \in S_{\mathrm{fin}}]$-points of $\mathbb{G}$, which one considers as a subgroup of the product of the real and p-adic groups by sending a matrix γ with coefficients in $\mathbb{Z}[\frac{1}{p} : p \in S_{\mathrm{fin}}]$ to the element $(\gamma, \gamma, \ldots, \gamma) \in G_\infty \times G_{p_1} \times \cdots \times G_{p_\ell}$. This embedding is called the *diagonal embedding*. It can easily be checked that (the image of) Γ forms a discrete subgroup. Often (e.g. when $\mathbb{G}$ is semisimple) Γ defined by this diagonal embedding will form a lattice in G.

A similar construction of arithmetically defined quotients $X = \Gamma\backslash G$ can be used in positive characteristic. Most of what we will discuss in this chapter (and possibly beyond) applies to either of these settings. However, so as to keep the notation at a minimum we will confine ourselves to the situation where $G = \mathbb{G}(k)$ is the group of k-points of an algebraic group $\mathbb{G}$ defined over k, where $k = \mathbb{R}$, $k = \mathbb{Q}_p$, or k equals a local field of positive characteristic. We will refer to this by briefly saying G is an algebraic group over a local field k. Also we only assume that $\Gamma < G$ is a discrete subgroup.

7.3. **The horospherical subgroup defined by a.** For the following fix some $a \in G$. Then we may define the *stable horospherical subgroup for a* by

$$G^- = \{g : a^n g a^{-n} \to e \text{ as } n \to \infty\},$$

[24] The reader who is familiar with adeles may want to consider them instead of finite products.

which in the setting described above is always a closed[25] subgroup of G. Similarly, one can define the *unstable horospherical subgroup* G^+ e.g. as the stable horospherical subgroup for a^{-1}. (We note, that in the theory of algebraic groups G^- and G^+ are also known as the unipotent radicals of the parabolic subgroups defined by a one-parameter subgroup containing a.)

Consider two points $x, xg \in X = \Gamma\backslash G$ for some $g \in G^-$. Then $a^n.x$ and $a^n.xg$ get closer and closer to one another as $n \to \infty$. In fact, $a^n.xg = xa^{-n}(a^n g a^{-n})$ and $a^n.x$ have distance $\leq d(a^n g a^{-n}, e) \to 0$. In that sense we will refer to $G^-.x$ as the *stable manifold through* x. Note that x may not be fixed or even periodic so the statement needs to be understood by the sequence of tuples of points as described. Also note that it is not clear that $G^-.x$ is necessarily the complete set of points y for which $d(a^n.y, a^n.x) \to 0$, but we will show that for all practical purposes it suffices to study $G^-.x$.

7.4. **Problem:** Suppose X is a compact quotient. Show that in this case $G^-.x \subset X$ is precisely the set of points $y \in X$ with $d(a^n.y, a^n.x) \to 0$ as $n \to \infty$.

7.5. **Entropy and the horospherical subgroup.** The following is one of the main results of this section.

7.6. Theorem. *Let μ be an a-invariant probability measure on $\Gamma\backslash G$. Let U be a closed subgroup of G^- normalized by a. Then:*

(i) *The* entropy contribution of U at x

$$D_\mu(a, U)(x) := \lim_{n\to\infty} \frac{\log \mu_x^U(a^{-n} B_1^U a^n)}{n}$$

exists for a.e. x and defines an a-invariant function on X.

(ii) *For a.e. x we have $D_\mu(a, U)(x) \leq h_{\mu_x^{\mathcal{E}}}(a)$, with equality if $U = G^-$. Here $\mathcal{E}$ denotes the σ-algebra of a-invariant sets as in §5.14.*

(iii) *For a.e. x we have $D_\mu(a, U)(x) = 0$ if and only if μ_x^U is finite, which again holds if and only if μ_x^U is trivial*[26].

In particular, the theorem shows that entropy must vanish for all invariant measures if the stable horospherical subgroup G^- is the trivial subgroup. This is the case for the horocycle flow (and all other unipotent flows), hence its entropy vanishes. Therefore, the most interesting case will be the study of the opposite extreme, namely, diagonalizable elements $a \in G$ (and in the proof we will restrict ourselves to this case). For instance, the theorem shows that entropy for the geodesic flow is determined precisely by the leaf-wise measure for the horocyclic subgroup, as for the time-one-map a_1 of the geodesic flow the stable horospherical subgroup is precisely the horocyclic subgroup.

7.7. Corollary. *The measure μ is G^--recurrent if and only if $h_{\mu_x^{\mathcal{E}}}(a) > 0$ a.e. Assume μ is additionally a-ergodic, then μ is G^--recurrent if and only if $h_\mu(a) > 0$.*

---

[25]This is not true for general Lie groups, hence our assumption that G should be a linear group or a simply connected Lie group. In the case of a linear group G^- can easily be defined by linear equations by first bringing a into its Jordan normal form.

[26]Recall that we consider the leaf-wise measures to be trivial if they equal the Dirac measure at the identity.

7.8. **Entropy and G^--invariance.** To state the second equally important theorem we ask first what is $h_{m_X}(a)$ where Γ is assumed to be a lattice and m_X denotes the Haar measure on X. The answer follows quickly from Theorem 7.6: Since m_X is invariant under G^-, its leaf-wise measures are Haar measures on G^-. Hence the expression in Theorem 7.6.(i) can be calculated and one obtains

$$D_{m_X}(a, G^-) = -\log\bigl|\det \mathrm{Ad}_a|_{\mathfrak{g}^-}\bigr|,$$

here Ad_a is the adjoint action of a on the Lie algebra $\mathfrak{g}$ and $\mathfrak{g}^-$ is the Lie algebra of G^- which is, by definition, invariant and being contracted by Ad_a.

The following theorem will characterize when a measure μ is invariant under G^- (or under $U \subset G^-$) in terms of the entropy $h_\mu(a)$ (or the entropy contribution of U). To state it most conveniently, let us define the *entropy contribution* of an a-normalized closed subgroup $U \subset G^-$ by

$$h_\mu(a, U) = \int D_\mu(a, U)\,d\mu$$

the integral of the entropy contributions at the various x. This way, the entropy contribution of G^- equals the entropy of a (cf. §3.5 and §5.14).

7.9. Theorem. *Let $U < G^-$ be an a-normalized closed subgroup of the horospherical subgroup G^- for some $a \in G$, and let $\mathfrak{u}$ denote the Lie algebra of U. Let μ be an a-invariant probability measure on $X = \Gamma\backslash G$. Then the entropy contribution is bounded by*

$$h_\mu(a, U) \le -\log\bigl|\det \mathrm{Ad}_a|_{\mathfrak{u}}\bigr|$$

and equality holds **if and only if** *μ is U-invariant.*

In many cases this theorem shows that the Haar measure on X is the unique measure of maximal entropy. For example the Haar measure on $\mathrm{SL}(2,\mathbb{Z})\backslash \mathrm{SL}(2,\mathbb{R})$ is the unique measure of maximal entropy as follows from Theorem 7.9: Since the stable horospherical subgroup is the upper unipotent subgroup in $\mathrm{SL}(2,\mathbb{R})$, we have that an a-invariant measure whose entropy equals that of the Haar measure must be invariant under the upper unipotent subgroup. Since $h_\mu(a) = h_\mu(a^{-1})$ we get the same for the lower unipotent subgroup. However, since the upper and the lower unipotent subgroups generate $\mathrm{SL}(2,\mathbb{R})$, we get that $h_\mu(a) = h_{m_X}(a)$ implies $\mu = m_X$. By the same argument one obtains the following more general corollary.

7.10. Corollary. *Suppose Γ is a lattice in G, and let $X = \Gamma\backslash G$. Suppose $a \in G$ is such that G is generated by G^+ and G^-. Then m_X is the unique measure of maximal entropy for the action of a on X, i.e., if μ is an a-invariant probability measure on X with $h_\mu(a) = h_{m_X}(a)$ then $\mu = m_X$.*

7.11. **Starting the proofs:** Let us start by discussing the technical assumption of the last section that a.e. orbit is embedded.

7.12. Lemma. *Let μ be an a-invariant probability measure on $X = \Gamma\backslash G$. Then for μ-a.e. x the map $u \in G^- \mapsto u.x$ is injective.*

7.13. Proof: Suppose $x = u.x$ for some nontrivial $u \in G^-$. Then $x_n = a^n.x = a^n u a^{-n}.x_n$ for all $n = 1, 2, \ldots$. However, $a^n u a^{-n} \to e$ so that the injectivity radius at x_n goes to 0 as $n \to \infty$. This shows that x does not satisfy Poincaré recurrence. Hence it belongs to a null set. □

7.14. **Semisimple elements and class A elements.** As before we assume that G is an algebraic group over a local field k (or that G is a simply connected real Lie group), $\Gamma < G$ a discrete subgroup, and $X = \Gamma\backslash G$. We say that $a \in G$ is *k-semisimple* if as an element of $\mathrm{SL}(n, k)$ its eigenvalues belong to k. In particular, this implies that the adjoint action Ad_a of a on the Lie algebra has eigenvalues in k and so is diagonalizable over k. (In the Lie group case the latter would be our assumption with $k = \mathbb{R}$.) We say furthermore that a is *class A* if the following properties hold:

- a is k-semisimple.
- 1 is the only eigenvalue of absolute value 1 for the adjoint action Ad_a.
- No two different eigenvalues of Ad_a have the same absolute value.

For class A elements a we have a decomposition of $\mathfrak{g}$, the Lie algebra of G, into subspaces

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_- \oplus \mathfrak{g}_+$$

where $\mathfrak{g}_0$ is the eigenspace for eigenvalue 1, $\mathfrak{g}_-$ is the direct sum of the eigenspaces with eigenvalues less than 1 in absolute value, and $\mathfrak{g}_+$ is the direct sum of the eigenspaces with eigenvalues greater than 1 in absolute value. These are precisely the Lie algebras of the corresponding subgroups

$$\begin{aligned} G^0 &= \{h : ah = ha\} = C_G(a), \\ G^- &= \{h : a^n h a^{-n} \to e \text{ as } n \to \infty\}, \\ G^+ &= \{h : a^{-n} h a^n \to e \text{ as } n \to \infty\}. \end{aligned}$$

We refer to G^0 as the *centralizer of a*, while G^- and G^+ are the horospherical subgroups of a.

If convenient we will assume[27] below that a is of class A as this gives us a convenient description of a neighborhood of $e \in G$ in terms of neighborhoods in the three subgroups G^0, G^-, and G^+. As before we will always assume that $U < G^-$ is a closed a-normalized subgroup of the stable horospherical subgroup.

7.15. **Problem:** Show that for any Lie group G and any $a \in G$ the Lie algebra generated by $\mathfrak{g}^-$ and $\mathfrak{g}^+$ is a Lie ideal in $\mathfrak{g}$. Deduce that the assumption regarding a in Corollary 7.10 is satisfied whenever G is a simple real Lie group and $\mathfrak{g}^-$ is nontrivial.

7.16. Lemma. *Let $U < G^-$ be a closed a-normalized subgroup for some $a \in G$, denote conjugation by a by $\theta(g) = aga^{-1}$ for $g \in G$. Let μ be an a-invariant probability measure on $X = \Gamma\backslash G$. Then $\mu^U_{a.x} \propto \theta_* \mu^U_x$ for a.e. x.*

7.17. **Proof:** As a normalizes U it maps an (r, U)-flower $(Y, \mathcal{A})$ with base E to another σ-algebra $a.\mathcal{A}$ of subsets of $a.Y$ whose atoms are still open U-plaques. More precisely, for $u \in U$ we have $aua^{-1} \in U$ and $a.(u.x) = \theta(u).(a.x)$. As a preserves the measure μ the conditional measures for $\mathcal{A}$ are mapped to those of $a.\mathcal{A}$. Combining this with Theorem 6.3.(iv) gives the lemma. □

[27] Replacing G^0 and $\mathfrak{g}^0$ with slightly more complicated versions this assumption can be avoided but in our applications a will always be of class A.

7.18. **Independence, a-Invariance.** From the definition of $D_\mu(a,U)(x)$ it follows that if the limit defining $D_\mu(a,U)(x)$ exists, then the original set B_1^U can be replaced by any bounded neighborhood O of $e \in U$ without affecting the limit $D_\mu(a,U)(x)$. In fact, if $a^k B_1^U a^{-k} \subset O \subset a^{-k} B_1^U a^k$ (and such a k exists as U is being contracted by a and both B_1^U and O are bounded neighborhoods) then $\mu(a^{-n+k} B_1^U a^{n-k}) \le \mu(a^{-n} O a^n) \le \mu(a^{-n-k} B_1^U a^{n+k})$ and this implies the claim (using the sandwich argument for sequences and $\frac{n\pm k}{n} \to 1$).

The a-invariance follows from Lemma 7.16: Replacing x by $a.x$ may be interpreted a.e. as replacing μ_x^U by a measure proportional to $\theta_* \mu_x^U$, and the latter replaces B_1^U by $O = a^{-1} B_1^U a$. Both the proportionality factor and the change to O does not affect the limit $D_\mu(a,U)(x)$ so that $D_\mu(a,U)(x) = D_\mu(a,U)(a.x)$ a.e.

7.19. **Preparing the reduction to the ergodic case:** Recall from §5.14 that for any a-invariant measure μ, we have the ergodic decomposition

$$\mu = \int \mu_x^{\mathcal{E}} d\mu(x)$$

where $\mathcal{E}$ is the σ-algebra of all a-invariant sets, and $\mu_x^{\mathcal{E}}$ is the conditional measure. Also recall from §3.5 that the entropy $h_\mu(a)$ equals the average of the entropies $h_{\mu_x^{\mathcal{E}}}(a)$ of the ergodic components. In what follows we wish to reduce the proof of Theorem 7.6 and 7.9 to the corresponding statements under the assumption of ergodicity. The reader who is willing to assume ergodicity[28] of a or to accept this, may continue reading with §7.25.

An important observation (the Hopf argument) is that we can choose the elements of $\mathcal{E}$ to be not only a-invariant, but in fact $\langle U, a\rangle$-invariant. This will allow us to reduce the proof of the main theorems to the case of a-ergodic invariant measures.

7.20. Lemma. *Let C be an a-invariant subset of X. Then there exists a $\langle G^-, a\rangle$-invariant set $\tilde{C}$ such that $\mu(C \triangle \tilde{C}) = 0$.*

7.21. **Proof (using the Hopf argument):** Let $\epsilon > 0$ and choose $f \in C_c(X)$ such that $\|f - 1_C\|_1 < \epsilon$. Set

$$C_\epsilon = \left\{ x : \lim_{n\to\infty} A(f,n)(x) > \frac{1}{2} \right\}$$

where $A(f,n) = \frac{1}{n} \sum_{i=0}^{n-1} f(a^i.x)$. Now

(7.21a) $\quad C \triangle C_\epsilon \subset \quad \{x : \lim_{n\to\infty} A(f,n)(x) \text{ does NOT exist }\}$

(7.21b) $\quad \cup\{x \in C : \lim_{n\to\infty} A(f,n)(x) \le \frac{1}{2}\}$

(7.21c) $\quad \cup\{x \notin C : \lim_{n\to\infty} A(f,n)(x) > \frac{1}{2}\}$

By the pointwise ergodic theorem, the set on the right of (7.21a) has measure 0. We are interested in showing that the measures of (7.21b) and (7.21c) are small. Since C is a-invariant, we have

$$\begin{aligned} x \in C \quad &\Rightarrow \quad A(f - 1_C, n)(x) = A(f,n)(x) - 1 \\ x \notin C \quad &\Rightarrow \quad A(f - 1_C, n)(x) = A(f,n)(x) \end{aligned}$$

(28)This assumption should not be confused with A-ergodicity which we will assume in the later sections but which in general does not imply a-ergodicity.

Let $M(f) = \sup_n |A(f,n)|$ be the maximal function as in the maximal ergodic theorem. Then (7.21b) $\cup$ (7.21c) $\subset \{x : M(f - 1_C)(x) \geq 1/2\}$. Therefore, $\mu\big(\text{(7.21b)}\cup\text{(7.21c)}\big) \leq 2||f - 1_C||_1 < 2\epsilon$ by the maximal ergodic theorem.

Furthermore, we claim that C_ϵ is G^--invariant. Notice that for any $h \in G^-$, we have that $a^n.(h.x) = a^n h a^{-n}.(a^n.x)$ and $a^n.x$ are asymptotic to one another. Since f has compact support it is uniformly continuous. Therefore, we have

$$\frac{1}{n}\sum_{i=0}^{n-1}\big[f(a^i.x) - f(a^i.(h.x))\big] \to 0$$

uniformly in x. This shows that C_ϵ is G^--invariant.

To finish the proof we may choose $\epsilon_n = 2^{-n}$ and

$$\tilde{C} = \overline{\lim_{n\to\infty}}\, C_{2^{-n}} = \bigcap_n \bigcup_{k\geq n} C_{2^{-n}}$$

to obtain a set as in the lemma. □

7.22. Proposition. *Let $\mu = \int \mu_x^{\mathcal{E}} d\mu(x)$ be an a-invariant probability measure, and $U < G^-$. Then for μ-a.e. x, for $\mu_x^{\mathcal{E}}$-a.e. y, we have $\mu_y^U = (\mu_x^{\mathcal{E}})_y^U$.*

In other words, by changing the leaf-wise measures for $\mu_x^{\mathcal{E}}$ at most on a $\mu_x^{\mathcal{E}}$-nullset, we may define $(\mu_x^{\mathcal{E}})_y^U$ to be equal to μ_y^U. With this definition in place, we also have $(\mu_x^{\mathcal{E}})_x^U = \mu_x^U$. (In the formulation of the proposition we avoided this formula as $\{x\}$ is a null set for $\mu_x^{\mathcal{E}}$ and so making claims for the leaf-wise measure at x would be irrelevant.)

7.23. **Proof:** Recall that the leaf-wise measures μ_x^U were determined by moving the conditional measures $\mu_x^{\mathcal{A}_i}$ to U and patching them together there. Here $(Y_i, \mathcal{A}_i)$ were U-flowers. By Lemma 7.20 (and Proposition 5.8) we may replace $\mathcal{E}$ by a countably generated σ-algebra consisting of a-invariant and U-invariant sets. In particular, this shows that the atoms of $\mathcal{E}|_{Y_i}$ are unions of the atoms of $\mathcal{A}_i$ (which are open U-plaques). However, using conditional measures for $\mathcal{A}_i$ it is easy to see that a measurable function that is constant on $\mathcal{A}_i$-atoms is in fact $\mathcal{A}_i$-measurable modulo μ. Therefore, we have $\mathcal{E}|_{Y_i} \subset \mathcal{A}_i$ modulo μ. However, this inclusion of σ-algebras implies that

$$E(E(f|\mathcal{A}_i)|\mathcal{E}|_{Y_i}) = E(f|\mathcal{E}|_{Y_i})$$

for any $f \in L^1$. In turn, using the defining properties of conditional measures (in terms of conditional expectations) this gives the following relation between the conditional measures: for μ-a.e. $x \in Y_i$ we have for $\mu_x^{\mathcal{E}|_{Y_i}}$-a.e. y that

$$\left(\mu_x^{\mathcal{E}|_{Y_i}}\right)_y^{\mathcal{A}_i} = \mu_y^{\mathcal{A}_i}.$$

Translating this to a property of leaf-wise measures we see that μ_y^U and $(\mu_x^{\mathcal{E}})_y^U$ agree on the subset of U corresponding to the atom $[x]_{\mathcal{A}_i}$ and the proposition follows by collecting the various null sets of Y_i. □

7.24. **Proof of reduction to ergodic case:** Working with double conditional measures as in the above proposition may be confusing, but it is useful for the following purpose: In the proof of Theorem 7.6 and 7.9 we are comparing the entropy of the ergodic components and the entropy contribution arising from the subgroup $U < G^-$. From §5.14 and §3.5 we know that

$$h_\mu(a) = \int h_{\mu_x^{\mathcal{E}}}(a) d\mu.$$

We would like to have a similar relationship between $D_\mu(a,U)(x)$ and $D_{\mu_x^{\mathcal{E}}}(a,U)(x)$. Using $(\mu_x^{\mathcal{E}})_x^U = \mu_x^U$ as in the discussion right after Proposition 7.22 we get

$$D_{\mu_x^{\mathcal{E}}}(a,U)(x) = D_\mu(a,U)(x).$$

Since $\mu_x^{\mathcal{E}}$ is a-invariant and ergodic for μ-a.e. x, and as we assume the statements of Theorem 7.6 and 7.9 in the ergodic case, the general case follows from this. □

7.25. Definition. *We say that a σ-algebra $\mathcal{A}$ is* subordinate to U (mod μ) *if for μ-a.e. x, there exists $\delta > 0$ such that*

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

We say that $\mathcal{A}$ is subordinate to U on Y *if and only if the above holds for a.e. $x \in Y$.*

We say that $\mathcal{A}$ is a-descending *if $a^{-1}.\mathcal{A} \subset \mathcal{A}$.*

Ignoring null sets to say that $\mathcal{A}$ is subordinate to U is basically equivalent to say that the $\mathcal{A}$-atoms are open U-plaques. Hence we have already established in the last section the existence of σ-algebras which are subordinate to U at least on some sets of positive measure. Also, it is rather easy to find an a-descending σ-algebra as $\bigvee_{n=0}^\infty a^{-n}.\mathcal{P}$ is a-descending for any countable partition (or even σ-algebra) $\mathcal{P}$. We note however, that the existence of an a-descending σ-algebra that is also subordinate is not trivial.

Recall that we may assume that μ is a-ergodic, so that the a-invariant function $D_\mu(a,U)(x)$ (whose existence we still have to show, see Prop. 7.34) must be constant a.e. If we are given an a-descending σ-algebra $\mathcal{A}$ that is subordinate to U, we will show the following properties (which, in particular, gives an independent meaning to the generic value of $D_\mu(a,U)(x)$):

(i) For a.e. x

$$\frac{\log \mu_x^U(a^{-n}B_1^U a^n)}{n} \to H_\mu(\mathcal{A}|a^{-1}.\mathcal{A}) = h_\mu(a,U)$$

as $n \to \infty$.

(ii) $h_\mu(a,U) \le h_\mu(a)$, with equality if $U = G^-$.

(iii) If $h_\mu(a,U) = 0$ then $a^{-1}.\mathcal{A} = \mathcal{A}$ (mod μ) and $\mu_x^U = \delta_e$ almost surely.

In other words, we will use the σ-algebra $\mathcal{A}$ as a gadget linking the two expressions $D_\mu(a,U)(x)$ and $h_\mu(a)$ appearing in the Theorem 7.6.

Recall that the "empirical entropy" $H_\mu(\mathcal{A}|a^{-1}.\mathcal{A})$ is the average of the "conditional information function"

$$I_\mu(\mathcal{A}|a^{-1}.\mathcal{A})(x) = -\log \mu_x^{a^{-1}.\mathcal{A}}([x]_{\mathcal{A}}).$$

7.26. **Hyperbolic torus automorphisms.** We first look at a particular example[29] where it is relatively easy to give an a-descending σ-algebra $\mathcal{A}$ that is subordinate to G^- and to see the connection to entropy. Let a be a hyperbolic automorphism of $\mathbb{T}^m = \mathbb{R}^m/\mathbb{Z}^m$. By this we mean that a is defined by a matrix with eigenvalues, real or complex, of absolute value different from one (this is the answer to Problem 3.10). We set $G = \mathbb{R}^m$, $\Gamma = \mathbb{Z}^m$, and write θ for the linear map on $\mathbb{R}^m$ defining a; for consistency we will still write $a.x$ for the action. It is easy to see that G^- is the sum[30] of all generalized eigenspaces for eigenvalues of absolute value less than one. By expansiveness we know that any partition $\mathcal{P}$ whose atoms have sufficiently small diameter will be a generator for a. By the same argument one easily shows that $\mathcal{A} = \bigvee_{n=0}^{\infty} a^{-n}.\mathcal{P}$ satisfies that the $\mathcal{A}$-atoms are of the form $[x]_{\mathcal{A}} = V_x.x$ for bounded subsets $V_x \subset G^-$. Also, $\mathcal{A}$ is a-descending. The remaining property that V_x contains the identity in the interior a.e. is not a general property (as it is likely not true if the boundaries of the partition elements are not null sets) but follows if we are a bit more careful in the choice of the partition $\mathcal{P}$. What we will need is the following quantitative strengthening of $\mu(\partial P) = 0$ for all $P \in \mathcal{P}$.

7.27. Lemma. *Let X be a locally compact metric space and let μ be a Radon measure on X. Then for every $x \in X$ and Lebesgue-a.e. $r > 0$ there exists a constant $c = c_{x,r}$ such that $\mu(\partial_\delta B_r(x)) \leq c\delta$ for all sufficiently small $\delta > 0$. Here we refer to*

$$\partial_\delta B = \{y \in X : \inf_{z \in B} d(y,z) + \inf_{z \notin B} d(y,z) < \delta\}$$

as the δ-neighborhood of the boundary[31] of a subset $B \subset X$.

7.28. Problem. *Prove Lemma 7.27 using the function $f(r) = \mu(B_r(x))$. A hint may be found in the footnote[32] on the next page.*

We say that a set B has *μ-thin boundary* if there exists some c such that $\mu(\partial_\delta B) \leq c\delta$ for all $\delta > 0$. It is clear that a set obtained from finitely many sets with μ-thin boundary via the set-theoretic operations of intersections, union, or complements also has μ-thin boundary. Hence by Lemma 7.27 any compact space has a partition $\mathcal{P}$ consisting of sets with μ-thin boundary and arbitrarily small diameter.

We also note another property, which is rather easy to verify for the Euclidean metric on $G = \mathbb{R}^m$ and the linear map θ defining the automorphism a.

7.29. Lemma. *There exists some $\alpha > 0$ and $d > 0$ depending on a and G such that for every $r \in (0,1]$ we have*

$$\theta^n(B_r^{G^-}) \subset B_{de^{-n\alpha}r}^{G}$$

for all $n \geq 1$.

7.30. Problem. *Let $a \in G$ be an element of class A. Prove Lemma 7.29 in the context of G being a real Lie group, assuming that G is endowed with a left invariant Riemannian metric.*

---

[29] This example almost fits into the framework under which we work, except that the automorphism we consider is not coming from an element of $G = \mathbb{R}^m$. We could use a bigger subgroup, namely a semidirect product of $\mathbb{Z}$ and $\mathbb{R}^m$, but this is not necessary and may be more confusing.

[30] This is always a real subspace even if some eigenvalues are complex.

[31] We use this phrase even though in general ∂B may be empty with $\partial_\delta B$ nonempty.

Prove Lemma 7.29 in the setting of G being an algebraic group defined over a p-adic field or a finite characteristic local field by first defining a metric on G. (If necessary it would not make a difference to our applications below to replace the upper bound 1 for r by some smaller quantity depending on a and G.)

We now show how the two properties in Lemma 7.27 and Lemma 7.29 can be used in combination.

7.31. Lemma. *Suppose $\mathcal{P}$ is a finite partition of $X = \Gamma\backslash G$ consisting of measurable sets with μ-thin boundary. Then for a.e. $x \in X$ there is some $\delta > 0$ such that*

$$B_\delta^{G^-}.x \subset [x]_{\bigvee_{n\geq 0} a^{-n}.\mathcal{P}}. \tag{7.31a}$$

7.32. **Proof:** Let c be the maximal constant as in the definition of μ-thin boundary for the elements of $\mathcal{P}$, and let α and d be as in Lemma 7.29. Also let $r = 1$. We write $\partial_\delta\mathcal{P}$ for the union of the δ-neighborhoods of the boundaries of the elements of $\mathcal{P}$.

Fix some $\delta > 0$ and define for $n \geq 0$ the set

$$E_n = a^{-n}.\partial_{de^{-n\alpha}\delta}\mathcal{P}.$$

By construction we have

$$\mu\left(\bigcup_{n\geq 0} E_n\right) \leq cd\left(\sum_{n\geq 0} e^{-n\alpha}\right)\delta,$$

which shows that for a.e. x there is some δ with $x \notin \bigcup_{n\geq 0} E_n$. Fix such an x and the corresponding δ, we claim that (7.31a) holds. Indeed let $h \in B_\delta^{G^-}$ (which in the case of $X = \mathbb{T}^m$ acts by addition $h.x = x + h$ on $x \in \mathbb{T}^m$) and suppose $h.x \notin [x]_{\bigvee_{n\geq 0} a^{-n}.\mathcal{P}}$. Then there would be some $n \geq 0$ such that $a^n.x$ and $a^n.(h.x)$ belong to different elements of the partition $\mathcal{P}$. However, θ contracts G^- and indeed $d(\theta^n(h), e) < de^{-n\alpha}\delta$ by Lemma 7.29. Therefore, $a^n.(h.x) = \theta^n(h).(a^n.x)$ and $a^n.x$ have distance less than $de^{-n\alpha}\delta$, which shows that both belong to $\partial_{de^{-n\alpha}\delta}\mathcal{P}$. However, this gives a contradiction to the definition of E_n and the choice of x and δ.□

7.33. **Hyperbolic torus automorphism concluded.** The discussion in §7.26 together with Lemma 7.31 shows that it is possible to choose $\mathcal{P}$ such that the σ-algebra $\mathcal{A} = \bigvee_{n=0}^\infty a^{-n}.\mathcal{P}$ is a-decreasing and subordinate to G^-. Recalling that $\mathcal{P}$ was constructed as a generator (c.f. §3.6) we also get

$$h_\mu(a) = h_\mu(a, \mathcal{P}) = H_\mu(\mathcal{A}|a^{-1}.\mathcal{A}).$$

This establishes the link between $H_\mu(\mathcal{A}|a^{-1}.\mathcal{A})$ and the entropy $h_\mu(a)$ in the case at hand; the link between $H_\mu(\mathcal{A}|a^{-1}.\mathcal{A})$ and the entropy contribution we now establish in great generality.

7.34. Proposition. *Suppose $\mathcal{A}$ is a countably generated σ-algebra subordinate to U, such that $\mathcal{A} \supset a^{-1}.\mathcal{A}$. Then*

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

In particular, the limit defining the entropy contribution of U at x exists.

---

(32) Notice that $f(r)$ is monotone and hence differentiable a.e.

7.35. **Proof:** We start by showing that

$$-\frac{1}{n}\log\mu_x^{a^{-n}.\mathcal{A}}([x]_{\mathcal{A}})\to H_\mu(\mathcal{A}|a^{-1}.\mathcal{A}).$$

Here notice first that by Proposition 5.17

$$\mu_x^{a^{-1}.\mathcal{A}}\big|_{[x]_{\mathcal{A}}}=\mu_x^{a^{-1}.\mathcal{A}}([x]_{\mathcal{A}})\mu_x^{\mathcal{A}}$$

for a.e. x since $[x]_{a^{-1}.\mathcal{A}}$ is a countable union of $\mathcal{A}$-atoms. More generally we obtain by the same argument that

$$\mu_x^{a^{-n}.\mathcal{A}}([x]_{\mathcal{A}})=\prod_{i=1}^{n}\mu_x^{a^{-i}.\mathcal{A}}([x]_{a^{-(i-1)}.\mathcal{A}}).$$

Also note that $\mu_{a.x}^{\mathcal{A}}=a_*\mu_x^{a^{-1}.\mathcal{A}}$ (as one may verify from the defining relation of $\mu_x^{\mathcal{A}}$ in terms of the conditional expectation). Combining these one gets by taking logarithms that

$$\begin{aligned}-\frac{1}{n}\log\mu_x^{a^{-n}.\mathcal{A}}([x]_{\mathcal{A}}) &= \sum_{i=1}^{n}\frac{-\log\mu_x^{a^{-i}.\mathcal{A}}([x]_{a^{-(i-1)}.\mathcal{A}})}{n}\\ &= \frac{1}{n}\sum_{i=0}^{n-1}I_\mu(\mathcal{A}|a^{-1}.\mathcal{A})(a^i.x)\\ &\to H_\mu(\mathcal{A}|a^{-1}.\mathcal{A})\end{aligned}$$

by the pointwise ergodic theorem (since μ is assumed to be a-ergodic).

We may also obtain in a similar manner that

$$\frac{\log\mu_x^U(a^{-n}B_1^Ua^n)}{n}\to\int\log\mu_x^U(a^{-1}B_1^Ua),$$

where we assume the normalization $\mu_x^U(B_1^U)=1$. Indeed by Lemma 7.16 we know $\mu_{a.x}^U(a^{-1}B_1^Ua)=\frac{\mu_x^U(a^{-2}B_1^Ua^2)}{\mu_x^U(a^{-1}B_1^Ua)}$, which easily generalizes to higher powers of a and then gives

$$\mu_x^U(a^{-n}B_1^Ua^n)=\prod_{i=0}^{n-1}\mu_{a^i.x}^U(a^{-1}B_1^Ua).$$

Taking the logarithm and using the pointwise ergodic theorem the above claim follows.

We outline the remainder of the proof of Proposition 7.34: Both of the above limits measure the growth rate of a dynamically expanded set in relation to a fixed set. By Theorem 6.3.(iv) the fact that in one expression we are using the conditional measure $\mu_x^{a^{-n}.\mathcal{A}}$ and in the other the leaf-wise measure μ_x^U seems irrelevant. However, what is unclear is the precise relationship between the shape $V_{n,x}\subset U$ of the atoms $[x]_{a^{-n}.\mathcal{A}}=V_{n,x}.x$ and the set $a^{-n}B_1^Ua^n$. We show below that as $n\to\infty$ the influence of the shape is negligible, thus obtaining the proposition.

Fix $\delta>0$ such that

(7.35a) $$Y:=\{x: B_\delta^U.x\subset[x]_{\mathcal{A}}\subset B_{\delta^{-1}}^U.x\}$$

has positive measure. By the argument in §7.18 (which only assumes the existence of the limit for $r=1$) we know that

(7.35b) $$\lim_{n\to\infty}\frac{\log\mu_x^U(a^{-n}B_r^Ua^n)}{n}$$

is independent of r for a.e. x. Moreover, for a.e. x there exists a sequence n_j of integers for which $a^{n_j}.x \in Y$. For those $n = n_j$ we therefore have

$$[x]_{a^{-n}.\mathcal{A}} = a^{-n}.[a^n.x]_{\mathcal{A}} \subset a^{-n} B^U_{\delta^{-1}} a^n.x$$

and similarly,

$$[x]_{a^{-n}.\mathcal{A}} \supset a^{-n} B^U_{\delta} a^n.x.$$

Therefore, $a^{-n} B^U_\delta a^n \subset V_{n,x} \subset a^{-n} B^U_{\delta^{-1}} a^n$. Also recall that $\mu_x^{a^{-n}.\mathcal{A}}$ is proportional to $\mu_x^U|_{V_{n,x}}.x$ by Theorem 6.3.(iv). Hence

$$\mu_x^{a^{-n}\mathcal{A}}([x]_{\mathcal{A}}) = \frac{c(x)}{\mu_x^U(V_{n,x})}$$

where $c(x) = \mu_x^U(V_{0,x})$. With this notation the above inclusions imply

$$\mu_x^U(a^{-n} B^U_\delta a^n) \le \mu_x^U(V_{n,x}) = c(x)\mu_x^{a^{-n}\mathcal{A}}([x]_{\mathcal{A}})^{-1} \le \mu_x^U(a^{-n} B^U_{\delta^{-1}} a^n)$$

for a.e. x. Taking the logarithm, letting $n = n_j \to \infty$, and using the independence of the limit in (7.35b) the proposition follows. □

7.36**. Returning to the general case.** Even though we used in the example of the hyperbolic torus automorphism certain special properties of the system, namely that X is compact and that a is expansive, it does give hope regarding the existence of a subordinate and a-descending σ-algebra in general. In fact, using somewhat similar methods (Lemma 7.29, 7.27, and 7.31 are general) as in the example we now establish the existence of the σ-algebra. However, linking the σ-algebras and the entropy (as we did in §7.33) will need more work.

7.37. Proposition. *Let μ be an a-invariant and ergodic probability measure on $\Gamma\backslash G$, and let $U < G^-$ be a closed subgroup normalized by a. Then there exists a countably generated σ-algebra $\mathcal{A}$ such that:*

(i) *$\mathcal{A}$ is subordinate to U.*
(ii) *$a^{-1}.\mathcal{A} \subset \mathcal{A}$, i.e., $\mathcal{A}$ is a-decreasing.*

We note that this establishes the existence of the limit in Theorem 7.6 (i) by using Proposition 7.34.

7.38**. Comment:** We note that without the assumption of ergodicity the proof below almost gives the claims of the proposition in the following sense: For every $\epsilon > 0$ there exists a set $Y \subset X$ of measure $\mu(Y) > 1 - \epsilon$ such that $\mathcal{A}$ is subordinate to U on Y and $\mathcal{A}$ is a-decreasing.

7.39**. Proof:** Applying Lemma 7.27 we can find some open $Y \subset X$ with compact closure such that $\mu(Y) > 1 - \epsilon$ and Y has μ-thin boundary, e.g. by letting $Y = B^X_r(x_0)$ for some large r. Below we will construct the σ-algebra $\mathcal{A}$ which will be subordinate to U on Y and a-decreasing. Note that under the assumption of ergodicity this gives the proposition: For a.e. $x \notin Y$ there exists some positive as well as some negative n with $a^n.x \in Y$ which together with $a^{-1}.\mathcal{A} \subset \mathcal{A}$ gives the correct upper, resp., lower bound for $[x]_{\mathcal{A}}$. More precisely, by correct upper bound we mean that $[x]_{\mathcal{A}} \subset B_x.x$ for some bounded subset $B_x \subset U$ and by correct lower bound we mean that $[x]_{\mathcal{A}} \supset O_x.x$ for some open $O_x \subset U$ containing the identity element.

Again applying Lemma 7.27 we can find a finite partition of Y into sets of small diameter (as specified below) and with μ-thin boundary — here Lemma 7.27

is applied to find for every $x \in \overline{Y}$ a small ball around x with μ-thin boundary and then a finite subcover is chosen using compactness. We add to this partition the set $X \setminus Y$ to obtain the partition $\mathcal{P}$. Since the boundaries of all elements of $\mathcal{P}$ are null sets, we may assume all elements of $\mathcal{P}$ are open (and ignore the remaining null set). By Lemma 7.31 we know that the atoms of $\bigvee_{n\geq 0} a^{-n}.\mathcal{P}$ contain a neighborhood of x in the direction of G^- almost surely, i.e., for a.e. $x \in X$ there is some $\delta > 0$ such that

$$[x]_{\bigvee_{n\geq 0} a^{-n}.\mathcal{P}} \supset B_\delta^{G^-}.x. \tag{7.39a}$$

We will replace $\mathcal{P}$ by a σ-algebra $\mathcal{P}^U$ in such a way that $\mathcal{A} = \bigvee_{n\geq 0} a^{-n}.\mathcal{P}^U$ will be subordinate to U (at least) on Y. Let P denote an element of $\mathcal{P}$ different from $X \setminus Y$. We may assume the diameter of P is smaller than the injectivity radius on Y, we get that P is the injective isometric image of an open subset $\tilde{P}$ of G. By assumption U is closed, so that the Borel σ-algebra $\mathcal{B}_{G/U}$ of the quotient G/U is countably generated. This induces a σ-algebra $\mathcal{C}_P$ first on $\tilde{P}$ and then also on P whose atoms are open U-plaques. We define $\mathcal{P}^U$ to be the countably generated σ-algebra whose elements are unions of elements of $\mathcal{C}_P$ for $P \in \mathcal{P}$ and possibly the set $X \setminus Y$, i.e., the atoms of x for $\mathcal{P}^U$ is either $X \setminus Y$ if $x \notin Y$ or an open U-plaque $V_x.x$ of x if $x \in Y$. We claim that for a.e. $x \in Y$ the atom $[x]_\mathcal{A}$ w.r.t. $\mathcal{A} = \bigvee_{n\geq 0} a^{-n}.\mathcal{P}^U$ is an open U-plaque. Indeed suppose x satisfies (7.39a) for some $\delta > 0$ (which we may assume is smaller than the injectivity radius) and $u \in B_\delta^U$. Then for all $n \geq 0$ we know that $a^n.x$ and $a^n u.x$ belong to the same element $P \in \mathcal{P}$. Fix some $n \geq 0$. If $P = X \setminus Y$, then $a^n.x$ and $a^n u.x$ still belong to the same atom of $\mathcal{P}^U$. If $P \neq X \setminus Y$, then we also claim that $a^n.x$ and $a^n u.x$ belong to the same atom of $\mathcal{P}^U$: The two elements $y = a^n.x, z = a^n u.x \in P$ correspond to two elements $\tilde{y}, \tilde{z} \in \tilde{P}$. Since P and $\tilde{P}$ are isometric and $a^n u a^{-n}$ is being contracted we conclude that these two points are still on the same U-coset $\tilde{y}U = \tilde{z}U$, for otherwise we would get a contradiction to the injectivity property at $z \in Y$. This shows that the atoms $[x]_\mathcal{A}$ for a.e. $x \in Y$ are indeed open U-plaques. $\square$

7.40. **Proof of Theorem 7.6.(iii):** Clearly if μ_x^U is finite, then the entropy contribution $h_\mu(a,U)$ vanishes (as it measures a growth rate). Assume now on the other hand $h_\mu(a,U) = H_\mu(\mathcal{A}|a^{-1}\mathcal{A}) = 0$, where $\mathcal{A}$ is as in Prop. 7.37 (cf. Prop. 7.34). Then

$$H_\mu(\mathcal{A}|a^{-1}.\mathcal{A}) = \int (-\log \mu_x^{a^{-1}.\mathcal{A}}([x]_\mathcal{A}))d\mu = 0$$

implies $\mu_x^{a^{-1}.\mathcal{A}}([x]_\mathcal{A}) = 1$ a.e. which is equivalent to $\mathcal{A} = a^{-1}.\mathcal{A} \bmod \mu$. Iterating this gives $a^m.\mathcal{A} = a^{-m}.\mathcal{A} \bmod \mu$ and $\mu_x^{a^{-m}\mathcal{A}}([x]_{a^m.\mathcal{A}}) = 1$ a.e. and for all $m \geq 1$. By Theorem 6.3.(iv) this says that $\mu_x^U(V_{-m,x} \setminus V_{m,x}) = 0$ a.e., where $V_{m,x}$ denotes the shape of the $a^m.\mathcal{A}$-atom of x. Using again the set Y in (7.35a) we see that the precise shapes do not matter as $V_{-m,x} \nearrow U$ and $V_{m,x} \searrow \{e\}$ as $m \to \infty$ for a.e. x. It follows that $\mu_x^U \propto \delta_e$. $\square$.

7.41. **Proof of the inequality $h_\mu(a,U) \leq h_\mu(a,U')$ for $U \subset U' \subseteq G^-$.** Assume both U and U' are closed a-normalized subgroups of G^- such that $U \subset U'$. By the construction of the σ-algebra we see that there exist two σ-algebras $\mathcal{A}$ and $\mathcal{A}'$ which are both a-decreasing and subordinate to U and to U', resp., such that additionally $\mathcal{A} \supset \mathcal{A}'$. In order to obtain these, one may use the same finite partition $\mathcal{P}$ and then carry the construction through with both groups.

We claim that $\mathcal{A}' \vee a^{-1}.\mathcal{A} = \mathcal{A} \pmod{\mu}$. We already know one inclusion, to see the other we describe the atoms for the σ-algebra $\mathcal{C} = \mathcal{A}' \vee a^{-1}.\mathcal{A}$. Suppose y and x are equivalent w.r.t. $\mathcal{C}$, then a.s. there exists some $u \in U$ with $y = u.x$ where u may be rather big because the $a^{-1}.\mathcal{A}$-atoms are in general bigger than the $\mathcal{A}$-atoms. To make this more precise, assume y, x belong to the set Y which was used in the constructions of the σ-algebras. Then we do not know that $d(e, u)$ is smaller than the injectivity radius (of Y). However, we know that $y = u'.x$ for some $u' \in U'$ (as the two points are also $\mathcal{A}'$-equivalent), and that $d(e, u')$ is less than the injectivity radius. Since for a.e. x the G^--leaf is embedded by Lemma 7.12, we must have $u = u'$. This implies that x and $y = u.x$ belong to the same atom of the σ-algebra $\mathcal{C}_P$ (for $x, y \in P \subset Y$) which was used in the construction of $\mathcal{P}^U$. This shows the two points are equivalent w.r.t. $\mathcal{A}$, first under the assumption that $x, y \in Y$ but the general case follows by the same argument and ergodicity by using the minimal n with $a^n.x, a^n.y \in Y$. As the atoms of the σ-algebra determine the σ-algebra at least mod μ the claim follows.

The claim implies the desired inequality since

$$h(a, U) = H_\mu(\mathcal{A}|a^{-1}.\mathcal{A}) = H_\mu(\mathcal{A}'|a^{-1}.\mathcal{A}) \leq H_\mu(\mathcal{A}'|a^{-1}.\mathcal{A}') = h(a, U')$$

by monotonicity of the entropy function with respect to the given (i.e., the second) σ-algebra. □

7.42**. First proof of the inequality in Theorem 7.9:** Recall that $U \subset G^-$ is a-normalized and that λ denotes the Haar measure on U, which is (unipotent and so) necessarily unimodular. We may normalize λ to have measure one on the unit ball of U. Then it is easy to see that $\lambda(a^{-n} B_1^U a^n) = c^n$, where c is the determinant of the adjoint representation of a^{-1} acting on the Lie algebra $\mathfrak{u}$ of U. Hence for a.e. x, we have by Theorem 6.30[33]

$$\begin{aligned} \overline{\lim_{n\to\infty}} \frac{\mu_x^U(a^{-n} B_1^U a^n)}{n^2 c^n} &= 0 \text{ and so} \\ h_\mu(a, U) = \lim_{n\to\infty} \frac{\log \mu_x^U(a^{-n} B_1^U a^n)}{n} &\leq \log c. \end{aligned}$$

This is the inequality in Theorem 7.9. In §7.55 we will give the proof of Theorem 7.9 in full including an independent proof of the inequality shown here. □

7.43**. Where we are.** To summarize we have shown Theorem 7.6.(i), the inequality in (ii), (iii), and the inequality in Theorem 7.9 (and also that it suffices to study ergodic measures). However, we still have to show the equality between the entropy contribution $h_\mu(a, G^-)$ and the entropy $h_\mu(a)$ and the relationship between invariance and equality in Theorem 7.9. We now turn to the former problem in general.

7.44. Proposition. *Let μ be an a-invariant and ergodic measure on $X = \Gamma\backslash G$. Then there exists a countable partition $\mathcal{P}$ with finite entropy which is a generator for*

[33] Strictly speaking the sets $a^{-n} B_1^U a^n$ may not be balls, but the proof can be adapted to allow for that and the additional thickening of the balls by the parameter 5 does not change the asymptotical behavior of $\lambda(a^{-n} B_1^U a^n)$.

$a \bmod \mu$. Moreover, the σ-algebra $\mathcal{A} = \bigvee_{n\geq 0} a^{-n}.\mathcal{P}$ is a-decreasing and subordinate to G^-.

Together with Proposition 7.34 this implies the last claim of Theorem 7.6.(ii). We will need a few more elementary lemmas.

7.45. Lemma. *There exists some $\alpha > 0$ depending on a and G such that for every $r > 0$ we have*

$$\theta^n(B^G_{e^{-|n|\alpha}r}) \subset B^G_r$$

for all $n \in \mathbb{Z}$. Here $\theta(g) = aga^{-1}$ for $g \in G$ again stands for conjugation by a.

Basically this lemma follows from the fact that conjugation is a Lipschitz map whose Lipschitz constant is the norm of the adjoint representation of a.

7.46. Lemma. *For every $\Omega \subset \Gamma\backslash G$ with compact closure, and for every α and $r > 0$, there exist $\kappa(G, \alpha), c(\Omega, r)$ such that for every n, the set Ω can be covered by $ce^{\kappa n}$ balls of radius $e^{-\alpha n}r$.*

For the proof of this lemma notice that the set Ω can be covered by finitely many small balls of fixed radius, and that in each one of these we may argue that the metric is basically flat (e.g. in characteristic zero the logarithm map would be bi-Lipshitz in a neighborhood of the identity and the claim is quite easy for a linear space). In a sense this lemma captures (in some weak way) the finite-dimensionality of the group in question.

7.47. **Proof:** Equipped with the lemmas above, we are ready to start the construction of our partition $\mathcal{P}$. Fix an open subset $\Omega \subset X = \Gamma\backslash G$ of compact closure, positive measure, and μ-thin boundary (see Lemma 7.27). We may assume Ω is a ball $B_{r/16}(x_0)$ where r is an injectivity radius at x_0.

7.48. **The partition $\mathcal{Q}$.** We define $\mathcal{Q} = \{\Omega, X \setminus \Omega\}$. By Lemma 7.31 we have that for a.e. x there exists some $\delta > 0$ with

$$B^{G^-}_\delta .x \subset [x]_{\bigvee_{n\geq 0} a^{-n}.\mathcal{Q}}. \tag{7.48a}$$

7.49. **The partition $\tilde{\mathcal{Q}}$.** Next we define $\tilde{\mathcal{Q}} = \{Q_i : i = 0, 1, 2, \ldots\}$, where we define $Q_0 = X \setminus \Omega$, resp. $Q_1 = \Omega \cap a^{-1}.\Omega$, $Q_2 = (\Omega \setminus a^{-1}.\Omega) \cap a^{-2}\Omega$, $\ldots$, in other words we split Ω into countably many sets according to when the points next visit Ω (under forward iterates of a). (Strictly speaking we should also add the set $Q_\infty = \Omega \cap \bigcap_{j=1}^\infty a^{-j}.X \setminus \Omega$ to the partition $\tilde{\mathcal{Q}}$, but by Poincaré Recurrence $\mu(Q_\infty) = 0$, so we may omit it from the discussion.)

We observe that $\tilde{\mathcal{Q}}$ is contained in the σ-algebra $\bigvee_{n=1}^\infty a^{-n}.\mathcal{Q}$. Therefore, $\bigvee_{n=1}^\infty a^{-n}.\mathcal{Q} = \bigvee_{n=1}^\infty a^{-n}.\tilde{\mathcal{Q}}$ and the above claim (7.48a) regarding the atoms remains true for $\tilde{\mathcal{Q}}$.

7.50. **Finite entropy.** We will now show that $H_\mu(\tilde{\mathcal{Q}}) < \infty$ (but we will need to refine it further to obtain the desired partition). First, note that $X \setminus \Omega$ can be partitioned according to how much time a point will spend (resp. has already spent) in $X \setminus \Omega$ before returning to (resp. since coming from) Ω, keeping in mind that the set of points which remain in $X \setminus \Omega$ forever (resp. have always been in $X \setminus \Omega$) has measure 0 by ergodicity. Moreover, the set of points that have spent time $t \geq 1$ in $X \setminus \Omega$ (including the current time) and will return to Ω in time $s \geq 1$ iterations of a the first time is exactly $a^t.Q_{t+s}$. This implies that

$X\setminus\Omega=\bigcup_{i=1}^\infty\bigcup_{t=1}^i a^tQ_{i+1}$ (with the union being disjoint), and since μ is a-invariant, we see that $\mu(X\setminus\Omega)=\sum_{i=1}^\infty i\mu(Q_{i+1})<1$. As the sets $Q_1,Q_2,\ldots$ partition Ω we also have $\mu(\Omega)=\sum_{i=1}^\infty\mu(Q_i)$, and so we conclude that

$$\sum_{i=1}^\infty i\mu(Q_i)=1.$$

We can therefore write

$$H_\mu(\tilde{\mathcal{Q}})=-\sum_{i=0}^\infty\mu(Q_i)\log\mu(Q_i)<\sum_{\mu(Q_i)>e^{-i}}\mu(Q_i)i+\sum_{\mu(Q_i)\le e^{-i}}e^{-i}i+c<\infty$$

by using monotonicity of $-\log t$ in the first case and the monotonicity of $-t\log t$ for small values of t in the second case (the constant c is there to handle the finitely many cases where the latter monotonicity may not apply).

7.51. **The partition $\mathcal{P}$.** We now apply Lemma 7.46 to Ω and conclude that for $i\ge1$ each of the sets $Q_i\subset\Omega$ may be covered with $\le ce^{i\kappa}$ many balls B_j of radius $e^{-\alpha i}r/8$. Here r is the injectivity radius at the center x_0 of the ball Ω and α is chosen as in Lemma 7.45. We will refine the partition $\tilde{\mathcal{Q}}$ by splitting each Q_i into smaller sets. However, so as not to destroy the property (7.48a) we will use instead of the original balls B_j some modified version of them that are "widened" or "smeared out" in the direction of G^-.

Fix some $Q_i\in\tilde{\mathcal{Q}}$ for $i\ge1$ and write $D=Q_i$ to simplify the notation, also let $B_1,B_2,\ldots,B_N$ with $N=N(i)\le ce^{i\kappa}$ be the cover obtained above. We split D into the sets $D_1,D_2,\ldots$ as follows:

$$\begin{aligned}D_1&=D\cap(B_{r/4}^{G^-}.B_1),\\ D_2&=D\cap(B_{r/4}^{G^-}.B_2)\setminus D_1,\ldots\end{aligned}$$

Roughly speaking, since the set $\Omega\supset D$ has small diameter (at most $r/8$) in comparison to its injectivity radius (r) and since the widening by $B_{r/4}^{G^-}$ is by a bigger radius, we should think of the splitting of D into the sets $D_1,\ldots$ as a splitting transversely to the G^--orbits.

This defines a partition of $D=Q_i$ into $\le ce^{i\kappa}$ many sets. Collecting these partitions for the various sets Q_i we obtain one partition $\mathcal{P}$ of X containing all of them and $Q_0=X\setminus\Omega$.

7.52. **Finite entropy.** Now for each n, we define $\mu|_{Q_n}$ to be the restricted measure normalized to be a probability measure. Then the entropy $H_{\mu|_{Q_n}}(\mathcal{P})\le\log c+\kappa n$ since the partition $\mathcal{P}$ restricted to Q_n has at most $\le ce^{n\kappa}$ many elements by construction. Also

$$H_\mu(\mathcal{P})=H_\mu(\tilde{\mathcal{Q}})+H_\mu(\mathcal{P}|\tilde{\mathcal{Q}}),$$

and the latter quantity may be expressed as the weighted average of the entropies $H_{\mu|_{Q_n}}(\mathcal{P})$ so that finally

$$H_\mu(\mathcal{P})\le H_\mu(\tilde{\mathcal{Q}})+\log c+\kappa\sum n\mu(Q_n)<\infty.$$

7.53**. Upper bound for atom.** We claim that the partition $\mathcal{P}$ has the property that for any $x, a^n.x \in \Omega$, we have

$$[x]_{\bigvee_{i=0}^{\infty} a^{-i}\mathcal{P}} \subset \left(\bigcap_{k=0}^{n} a^{-k} B_r^G a^k \right) . x,$$

which is quite similar to what we proved in the case of a hyperbolic torus automorphism. The idea is that, although we do not learn much information about the orbits during the time it spends near the cusp (our partition element $Q_0 = X \setminus \Omega \in \mathcal{P}$ is rather crude there and moreover the injectivity radius is not uniform there), we compensate by learning a great deal about the point at the time at which it leaves Ω.

To prove the claim assume $x, a^n.x \in \Omega$ and $y = g.x \in [x]_{\bigvee_{i=0}^{\infty} a^{-i}\mathcal{P}}$ for some $g \in B_r$. Then $x \in Q_i = D$ for some i — this means that $a^i.x \in \Omega$ — and $x \in D_j$ for some $j \le N(i)$. We will first show the claim for $n = i$. By equivalence of $y = g.x$ to x and by construction of the set D_j we get that $x = u_x h_x.z_j$ with $u_x \in B_{r/4}^{G^-}$ and $h_x \in B_{e^{-\alpha n} r/8}^{G}$ and similarly for y, where $z_j \in D$ is the center of the ball B_j used to define D_j. We may remove z_j from the formulas and obtain first that $y = g.x = u_y h_y h_x^{-1} u_x^{-1}.x$ which implies $g = u_y h_y h_x^{-1} u_x^{-1} \in B_r^G$ as r is an injectivity radius. If r is sufficiently small we obtain from this $g = uh$ with $u = u_y u_x^{-1} \in B_{r/2}^{G^-}$ and $h = u_x (h_y h_x^{-1}) u_x^{-1} \in B_{e^{-\alpha n} r/2}$ as conjugation by a small element does not change the metric much. This shows that $a^k g a^{-k} = (a^k u a^{-k})(a^k h a^k) \in B_{r/2}^{G^-} B_{r/2}^{G}$ for $k = 1, \ldots, i$ by Lemma 7.45, which proves the claim in the case of $n = i$.

If $n > i$ we obtain from the above that $a^k g a^{-k} \in B_r^G$ for $k = 1, \ldots, i$ and then we may repeat the argument with $x, y = g.x$, and g replaced by $a^i.x, a^i.y$ and $a^i g a^{-i}$ resp., and with n replaced by $n - i$. Repeating the argument as needed shows the claim.

The claim implies that

$$[x]_{\mathcal{A}} \subset B_r^{G^- G^0}.x$$

for a.e. $x \in \Omega$, where we define $\mathcal{A} = \bigvee_{n=0}^{\infty} a^{-n}\mathcal{P}$ and we recall that $G^0 = C_G(a)$. Indeed for a.e. $x \in \Omega$ we have infinitely many n with $a^n.x \in \Omega$ by Poincarè recurrence.

Moreover, if μ is not compactly supported, then $\mu(Q_n) \neq 0$ for infinitely many n which implies that the above atom is actually contained in $B_r^{G^-}.x$ for a.e. $x \in \Omega$. In fact, suppose $\mu(Q_{n_0}) > 0$ then for a.e. $x \in \Omega$ we know that there are infinitely many n with $a^n.x \in Q_{n_0}$. Take one such x and assume that $g.x$ is equivalent to x and $g = uh$ with $u \in G^-$ and $h \in G^0$. This implies that $a^n g a^{-n} = a^n u a^{-n} h = u'h$ is the displacement between $a^n.x$ and $a^n g.x$ which implies $h \in B_{e^{-\alpha n_0} r/2}^{G}$. As we know this for infinitely many n_0 we obtain $h = e$.

If however, we have $\mu(Q_n) = 0$ for all but finitely many n, then $\mathcal{P}$ is actually a finite partition $\bmod \mu$ and the last statement may not hold. However, in this case we may artificially split one of the sets of positive measure into countably many sets of positive measures such that for every ϵ we have a partition element of positive measure contained in a set of the form $B_{r/4}^{G^-} B_\epsilon.x_\epsilon$. Making these new partition elements small enough, we may assume that their measure decays rapidly which ensures that the resulting partition still has finite entropy. With this refined partition the above holds also in this case.

Notice that the above statements regarding the upper bound $B_r^{G^-}.x$ of the atom were stated for $x \in \Omega$, but that a slightly weaker form also holds for a.e. $x \in X$. In fact, if $x \in X$ and $n \geq 1$ is such that $a^n.x \in \Omega$ satisfies the inclusion $[a^n.x]_{\mathcal{A}} \subset B_r^{G^-} a^n.x$ then we have that $[x]_{\mathcal{A}} \subset B_s^{G^-}.x$ for some s that depends on n.

7.54**. Lower bound for atom.** To finish the proof we wish to show that (7.48a) also holds for the partition $\mathcal{P}$. So suppose $x \in \Omega$ and $\delta > 0$ satisfies (7.48a). Here we will use the fact that we "widened" the balls B_j in the direction of G^- to obtain the sets $D_j \subset Q_i$. We may assume $\delta < r/8$, and pick some $u \in B_\delta^{G^-}$. As $\tilde{\mathcal{Q}}$ is contained in the σ-algebra generated by $\bigvee_{n\geq 0} a^{-n}.\mathcal{Q}$ we know that x and $u.x$ belong to the same $Q_i = D \in \tilde{\mathcal{Q}}$. Suppose $x \in D_j$ which shows $x = u_x h_x.z_j$ with $u_x \in B_{r/4}^{G^-}$ and $h_x \in B_{e^{-\alpha_i}r/8}^{G}$ where $z_j \in B_j \cap \Omega$ is the center of the ball B_j that was used to construct D_j. Now clearly $u.x = (uu_x)h_x.z_j$ and $uu_xh_x \in B_{r/2}^{G}$. As the diameter of Ω is at most $r/8$ by definition, we obtain $uu_xh_x \in B_{r/8}^G$ since r is an injectivity radius on Ω. Together with $h_x \in B_{r/8}^G$ this implies $uu_x \in B_{r/4}^{G^-}$. (To see this notice that by left invariance of the metric we have $d(g,e) = d(e,g^{-1}) \leq d(e,h) + d(h,g^{-1}) = d(e,h) + d(gh,e)$ for all $g,h \in G$.) This implies that $u.x$ also belongs to $B_{r/4}^{G^-}.B_j$ and D. In fact this shows $u.x \in D_j$, for if $u.x \notin D_j$ then necessarily $u.x \in D_{j'}$ for some $j' < j$ but then by symmetry of the argument between x and $u.x$ we would have also $x \notin D_j$. Therefore, x and $u.x$ belong to the same element of $\mathcal{P}$. Repeating the argument as needed starting with $a^i.x$ and $a^iu.x$ shows that x and $u.x$ are equivalent with respect to $\bigvee_{n=0}^\infty a^{-n}\mathcal{P}$. The points $x \in X \setminus \Omega$ are dealt with in the same manner as before by choosing a minimal n with $a^n.x \in \Omega$. This finishes the proof of Proposition 7.44. □

7.55**. Proof of Theorem 7.9.** Let $U < G^-$ be a closed a-normalized subgroup. Let μ be an a-invariant and ergodic probability measure on $X = \Gamma\backslash G$. We wish to show that the entropy contribution is bounded by $h_\mu(a,U) \leq J$ where $J = -\log\bigl|\det \mathrm{Ad}_a|_{\mathfrak{u}}\bigr|$ is the negative logarithm of the absolute value of the determinant of the adjoint representation of a restricted to the Lie algebra $\mathfrak{u}$ of U. As we will show we only have to use convexity of $\log t$ for $t \in \mathbb{R}$. However, we will have to use it on every atom $[x]_{a^{-1}.\mathcal{A}}$ for an a-decreasing σ-algebra which is subordinate to U.

We fix a Haar measure m_U on U, and note that

$$m_U(a^{-1}Ba) = e^J m_U(B) \text{ for any measurable } B \subset U. \tag{7.55a}$$

For $x \in X$ we write $V_x \subset U$ for the shape of the $\mathcal{A}$-atom so that $V_x.x = [x]_{\mathcal{A}}$ a.e. Recall that $\mu_x^{a^{-1}.\mathcal{A}}$ is a probability measure on $[x]_{a^{-1}.\mathcal{A}} = a^{-1}V_{a.x}a.x$ which is used in the definition of

$$H_\mu(\mathcal{A}|a^{-1}.\mathcal{A}) = -\int \log \mu_x^{a^{-1}.\mathcal{A}}([x]_{\mathcal{A}}).$$

We wish to compare this to a similar expression where we use (in a careful manner) the Haar measure m_U on U as a replacement for the conditional measures. We note however, that we will always work with the given measure μ on X, so our notion of "a.e." is here always meant w.r.t. μ. We define τ_x^{Haar} to be the normalized push forward of $m_U|_{a^{-1}V_{a.x}a}$ under the orbit map, i.e., we define

$$\tau_x^{\mathrm{Haar}} = \frac{1}{m_U(a^{-1}V_{a.x}a)} m_U\big|_{a^{-1}V_{a.x}a}.x,$$

which again is a probability measure on $[x]_{a^{-1}\mathcal{A}}$.

We define

$$p(x) = \mu_x^{a^{-1}.\mathcal{A}}([x]_{\mathcal{A}})$$

which appears in the definition of $H_\mu(\mathcal{A}|a^{-1}.\mathcal{A})$. By analogy we also define

$$p^{\mathrm{Haar}}(x) = \tau_x^{\mathrm{Haar}}([x]_{\mathcal{A}}) = \frac{m_U(V_x)}{m_U(a^{-1}V_{a.x}a)} = \frac{m_U(V_x)}{m_U(V_{a.x})} e^{-J}$$

where we used (7.55a). Taking the logarithm and applying the ergodic theorem (check this) we see that $-\int \log p^{\mathrm{Haar}} d\mu = J$.

Now we recall that both $\mathcal{A}$ and $a^{-1}.\mathcal{A}$ are subordinate to U, which means that after removing a null set they must be countably equivalent. In other words, there exists a null set N such that for $x \notin N$ the $\mathcal{A}$-atom of x contains an open neighborhood of x in the U-orbit. We may also assume that for $x \notin N$ there are infinitely many positive and negative n with $a^n.x \in Y$ where Y is as in (7.35a). Since U is second countable, this implies that

$$[x]_{a^{-1}.\mathcal{A}} \setminus N = \bigcup_{i=1}^{\infty} [x_i]_{\mathcal{A}} \setminus N$$

where the union is disjoint. For a.e. x we wouldn't have to be too careful about the null set N as it is also a null set for the conditional measure, but note that it may not be a null set for τ_x^{Haar}. Therefore, we write

$$[x]_{a^{-1}.\mathcal{A}} = \bigcup_{i=1}^{\infty} [x_i]_{\mathcal{A}} \cup N_x$$

where N_x is a null set for $\mu_x^{a^{-1}.\mathcal{A}}$ but maybe not for τ_x^{Haar}. We may assume $\mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}}) > 0$, otherwise we just remove this atom from the list and increase N_x accordingly. This shows

$$\sum_{i=1}^{\infty} \mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}}) = 1$$

but only

$$\sum_{i=1}^{\infty} \tau_x^{\mathrm{Haar}}([x_i]_{\mathcal{A}}) \leq 1.$$

We now integrate $\log p^{\mathrm{Haar}} - \log p$ over the atom $[x]_{a^{-1}.\mathcal{A}}$ to get

$$\int \log p^{\mathrm{Haar}} d\mu_x^{a^{-1}.\mathcal{A}} - \int \log p \, d\mu_x^{a^{-1}.\mathcal{A}}, \tag{7.55b}$$

but as both functions are constant on the $\mathcal{A}$-atoms (and as N_x is a null set w.r.t. the measure w.r.t. which we integrate) this integral is nothing but the countable sum

$$= \sum_{i=1}^{\infty} \left(\log \frac{\tau_x^{\mathrm{Haar}}([x_i]_{\mathcal{A}})}{\mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}})} \right) \mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}})$$

Using now convexity of $\log t$ for $t \in \mathbb{R}$ with $\mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}})$ as the weights at $t_i = \frac{\tau_x^{\mathrm{Haar}}([x_i]_{\mathcal{A}})}{\mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}})}$ we get

$$(7.55c) \quad = \sum_{i=1}^{\infty} \log(t_i) \mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}}) \leq \log\left(\sum_{i=1}^{\infty} t_i \mu_x^{a^{-1}.\mathcal{A}}([x_i]_{\mathcal{A}})\right)$$
$$= \log\left(\sum_{i=1}^{\infty} \tau_x^{\mathrm{Haar}}([x_i]_{\mathcal{A}})\right) = \log \tau_x^{\mathrm{Haar}}\left(\bigcup_{i=1}^{\infty}[x_i]_{\mathcal{A}}\right) \leq 0.$$

Integrating this inequality over all of X and recalling the relation of the function p with the entropy contribution $h_\mu(a,U) = H_\mu(\mathcal{A}|a^{-1}.\mathcal{A})$ and of the function p^{Haar} with J gives the desired inequality.

In case of equality we use strict convexity of $\log t$: If $h_\mu(a,U) = J$, then the integral of the non-positive (due to (7.55c)) expression in (7.55b) vanishes. Therefore, for a.e. atom (7.55b) vanishes, or equivalently we must have 0 on both sides of (7.55c). However, this means that $\tau_x^{\mathrm{Haar}}(N_x) = 0$ and that $t_i = 1$ for all i by strict convexity of $\log t$. Notice that $t_i = 1$ means that the conditional measure $\mu_x^{a^{-1}.\mathcal{A}}$ gives the same weight to the $\mathcal{A}$-atoms $[x_i]_{\mathcal{A}}$ as does the normalized Haar measure τ_x^{Haar} on the $a^{-1}.\mathcal{A}$-atom.

Using that $H_\mu(a^k.\mathcal{A}|a^{-\ell}.\mathcal{A}) = (k+\ell)h_\mu(a,U) = (k+\ell)J$ for any $k,\ell \geq 0$ together with the same argument we obtain that the conditional measure $\mu_x^{a^{-\ell}.\mathcal{A}}$ gives the same weight to the $a^k.\mathcal{A}$-atoms as does the normalized Haar measure on the $a^{-\ell}.\mathcal{A}$-atom. For a.e. x the $a^{-\ell}.\mathcal{A}$-atom can be made arbitrarily large as there is a sequence $\ell_n \to \infty$ with $a^{\ell_n}.x \in Y$. Now fix ℓ, then the various $a^k.\mathcal{A}$-atoms for all $k \geq 0$ generate the Borel σ-algebra on the $a^{-\ell}.\mathcal{A}$-atom, at least on the complement of N which is a null set both for $\mu_x^{a^{-\ell}.\mathcal{A}}$ and for the normalized Haar measure on the atom. This follows as for $\mu_x^{a^{-\ell}.\mathcal{A}}$-a.e. y the $a^k.\mathcal{A}$ atom can be made to have arbitrarily small diameter since for $y \notin N$ there is a sequence $k_n \to \infty$ with $a^{-k_n}.y \in Y$. This shows that $\mu_x^{a^{-\ell}.\mathcal{A}}$ equals the normalized Haar measure on the atom $[x]_{a^{-\ell}.\mathcal{A}}$. Using this for all ℓ we see that the leaf-wise measure μ_x^U is the Haar measure on U, and so that μ is U-invariant (c.f. Problem 6.28). This concludes the proof of Theorem 7.9. □

8. The product structure

8.1. **Assumptions.** In the previous chapter we considered a measure μ on $\Gamma\backslash G$ invariant under the action of a diagonalizable element $a \in G$, and studied in some detail the leafwise measures induced by μ on orbits of unipotent groups U contracted by a. When considering the action of a multiparameter diagonalizable group $A \subset G$, it is often possible to find some $a \in A$ which contracts some nontrivial unipotent group U but which acts isometrically on orbits of some other group T (which may well be contracted by some other element $a' \in A$). In this case there is a surprisingly simple relation between the leafwise measures of the group generated by U and T (which we assume to be simply the product group) and the leafwise measures for each of these groups: essentially, the leafwise measures for TU will be the product of the leafwise measures for T and U!

Even though a key motivation to looking at these conditional measures is our desire to understand action of multiparameter diagonal groups, we will make use of a single diagonalizable $a \in G$ (more precisely — an element of the class A

defined in §7.14). Let $U \subset G^-$ be a-normalized, and contracted by a. Finally let $T \subset G^0 = C_G(a)$ centralize a and assume that T normalizes U (this is not a very restrictive condition: in particular, the reader can easily verify that the full contracting subgroup for a is normalized by T). We define $H = TU \subset G$, which we can identify with $T \ltimes U$, and will show below that the leaf-wise measure for H is proportional to the product of the leaf-wise measures for T and U.

8.2**.** This simple relation was discovered by M.E. and A. Katok and is one of the key ingredients in the paper [**EK03**] and extended in [**EK05**] (cf. also [**Lin06**, §6]). A weaker form of this relation can be derived from the work of H. Hu [**Hu93**] on entropy of smooth $\mathbb{Z}^d$-actions, and the relation between entropy and leafwise measures, and was used by Katok and Spatzier in [**KS96**].

8.3**. Example.** The following is an example (for $G = \mathrm{SL}(3, \mathbb{R})$) to have in mind: Let

$$a = \begin{pmatrix} e^{-2} & & \\ & e & \\ & & e \end{pmatrix}, \quad U = \left\{ \begin{pmatrix} 1 & * & * \\ & 1 & 0 \\ & & 1 \end{pmatrix} \right\}, \quad T = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ & 1 & * \\ & & 1 \end{pmatrix} \right\}$$

so that

$$H = \left\{ \begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix} \right\}.$$

We note that the Haar measure of H is the three-dimensional Lebesgue measure and so is also the direct product of the Haar measures on T and U.

8.4**. Short reminder.** Recall that the equivalence classes by proportionality of the leaf-wise measures live in a compact metric space, because of the growth property from Theorem 6.30. More precisely, recall that we have a function $\rho > 0$ such that $\int \rho d\mu_x^T < \infty$ a.e. Taking a sequence $\{0 \leq f_i \leq \rho\}_{i=1}^\infty \subset C_c(T)$ spanning a dense subset, we may define

$$d([\nu_1], [\nu_2]) := \sum_{i=1}^\infty 2^{-i} \left| \frac{\int f_i d\nu_1}{\int \rho d\nu_1} - \frac{\int f_i d\nu_2}{\int \rho d\nu_2} \right|$$

for any two equivalence classes of Radon measures with $\int \rho d\nu_i < \infty$. If we chose a representative of the equivalence class we may assume $\int \rho d\nu_i = 1$. This way, the metric just defined corresponds to the weak* topology in the space of Radon measures

$$\Big\{ \nu : \int \rho d\nu = 1 \Big\} \subset \Big\{ \nu : \int \rho d\nu \leq 1 \Big\}.$$

This way the leaf-wise measure μ_x^T can be interpreted as a measurable function with values in a compact metric space.

We also recall the property of leaf-wise measures (Theorem 6.3.(iii)):

$$[\mu_x^T] = [(\mu_{t.x}^T).t] \tag{8.4a}$$

whenever $t \in T$, and $x, t.x \in X'$ (a set of full measure). The following proposition extends this by explaining how μ_x^T transforms under the bigger group $H = TU_-$.

8.5. Proposition. *There exists $X' \subset X$ of full measure, such that for every $x \in X'$ and $h \in H$ such that $h.x \in X'$, we have*

$$[\mu_x^T] = [(\mu_{h.x}^T)t]$$

where $h = tu' = u''t$ for some $u'', u' \in U$ and $t \in T$.

The special case of $t = e$ of this proposition implies the following (note, however that Proposition 8.5 is a substantially stronger statement — cf. §8.9):

8.6. Corollary. *Let $u \in U$. Then $x, u.x \in X'$ implies $[\mu_x^T] = [\mu_{u.x}^T]$.*

8.7. **Proof of Proposition 8.5:** As explained above the map $x \mapsto [\mu_x^T]$ is Borel measurable, from $\Gamma\backslash G$ to a compact metric space. By Luzin's Theorem, for any $\epsilon > 0$, there exists a compact $K_\epsilon \subset \Gamma\backslash G$ such that:

- $\mu(K_\epsilon) > 1 - \epsilon$,
- $x \mapsto [\mu_x^T]$ is continuous on K_ϵ, and
- (8.4a) holds whenever $x, t.x \in K_\epsilon$.

Define

$$X_\epsilon = \left\{x \in K_\epsilon : \sup_n \frac{1}{n} \sum_{i=0}^{n-1} 1_{X\setminus K_\epsilon}(a^i.x) < 1/2\right\}$$

Then using the maximal ergodic theorem one easily verifies that $\mu(X_\epsilon) > 1 - 2\epsilon$.

If $x, h.x \in X_\epsilon$, then there is a sequence $n_i \to \infty$ such that $a^{n_i}x, a^{n_i}h.x \in K_\epsilon$. Passing to a subsequence if necessary, we may assume that $a^{n_i}x \to x_0$. We note that a commutes with the elements of T by definition and so Lemma 7.16 implies that $\mu_{a^n.x}^T = \mu_x^T$ for every n and a.e. x. We may assume this holds for any $x \in X_\epsilon$. By continuity on K_ϵ we have

$$[\mu_x^T] = [\mu_{a^{n_i}.x}^T] \to [\mu_{x_0}^T].$$

For $h.x$, we can rewrite and get

$$a^{n_i}h.x = a^{n_i}tu'.x = t(a^{n_i}u'a^{-n_i})a^{n_i}.x \to t.x_0,$$

since the term in parentheses $a^{n_i}u'a^{-n_i} \to e$ as $n_i \to \infty$. So again by continuity we have

$$[\mu_{h.x}^T] = [\mu_{a^{n_i}h.x}^T] \to [\mu_{t.x_0}^T].$$

Together, we have for $x, h.x \in X_\epsilon$

$$[\mu_x^T] = [\mu_{x_0}^T] = [(\mu_{t.x_0}^T)t] = [\mu_{a^{n_i}h.x}^T t] = [\mu_{h.x}^T t]$$

as desired. We conclude the proof by letting $\epsilon = \frac{1}{n} \searrow 0$, choosing K_ϵ increasing, and defining X' to be the union of the $X_{\frac{1}{n}}$. $\square$

8.8. Corollary (Product structure)**.** *Let $H = T \ltimes U$ be as in §8.1. There exists $X' \subset X$ of full measure, such that for every $x \in X'$ we have*

$$\mu_x^H \propto \iota(\mu_x^T \times \mu_x^U),$$

where $\iota : (t, u) \in T \times U \mapsto tu \in H$.

8.9. **Remark:** All the essential facts for this corollary have already been proved as we explain now. Suppose for a minute that the product formula as in the last corollary holds, then indeed the leaf-wise measure for T of a point $u.x$ with $u \in U$ should be the same as for x, but as usual we should allow for null sets and just state this for a.e. $u \in U$ (w.r.t. the natural measure μ_x^U there). This is what we proved in Corollary 8.6 for a.e. x. (Recall that $\mu(X \setminus X') = 0$ implies that for μ-a.e. $x \in X$ and μ_x^U-a.e. u we have $u.x \in X'$.)

However, this property does not imply that μ_x^H is a product measure: E.g. if μ_x^H were a measure supported on a measurable graph from T to U then μ_x^U would typically be trivial and this would make the above true while μ_x^H may not be a product measure.

Proposition 8.5, on the other hand, does contradict the prevalence of this type of leafwise measures. We may rephrase Proposition 8.5 as follows: for μ-a.e. x and μ_x^H-a.e. $h = tu$ (with $t \in T$ and $u \in U$) we know that the leaf-wise measure $\mu_{h.x}^T$ is, apart from the shift by t (and possibly a proportionality factor), the same as μ_x^T. This property is incompatible with a graph-like measure we described above unless the graph describes a constant map (which is compatible with the product structure we claim). To convert this heuristic into an argument we need to prove another lemma regarding leaf-wise measures.

8.10. Lemma. *Let H be a locally compact second countable group acting nicely on X (say, locally and measure-theoretically free), and let μ be a Radon (i.e., locally finite) measure on X. Assume $H = LM = \iota(L \times M)$ is topologically isomorphic (under the product map $\iota(\ell, m) = \ell m$ for $\ell \in L, m \in M$) to the product of two closed subgroups $L, M < H$. Then L acts by restriction on X and on H by left translation, and so gives rise to families of leaf-wise measures μ_x^L and $(\mu_x^H)_h^L$ for $x \in X$ and $h \in H$. Then there exists $X' \subset X$ of full measure such that whenever $x \in X'$ we have $[(\mu_x^H)_h^L] = [\mu_{h.x}^L]$ for μ_x^H-a.e. $h \in H$.*

Roughly speaking the above is what we should expect: μ_x^H is the measure on H such that $\mu_x^H.x$ describes μ along on the orbit $H.x$. Similarly, $(\mu_x^H)_h^L$ is the measure on L for which $(\mu_x^H)_h^L h$ describes μ_x^H on the coset Lh, and so we expect that $(\mu_x^H)_h^L$ will be such that $(\mu_x^H)_h^L h.x$ describes μ on the orbit $Lh.x$ which suggests the conclusion.

8.11. **Proof:** Let $\Xi \subset X$ be an R-cross-section for the action of H on some set of positive measure (see Definition 6.6). Let $\tilde{\mathcal{A}}_H$ be the σ-algebra $\{B_R^H, \emptyset\} \otimes \mathcal{B}(\Xi)$ on $B_R^H \times \Xi$, where $\mathcal{B}(\Xi)$ is the Borel σ-algebra on Ξ. The map $\iota(h, x) = h.x$ is injective on $B_R^H \times \Xi$ by definition, and so $\mathcal{A}_H = \iota(\tilde{\mathcal{A}}_H)$ is a countably generated σ-algebra of Borel sets. The atom $[x]_{\mathcal{A}_H}$ is an open H-plaque for any $x \in \iota(B_R^H \times \Xi) = B_R^H.\Xi$ (namely equal to $B_R^H.z$ for some $z \in \Xi$).

We further define $\tilde{\mathcal{A}}_L := \{LB \cap B_R^H : B \in \mathcal{B}(M)\}$ where $\mathcal{B}(M)$ is the Borel σ-algebra on M, which is by assumption a global cross-section of L in H. The σ-algebra $\mathcal{A}_L = \iota(\tilde{\mathcal{A}}_L \times \mathcal{B}(\Xi))$ is countably generated, and $[x]_{\mathcal{A}_L}$ is an open L-plaque for all $x \in \iota(B_R^H \times \Xi)$. Note that $\mathcal{A}_L \supset \mathcal{A}_H$.

The measures μ_x^H and μ_x^L can be defined by the values of conditional measures with respect to a countable collection of σ-algebras $\mathcal{A}_H^{(i)}$ and $\mathcal{A}_L^{(i)}$ constructed as above. On each of these σ-algebras a corresponding compatibility condition is satisfied due to the inclusion $\mathcal{A}_L^{(i)} \supset \mathcal{A}_H^{(i)}$; this implies the lemma. $\square$

8.12. **Proof of Corollary 8.8:** Take $x \in X$ to be typical (i.e., outside of the union of bad null sets from Proposition 8.5, and Lemma 8.10 applied to both $L = T$, $M = U$ and $L = U$, $M = T$). We are going to combine these statements, but for this it will be easier to restrict μ_x^H to the bounded product set $Q = B_r^T B_r^U \subset H$ for some $r > 0$, which we may envision as a rectangle with sides B_r^T and B_r^U.

Using $L = T$ Lemma 8.10 is telling us that the conditional measures for $\mu_x^H|_Q$ with respect to the σ-algebra $\mathcal{A} = \{B_r^T, \emptyset\} \otimes \mathcal{B}(B_r^U)$ can be obtained from the leaf-wise measures $\mu_{h.x}^T$ (for μ_x^H-a.e. $h \in Q$). As usual, we have to shift the leaf-wise measure for T back to the space in question, which after applying the lemma may be taken to be H, and restrict to the atoms of the σ-algebra $\mathcal{A}$ in question. This gives

$$\big(\mu_x^H\big)_h^{\mathcal{A}} \propto \big(\mu_{h.x}^T h\big)\big|_Q. \tag{8.12a}$$

However, Proposition 8.5 gives

$$\mu_{tu.x}^T t \propto \mu_x^T. \tag{8.12b}$$

for μ_x^H-a.e. $h = tu \in Q$ (which is a form of independence of $\mu_{tu.x}^T$ in terms of $u \in U$). Using (8.12a)-(8.12b) together, we obtain that the conditional measures of $\mu_x^H|_Q$ with respect to $\mathcal{A} = \{B_r^T, \emptyset\} \otimes \mathcal{B}(B_r^U)$ at $h = tu \in Q$ is equal to $\mu_x^T|_{B_r^T} \times \delta_u$ normalized to be a probability measure. However, this just says that $\mu_x^H|_Q$ is a product measure which is proportional to $\iota(\mu_x^T \times \nu_r)$ for some finite measure ν_r on B_r^U.

Varying r it is easy to check that one can patch these measures ν_r together (i.e. that they extend each other up to a proportionality factor) to obtain a Radon measure ν on U and that μ_x^H is in fact proportional to $\iota(\mu_x^T \times \nu)$. We wish to show that $\nu \propto \mu_x^U$.

As $\iota(\mu_x^T \times \nu)$ is a product measure it is clear what the conditional measures for it are with respect to a σ-algebra whose atoms are of the form tV for open subsets $V \subset U$. However, this corresponds really to the right action of U on H while we have to use the left action if we want to apply Lemma 8.10 for $L = U$ and $M = T$. Luckily U is a normal subgroup, so at least the orbits of these two actions are the same even though the way these two actions identify the orbit with the group differs. We now analyze this in more detail.

Restrict again to $Q = B_r^T B_r^U \subset H$ and consider the σ-algebra $\mathcal{A}' = \mathcal{B}(B_r^T) \otimes \{B_r^U, \emptyset\}$, whose atoms are tB_r^U for $t \in B_r^T$. We know that the conditional measure of $\mu_x^H|_Q$ at $h = tu$ equals $\iota(\delta_t \times \nu_r)$. Considering now the action of U by left multiplication on H we see that the atom of $h = tu \in Q$ corresponds to the set $V_h = tB_r^U h^{-1} \subset U$. Using these σ-algebras for all positive integers r we characterize the leaf-wise measures of μ_x^H with respect to the U-action and obtain that $(\mu_x^H)_h^U$ must be proportional to (the push forward) $t\,\nu\,h^{-1}$ for μ_x^H-a.e. h.

To summarize we know

$$\mu_{h.x}^U \propto (\mu_x^H)_h^U \propto t\,\nu\,h^{-1} \text{ where } h = tu. \tag{8.12c}$$

Clearly the above gives the desired statement if we just set $h = e$. However, strictly speaking we are not allowed to use $h = e$ as we only know these two formulas for μ_x^H-a.e. $h \in H$. Instead we may show the corresponding claim not for the x we started with but for $h.x$ for μ_x^H-a.e. $h \in H$. This will show that the corollary holds a.e. In fact, for μ_x^H-a.e. $h = tu$ we know

$$\mu_{h.x}^H \propto \mu_x^H h^{-1} \propto \iota(\mu_x^T \times \nu)h^{-1} \propto \iota(\mu_{h.x}^T t \times \nu)h^{-1} \propto \\ \propto \iota(\mu_{h.x}^T \times t\nu h^{-1}) \propto \iota(\mu_{h.x}^T \times \mu_{h.x}^U)$$

by combining Theorem 6.3.(iii) for the action of H on X, with the product structure at μ_x^H already obtained, with Proposition 8.5, with the definition of $\iota(t,u) = tu$, and finally with (8.12c). □

From the product structure just proven we can read off an analogue of Corollary 8.6 for the action of T. We note however that, since T and U may not commute, a full analogue of Proposition 8.5 with the roles of T and U reversed will in general not hold (unless one allows conjugation as in the proof above).

8.13. Corollary. *Let $t \in T$. Then $x, t.x \in X'$ implies $[\mu_x^U] = [\mu_{t.x}^U]$.*

We refer to Example 6.5.2 for a discussion showing that this mild coincidence of leaf-wise measures may indeed be a very special property.

9. Invariant measures and entropy for higher rank subgroups A, the high entropy method

9.1. As before we consider the space $X = \Gamma\backslash G$, where G is an algebraic group over a characteristic zero local field k (say $k = \mathbb{R}$ or $\mathbb{Q}_p$ for simplicity). We fix an algebraic subgroup $A \subset G$ which is diagonalizable over the ground field k. In algebraic terms A is the group of k-points of a k-split torus, but we may simply refer to A as a *torus* and to its action on X as a *torus action.* We will assume that we have a homomorphism $\alpha : (k^\times)^n \hookrightarrow G$ which is defined by polynomials with coefficients in k and whose range equals A. We may often suppress the isomorphism and use A and $(k^\times)^n$ interchangeably. For example, if $G = \mathrm{SL}(4, \mathbb{R})$, we could have

$$\alpha : (t,s) \mapsto \begin{pmatrix} t^2 & & & \\ & ts & & \\ & & s & \\ & & & t^{-3}s^{-2} \end{pmatrix}.$$

Let $\mathfrak{g}$ be the Lie algebra of G. Recall that in zero characteristic, the functions exp and log are homeomorphisms between neighborhoods of $0 \in \mathfrak{g}$ and $e \in G$. Hence, the restriction of the adjoint action of A on the Lie algebra $\mathfrak{g}$ gives a good description of the behavior of conjugation on G which as we have seen is crucial in the study of the action of the elements of A on X.

9.2. A character λ is a homomorphism $\lambda : (k^\times)^n \simeq A \to k^\times$ defined by polynomials with coefficients in k; these polynomials necessarily have the form $\lambda(t_1, \ldots, t_n) = t_1^{\ell_1} \cdots t_n^{\ell_n}$ where $\ell_1, \ldots, \ell_n \in \mathbb{Z}$.

We say that a character λ is a *weight* (which one also may refer to as eigenvalue, Lyapunov weight, or root) for the action of A if there is some nonzero $\underline{x} \in \mathfrak{g}$ such that for every $a \in A$, we have $Ad_a(\underline{x}) = \lambda(a)\underline{x}$. The set of all such $\underline{x} \in \mathfrak{g}$ is the *weight space* $\mathfrak{g}^\lambda$. By the assumption that A is diagonalizable we get a decomposition $\mathfrak{g} = \bigoplus_{\lambda \in \Phi} \mathfrak{g}^\lambda$ where Φ is the set of all weights.

For $a \in A$ the subspace $\mathfrak{g}_a^- = \bigoplus_{|\lambda(a)|<1} \mathfrak{g}^\lambda$ is a nilpotent subalgebra, and exp gives a global homeomorphism from $\mathfrak{g}_a^-$ to the horospherical group $G_a^- < G$. Here, the absolute value comes from the Archimedean norm on $\mathbb{R}$ resp. the p-adic norms on $\mathbb{Q}_p$. Also, we introduced the subscript in the notation G_a^- to explicate the dependence of the horospherical subgroup on the element $a \in A$ used.

Note that $[\mathfrak{g}^\lambda, \mathfrak{g}^\eta] \subset \mathfrak{g}^{\lambda\eta}$ which follows easily from the formula

$$Ad_a([\underline{x}, \underline{y}]) = [Ad_a(\underline{x}), Ad_a(\underline{y})] \qquad \underline{x}, \underline{y} \in \mathfrak{g}.$$

Note that in general $\mathfrak{g}^\lambda$ is not a sub-Lie-algebra if λ^2 is also a weight.

9.3. Define an equivalence relation on Φ by $\lambda \sim \eta$ if there exist positive integers ℓ, m such that $\lambda^\ell = \eta^m$. This means that λ and η are weights in the same "direction" — characters are in a one-to-one correspondence with $\mathbb{Z}^n$ and under this correspondence $\lambda \sim \eta$ if and only if they are on the same ray from the origin. For a nontrivial weight $\lambda \in \Phi$ we define the *coarse Lyapunov subalgebra*

$$\mathfrak{g}^{[\lambda]} := \bigoplus_{\eta \sim \lambda} \mathfrak{g}^\eta.$$

We note that exp gives a globally defined homeomorphism between $\mathfrak{g}^{[\lambda]}$ and a unipotent subgroup $G^{[\lambda]}$ which we will refer to as the *coarse Lyapunov subgroup*.

Note that λ is nontrivial (i.e., not the constant homomorphism) implies that $\mathfrak{g}^\lambda$ can be made part of some $\mathfrak{g}_a^-$ for some correctly chosen $a \in A$. Moreover, two weights λ and η are equivalent if and only if their corresponding weight spaces are contained in $\mathfrak{g}_a^-$ for the same set of $a \in A$. In this sense, one might say that weights are equivalent if they cannot be distinguished by any elements of a in terms of whether or not the weight space is being contracted.

Similarly, the coarse Lyapunov subgroup $G^{[\lambda]}$ is the intersection of stable horospherical subgroups for various elements of A and is a smallest nontrivial such subgroup. Dynamically speaking, we may say that the orbits of the coarse Lyapunov subgroups are the smallest nontrivial intersections one can obtain by intersecting stable manifolds of various elements of A.

9.4. In this section, we study the structure of the leaf-wise measures on these coarse Lyapunov groups. This study due to Einsiedler and Katok [**EK03, EK05**] by itself gives sufficient information to yield the following measure classification theorem:

9.5. Theorem (Einsiedler and Katok [**EK03**])**.** *Let Γ be a discrete subgroup in $G = \mathrm{SL}(3,\mathbb{R})$ and define $X = \Gamma\backslash G$. Let A be the full diagonal subgroup of G and suppose μ is an A-invariant and ergodic probability measure on X. Let*

$$a = \begin{pmatrix} t & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & t^{-1}s^{-1} \end{pmatrix} \in A$$

and suppose that

$$h_\mu(a) > \frac{1}{2}\big(\big|\log|\tfrac{t}{s}|\big| + \big|\log|t^2 s|\big| + \big|\log|ts^2|\big|\big).$$

Then μ is the Haar measure m_X on X and in particular Γ is a lattice.

We note that the expression in the parenthesis is the entropy of the Haar measure m_X. Hence, the theorem (as well as its generalizations below) says that an ergodic measure whose entropy is close to that of the Haar measure must be the Haar measure.

In the next section we present a completely different technique (the low entropy method) that will allow us to sharpen the above theorem, treating all positive entropy measures.

9.6. Fixing some $a \in A$ for which $\mathfrak{g}_a^-$ is nontrivial (equivalently, there is some $\lambda \in \Phi$ so that $|\lambda(a)| < 1$) we obtain a decomposition $\mathfrak{g}_a^- = \bigoplus_{i=1,\dots,\ell} \mathfrak{g}^{[\lambda_i]}$ into finitely many of these "coarse" Lyapunov subalgebras (corresponding to the subgroups $G^{[\lambda_i]}$).

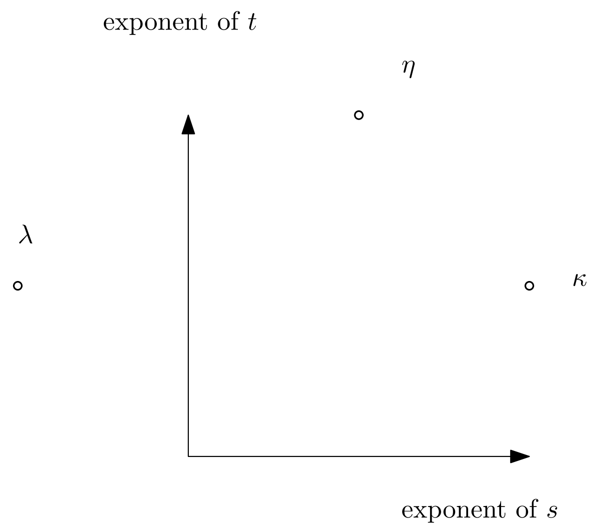

Figure 3. Weights for the Heisenberg group

We order these coarse Lyapunov weights $[\lambda_1], [\lambda_2], \dots, [\lambda_\ell]$ so that for each i, the weight λ_i, or more precisely the corresponding point in $\mathbb{Z}^k$ is not in the convex cone generated by the points corresponding to $\lambda_{i+1}, \dots, \lambda_\ell$ — we refer to this by saying that λ_i is *exposed*. This ordering guarantees that for every i there will be an element $a \in A$ so that $\lambda_i(a) = 1$ but $|\lambda_j(a)| < 1$ for $i < j < \ell$.

9.7. **Example:** Take $G = \mathrm{SL}(3, \mathbb{R})$, and A the full diagonal subgroup. Let α be the homomorphism $\alpha : (t, s) \mapsto \begin{pmatrix} t & & \\ & s & \\ & & t^{-1}s^{-1} \end{pmatrix}$. Here $\mathfrak{g}$ is the algebra of traceless matrices. Suppose now $a = \alpha(t, s)$ with $|t| < |s| < |t^{-1}s^{-1}|$. Then $\mathfrak{g}_a^-$ is the algebra of upper triangular nilpotent matrices. Moreover, the coarse Lyapunov subalgebras are the 1-dimensional spaces

$$\mathfrak{g}^{[\lambda]} = \begin{pmatrix} 0 & * & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix}, \quad \mathfrak{g}^{[\eta]} = \begin{pmatrix} 0 & 0 & * \\ & 0 & 0 \\ & & 0 \end{pmatrix}, \text{ and } \quad \mathfrak{g}^{[\kappa]} = \begin{pmatrix} 0 & 0 & 0 \\ & 0 & * \\ & & 0 \end{pmatrix},$$

with the corresponding weights $\lambda = ts^{-1}, \eta = t^2 s$ and $\kappa = ts^2$. As Figure 3 shows, $\lambda_1 = \lambda$, $\lambda_2 = \eta$ and $\lambda_3 = \kappa$ is a legitimate ordering (as would be the reverse ordering and λ, κ, η, but not η, λ, κ). We also have the corresponding subgroups

$$G^{[1]} = \begin{pmatrix} 1 & * & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \quad G^{[2]} = \begin{pmatrix} 1 & 0 & * \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \text{ and } \quad G^{[3]} = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & * \\ & & 1 \end{pmatrix}.$$

where $G^{[1]}$ and $G^{[3]}$ each commute with $G^{[2]}$, and the commutator $[G^{[1]}, G^{[3]}] = G^{[2]}$.

9.8. Theorem. *Let $A = \alpha((k^\times)^n)$, and suppose μ is an A-invariant measure on $X = \Gamma\backslash G$. Fix some $a \in A$, and choose an allowed order of the coarse Lyapunov subalgebras contracted by a (as described above). Then for μ-a.e. $x \in X$, we have*

$$\mu_x^{G^-} \propto \iota(\mu_x^{G^{[1]}} \times \mu_x^{G^{[2]}} \times \cdots \times \mu_x^{G^{[k]}})$$

where $\iota(g_1, g_2, \ldots, g_k) = g_1 g_2 \cdots g_k$ is the product map.

9.9**. Proof:** By assumption $[\lambda_1]$ is exposed within the set of all Lyapunov weights appearing in G_a^- so that there exists some $a' \in A$ with $G^{[1]} \subset G_{a'}^0$ and $U = G^{[2]} \cdots G^{[k]} = G_a^- \cap G_{a'}^-$. It follows easily that U is a normal subgroup of G_a^- and that $G_a^- \simeq G^{[1]} \ltimes U$ where the isomorphism is just the map ι taking the product. From Corollary 8.8 we deduce that $\mu_x^{G_a^-} \propto \iota(\mu_x^{G^{[1]}} \times \mu_x^U)$. Repeating the argument, starting with $G^{[2]}$ inside U, the theorem follows. □

9.10. Corollary.

$$h_\mu(a, G^-) = \sum_{i=1}^{k} h_\mu(a, G^{[i]})$$

This follows from Theorem 7.6 and Theorem 9.8 since the left hand side is

$$\lim_{n\to\infty} \frac{\log \mu_x^{G^-}(a^n B_1^{G^-} a^{-n})}{n}$$

and we have already shown in §7.18 that the particular shape of the set $B_1^{G^-}$ used in the definition does not matter. Using the product set

$$B_1^{G^{[1]}} B_1^{G^{[2]}} \cdots B_1^{G^{[k]}}$$

instead we obtain with the theorem that the left hand side splits into the corresponding expression for $G^{[i]}$. Hence in this setting our term 'entropy contribution' is quite accurate. We note, however, that in general such a formula does not hold for a finer foliation than the coarse Lyapunov subalgebras.

9.11**. Getting invariance.** In fact, more is true. Let us for now continue Example 9.7 (which will lead to the proof of Theorem 9.5), and consider $f \in C_c(G_a^-)$, and observe that

$$\begin{aligned} \int f(g) d\mu_x^{G^-} &= \int f(g_1 g_2 g_3) d\mu_x^{[1]}(g_1) d\mu_x^{[2]}(g_2) d\mu_x^{[3]}(g_3) \\ &= \int f(g_3 g_2 g_1) d\mu_x^{[3]}(g_3) d\mu_x^{[2]}(g_2) d\mu_x^{[1]}(g_1) \end{aligned}$$

where $\mu_x^{[i]} := \mu_x^{G^{[i]}}$. This follows from Theorem 9.8 by using the two allowed orders $1, 2, 3$ resp. $3, 2, 1$. Notice that, since both $G^{[1]}$ and $G^{[3]}$ commute with $G^{[2]}$, we can rewrite $g_1 g_2 g_3 = g_2 g_1 g_3$, and $g_3 g_2 g_1 = g_2 g_3 g_1 = (g_2[g_3, g_1]) g_1 g_3$. Inserting this above, and taking the leaf-wise measure for the $G^{[2]}$-action on G_a^- we find that $\mu_x^{[2]} \propto \mu_x^{[2]}[g_3, g_1]$ for $\mu_x^{[1]}$-a.e. g_1 and $\mu_x^{[3]}$-a.e. g_3 (by using Lemma 8.10 and by recalling that $[g_3, g_1] \in G^{[2]}$).

Now, if $[g_3, g_1]$ has infinite order, in other words if the element $[g_3, g_1]$ is nontrivial, then $\mu_x^{[2]}$ must be $[g_3, g_1]$-invariant; since otherwise, successive translations by $[g_3, g_1]$ would cause $\mu_2(B_r^{G^{[2]}})$ to grow exponentially, contradicting Theorem 6.30. Since the set $\{g_2 : (\mu_x^{[2]}).g_2 = \mu_x^{[2]}\}$ is closed, it follows that if both $\mu_x^{[1]}$ and $\mu_x^{[3]}$ are non-atomic, we must have $\mu_x^{[2]}$ invariant under $[\operatorname{supp}\mu_x^{[1]}, \operatorname{supp}\mu_x^{[3]}]$.

This is a significant restriction. By Poincaré Recurrence, we know that $\mu_x^{[1]} = \delta_e$ or $\operatorname{supp}\mu_x^{[1]}$ contains arbitrarily small (and large) elements; and similarly for $\mu_x^{[2]}$. In the example of the Heisenberg group above, there are only three possible cases: either $\mu_x^{[1]}$ or $\mu_x^{[3]}$ is trivial, or else the closed group generated by $[\operatorname{supp}\mu_x^{[1]}, \operatorname{supp}\mu_x^{[3]}]$ equals $G^{[2]}$, and so $\mu_x^{[2]}$ is a Haar measure on $G^{[2]}$.

9.12**. Proof of Theorem 9.5:** The above shows (in the notation of Example 9.7) that if $\mu_x^{[1]}$ and $\mu_x^{[3]}$ are both nontrivial at x, then a.s. μ_2 is the Haar measure on $G^{[2]}$. Lemma 7.16 shows that the set of points where $\mu_x^{[i]}$ is trivial is A-invariant, and so has either measure zero or one by ergodicity. Supposing that $\mu_x^{[1]}$ and $\mu_x^{[3]}$ are both nontrivial a.e., we get that $\mu_2 = \mu_x^{G^{[2]}}$ equals the Haar measure on $G^{[2]}$ a.e. and so that μ is invariant under $G^{[2]}$ by Problem 6.28. We now bring in entropy and the assumption to the theorem to justify the assumptions to this 'commutator argument'.

Let now $a \in A$ be as in the theorem. There are essentially two cases for elements of A: An element $a \in A$ is called *regular* if all of its eigenvalues are different, and is called *singular* if two eigenvalues are the same. If a is regular, then we may assume it is as in Example 9.7, for otherwise we get a group isomorphic to the Heisenberg group embedded in some other way into $\mathrm{SL}(3,\mathbb{R})$. If a is singular we may assume (again in the notation of Example 9.7) that $t = s$ with $|t| < 1$.

We define the opposite weight spaces

$$\mathfrak{g}^{[-1]} = \begin{pmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathfrak{g}^{[-2]} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ * & 0 & 0 \end{pmatrix}, \text{ and } \quad \mathfrak{g}^{[-3]} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & * & 0 \end{pmatrix},$$

and similarly the coarse Lyapunov subgroups.

In the singular case $G_a^- = G^{[2]}G^{[3]}$ and

$$h_\mu(a) = h_\mu(a, G_a^-) = h_\mu(a, G^{[2]}) + h_\mu(a, G^{[3]})$$

by Theorem 7.6 and Corollary 9.10. By Theorem 7.9 each summand on the right is bounded by $3\big|\log|t|\big|$ (which is precisely the entropy contribution for the Haar measure). By assumption on the entropy we have $h_\mu(a) > 3\big|\log|t|\big|$ (i.e., entropy is more than one half of the maximal entropy), so that both entropy contributions are positive. In turn, this shows that both leaf-wise measures $\mu_x^{[2]}$ and $\mu_x^{[3]}$ are nontrivial a.e. By symmetry of entropy $h_\mu(a) = h_\mu(a^{-1})$ we also get that both $\mu_x^{G^{[-2]}}$ and $\mu_x^{G^{[-3]}}$ are nontrivial a.e. However, the two subgroups $G^{[2]}$ and $G^{[-3]}$ do not commute and have commutator $G^{[1]}$. Moreover, the three groups $G^{[2]}$, $G^{[1]}$, and $G^{[-3]}$ generate a stable horospherical subgroup $G_{a'}^-$ (for some regular $a' \in A$) which is isomorphic to the Heisenberg group studied so far. By the above commutator argument we get that μ is invariant under $G^{[1]}$. Note that we could also have used the triple $G^{[3]}$, $G^{[-1]}$, and $G^{[-2]}$ to obtain invariance under $G^{[-1]}$. So we obtain in the singular case that in fact all leaf wise measures of the coarse Lyapunov subgroups are nontrivial a.e. (and some of them are Haar measures). This is enough to imply that μ is invariant under all coarse Lyapunov subgroups (and so must be the Haar measure m_X) by the commutator argument: Any of the coarse Lyapunov subgroups $G^{[i]}$ is the commutator of two other coarse Lyapunov subgroups $G^{[j]}$ and $G^{[k]}$ such that all three of them generate a stable horospherical subgroup (isomorphic to the Heisenberg group).

In the regular case there are a few more possibilities. We know that $h_\mu(a) = h_\mu(a, G^{[1]}) + h_\mu(a, G^{[2]}) + h_\mu(a, G^{[3]})$. In this case the upper bounds coming from Theorem 7.9 are for the three summands $\log|\frac{s}{t}|$, $-\log|t^2 s|$, resp. $-\log|ts^2|$. Note that the second term equals the sum of the other two, so that our assumption translates to the assumption that at least two out of the three entropy contributions must be positive — any particular entropy contribution coming from one coarse Lyapunov subgroup cannot give more than one half of the maximal entropy. Hence we conclude that at least two of the three leaf-wise measure $\mu_x^{G^{[1]}}$, $\mu_x^{G^{[2]}}$, or $\mu_x^{G^{[3]}}$ must be nontrivial a.e. From the above we know that if $\mu_x^{G^{[1]}}$ and $\mu_x^{G^{[3]}}$ are nontrivial a.e., then $\mu_x^{G^{[2]}}$ is actually the Haar on $G^{[2]}$ a.e. and so again nontrivial a.e. Using again symmetry of entropy and the same commutator argument within various stable horospherical subgroups the theorem follows easily. □

9.13**. Problem.** Prove the following version of the high entropy theorem for quotients of $G = \mathrm{SL}(n, \mathbb{R})$ (starting with $n = 3$). Suppose μ is an A-invariant and ergodic probability measure on $X = \Gamma\backslash G$ such that all nontrivial elements of A have positive entropy. Deduce that μ is the Haar measure on X.

Generalizing the commutator argument leads to the following theorem.

9.14. Theorem. *(High entropy theorem) Let μ be an A-invariant and ergodic probability measure on $X = \Gamma\backslash G$. Let $[\zeta]$ and $[\eta]$ be coarse Lyapunov weights such that $[\zeta] \neq [\eta] \neq [\zeta^{-1}]$. Then for a.e. x, μ is invariant under the group generated by $[\operatorname{supp}\mu_x^{G^{[\zeta]}}, \operatorname{supp}\mu_x^{G^{[\eta]}}]$. In fact the same holds with $\operatorname{supp}\mu_x^{G^{[\zeta]}}$ and $\operatorname{supp}\mu_x^{G^{[\eta]}}$ replaced by the smallest Zariski closed A-normalized subgroups containing the supports.*

To prove this in general we need a few more preparations.

9.15**. Invariance subgroups.** Let $a \in A$ and assume $U \subset G_a^-$ is a-normalized. We define for any x the closed subgroup

$$\mathrm{Stab}_x^U = \{u \in U : u\mu_x^U = \mu_x^U\} < U.$$

Over $\mathbb{R}$, there are — in some sense — very few closed subgroups, which restricts the possibilities for Stab_x^U. More precisely, we claim that Stab_x^U equals the connected component $(\mathrm{Stab}_x^U)^0$ of the identity in Stab_x^U, at least for a.e. x.

To see this let $d(x)$ be the distance from $(\mathrm{Stab}_x^U)^0$ to $\mathrm{Stab}_x^U \setminus (\mathrm{Stab}_x^U)^0$ (using a left invariant metric) and define $d(x) = 0$ if the claim holds for x.

Now $a^n \mathrm{Stab}_x^U a^{-n} = \mathrm{Stab}_{a^n.x}^U$ (see Lemma 7.16), and since a contracts U, we must have $d(a^n x) \to 0$ as $n \to \infty$. Thus, we see that $d(x) = 0$ for a.e. x by Poincaré recurrence. Therefore, Stab_x^U is connected.

9.16**. Problem.** Over $\mathbb{Q}_p$ it doesn't make sense to speak of the connected component (as it would be the trivial group in any case), but we can speak of the maximal algebraic subgroup contained in Stab_x^U. For this recall that exp and log are polynomial isomorphisms between the Lie algebra of U and U. Also sub-Lie algebras are mapped under this map to Zariski closed subgroups of U. We may define $(\mathrm{Stab}_x^U)^0$ to be the exponential image of the largest subalgebra contained in the logarithmic image of Stab_x^U. (The reader may verify that over $\mathbb{R}$ this defines the connected component.) Show that $(\mathrm{Stab}_x^U)^0 = \mathrm{Stab}_x^U$ a.e. (The situation is in a sense opposite to the real case where one had to apply the contraction by a to obtain small elements — over $\mathbb{Q}_p$ we can simply take a power of an element of Stab_x^U to obtain small elements but one has to apply the expansion a^{-1} and Poincaré recurrence to obtain big elements of Stab_x^U.)

9.17**.** Stab_x^U **is normalized by** A**.** If U is one-dimensional, then this follows simply from $\mathrm{Stab}_x^U = (\mathrm{Stab}_x^U)^0$. However, in general this is a special property which again is a result of Poincaré recurrence.

In fact, as $\mathrm{Stab}_x^U = (\mathrm{Stab}_x^U)^0$ it is uniquely determined by its Lie algebra $\mathfrak{s}_x$. Notice that $\mathfrak{s}_{a^n.x} = \mathrm{Ad}_a^n \mathfrak{s}_x$ a.e. However, A is generated by elements $a \in A$ whose eigenvalues are all powers of a single number t. For these it follows that either $\mathfrak{s}_{a.x} = \mathfrak{s}_x$ or that $\mathfrak{s}_{a^n.x}$ approaches a sub Lie algebra $\mathfrak{h}$ for which $\mathrm{Ad}_a \mathfrak{h} = \mathfrak{h}$. In fact, this follows from considering the alternating tensor product of the Lie algebra of U of degree equal to the dimension of $\mathfrak{s}_x$ (which is independent of x for a.e. x by ergodicity of A): The action of the class A element a on that space still has all eigenvalues equal to a power of t and either the point corresponding to $\mathfrak{s}_x$ is an eigenvector for that action or it approaches projectively one when the iterates of a are applied to it. By Poincaré recurrence the latter is not possible for a.e. x, hence the conclusion.

In particular, the above shows that $\mathrm{Stab}^U = \mathrm{Stab}_x^U$ is independent of x for a.e. x as μ is A-ergodic. This makes the following lemma useful for $H = U$ and $L = \mathrm{Stab}^U$.

9.18. Lemma. *Let H act on X, and $L < H$ be a subgroup. Suppose that for every $l \in L$, we have $l\mu_x^H = \mu_x^H$ for a.e. $x \in X$. Then μ is L-invariant.*

This follows literally from Problem 6.28 and Lemma 8.10, but also purely from the argument behind Problem 6.28.

9.19**. Proof of Theorem 9.14.** We take two coarse Lyapunov weights $[\zeta]$ and $[\eta]$ satisfying $[\zeta] \neq [\eta] \neq [\zeta^{-1}]$ as in the theorem. Then there exists an A-normalized subgroup $H \subset G_a^-$ (for some $a \in A$) which is a product of coarse Lyapunov subgroups for which $[\zeta], [\eta]$ are both exposed. This implies as in Theorem 9.8 that

$$\mu_x^H \propto \iota(\mu_x^{G^{[\zeta]}} \times \mu_x^{G^{[\eta]}} \times \mu_x^U) \propto \iota(\mu_x^{G^{[\eta]}} \times \mu_x^{G^{[\zeta]}} \times \mu_x^U)$$

where U is the product of all coarse Lyapunov subgroups that are contained in H except for $G^{[\zeta]}$ and $G^{[\eta]}$. The argument in 9.11 now shows that μ_x^U must be invariant under $[\mathrm{supp}\,\mu_x^{G^{[\zeta]}}, \mathrm{supp}\,\mu_x^{G^{[\eta]}}] \in \mathrm{Stab}^U$. Together with the above discussion, this implies that μ is invariant under $[\mathrm{supp}\,\mu_x^{G^{[\zeta]}}, \mathrm{supp}\,\mu_x^{G^{[\eta]}}]$ for a.e. x as claimed.

We now wish to prove the additional claim that μ is, for a.e. x, also invariant under the commutators $[h^{[\zeta]}, h^{[\eta]}]$ of elements of the smallest Zariski closed

a-normalized subgroups $P^{[\zeta]} \ni h^{[\zeta]}$ and $P^{[\eta]} \ni h^{[\eta]}$ containing $\operatorname{supp} \mu_x^{G^{[\zeta]}}$ resp. $\operatorname{supp} \mu_x^{G^{[\eta]}}$. We may assume ζ (and similarly η) is an indivisible weight, i.e., all other weights which are coarsely equivalent to ζ are powers of ζ. Now notice that Zariski closed subgroups of the unipotent group $G^{[\zeta]}$ are precisely the exponential images of subalgebras of the Lie algebra of $G^{[\zeta]}$. To prove the above we are first claiming that if $g^{[\zeta]} \in \operatorname{supp} \mu_x^{G^{[\zeta]}}$ and $h^{[\eta]} \in \operatorname{supp} \mu_x^{G^{[\eta]}}$ and we write

$$\begin{aligned} \log g^{[\zeta]} &= u^{[\zeta]} = u_\zeta + u_{\zeta^2} + \cdots \\ \log h^{[\eta]} &= v^{[\eta]} = v_\eta + v_{\eta^2} + \cdots \end{aligned}$$

with $u_\xi, v_\xi \in \mathfrak{g}^\xi$, then a.s. $\exp[u_{\zeta^m}, v_{\eta^n}]$ preserves the measure μ (or equivalently $\mu_x^{G^{[\zeta^m\eta^n]}}$). For this we have to proceed by induction on the complexity of the subgroup U. If U is the trivial subgroup, there is nothing to prove as in this case $G^{[\zeta]}$ and $G^{[\eta]}$ commute. For the general case we have to compare the group theoretic commutator

$$[g^{[\zeta]}, h^{[\eta]}] = (g^{[\zeta]})^{-1}(h^{[\eta]}])^{-1} g^{[\zeta]} h^{[\eta]}$$

with the Lie theoretic commutator in the Lie algebra $\mathfrak{g}$. By the Campbell-Baker-Hausdorff formula the former equals

$$[g^{[\zeta]}, h^{[\eta]}] = \exp\bigl([u^{[\zeta]}, v^{[\eta]}] + \cdots\bigr), \tag{9.19a}$$

where the dots indicate a finite sum of various iterated commutators of $u^{[\zeta]}$ and $v^{[\eta]}$ with $[u^{[\zeta]}, v^{[\eta]}]$. Let us refer to $[u^{[\zeta]}, v^{[\eta]}]$ as the main term. Note that in $\log g$ the only term of weight $\zeta\eta$ is $[u_\zeta, v_\eta]$ (which is part of the main term), as all terms indicated by the dots only contain terms of weight $\zeta^k\eta^\ell$ with $k + \ell \geq 3$. As $g \in \operatorname{Stab}^U$ and this group is A-normalized and equals the exponential image of its Lie algebra, we see that $[u_\zeta, u_\eta]$ belongs to the Lie algebra of Stab^U. We note that this implies that $\exp[u_\zeta, v_\eta]$ preserves $\mu_x^{G^{[\zeta\eta]}}$ which implies that $\exp[u^\zeta, v^\eta] \in \operatorname{supp} \mu_x^{G^{[\zeta\eta]}}$. If we replace η by $\zeta\eta$ and $h^{[\eta]}$ by $\exp[u_\zeta, v_\eta]$, we obtain a situation as before but with a smaller dimensional subgroup U' replacing U. By the inductive hypothesis we conclude that all terms of the form $\exp[u_{\zeta^m}, [u_\zeta, v_\eta]]$ preserve the measure. However, this now shows that the term inside the exponential in (9.19a) corresponding to weight $\zeta^2\eta$ is the sum of $[u^{\zeta^2}, v^\eta]$ (which is part of the main term) and of a multiple of $[u_\zeta, [u_\zeta, v_\eta]]$. As before we conclude that this sum belongs to the Lie algebra of Stab^U, which in return shows the same for $[u^{\zeta^2}, v^\eta]$ (and similarly for $[u^\zeta, v^{\eta^2}]$). Proceeding inductively one shows in the same manner that all components $[u_{\zeta^m}, u_{\eta^n}]$ of the main term belongs to the Lie algebra of Stab^U.

As the Lie bracket is bilinear, it is clear that we may multiply the various components u_{ζ^m} and u_{η^n} by scalars without affecting the conclusion. It remains to show that if $[u_1, v]$ and $[u_2, v]$ for $u_1, u_2 \in \mathfrak{g}^{[\zeta]}$ and $v \in \mathfrak{g}^{[\eta]}$ belong to the Lie algebra of Stab^U, then the same is true for $[[u_1, u_2], v]$. However, this follows from the Jacobi identity

$$[[u_1, u_2], v] = -[[v, u_1], u_2] - [[u_2, v], u_1]$$

where the terms on the right belong to the Lie algebra of Stab^U by what we already established.

The above together shows that we may take u, v in the Lie algebra generated by $\log \operatorname{supp} \mu_x^{[\zeta]}$ resp. generated by $\log \operatorname{supp} \mu_x^{[\eta]}$ and obtain that $[u, v]$ belongs to the

Lie algebra generated by Stab^U. This together with the Campbell-Baker-Hausdorff formula is the desired result. □

As a corollary, we have the following entropy gap principle.

9.20. Theorem. *Let G be a* simple *algebraic group defined over $\mathbb{R}$ and connected in the Hausdorff topology. Let $\Gamma < G$ be a discrete subgroup. Say $A \subset G$ is a split torus of rank at least 2. Suppose μ is an A-invariant and ergodic probability measure on $X = \Gamma\backslash G$. Then for every a, there exists $h_0 < h_{m_X}(a)$ such that $h_\mu(a) > h_0$ implies that $\mu = m_X$ is the Haar measure on X.*

We note that the entropy $h_{m_X}(a)$ of the Haar measure m_X on X is determined by a concrete formula involving only Ad_a, and so is independent of Γ. If Γ is not assumed to be a lattice, we still write $h_{m_X}(a)$ for this expression. With this in mind, we do not have to assume in the above theorem that Γ is a lattice, rather obtain this as part of the conclusion if only $h_\mu(a) > h_0$.

9.21. **Example:** We illustrate Theorem 9.20 as well as another formulation of the high entropy theorem in the case of $G = \mathrm{SL}(3,\mathbb{R})$ (as in Problem 9.13).

Say

$$a = \begin{pmatrix} e^{-t} & & \\ & 1 & \\ & & e^t \end{pmatrix} \qquad G_a^- = \begin{pmatrix} 1 & G^{[1]} & G^{[2]} \\ & 1 & G^{[3]} \\ & & 1 \end{pmatrix}$$

We have

$$h_{m_X}(a) = \sum_{i=1}^{3} h_\lambda(a, G^{[i]}) = t + 2t + t.$$

If we take $h_0 = 3t$, then $h_\mu(a) > 3t$ implies that there is an entropy contribution from all 3 expanding directions, and so all three leaf-wise measures are non-trivial almost everywhere. Therefore the support of each $\mu_x^{G^{[i]}}$ is all of $G^{[i]}$, and the high-entropy method then implies that μ is invariant under all $G^{[i]}$, and therefore invariant under G, so μ is the Haar measure on X.

Now suppose

$$a = \begin{pmatrix} e^{-2t} & & \\ & e^t & \\ & & e^t \end{pmatrix} \qquad G_a^- = \begin{pmatrix} 1 & G^{[1]} & G^{[2]} \\ & 1 & \\ & & 1 \end{pmatrix}$$

we have central directions that are neither expanded nor contracted by a. Here, we have

$$h_{m_X}(a) = 3t + 3t$$

and $h_\mu(a) > h_0 = 3t = \frac{1}{2} h_{m_X}(a_t)$ implies that the Zariski closure of $\operatorname{supp} \mu_x^{G^{[i]}}$ is a.s. all of $G^{[i]}$, and so by taking the commutator we get invariance of μ under the central direction as well (eg., since $[G^{[1]}, G^{[-2]}]$ is the lower central direction, μ is invariant under this direction as well.)

Now suppose we know that, for *every* a, we have $h_\mu(a) > 0$. By examining the element $a = \begin{pmatrix} e^{-2t} & & \\ & e^t & \\ & & e^t \end{pmatrix}$ as above, we find that either $\mu_x^{G^{[1]}}$ or $\mu_x^{G^{[2]}}$ is nontrivial almost everywhere. If we assume that, say, $\mu_x^{G^{[2]}}$ is trivial (and hence that $\operatorname{supp} \mu_x^{G^{[1]}}$

is Zariski dense in $G^{[1]}$ a.s.), then we can use the element $a = \begin{pmatrix} e^{-t} & & \\ & e^{-t} & \\ & & e^{2t} \end{pmatrix}$ to show that $\mu_x^{G^{[3]}}$ has Zariski dense support in $G^{[3]}$ a.s., and we get invariance under $G^{[2]}$ anyway. By similar arguments using other singular elements a, we can get invariance under any $G^{[i]}$, and so μ must be the Haar measure on X.

9.22. Lemma. *Let $V < U$ be a-normalized closed subgroups of the stable horospherical subgroup G_a^-. Suppose that* $\operatorname{supp} \mu_x^U \subset V$ *for a.e. x. Then*

$$h_\mu(a, U) \leq h_{m_X}(a, V) \leq h_{m_X}(a, U)$$

In fact, the second inequality is strict (and uniformly so) if V is a proper *subgroup of U.*

Note that the assumption on the support of μ_x^U implies that $h_\mu(a, V) = h_\mu(a, U)$. With this in mind the lemma follows from Theorem 7.9.

9.23. Lemma. *Let $a \in A$ and let $[\eta]$ be coarse Lyapunov weight contracted by a. Under the hypotheses of 9.20, for h_0 large enough and μ-a.e. x, we have that $G^{[\eta]}$ is the smallest a-normalized Zariski closed subgroup containing the support of $\mu_x^{G^{[\eta]}}$.*

This follows by combining Lemma 9.22 and Corollary 9.10.

9.24. Proposition. *For any nontrivial $a \in A$, the simple group G is generated by the set of commutators $[G^{[\lambda]}, G^{[\eta]}]$ of all pairs of coarse Lyapunov subgroups which satisfy $[\eta] \neq [\lambda] \neq [\eta^{-1}]$ and $\eta(a) \neq 1 \neq \lambda(a)$.*

We see that Theorem 9.20 follows from Theorem 9.14 together with Proposition 9.24 and Lemma 9.23.

9.25**. Proof of Proposition 9.24.** Let V be a lower dimensional subgroup of the group of characters of A. Let

$$\mathfrak{w} = \operatorname{span}\{\mathfrak{g}^\lambda, [\mathfrak{g}^\eta, \mathfrak{g}^\lambda] : \lambda, \eta \notin V\}$$

We claim that $\mathfrak{w}$ is a Lie ideal of $\mathfrak{g}$. To check this, we first take $x \in \mathfrak{g}^\delta \subset \mathfrak{w}$ (first type of elements) for some $\delta \notin V$, and some $z \in \mathfrak{g}^\zeta$ and look at $[x, z]$. There are two cases:

(i) If $\zeta \notin V$, then $[x, z] \in \mathfrak{w}$ by definition due to the second type of elements of $\mathfrak{w}$.
(ii) If $\zeta \in V$, then $[x, z] \in \mathfrak{g}^{\delta\zeta} \subset \mathfrak{w}$ due to the first type of elements since $\delta\zeta \notin V$.

Assume now $[x, y] \in \mathfrak{w}$ with $x \in \mathfrak{g}^\lambda$, $y \in \mathfrak{g}^\eta$, and $\lambda, \eta \notin V$ as in the second type of elements of $\mathfrak{w}$. Also let as before $z \in \mathfrak{g}^\zeta$. There are again two cases:

(i) If $\zeta \notin V$ then by the above cases $[[x, y], z] \in \mathfrak{w}$.
(ii) In the remaining case of $\zeta \in V$ we use the Jacobi identity

$$[[x, y], z] = -[[y, z], x] - [[z, x], y]$$

which leads to the expressions $[[y, z], x]$ and $[[z, x], y]$. However, $[y, z] \in \mathfrak{g}^{\eta\zeta}$ with $\eta\zeta \notin V$ and $\lambda \notin V$ shows that $[[y, z], x]$ is an expression of the second type in the definition of $\mathfrak{w}$. The same holds for $[[z, x], y]$ which shows that $[[x, y], z] \in \mathfrak{w}$ as claimed.

As V is assumed to be lower dimensional and the weights span, $\mathfrak{w}$ is nontrivial and hence equal to $\mathfrak{g}$ by the assumption that G is simple.

Now we let V be the kernel of the evaluation map $\lambda \mapsto \lambda(a)$ and let $\zeta \in V$ be a nontrivial weight. Then the above claim shows that all elements of the Lie algebra $\mathfrak{g}^\zeta$ can be written as sums of Lie commutators of elements of $\mathfrak{g}^\eta, \mathfrak{g}^\lambda$ with $\eta, \lambda \notin V$. Here we cannot have $[\eta] = [\lambda]$ or $[\eta] = [\lambda^{-1}]$ as otherwise the commutator would belong to a weight space $\mathfrak{g}^\zeta$ also satisfying that ζ is trivial, $[\eta] = [\zeta]$, or that $[\eta] = [\zeta^{-1}]$ which is impossible as $\eta \notin V$ but $\zeta \in V$.

Similarly we may now set V equal to the subgroup of characters equivalent to a given nontrivial λ or its inverse λ^{-1}. Applying again the above we see that the elements of the weight space $\mathfrak{g}^\lambda$ can be written as sums of Lie commutators of elements of $\mathfrak{g}^\eta, \mathfrak{g}^\lambda$ with $[\eta] \neq [\lambda] \neq [\eta^{-1}]$. (By the argument above we do not have to restrict ourselves any longer to weight spaces that do not commute with a). Therefore, all elements of all nonzero weight spaces can be generated by the Lie brackets that we consider.

Finally, note that the Lie algebra generated (set V equal to the trivial group) by all nonzero weight spaces is the whole of $\mathfrak{g}$, so that $\mathfrak{g}$ is generated indeed by the Lie brackets that we consider. □

10. Invariant measures for higher rank subgroups A, the low entropy method

10.1. In this section we sketch the proof of a theorem regarding A-invariant measures where only positivity of entropy is assumed (instead of entropy close to being maximal). In addition to the ideas we already discussed they use one more method, which we refer to as the low entropy method. This method was first used in [**Lin06**]; one of the main motivations being the Arithmetic Quantum Unique Ergodicity Conjecture which is partially resolved in that paper — see §13.

10.2. A basic feature of this method is that it gives a prominent role to the dynamics of the unipotent groups normalized by A, even though these unipotent groups a-priori do not preserve the measure in any way. Ideas of Ratner, particularly from her work on the horocycle flow [**Rt82a, Rt82b, Rt83**] are used in an essential way.

We first present this method which has been extended to fairly general situations in [**EL08**] in one particular case (the reader who is averse to p-adic numbers is welcome to replace $\mathrm{SL}(2,\mathbb{Q}_p)$ by $\mathrm{SL}(2,\mathbb{R})$ in the theorem and its proof below).

10.3. Theorem. *Let $X = \Gamma\backslash \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{Q}_p)$, where Γ is an irreducible lattice in $G = \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{Q}_p)$. Let $A = \left(\begin{pmatrix} * & \\ & * \end{pmatrix} \times e\right)$ be the (one parameter) diagonal subgroup in the $\mathrm{SL}(2,\mathbb{R})$ factor. Suppose μ is an A-invariant probability measure such that*

- *μ is $\mathrm{SL}(2,\mathbb{Q}_p)$-recurrent.*
- *Almost all A-ergodic components of μ have positive entropy under the A-flow.*

Then μ is the Haar measure on X.

10.4. We recall that a lattice Γ in $G_1 \times G_2$ is said to be irreducible if the kernel of the projection to each factor is finite. For the case at hand of $G_1 = \mathrm{SL}(2,\mathbb{R})$ and $G_2 = \mathrm{SL}(2,\mathbb{Q}_p)$ this is equivalent to both projections being dense. This assumption of irreducibility is clearly necessary; the assumption that Γ is a lattice (i.e., that it has finite covolume) is not, though it is not clear that classifying *probability* measures in the non-compact case is a very natural question. An example of an irreducible lattice in $\mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{Q}_p)$ is $\mathrm{SL}(2,\mathbb{Z}[\frac{1}{p}])$ (embedded diagonally).

10.5. We remark that unlike many measure classification theorems it is not possible to reduce Theorem 10.3 to the case of μ being A-ergodic. This is because if one takes an arbitrary measure μ satisfying the condition of the theorem and take its ergodic decomposition with respect to the A action there is no reason to expect the ergodic components to remain $\mathrm{SL}(2,\mathbb{Q}_p)$-recurrent. The fact that we are considering general invariant measures requires us to demand that not only does μ have positive entropy under A, but that each ergodic component has positive entropy.

10.6. The requirement that μ be $\mathrm{SL}(2,\mathbb{Q}_p)$-recurrent is clearly necessary, there are plenty of A invariant and ergodic measures on $\Gamma\backslash\mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{Q}_p)$ with positive entropy. E.g., when $\Gamma = \mathrm{SL}(2,\mathbb{Z}[\frac{1}{p}])$ as above, $\Gamma\backslash\mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{Q}_p)$ is a compact extension of $\mathrm{SL}(2,\mathbb{Z})\backslash\mathrm{SL}(2,\mathbb{R})$ (with the action of A respected by the corresponding projection map) and hence any A-invariant and ergodic measure on $\mathrm{SL}(2,\mathbb{Z})\backslash\mathrm{SL}(2,\mathbb{R})$ can be lifted to an invariant and ergodic measure on $\Gamma\backslash\mathrm{SL}(2,\mathbb{R})\times \mathrm{SL}(2,\mathbb{Q}_p)$ with exactly the same entropy[(34)].

10.7. **Outline of Proof. The starting point.** Let $T = \mathrm{SL}(2,\mathbb{Q}_p)$ and let $U = \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix}$ be the real upper unipotent subgroup; then $U = G_a^-$ and $T < C_G(a) \cap C_G(U)$ for e.g.[(35)] $a = \left(\begin{pmatrix} e^{-1} & \\ & e \end{pmatrix}, \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \right) \in A$. In particular, the assumptions to Corollary 8.8 are satisfied and the leaf-wise measures for the subgroup $H = TU$ are product measures a.s. By Corollary 8.13 there is a subset of full measure $X' \subset X$ such that we have $\mu_x^U = \mu_y^U$ whenever $x, y \in X'$ belong to the same $\mathrm{SL}(2,\mathbb{Q}_p)$-orbit. This shows together with the assumed recurrence that we can find many close-by points with the same leaf-wise measures, i.e., $y = (g_1, g_2).x$ with the displacements $g_1 \in \mathrm{SL}(2,\mathbb{R})$ and $g_2 \in \mathrm{SL}(2,\mathbb{Q}_p)$ both close to the identity and $\mu_x^U = \mu_y^U$.

As we have already observed in Example 6.5.2 the coincidence of leaf-wise measures can have strong implications. This is the case here. By replacing both x and y by $u.x$ and $u.y$ for some (in a certain sense) typical $u \in U$, we bring the polynomial shearing properties of the U-flow in the picture.

---

(34) This last claim requires some justification; what is immediate and is sufficient for our purpose is that the lifted measure would have at least the same entropy as the original measure. Also if the lifted measure is not ergodic, then one can take a typical ergodic component of it which will also be a lift of the original measure.

(35) Here e is the constant $2.71828\ldots$; below and above e is also used to denote the identity element of G.

10.8**. Polynomial divergence.** If, starting with $y = (g_1, g_2).x$, one moves along the U-orbit, the displacement of $x' = u.x$ and $y' = u.y$ is the conjugate (ug_1u^{-1}, g_2) and the U-action by conjugation is shearing depending polynomially on the time parameter in U. More precisely, given x and $y = (g_1, g_2).x$ with $g_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \notin U$ (an assumption which we will need to justify), we apply $u(s) = \begin{pmatrix} 1 & s \\ & 1 \end{pmatrix}$ to both to get

$$x_s = \begin{pmatrix} 1 & s \\ & 1 \end{pmatrix}.x$$

$$y_s = \begin{pmatrix} 1 & s \\ & 1 \end{pmatrix}(g_1, g_2).x = (g_{(s)}, g_2)x_s$$

We compute

$$g_{(s)} = u(s)g_1u(-s) = \begin{pmatrix} a + cs & b + (d-a)s - cs^2 \\ c & d - cs \end{pmatrix}. \tag{10.8a}$$

Hence the terms contributing to the divergence are $|d - a|$ and $\sqrt{|c|}$. (As it turns out using the square root puts the two terms on more equal footing). For $S = \min\left(\frac{1}{|d-a|}, \frac{1}{\sqrt{|c|}}\right)$, we expect to have $g_{(S)} \approx \begin{pmatrix} 1 & r \\ & 1 \end{pmatrix} = u(r)$ for some nontrivial r. Let us explain this more precisely: starting with the top right corner it is possible that by some coincidence for $s = S$ the terms $(d - a)s$ and cs^2 might cancel each other, but this would be an exception: for most $s \in [-S, S]$ the top right coordinate $b + (d - a)s - cs^2$ in (10.8a) will be of the order 1 (for a linear polynomial this is obvious, for the general quadratic polynomial one only needs that s is sufficiently far from both roots). Also $g_{(s)} = u(s)g_1u(-s)$ is bounded for $s \in [-S, S]$, and except for the top right entry, for s in that range, all other entries will be very close to the corresponding entries in the identity matrix (i.e., to one for the diagonal elements, and to zero for the bottom left corner): indeed, the diagonal entries are $a + cs$ and $d - cs$ which are close to 1 as $|cs| \leq |c|S \leq \sqrt{|c|}$ and the bottom left entry c is close to zero (and unchanged).

Hence $g_{(s)}$ will indeed be approximately $u(r)$ for some nontrivial and bounded r for most $s \in [-S, S]$.

10.9**. Choosing the correct** $u(s)$**, two conditions.** We will need to choose some $s \in [-S, S]$ such that $g_{(s)}$ has significant size. By the above discussion this is quite easy and a purely **algebraic condition**. At the same time $x' = x_s$ and $y' = y_s$ should have good properties with respect to the measure μ, which is a **measure theoretic condition**, the verification of which requires more work.

Clearly some condition on the points x' and y' is needed for them to give any meaningful information about the measure μ — the most important for us will be that both x' and y' belong to some compact set K on which the map sending a point z to the leafwise measure μ_z^U is continuous. We can also assume that this set K is contained in the conull set on which Theorem 6.3 (iii) holds so that (by applying that theorem to both x and y) we can deduce from our original assumption $\mu_x^U = \mu_y^U$ that $\mu_{x'}^U = \mu_{y'}^U$.

If we could take a limit, taking the original points x, y from a sequence x_k, y_k ever closer together (which forces $r_k \to \infty$), the limit points x'', y'' of a common subsequence of two points x'_k, y'_k will end up being different points on the same U-orbit. Restricting everything to a large compact set K where $z \mapsto \mu_z^U$ is continuous we would obtain $\mu_{x''}^U = \mu_{y''}^U$ and since $y'' = u.x''$ for some nontrivial $u \in U$ also $\mu_{x''}^U \propto \mu_{y''}^U u$ by Theorem 6.3 (iii). This leads to U-invariance of μ almost[36] as in the last section.

For this to work we need to ensure that given the two close-by points x_k, y_k with the same leaf-wise measures (and other good properties that hold on sets of large measure) we can find some $u(s_k) \in U$ such that $x'_k = u(s_k).x_k, y'_k = u(s_k).y_k \in K$ and the displacement is significant but of bounded size. As explained above it is easy to find some S_k depending on the displacement $g_1^{(k)}$ such that $u(s)g_1^{(k)}u(s)^{-1}$ is significant but bounded for all $s \in [-S_k, S_k]$ except those belonging to two small subintervals of $[-S_k, S_k]$. So basically we have two requirements for s_k, it shouldn't belong to one of two small subintervals which have been found using purely algebraic properties, and we also want both points $x'_k = u(s_k).x_k, y'_k = u(s_k).y_k$ to belong to the compact set K on which everything behaves nicely — which is a measure theoretic property involving μ since all we know about K is that it has large μ-measure.

10.10**. A maximal ergodic theorem.** To prove the latter property we need a kind of ergodic theorem for the U-action with respect to μ, even though we do not know invariance under U. A maximal ergodic theorem for the U-action would imply that for a given set of large measure K, the set of points x, for which there is some scale S for which it is not true that for *most* $s \in [-S, S]$ we have $u(s).x \in K$, has small μ-measure (and so can be avoided in the argument). However, here the correct notion of *most* must come from the measure μ_x^U instead of the Lebesgue measure as μ is not known to be invariant under U.

There are several versions of such maximal ergodic theorems in the literature starting from Hurewicz [**Hur44**]; see also [**Bec83**]. In [**Lin06**] a variant proved in the appendix to that paper jointly with D. Rudolph was used. An alternative approach which we have employed in [**EL08**] is to use the decreasing Martingale theorem by using the sequence of σ-algebras $a^{-n}\mathcal{A}$, where $\mathcal{A}$ is subordinate to U on a set of large measure and a-decreasing as in Definition 7.25. The latter approach has the advantage of working in greater generality, see Comment 7.38.

10.11**. Compatibility issue.** Assume now that a sufficient form of such a maximal ergodic theorem holds for the U-action. This then implies, starting with sufficiently well behaved initial points x_k, y_k (with $\mu_{x_k}^U = \mu_{y_k}^U$), that for $\mu_{x_k}^U$-most $s \in [S_k, S_k]$ (say for 90%) we have $u(s).x_k, u(s).y_k \in K$. Even so there is still a gap in the above outline: Can we ensure that the two subintervals of $[-S_k, S_k]$, where $u(s)g_1^{(k)}u(s)^{-1}$ is too little, have also small mass with respect to $\mu_{x_k}^U$? This is desired as it would ensure the compatibility of the algebraic and measure-theoretic properties needed, since in this case for $\mu_{x_k}^U$-most $s \in [-S_k, S_k]$ both properties would hold. However, if e.g. $\mu_{x_k}^U$ is trivial, i.e., is supported on the identity only, this is not the case. Luckily by assumption entropy is positive for a.e. ergodic component which translates to $\mu_{x_k}^U$ being nontrivial a.e. Even so, $\mu_{x_k}^U$ could give

[36]The cautious reader may be concerned about the lack of ergodicity assumption here. Indeed one first only obtains that some ergodic component is U-invariant, but one may apply the whole argument to the measure restricted to the subset where μ_x^U is not the Lebesgue measure to obtain a contradiction.

large mass to very small subintervals and the compatibility of the two conditions does not seem automatic.

10.12**. Self-similarity of leaf-wise measures.** What rescues the argument is a kind of self-similarity of the measures μ_x^U. E.g. if one assumes a doubling condition of the form that there exists some $\rho \in (0,1)$ for which

$$\mu_x^U(B_{\rho S}^U) < \frac{1}{2}\mu_x(B_S^U) \text{ for all } S > 0, \tag{10.12a}$$

then sufficiently small symmetric subintervals of a given interval $[-S,S]$ also get small μ_x^U-mass. (Given such a ρ we then would adjust the meaning of 'significant' in the discussion of §10.8 and §10.9.) There is no reason why such a strong regularity property of the conditional measures should hold. However, the A-action on X together with Lemma 7.16 implies some regularity properties: E.g. by Poincaré recurrence there are infinitely many S such that (after rescaling) μ_x^U restricted to B_S^U is very similar to μ_x^U restricted to B_1^U. To obtain something similar to (10.12a) we notice first that there is some $\rho > 0$ such that

$$\mu_x^U(B_\rho^U) < \frac{1}{2} \tag{10.12b}$$

except possibly on a set Z of small μ-measure. Then one can apply the standard maximal ergodic theorem for the action of

$$a^r = \left(\begin{pmatrix} e^{-r} & \\ & e^r \end{pmatrix}, \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \right) \in A$$

to show that for μ-most x and for any given K of large μ-measure, most $r \in [0,R]$ satisfy that

$$\mu_x^U(B_{\rho e^{2r}}^U) < \frac{1}{2}\mu_x^U(B_{e^{2r}}^U) \tag{10.12c}$$

(which is equivalent by Lemma 7.16 to $a^r.x$ satisfying (10.12b)). As it turns out the weaker (10.12c) is sufficient and one does not need (10.12a).

10.13**. The heart of the argument, choosing t.** Given the two points x_k, y_k and with them the parameter S_k we would need the regularity (10.12c) for $r = \frac{1}{2}\log S_k$ in order to apply the arguments from above. This may or may not happen but we can increase our chance of succeeding by looking not only at the given points x_k, y_k but also at all the points $a^t x_k, a^t y_k$ for some $t \in [0, T_k]$ for the appropriate choice of T_k (which in this case turns out to be $T_k = \frac{1}{2}\log S_k$).

This is the technical heat of the argument, and we sketch the proof below. To simplify matters, we assume that either $|d-a| \gg \sqrt{|c|}$ or $|d-a| \ll \sqrt{|c|}$ (with $\gg$ used here in a somewhat loose sense that we refrain from making more precise in this sketch).

We will only chose values of t for which the new points $a^t x_k, a^t y_k$ have good properties with respect to μ (i.e., belong to a previously defined set of points with good properties etc.), which in view of the (standard) maximal inequality holds for most $t \in [0, T_k]$ if the original points x_k, y_k were chosen from a suitable set of large measure.

Suppose first $|d-a| \gg \sqrt{|c|}$. In this case, the parameters a, d and with it S_k are unchanged when x_k and y_k are replaced by $a^t.x_k$ and $a^t.y_k$. Therefore, the regularity property (10.12c) is needed for the point $a^t.x_k$ and scale $r = \frac{1}{2}\log S_k$, which is equivalent to (10.12c) holding at the original point x for $r' = t + \frac{1}{2}\log S_k$.

At this stage we still have the freedom to chose t almost arbitrarily in the range $0 \leq t \leq \frac{1}{2}\log S_k$. As (10.12c) can be assumed to hold at x for most $r' \in [0, \log S_k]$ we can indeed chose t so that at $a^t.x_k$ (10.12c) holds for precisely the value of r we need.

In the second case $|d-a| \ll \sqrt{|c|}$, the important parameter $\sqrt{|c|}$ and with it S_k do change when x_k and y_k are replaced as above. The danger here is that if the parameter S_k changes in a particular way, it may be that one is still interested in the regularity property (10.12c) for x and the very same $r = \frac{1}{2}\log S_k$ even after introducing t. The reader may verify that this is **not** the case, after calculating the parameter $S_k(t)$ for the points $a^t.x_k$ and $a^t.y_k$ as a function of t one sees that $t + S_k(t)$ is affine with a linear component $\frac{1}{2}t$. As before a density argument gives that it is possible to find t as required. In the general case, the function one studies may switch between having linear part t and having linear part $\frac{1}{2}t$, i.e., may be only piecewise linear, but this does not alter the density argument for finding t. Moreover, one easily checks that $a^t.x_k$ and $a^t.y_k$ are still close together.

Having found t, one has the required regularity property to apply the density argument for $s \in [-S_k, S_k]$ and obtains $x'_k, y'_k \in K$ which differ mostly by some element of U of bounded but significant size. As mentioned before, taking the limit along some subsequence concludes the argument.

10.14**. Justification for $g_1 \notin U$.** Let us finish the outline of the proof of Theorem 10.3 by justifying the assertion in §10.8 that one can find $x, y = (g_1, g_2).x$ with $g_1 \notin U$ and the same U leaf-wise measure using the recurrence of the $\mathrm{SL}(2, \mathbb{Q}_p)$-action. By construction, $y = (e, h).x$ for some big $h \in \mathrm{SL}(2, \mathbb{Q}_p)$, and we have already verified that the U leafwise measure at x and y are the same using the product lemma. What remains is to explain why we can guarantee that $g_1 \notin U$.

By Poincaré recurrence we may assume that our initial point x satisfies that there is a sequence $t_n \to \infty$ with $a^{t_n}.x \to x$. If now $g_1 \in U$, then $a^{t_n} g_1 a^{-t_n} \to e$ and applying a^{t_n} to $(e, h).x = y = (g_1, g_2).x$ we would obtain $(e, h).x = (e, g_2).x$. As h is big, but g_2 is small, we obtain the nontrivial identity $x = (e, h^{-1}g_2).x$ which is impossible as the lattice Γ is irreducible.

11. Combining the high and low entropy methods.

11.1. Consider now the action of the diagonal group A on the space $X_n = \mathrm{SL}(n, \mathbb{Z})\backslash \mathrm{SL}(n, \mathbb{R})$. The method of proof of Theorem 10.3 can be adapted to study the A-invariant measures also in this case, but there are some extra twists; specifically we will need to combine in the low entropy method we have developed in the previous section with the high entropy method presented in §9. This has been carried out in the paper [**EKL06**] of the authors and A. Katok, and the results of this section are taken from that paper.

11.2. We recall the following conjecture regarding invariant measures on $X_n = \mathrm{SL}(n, \mathbb{Z})\backslash \mathrm{SL}(n, \mathbb{R})$, which is due to Margulis, Katok and Spatzier, and Furstenberg (cf. [**Mar00**]):

11.3. Conjecture. *Let A be the group of diagonal matrices in $\mathrm{SL}(n, \mathbb{R})$, $n \geq 3$. Then any A-invariant and ergodic probability measure μ on X_n is homogeneous.*

It is not hard to classify the possible homogeneous measures (see e.g. [**LW01**]). For n prime, the situation is particularly simple: any A-invariant homogeneous probability measure on X_n is either the natural measure on a periodic A-orbit, or the $\mathrm{SL}(n,\mathbb{R})$ invariant measure m on X_n.

11.4. In [**EKL06**] we give together with A.Katok the following partial result towards Conjecture 11.3:

11.5. Theorem ([**EKL06**, Theorem 1.3]). *Let A be the group of diagonal matrices as above and $n \geq 3$. Let μ be an A-invariant and ergodic probability measure on X_n. Then one of the following holds:*

(i) *μ is an A-invariant homogeneous measure which is not supported on a periodic A-orbit.*

(ii) *for every one-parameter subgroup $\{a_t\} < A$, $h_\mu(a_t) = 0$.*

By the classification of A-invariant homogeneous measures alluded to in §11.3, if (i) holds μ is not compactly supported.

11.6. For $1 \leq i \neq j \leq n$, let U_{ij} denote the one parameter unipotents subgroup of $\mathrm{SL}(n,\mathbb{R})$ which consists of all matrices that have 1 on the diagonal and 0 at all other entries except the (i,j) entry, and let μ be an A-invariant and ergodic probability measure on X_n which has positive entropy with respect to some $a_0 \in A$. By Theorem 7.6 and our assumption regarding positive entropy of μ it follows that the leaf-wise measure $\mu_x^{G^-_{a_0}}$ are nontrivial almost everywhere (this requires a bit of explanation, as μ is A-ergodic but not necessarily a_0-ergodic; however if one takes the ergodic decomposition of μ with respect to a_0 one gets from the ergodicity of μ under A that each ergodic component has the same entropy with respect to a_0 and one can apply Theorem 7.6 to each components separately). Using the product structure of $\mu_x^{G^-}$ given by Corollary 8.8 and the ergodicity under A it follows that there is some $i \neq j$ so that $\mu_x^{U_{ij}} =: \mu_x^{ij}$ is nontrivial almost everywhere. For notational simplicity suppose this happens for $(i,j) = (1,n)$.

11.7. One can now apply the argument described in §10.7– §10.14 to the group U_{1n} and an appropriate $a = \mathrm{diag}(\alpha_1, \dots, \alpha_n) \in A$ (we assume all the $\alpha_i > 0$) . One obvious requirement for a is that it contracts U_{1n}, i.e., that $\alpha_1 < \alpha_n$. It turns out though that in the proof (specifically, in §10.13) additional more subtle conditions on a need to be imposed that are nonetheless easy to satisfy: indeed in this case what one needs is simply that

$$\alpha_1 < \min_{1<i<n} \alpha_i \leq \max_{1<i<n} \alpha_i < \alpha_n.$$

For example, we can take $a = \mathrm{diag}(e^{-1}, 1, \dots, 1, e)$ which together with U_{1n} and U_{n1} form a subgroup of $\mathrm{SL}(n,\mathbb{R})$ isomorphic to $\mathrm{SL}(2,\mathbb{R})$.

11.8. We recall what was the outcome of the argument given in §10.7– §10.14 for $G = \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{Q}_p)$. The end result of that long argument was finding two distinct "μ-typical" points x, y with the same leafwise measures (i.e., $\mu_x^U = \mu_y^U$) with $y = u.x$ for some nontrivial $u \in U$.

An appropriate adaptation of this argument to the case at hand (i.e., $G = \mathrm{SL}(n,\mathbb{R})$) will yield at the end two μ-typical points x, y with the same U_{1n}-leafwise measures which differ by some element u obtained in a limiting procedure involving the shearing properties of U_{1n}. It turns out that in this case these limiting directions u may not belong to U_{1n} but rather one has $u \in C_G(U_{1n}) \cap G^-$; note that this group

$C_G(U_{1n}) \cap G^-$ is precisely the group generated by the 1-parameter unipotent groups U_{ij} with either $i = 1$ or $j = n$ (or both).

11.9**.** Playing around with leaf-wise measures, one can show that the measure μ must satisfy one of the following two possibilities:

(i) One can find a subset $X' \subset X_n$ of full measure such that every two points $x, y \in X'$ on the same $C_G(U_{1n}) \cap G^-$-orbit are in fact on the same U_{1n} orbit.

(ii) There are $(i, j) \neq (1, n)$ with $i = 1$ or $j = n$ so that μ_x^{ij} is nontrivial a.s. .

If (i) holds, then the points x, y obtained in §11.8 in fact differ along U_{1n} from which one can deduce, exactly as in the proof of Theorem 10.3, that μ is U_{1n} invariant where we are clearly at the endgame; e.g. one can apply Ratner's measure classification theorem, though it is better to first get some more information out of the proof, specifically invariance along U_{n1}. Ratner's measure classification theorem for semisimple groups (such as the group generated by U_{1n} and U_{n1} which is isomorphic to $\mathrm{SL}(2, \mathbb{R})$) is substantially simpler than the general case (for a simple proof see [**Ein06**]). Moreover, also the analysis of all possible cases is much simpler if one first establishes invariance under this bigger group.

If (ii) holds, by using the time-symmetry of entropy for the element $a = \mathrm{diag}(e^{-1}, 1, \dots, 1, e)$ we obtained that there are some (i', j') with $i' = 1$ or $j' = n$ or both so that $\mu_x^{j'i'}$ is nontrivial a.s. (note the switch in the order of the indices!). If $(i', j') \neq (1, n)$ we can apply Theorem 9.14 to obtain that μ is invariant under the group $[U_{j'i'}, U_{1n}]$ (which is either $U_{j'n}$ or $U_{1i'}$): again arriving at the endgame of the proof. If $(i', j') = (1, n)$ we obtain similarly that μ is invariant under $[U_{n1}, U_{ij}]$.

11.10**.** The above simplified discussion neglects to mention one crucial point. In Theorem 10.3, an important assumption was that Γ is irreducible, an assumption which only entered in order to show that there are nearby "typical" points x and y which differ in a shearable direction (i.e., not by an element in $U \times \mathrm{SL}(2, \mathbb{Q}_p)$) — c.f. §10.14.

The same issue arises also in the case of $\mathrm{SL}(n, \mathbb{R})$. For the particular lattice we are considering, namely $\mathrm{SL}(n, \mathbb{Z})$, one can show such nearby "shearable" pairs exist; but for a general lattice, even in $\mathrm{SL}(n, \mathbb{R})$, this problem can actually happen, and is precisely the source of an important class of counterexamples discovered by M. Rees to the most optimistic plausible measure classification conjecture for multidimensional diagonalizable groups [**Ree82**] (for a more accessible source, see [**EK03**, Section 9]; the same phenomena has been discovered independently in a somewhat different context by S. Mozes [**Moz95**]).

12. Application towards Littlewood's Conjecture

12.1**.** In this section we present an application of the measure classification results we have developed in the previous sections towards the following conjecture of Littlewood:

12.2. Conjecture (Littlewood (c. 1930))**.** *For every $\alpha, \beta \in \mathbb{R}$,*

$$\varliminf_{n \to \infty} n \|n\alpha\| \|n\beta\| = 0, \tag{12.2a}$$

where $\|w\| = \min_{n \in \mathbb{Z}} |w - n|$ is the distance of $w \in \mathbb{R}$ to the nearest integer.

12.3. The work we present here toward this conjecture was first presented in the paper [**EKL06**] which is joint paper of A. Katok and us. The presentation of this work is taken essentially verbatim from [**Lin07**, Sec. 6].

12.4. It turns out that Littlewood's conjecture would follow from the Conjecture 11.3. The reduction is not trivial and is essentially due to Cassels and Swinnerton-Dyer [**CSD55**], though there is no discussion of invariant measures in that paper[(37)]. A more recent discussion of the connection highlighting Cassels' and Swinnerton-Dyer's work can be found in [**Mar97**].

We need the following criterion for when α, β satisfy (12.2a):

12.5. Proposition. *(α, β) satisfy (12.2a) if and only if the orbit of*

$$x_{\alpha,\beta} = \mathrm{SL}(3,\mathbb{Z}) \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

under the semigroup

$$A^+ = \{a(s,t) : s,t \geq 0\} \qquad a(s,t) = \begin{pmatrix} e^{s+t} & 0 & 0 \\ 0 & e^{-s} & 0 \\ 0 & 0 & e^{-t} \end{pmatrix}$$

is unbounded[(38)] *in $X_3 = \mathrm{SL}(3,\mathbb{Z})\backslash \mathrm{SL}(3,\mathbb{R})$. Moreover, for any $\delta > 0$ there is a compact $C_\delta \subset X_3$, so that if $\underline{\lim}_{n\to\infty} n \|n\alpha\| \|n\beta\| \geq \delta$ then $A^+.x_{\alpha,\beta} \subset C_\delta$.*

12.6. Before we prove Proposition 12.5 we need to understand better what it means for a set $E \subset X_3$ to be bounded. We write $\pi_\Gamma : \mathrm{SL}(n,\mathbb{R}) \to X_n$ for the natural map that sends $g \in \mathrm{SL}(n,\mathbb{R})$ to $\mathrm{SL}(n,\mathbb{Z})g \in X_n$. We have the following important criterion (see e.g. [**Rag72**, Chapter 10]):

12.7. Proposition (Mahler's compactness criterion)**.** *Let $n \geq 2$. A set $E \subset X_n = \mathrm{SL}(n,\mathbb{Z})\backslash \mathrm{SL}(n,\mathbb{R})$ is bounded if and only if there is some $\epsilon > 0$ so that for any $x = \pi_\Gamma(g) \in E$ there is no vector v in the lattice spanned by the rows of g with $\|v\|_\infty < \epsilon$.*

12.8. **Proof of Proposition 12.5.** We prove only that $A^+.x_{\alpha,\beta}$ unbounded implies that (α, β) satisfies (12.2a); the remaining assertions of this proposition follow similarly and are left as an exercise to the reader.

Let $\epsilon \in (0, 1/2)$ be arbitrary and write

$$g_{\alpha,\beta} = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By Mahler's compactness criterion (see §12.7), if $A^+.x_{\alpha,\beta}$ is unbounded, there is a $a \in A^+$ such that in the lattice generated by the rows of $g_{\alpha,\beta}a^{-1}$ there is a nonzero vector v with $\|v\|_\infty < \epsilon$. This vector v is of the form

$$v = (ne^{-s-t}, (n\alpha - m)e^s, (n\beta - k)e^t)$$

(37) It is worthwhile to note that this remarkable paper appeared in 1955, many years before Conjecture 11.3 was made, and even before 1967 when Furstenberg made his related discoveries about scarcity of invariant sets and measures for the maps $x \mapsto 2x \bmod 1$ and $x \mapsto 3x \bmod 1$ on $\mathbb{R}/\mathbb{Z}$! The same paper also implicitly discusses the connection between Oppenheim's conjecture and the action of $\mathrm{SO}(2,1)$ on X_3.

(38) I.e. $\overline{A^+.x_{\alpha,\beta}}$ is not compact.

where n, m, k are integers at least one of which is nonzero, and $s, t \geq 0$. Since $\|v\|_\infty < 1/2$, $n \neq 0$ and $\|n\alpha\| = (n\alpha - m)$, $\|n\beta\| = (n\beta - k)$. Without loss of generality $n > 0$ and

$$n\,\|n\alpha\|\,\|n\beta\| \leq \|v\|_\infty^3 < \epsilon^3.$$

As ϵ was arbitrary, (12.2a) follows. □

12.9. We now turn to answering the following question: With the partial information given in Theorem 11.5, what information, if any, do we get regarding Littlewood's conjecture?

12.10. Theorem ([**EKL06**, Theorem 1.5])**.** *For any $\delta > 0$, the set*

$$\Xi_\delta = \left\{(\alpha, \beta) \in [0,1]^2 : \varliminf_{n\to\infty} n\,\|n\alpha\|\,\|n\beta\| \geq \delta\right\}$$

has zero upper box dimension[39][40].

12.11. We present a variant of the proof of this theorem given in [**EKL06**]. The first step of the proof, which is where Theorem 11.5 is used, is an explicit sufficient criterion for a single point α, β to satisfy Littlewood's conjecture (§12.2). This criterion is based on the notion of topological entropy; see §3.18 for the definition and basic properties of this entropy.

Let $a_{\sigma,\tau}(t) = a(\sigma t, \tau t)$, with $a(s,t)$ as in §12.5.

12.12. Proposition. *Suppose that $(\alpha, \beta) \in \mathbb{R}^2$ does not satisfy (12.2a) or equivalently that $A^+.x_0$ is bounded. Then for any $\sigma, \tau \geq 0$, the topological entropy of $a_{\sigma,\tau}$ acting on the compact set*

$$\overline{\{a_{\sigma,\tau}(t).x_{\alpha,\beta} : t \in \mathbb{R}^+\}}$$

vanishes.

12.13. **Proof.** Let x_0 be as in the proposition such that $A^+.x_0$ is bounded. If the topological entropy were positive, then by the variational principal in §3.21, there is an $a_{\sigma,\tau}$-invariant measure μ supported on $\overline{\{a_{\sigma,\tau}(t).x_0 : t \in \mathbb{R}^+\}}$ with $h_\mu(a_{\sigma,\tau}) > 0$.

Define for any $S > 0$

$$\mu_S = \frac{1}{S^2} \iint_0^S a(s,t)_*\mu \, ds\, dt,$$

with $a(s,t)_*\mu$ denoting the push forward of μ under the map $x \mapsto a(s,t).x$. Since $a(s,t)$ commutes with the one parameter subgroup $a_{\sigma,\tau}$, for any $a_{\sigma,\tau}$-invariant measure μ' the entropies satisfy

$$h_{\mu'}(a_{\sigma,\tau}) = h_{a(s,t)_*\mu'}(a_{\sigma,\tau}).$$

If μ has the ergodic decomposition $\int \mu_\xi \, d\nu(\xi)$, the measure μ_S has ergodic decomposition $S^{-2} \iint_0^S \int a(s,t)_*\mu_\xi \, d\nu(\xi)\, ds\, dt$ and so by §3.5, for every S

$$h_{\mu_S}(a_{\sigma,\tau}) = h_\mu(a_{\sigma,\tau}).$$

[39] I.e., for every $\epsilon > 0$, for every $0 < r < 1$, one can cover Ξ_δ by $O_{\delta,\epsilon}(r^{-\epsilon})$ boxes of size $r \times r$.

[40] Since (12.2a) depends only on α, β mod 1 it is sufficient to consider only $(\alpha, \beta) \in [0,1]^2$.

All μ_S are supported on the compact set $\overline{A^+.x_0}$, and therefore there is a subsequence converging weak* to some compactly supported probability measure μ_∞, which will be invariant under the full group A. By semicontinuity of entropy (§3.15),

$$h_{\mu_\infty}(a_{\sigma,\tau}) \geq h_\mu(a_{\sigma,\tau}) > 0,$$

hence by Theorem 11.5 the measure μ_∞ is not compactly supported[(41)] — a contradiction. □

12.14. Fix $\sigma, \tau \geq 0$. For $\alpha, \beta \in \mathbb{R}$ we define $X_{\alpha,\beta} = \overline{\{a_{\sigma,\tau}(t).x : t \in \mathbb{R}^+\}}$. Proposition 12.12 naturally leads us to the question of the size of the set of $(\alpha, \beta) \in [0,1]^2$ for which $h_{\mathrm{top}}(X_{\alpha,\beta}, a_{\sigma,\tau}) = 0$. This can be answered using the following general observation:

12.15. Proposition. *Let X' be a metric space equipped with a continuous $\mathbb{R}$-action $(t,x) \mapsto a_t.x$. Let X'_0 be a compact a_t-invariant[(42)] subset of X' such that for any $x \in X'_0$,*

$$h_{\mathrm{top}}(Y_x, a_t) = 0 \qquad Y_x = \overline{\{a_t.x : t \in \mathbb{R}^+\}}.$$

Then $h_{\mathrm{top}}(X'_0, a_t) = 0$.

12.16. **Proof.** Assume for contradiction that $h_{\mathrm{top}}(X'_0, a_t) > 0$. By the variational principle (§3.21), there is some a_t-invariant and ergodic measure μ on X'_0 with $h_\mu(a_t) > 0$.

By the pointwise ergodic theorem, for μ-almost every $x \in X'_0$ the measure μ is supported on Y_x. Applying the variational principle again (this time in the opposite direction) we get that

$$0 = h_{\mathrm{top}}(Y_x, a_t) \geq h_\mu(a_t) > 0$$

a contradiction. □

12.17. Corollary. *Consider, for any compact $C \subset X_3$ the set*

$$X_C = \left\{x \in X_3 : A^+.x \subset C\right\}.$$

Then for any $\sigma, \tau \geq 0$, it holds that $h_{\mathrm{top}}(X_C, a_{\sigma,\tau}) = 0$.

12.18. **Proof.** By Proposition 12.12, for any $x \in X_C$ the topological entropy of $a_{\sigma,\tau}$ acting on $\overline{\{a_{\sigma,\tau}(t).x : t \in \mathbb{R}^+\}}$ is zero. The corollary now follows from Proposition 12.15. □

12.19. We are now in position to prove Theorem 12.10, or more precisely to deduce the theorem from Theorem 11.5:

---

(41) Notice that a priori there is no reason to believe μ_∞ will be A-ergodic, while Theorem 11.5 deals with A-ergodic measures. So an implicit exercise to the reader is to understand why we can still deduce from $h_{\mu_\infty}(a_{\sigma,\tau}) > 0$ that μ_∞ is not compactly supported.

(42) Technical point: we only use that $a_t.X' \subset X'$ for $t \geq 0$. The variational principle (§3.21) is still applicable in this case.

12.20. **Proof of Theorem 12.10.** To show that Ξ_δ has upper box dimension zero, we need to show, for any $\epsilon > 0$ and for any $r \in (0,1)$ that the set Ξ_δ can be covered by $O_\epsilon(r^{-\epsilon})$ boxes of side r, or equivalently that any r-seperated set (i.e., any set S such that for any $x, y \in S$ we have $\|x-y\|_\infty > r$) is of size $O_{\delta,\epsilon}(r^{-\epsilon})$.

Let C_δ be as in Proposition 12.5. Let d denote a left invariant Riemannian metric on $G = \mathrm{SL}(3,\mathbb{R})$. Then d induces a metric, also denote by d on X_3. For $a, b \in \mathbb{R}$ let

$$g_{a,b} = \begin{pmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since C_δ is compact and d is induced from a left invariant Riemannian metric (so that there is an injectivity radius on C_δ), there will be r_0, c_0 such that for any $x \in C_\delta$ and $|a|, |b| < r_0$

$$d(x, g_{a,b}.x) \geq c_0 \max(|a|, |b|).$$

For any $\alpha, \alpha', \beta, \beta' \in \mathbb{R}$ we have that

$$x_{\alpha,\beta} = g_{\alpha'-\alpha,\beta'-\beta}.x_{\alpha',\beta'}$$

and more generally for any n

$$a_{1,1}^n.x_{\alpha,\beta} = g_{e^{3n}(\alpha'-\alpha),e^{3n}(\beta'-\beta)}.a_{1,1}^n.x_{\alpha',\beta'}.$$

It follows that if $S \subset \Xi_\delta$ is r-separated for $r = e^{-3n}r_0 \in (0, r_0)$ then

$$S' = \{x_{\alpha,\beta} : (\alpha,\beta) \in S\}$$

is (n, c_0r_0)-separated for $a_{1,1}$ in the sense of §3.18. By definition of C_δ and Ξ_δ, we have that (in the notations of §12.17) the set $S' \subset X_{C_\delta}$, a set which has zero topological entropy with respect to the group $a_{1,1}$. It follows that the cardinality of a maximal (n, c_0r_0)-separated set in S' is at most $O_{\delta,\epsilon}(\exp(\epsilon n))$; hence for $r < r_0$ the cardinality of a maximal r-separated subset of Ξ_δ is $O_{\delta,\epsilon}(r^{-\epsilon})$. $\square$

13. Application to Arithmetic Quantum Unique Ergodicity

13.1. We begin by recalling some basic facts about harmonic analysis on $\Gamma\backslash\mathbb{H}$. Here, $\mathbb{H} := \{x + iy : y > 0\}$ is the upper-half plane model of the hyperbolic plane (for more details, see [**Lan75**]). It is isomorphic to $G/K = \mathrm{SL}(2,\mathbb{R})/SO(2,\mathbb{R})$, and carries the Riemannian G-invariant metric $ds^2 = \frac{dx^2+dy^2}{y^2}$, where the action of G is given by fractional linear transformations in the usual way. This metric gives us the invariant area form $d\,\mathrm{area} = \frac{dxdy}{y}$.

We have the Laplacian operator

$$\Delta = y^2[\partial_x^2 + \partial_y^2]$$

which is also invariant under G. We wish to study $L^2(\Gamma\backslash\mathbb{H}, \text{area}) \cong L^2(\Gamma\backslash G, \mu_{\text{Haar}})_K$ for Γ a lattice in $G = \mathrm{SL}(2,\mathbb{R})$. Here $L^2(\Gamma\backslash G, \mu_{\text{Haar}})_K$ denotes the space of K-invariant L^2-functions on $\Gamma\backslash G$.

13.2. For any $\phi \in C_c(G)$, we can write a convolution operator

$$\phi * f := \int_G \phi(g) f(xg^{-1}) dg$$

but we will restrict ourselves to K-bi-invariant functions $\phi \in C_c(K\backslash G/K)$. In this case, we have the nice property that

$$\phi * \psi * f = (\phi * \psi) * f = (\psi * \phi) * f = \psi * \phi * f$$

and so these operators form a large commutative algebra which is called the **Hecke ring**. Since the Laplacian can be written as a limit of such convolution operators, it commutes with this algebra as well.

13.3. We begin with Γ cocompact (harmonic analysis is much easier in this case). The mapping $A_\phi : f \mapsto \phi * f$ is a compact, normal operator (this is false if Γ is not cocompact!). Therefore $L^2(M)$ is spanned by an orthonormal set of eigenfunctions of A_ϕ. Since Δ and A_ϕ commute, we can find an orthonormal basis of joint eigenfunctions.

In fact, if f is an eigenfunction of Δ, say $\Delta f = \lambda f$, it will automatically be an eigenfunction of the convolution operator A_ϕ, and the corresponding eigenvalue is by definition equal to the spherical transform of ϕ.

Weyl's Law (true for general compact surfaces) gives the asymptotic number of eigenvalues of Δ:

$$\#\{\text{eigenvalues of } \Delta \leq T\} \sim \pi \cdot \text{area}(M) \cdot T$$

13.4. In many cases we will be interested in Γ NOT cocompact; eg. $\Gamma_1 = \mathrm{SL}(2, \mathbb{Z})$ or one of the "principle congruence subgroups" $\Gamma_N = \{\gamma \equiv I \mod N\}$. For simplicity, let us assume for the moment that M has one cusp (at ∞).

We have some explicit eigenfunctions of Δ on $\mathbb{H}$. For example,

$$\Delta(y^s) = s(s-1)y^s = -(1/4 + t^2)y^s$$

where we make the convenient substitution $s = 1/2 + it$. These eigenfunctions correspond to planar waves going up; note that they are not Γ-invariant. Closely related are the Eisenstein series, which are Γ-invariant eigenfunctions of the Laplacian, not in $L^2(\Gamma\backslash\mathbb{H})$, satisfying

$$E_{1/2+it}(z) = y^{1/2+it} + \theta(1/2 + it) y^{1/2-it} + \text{ (rapidly decaying terms)}$$

We have the following spaces [**CS80**, Ch. 6-7]:

- $L^2_{\text{Eisenstein}}$ spanned by the Eisenstein series, and constituting the continuous part of the spectrum.
- $L^2_{\text{constants}}$ of constant functions.
- L^2_{cusp} of **cusp forms**, the orthogonal complement of the others. This consists of the functions f on $\Gamma\backslash\mathbb{H}$ whose integral along all periodic horocycles vanishes, i.e., (identifying the functions on $\Gamma\backslash\mathbb{H}$ with Γ-invariant functions on $\mathbb{H}$) functions f so that $\int_0^1 f(x + iy) dx = 0$ for all $y > 0$.

13.5. Selberg [**Sel56**] proved that if Γ is a **congruence subgroup**, i.e., $\Gamma_N < \Gamma < \Gamma_1$ for some N, then Weyl's Law holds for *cusp* forms

$$\#\{\text{eigenvalues of cusp forms } \leq T\} \sim \pi \cdot \text{area}(M) \cdot T$$

This is very far from the generic picture, where Phillips and Sarnak have conjectured that L^2_{cusp} is *finite dimensional* for generic Γ. While this remains to present an open question, significant results in this direction have been obtained by them [**PS92**] and Wolpert [**Wol94**].

Why is the case of congruence lattices so special? They carry a lot of extra symmetry, which makes it a lot easier for cusp forms to arise. We will now discuss these symmetries.

13.6. **Hecke Correspodence.**

We will define the Hecke correspondence (for a given prime p) in several equivalent ways. For simplicity we work with $\Gamma = \Gamma_1$.

13.7. First, associate to $z \in \mathbb{H}$ the $p+1$ points in $\Gamma\backslash\mathbb{H}$

$$T_p(z) := \Gamma\backslash \left\{ pz, \frac{z}{p}, \frac{z+1}{p}, \dots, \frac{z+p-1}{p} \right\}$$

each of these points is a fractional linear image of z, so each branch of this mapping is an isometry. One needs to check that this passes to the quotient by Γ; i.e., that if $z = \gamma z'$ then $T_p(z) = T_p(z')$. Since Γ is generated by the maps $z \mapsto z+1$ and $z \mapsto -\frac{1}{z}$, and $T_p(z)$ is obviously invariant under the former, it remains to check that $T_p(z) = T_p(-\frac{1}{z})$, which is left to the reader.

13.8. We will now give an alternate way to define T_p. It will be more convenient for us to work in $\mathrm{PGL}(2,\mathbb{R})$ instead of $\mathrm{SL}(2,\mathbb{R})$, and so we take $\Gamma = \mathrm{PGL}(2,\mathbb{Z})$ [this is not quite $\mathrm{SL}(2,\mathbb{Z})$ because matrices with determinant -1 are allowed on the one hand, but on the other hand $\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ and $\begin{pmatrix} -1 & \\ & -1 \end{pmatrix}$ which were distinct elements of $\mathrm{SL}(2,\mathbb{R})$ are identified in $\mathrm{PGL}(2,\mathbb{Z})$, but for our purposes this difference is very minor]. The matrix $\gamma_p = \begin{pmatrix} p & \\ & 1 \end{pmatrix} \in \mathit{comm}(\Gamma)$, where $\mathit{comm}(\Gamma)$ denotes the **commensurator** of Γ — the set of $\gamma \in G$ such that $[\Gamma : \gamma\Gamma\gamma^{-1} \cap \Gamma] < \infty$.

Note that

$$\Gamma\gamma_p\Gamma = \Gamma \begin{pmatrix} p & \\ & 1 \end{pmatrix} \sqcup \bigsqcup_{i=0}^{p-1} \Gamma \begin{pmatrix} 1 & i \\ 0 & p \end{pmatrix}$$

The mapping $T_p : \Gamma g \mapsto \Gamma\gamma_p\Gamma g$ gives the same correspondence as above.

Because we defined this correspondence by left multiplication, we can still quotient by K on the right, to get a correspondence on $\Gamma\backslash\mathbb{H}$.

13.9. Here is a third way to define the same correspondence. Since we can identify $X_2 = \mathrm{PGL}(2,\mathbb{Z})\backslash \mathrm{PGL}(2,\mathbb{R})$ with the space of lattices in $\mathbb{R}^2$ (up to homothety), we can define for $x \in X_2$ the set $T_p(x)$ to be the set of all lattices $y \in X_2$ homothetic to a sublattice of x of index p; or equivalently as the set of all $y \in X_2$ which contain a lattice homothetic to x as a sublattice of index p.

One should check that this agrees with the previous definitions (in particular, that $T_p(x)$ consists of $p+1$ points, which is not obvious from this defintion).

13.10. Lastly, we consider $\mathrm{PGL}(2,\mathbb{Z}[\frac{1}{p}])\backslash\mathrm{PGL}(2,\mathbb{R})\times\mathrm{PGL}(2,\mathbb{Q}_p)$, the space of $\mathbb{Z}[\frac{1}{p}]$-modules that are lattices in $\mathbb{R}^2\times\mathbb{Q}_p^2$, again up to homothety. By this we mean that an element of this space looks like $\mathbb{Z}[\frac{1}{p}](v_1,w_1)\oplus\mathbb{Z}[\frac{1}{p}](v_2,w_2)$ where $\{v_1,v_2\}$ is an $\mathbb{R}$-basis for $\mathbb{R}^2$, and (w_1,w_2) is a $\mathbb{Q}_p$-basis for $\mathbb{Q}_p^2$; and the two points $\mathbb{Z}[\frac{1}{p}](v_1,w_1)\oplus\mathbb{Z}[\frac{1}{p}](v_2,w_2)$ and $\mathbb{Z}[\frac{1}{p}](\lambda v_1,\theta w_1)\oplus\mathbb{Z}[\frac{1}{p}](\lambda v_2,\theta w_2)$ are identified for any $\lambda\in\mathbb{R}$ and $\theta\in\mathbb{Q}_p$.

Let $\pi:\mathbb{R}^2\times\mathbb{Q}_p^2\to\mathbb{R}^2$ be the natural projection, and consider the map $\pi_1: x\mapsto\pi(x\cap\mathbb{R}^2\times\mathbb{Z}_p^2)$ for x a lattice as above. Then $\pi_1(x)$ is a lattice in $\mathbb{R}^2$ and it respects equivalence up to homothety. Moreover, for every lattice $y\in X_2$, the inverse image $\pi_1^{-1}(y)=\mathrm{PGL}(2,\mathbb{Z}_p).x$ for some x. We've shown that

$$\mathrm{PGL}(2,\mathbb{Z}[\frac{1}{p}])\backslash\mathrm{PGL}(2,\mathbb{R})\times\mathrm{PGL}(2,\mathbb{Q}_p)/\mathrm{PGL}(2,\mathbb{Z}_p)\cong X_2$$

Using (a p-adic version of) the KAK-decomposition, we can write any $g_p\in\mathrm{PGL}(2,\mathbb{Q}_p)$ as $k_1\begin{pmatrix}p^n & \\ & 1\end{pmatrix}k_2$ for some $k_1,k_2\in K$ and some integer n. Then the map $x\mapsto\pi_1(g_p\cdot\pi_1^{-1}(x))$ yields a finite collection of points: x if $n=0$, the set $T_p(x)$ if $n=1$, and a finite set which we will denote by $T_{p^k}(x)$ if $n=k>1$. This gives our fourth equivalent definition of the Hecke correspondence.

13.11. The Hecke correspondence allows us to define an operator, also denoted by T_p, on $L^2(\Gamma\backslash G)$ (resp. on $L^2(\Gamma\backslash\mathbb{H})$) by

$$T_pf(x)=\frac{1}{\sqrt{p}}\sum_{y\in T_p(x)}f(y)$$

As a side remark, we note that for the Eisenstein series, the eigenvalues of T_p can be computed explicitly, and we have

$$\begin{aligned}T_pE_{1/2+it}&=\cos(t\log p)E_{1/2+it}\\&=(p^{\sqrt{\Delta+1/4}}+p^{-\sqrt{\Delta+1/4}})E_{1/2+it}\end{aligned}$$

The operator $(p^{\sqrt{\Delta+1/4}}+p^{-\sqrt{\Delta+1/4}})$ is essentially the propagating operator of the wave equation.

This property equating two operators which are defined by completely different means (eg., one by global symmetries and one by local differential structure) should be quite rare. This is one indication that $L^2_{\mathrm{Eisenstein}}$ should be very small in the arithmetic situations, and hence L^2_{cusp} should contain the vast majority of the eigenfunctions. This idea can be used to give an alternative elementary proof of the existence of cusp forms [**LV07**].

One should also note that there are compact surfaces $\Gamma\backslash\mathbb{H}$ with Hecke symmetries; one way to construct such lattices Γ is via quaternion algebras (see e.g. [**Mor**, Ch. 7]), for example

$$\Gamma=\left\{\begin{pmatrix}x+\sqrt{2}y & z+\sqrt{2}w\\ 5(z-\sqrt{2}w) & x-\sqrt{2}y\end{pmatrix}:x,y,z,w\in\mathbb{Z},\right.\qquad\left.x^2-2y^2-5t^2+10w^2=1\right\}.$$

13.12**.** We now discuss the quantum unique ergodicity conjecture, in particular in the arithmetic case. We begin with a general compact Riemannian manifold M, on which we have the Laplacian Δ_M, and we wish to understand the distribution properties of eigenfunctions of Δ_M.

According to Schroedinger, the motion of a free (spinless, non-relativistic) quantum particle flowing in the absence of external forces on M is given by the equation

$$i\frac{\partial\psi}{\partial t} = \Delta_M\psi$$

This defines a unitary evolution, i.e., the norm $||\psi(\cdot,t)||_{L^2}$ is independent of t. We will always take $||\psi||_{L^2} = 1$.

The Born interpretation of the "wave function" ψ is that the function $|\psi|^2 d(\mathrm{vol})$ defines a probability measure on M, representing the average position of a particle in the state ψ; i.e., for any (measurable) region $A \subset M$, the probability of finding our particle in A at time t is given by $\int_A |\psi(x,t)|^2 d\,\mathrm{vol}(x)$, where $d\,\mathrm{vol}$ is the Riemannian volume on M. Note that if ψ is an eigenfunction of Δ_M, then the time dependence of ψ only appears as a phase; i.e., $\psi(x,t) = e^{-i\lambda t}\psi(x,0)$. Hence eigenfunctions give rise to steady states, or invariant quantum distributions, $d\tilde{\mu}_\psi = |\psi|^2 d\,\mathrm{vol}$.

Let $\pi : S^*M \to M$ be the canonical projection. One can (see below) lift these $\tilde{\mu}_\psi$ to measures μ_ψ on the unit cotangent bundle S^*M which satisfy:

(i) $\left|\int \tilde{f} d\pi_*\mu_\psi - \int \tilde{f} d\tilde{\mu}_\psi\right| < \lambda^{-0.1}$ for any $\tilde{f} \in C^\infty(M)$

(ii) $\left|\int Hf d\mu_\psi\right| < \lambda^{-0.1}$ for any $f \in C^\infty(S^*M)$, where H is differentiation along the geodesic flow.

Suppose now that $\{\psi_i\}$ is a sequence of (normalized) eigenfunctions whose eigenvalues $\lambda_i \to \infty$, denote by $\tilde{\mu}_i = \tilde{\mu}_{\psi_i}$ the corresponding measures, and let μ_i be the corresponding lifts. The above conditions guarantee that any weak* limit point μ_∞ of the μ_i will satisfy

- $\pi_*\mu_\infty = \tilde{\mu}_\infty$ (the weak* limit of the corresponding $\tilde{\mu}_i$).
- $\int Hf d\mu_\infty = 0$, i.e., $\frac{\partial}{\partial t}\int f(g_t.x)d\mu_\infty = 0$. This means that μ_∞ is g_t invariant.

We call the μ's "microlocal lifts" of the $\tilde{\mu}$'s.

13.13. Definition. *Any weak* limit μ_∞ of $\{\mu_i\}$ as above is called a* **quantum limit**.

13.14**.** Here we will be interested in the special case of $M = \Gamma\backslash\mathbb{H}$ for Γ an arithmetic lattice; e.g. Γ a congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$, or one of the arithmetic compact quotients mentioned earlier. These manifolds carry the extra symmetry of the Hecke operators, and since all of these operators commute, we can find a basis of L^2 (or L^2_{cusp} in the non-compact case) consisting of joint eigenfunctions of Δ_M and all of the T_p, such joint eigenfunctions are called **Hecke-Maass forms**. Any weak* limit of μ_{ψ_i}, where all of the ψ_i are Hecke-Maass forms, is called an **arithmetic quantum limit**.

13.15**.** For now, we assume that M is compact. Snirlman, Colin de Verdiere, and Zelditch have shown that if $\{\psi_i\}_{i=1}^\infty$ is a full set of (normalized) eigenfunctions ordered by eigenvalue, then the average $\frac{1}{N}\sum_{i=1}^N \mu_{\psi_i}$ converges to the Liouville measure on S^*M. If we assume that the geodesic flow on M is ergodic with respect to Liouville measure (satisfied e.g. if M has negative sectional curvature), then outside a set E of indices of density zero (i.e., $\lim_{N\to\infty} \frac{1}{N}\#\{i \in E : i < N\} = 0$), the sequence $\{\mu_i\}_{i\notin E}$ converges to Liouville measure; this is because an ergodic measure cannot be written as a proper convex combination of other invariant measures.

13.16. Conjecture (Rudnick-Sarnak [**RS94**]). *Let M be a compact, Riemannian manifold of negative sectional curvature. Then the Liouville measure on S^*M is the unique quantum limit.*

13.17. Theorem ([**Lin06**, Theorem 1.4]). *Say $M = \Gamma\backslash\mathbb{H}$, for Γ arithmetic (of finite covolume, but not necessarily co-compact). Then the only arithmetic quantum limits are scalar multiples of Liouville measure (i.e., the measure is a Haar measure).*

13.18. Corollary. *Let $f \in C_c(M)$ be such that $\int_M f = 0$. Then for a sequence $\{\psi_i\}$ of Hecke-Maass forms, we have*

$$\int_M f|\psi_i|^2 d\,\mathrm{area}(x) \to 0$$

as $i \to \infty$.

Note that in Theorem 13.17 we do not know that the limit measure is a probability measure. If Γ is cocompact, then this is immediate; but in the case of Γ a congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$, there remains the possibility that some (or possibly all) of the mass escapes to the cusp in the limit. We note that since the summer school this problem has been solved: Soundararajan [**Sou09**] proved, by purely number theoretic methods, that the arithmetic quantum limits are probability measures. Together this proves the arithmetic quantum unique ergodicity conjecture.

13.19**.** For example, we could take f in Corollary 13.18 to also be a Hecke-Maass form (recall these are orthogonal to constants, so the hypothesis is satisfied). In fact, since these span L^2_{cusp}, the statement of Corollary 13.18 will hold for all such f if and only if it holds for all Hecke-Maass forms.

An identity of Watson shows that the quantity $\int \psi_1\psi_2\psi_3 d\,\mathrm{area}$ can be expressed in terms of L-functions, specifically

$$\left|\int \psi_1\psi_2\psi_3 d\,\mathrm{area}\right|^2 = \frac{\pi^4\Lambda(\frac{1}{2}, \psi_1 \times \psi_2 \times \psi_3)}{\Lambda(1, \mathrm{sym}^2\psi_1)\Lambda(1, \mathrm{sym}^2\psi_2)\Lambda(1, \mathrm{sym}^2\psi_3)}.$$

Hence good estimates on the completed L-function $\Lambda(\frac{1}{2}, \psi_1 \times \psi_2 \times \psi_3)$ would imply Arithmetic QUE. Unfortunately, the best estimates we have for this L-functions gives only a trivial bound, and so Theorem 13.17 does not follow from existing technology in this direction. The Generalized Riemann Hypothesis (GRH) would imply a bound of $\lesssim \lambda_i^{-1/4}$, which would not only imply Theorem 13.17, but would give an optimal rate of convergence. Further discussion of many of these topics can be found in the survey [**Sar03**].

13.20**.** We also note that Theorem 13.17 has been extended to other $\Gamma\backslash G$ by Silberman and Venkatesh; e.g. for $G = \mathrm{SL}(p,\mathbb{R})$ and Γ a congruence lattice therein [**SV04, SV06**].

13.21. As a first step, we wish to construct the measures μ_i from the $\tilde{\mu}_i$, which satisfy the conditions of 13.12. The "standard" way to do this is via pseudodifferential calculus (see e.g. [**Ana**]), but we wish to give a representation-theoretic construction, which will respect the Hecke symmetries that we wish to exploit.

Given an eigenfunction $\phi \in L^2(\Gamma\backslash\mathbb{H}) = L^2(\Gamma\backslash \mathrm{SL}(2,\mathbb{R}))_K$, we can translate the function via $g.\phi(x) = \phi(xg^{-1})$ for any $g \in \mathrm{SL}(2,\mathbb{R})$. Taking all possible translates, we get a representation

$$V_\phi = \overline{\langle g.\phi : g \in G\rangle} = \overline{\langle \psi * \phi : \psi \in C_c(G)\rangle}$$

where the action of ψ is by convolution as in §13.2. This representation is unitary (since the Riemannian measure is G-invariant), and also irreducible (this is not quite obvious, but follows from the general theory). Moreover, the isomorphism class of this representation is completely determined by the eigenvalue of ϕ.

In fact, we can write down an explicit model $\tilde{V}_t \cong V_\phi$ for this representation (where t is determined by $\Delta_M \phi = (\frac{1}{4} + t^2)\phi)$. The Hilbert space on which the representation acts will be simply $L^2(K)$. To understand the action of $G = \mathrm{SL}(2,\mathbb{R})$, extend any function f on K to a function $\tilde{f}$ on G using the NAK decomposition of $\mathrm{SL}(2,\mathbb{R})$,

$$g = n\hat{a}k = \begin{pmatrix} 1 & s \\ & 1 \end{pmatrix} \begin{pmatrix} a & \\ & a^{-1} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

via

$$\tilde{f}(n\hat{a}k) = a^{1+2it} f(k)$$

now define $g.f$ as the restriction of the left translation of $\tilde{f}$ by g to K. It can be shown by an explicit calculation that this representation is unitary as long as $t \in \mathbb{R}$, i.e., as long as the eigenvalue of ϕ under the Laplacian is $\geq 1/4$ (for our purposes, this is all we care about).

For every $n \geq 0$ we may choose the vector $\Phi^{(n)} \in V_\phi$ that corresponds to the (normalized) Dirichlet kernel

$$f(k(\theta)) = \frac{1}{\sqrt{2n+1}} \frac{\sin{(n+\frac{1}{2})\theta}}{\sin\frac{\theta}{2}}$$

which is the n-th order approximation to the δ-function on K. We then set $\mu_i = |\Phi_i^{(n)}|^2 d\mathrm{vol}$ (here $d\,\mathrm{vol}$ is the Riemannian volume on $\Gamma\backslash G$ and n will be chosen as a function of i below), and we will see that the μ_i are close to being invariant under the geodesic flow.

What is the role of n in all of this? There are two competing properties:

- The larger the value of n, the closer f is to a δ function, and the better the invariance properties of μ_i. The problem is that then μ_i loses much of its relation to $\tilde{\mu}_i$ (i.e., $\pi_*\mu_i$ and $\tilde{\mu}_i$ become farther apart).
- The smaller the value of n, the closer f is to a constant function, which means that μ_i agrees well with $\tilde{\mu}_i$, but μ_i loses its invariance properties.

However, as $i \to \infty$ (and simultaneously also $t \to \infty$), both approximations improve; hence if we "split the difference" by choosing an appropriate value of n for each t, we will get both desired properties in the limit.

13.22. We now wish to explain why large n values make μ_i more invariant. We define the following differential operators:

$$\begin{aligned} Hf &= \frac{\partial}{\partial s} f\left(g\begin{pmatrix} e^s & \\ & e^{-s}\end{pmatrix}\right) \\ Vf &= \frac{\partial}{\partial s} f\left(g\begin{pmatrix} \cosh(s) & \sinh(s) \\ \sinh(s) & \cosh(s)\end{pmatrix}\right) \\ Wf &= \frac{\partial}{\partial \theta} f(gk(\theta)) \end{aligned}$$

H is differentiation along the geodesic flow, V is differentiation along the perpendicular direction to the geodesic flow, and W is differentiation in the rotational direction (i.e., fixing the point in M and letting the direction of the tangent vector vary).

We have the **Casimir** element

$$\omega = H^2 + V^2 - W^2$$

which is self-adjoint, commutes with translations by any $g \in G$, and coincides (up to a scalar) with Δ_M on the subspace $\{f : Wf = 0\} = L^2(\Gamma\backslash G)_K$. Every vector $\psi \in V_\phi$ is an eigenfunction of ω, with eigenvalue $1 + 4t^2$.

For t large, consider

$$\begin{aligned} \langle \omega\Phi, \Phi f\rangle &= \langle \Phi, \omega(\Phi f)\rangle \\ &= \langle \Phi, (\omega\Phi) f\rangle + \langle \Phi, \Phi(\omega f)\rangle + \\ &\qquad + 2\langle \Phi, H\Phi H f\rangle + 2\langle \Phi, V\Phi V f\rangle + 2\langle \Phi, W\Phi W f\rangle, \end{aligned}$$

which follows from ω being self-adjoint and from the product rule for differentiation. Note that the first terms in both lines are equal:

$$\langle \omega\Phi, \Phi f\rangle = \langle \Phi, (\omega\Phi) f\rangle = (1 + 4t^2)\langle \Phi, \Phi f\rangle,$$

and also that for fixed f, the quantity $\langle \Phi, \Phi(\omega f)\rangle = O_f(1)$ as the eigenvalue t tends to infinity. On the other hand, if n is large (but much smaller than t), Φ is close to being an eigenfunction of H of large eigenvalue ($\sim it$), and both $||V\Phi||$ and $||W\Phi||$ are much less than $t||\Phi||$. Dividing by the "eigenvalue for H" we must have $\langle \Phi, \Phi H f\rangle = o(1)$.

But what is $\langle \Phi, \Phi H f\rangle$? By definition, it is the integral $\int H f d\mu_i$ of the derivative of f along the geodesic flow. Since this tends to 0 as t gets large, we get

$$\frac{\partial}{\partial t}\int f(\cdot a_t) d\mu_\infty = 0$$

if μ_∞ is a weak* limit point of the μ_i; i.e., we have that μ_∞ is a_t-invariant.

13.23. We have shown that any weak* limit point of the measures $\tilde{\mu}_i$ is a projection of a measure μ_∞ on $\Gamma\backslash \mathrm{SL}(2,\mathbb{R})$ which is a_t-invariant. But as we know well, there are many a_t-invariant measures here, even with positive entropy!

Thus in order to classify quantum limits, we will have to use additional information about these limits. At this stage, we will abandon the properties coming from the ϕ_i being eigenfunctions of Δ (though we have not harnessed the full power of this assumption), and use properties of Hecke eigenfunctions.

The fact that the ϕ_i are Hecke eigenfunctions implies (since the Hecke operators are defined by translations) that Φ_i (indeed, any vector in V_{ϕ_i}) is a Hecke eigenfunction. Now, one certainly cannot expect $|\Phi_i|^2$ to be a Hecke eigenfunction, but traces of this symmetry do survive in the measures μ_i as well as their limit μ_∞.

13.24. Recall the Hecke correspondence (fourth formulation) given via the projection map

$$\mathrm{PGL}(2,\mathbb{Z}\left[\frac{1}{p}\right])\backslash\mathrm{PGL}(2,\mathbb{R})\times\mathrm{PGL}(2,\mathbb{Q}_p)\to\mathrm{PGL}(2,\mathbb{Z})\backslash\mathrm{PGL}(2,\mathbb{R}).$$

For each x we have a set of points $T_p(x)$, and its iterates, giving a Hecke tree which is the projection of a full $\mathrm{PGL}(2,\mathbb{Q}_p)$-orbit of x.

13.25. Definition. *A measure μ on $\mathrm{PGL}(2,\mathbb{Z})\backslash\mathrm{PGL}(2,\mathbb{R})$ is p-Hecke recurrent if there is a measure $\tilde{\mu}$ on*

$$\mathrm{PGL}(2,\mathbb{Z}\left[\frac{1}{p}\right])\backslash\mathrm{PGL}(2,\mathbb{R})\times\mathrm{PGL}(2,\mathbb{Q}_p)$$

such that $\pi_*\tilde{\mu}=\mu$ and $\tilde{\mu}$ is $\mathrm{PGL}(2,\mathbb{Q}_p)$-recurrent.

13.26. **Problem:** Show that the property of p-Hecke recurrence is independent of the lifting; i.e., μ is p-Hecke recurrent if and only if *any* lifting measure $\tilde{\mu}$ is $\mathrm{PGL}(2,\mathbb{Q}_p)$-recurrent.

13.27. Let $\mathcal{G}$ be an abstract $p+1$-regular tree, with a distinguished base point. For a more direct definition of p-Hecke recurrence, we can define leafwise measures $\mu_x^{\mathcal{G}}$ on these Hecke trees (our space is foliated into Hecke orbits), and then as before Hecke recurrence will hold whenever these leafwise measures are infinite a.e.

Note that unlike the case of group actions, there is no canonical labeling on the p-Hecke tree of a point $x\in X$ in terms of the nodes of $\mathcal{G}$. The only inherent structure on these p-Hecke trees is the (discrete) tree metric; and a construction of leafwise measures in such cases is given in [**Lin06**].

To avoid having to introduce this formalism we can consider instead the corresponding non-locally finite measure $\mu_{x,p}=\mu_x^{\mathcal{G}}.x$ on $\mathrm{PGL}(2,\mathbb{Z})\backslash\mathrm{PGL}(2,\mathbb{R})$.

13.28. These leafwise measures (more precisely, their image under the embedding of the abstract $p+1$-regular tree $\mathcal{G}$ to p-Hecke trees in $\mathrm{PGL}(2,\mathbb{Z})\backslash\mathrm{PGL}(2,\mathbb{R})$ satisfy a.s. that

$$\frac{\mu_{x,p}(y)}{\mu_{x,p}(x)}=\lim_{r\to 0}\frac{\mu(B_r(y))}{\mu(B_r(x))}$$

where $B_r(x)=x.B_r^G(1)$ is a small ball around x in the group G.

Now, since Φ_i are Hecke eigenfunctions, the restriction of Φ_i to each Hecke tree will give an eigenfunction of the tree Laplacian. Hecke recurrence will then follow (after a short argument that can be found in e.g. [**Lin06**, Sec. 8]) from

13.29. Lemma. *Let $\mathcal{G}$ be a $p+1$-regular tree, and $\phi:\mathcal{G}\to\mathbb{C}$ a function such that $\Delta_{\mathcal{G}}\phi=\lambda_p\phi$. Then $\phi\notin L^2(\mathcal{G})$; in fact, there exists a (universal) constant independent of λ_p, such that*

$$\sum_{d(x,y)\leq R}|\phi|^2\geq cR|\phi(x)|^2$$

13.30. This implies that our quantum limit will be both a_t-invariant and Hecke recurrent. By Theorem 10.3, if a.e. ergodic component of μ has positive entropy (this was shown for arithmetic quantum limits by Bourgain-Lindenstrauss [**BL03**]), then μ is G-invariant; i.e., μ is a multiple of Haar measure.

References

[Ana] Nalini Anantharaman. Eigenfunctions of the Laplacian on negatively curved manifolds: a semiclassical approach. in this volume.

[Bec83] María E. Becker. A ratio ergodic theorem for groups of measure-preserving transformations. *Illinois J. Math.*, 27(4):562–570, 1983.

[BL03] Jean Bourgain and Elon Lindenstrauss. Entropy of quantum limits. *Comm. Math. Phys.*, 233(1):153–171, 2003.

[Buz97] Jérôme Buzzi. Intrinsic ergodicity of smooth interval maps. *Israel J. Math.*, 100:125–161, 1997.

[CS80] Paul Cohen and Peter Sarnak. Notes on the Selberg trace formula. chapters 6 and 7 posted at `http://www.math.princeton.edu/sarnak/`, 1980.

[CSD55] J. W. S. Cassels and H. P. F. Swinnerton-Dyer. On the product of three homogeneous linear forms and the indefinite ternary quadratic forms. *Philos. Trans. Roy. Soc. London. Ser. A.*, 248:73–96, 1955.

[Ein06] Manfred Einsiedler. Ratner's theorem on SL(2,R)-invariant measures, 2006.

[EK03] Manfred Einsiedler and Anatole Katok. Invariant measures on G/Γ for split simple Lie groups G. *Comm. Pure Appl. Math.*, 56(8):1184–1221, 2003. Dedicated to the memory of Jürgen K. Moser.

[EK05] Manfred Einsiedler and Anatole Katok. Rigidity of measures—the high entropy case and non-commuting foliations. *Israel J. Math.*, 148:169–238, 2005. Probability in mathematics.

[EKL06] Manfred Einsiedler, Anatole Katok, and Elon Lindenstrauss. Invariant measures and the set of exceptions to Littlewood's conjecture. *Ann. of Math. (2)*, 164(2):513–560, 2006.

[EL06] Manfred Einsiedler and Elon Lindenstrauss. Diagonalizable flows on locally homogeneous spaces and number theory. In *International Congress of Mathematicians. Vol. II*, pages 1731–1759. Eur. Math. Soc., Zürich, 2006.

[EL08] Manfred Einsiedler and Elon Lindenstrauss. On measures invariant under diagonalizable actions the rank one case and the general low entropy method. *Journal of Modern Dynamics*, 2(1):83–128, 2008.

[ELMV06] Manfred Einsiedler, Elon Lindenstrauss, Philippe Michel, and Akshay Venkatesh. The distribution of periodic torus orbits on homogeneous spaces, 2006. submitted (42 pages).

[ELMV07] Manfred Einsiedler, Elon Lindenstrauss, Philippe Michel, and Akshay Venkatesh. Distribution of periodic torus orbits and Duke's theorem for cubic fields, 2007. submitted (58 pages).

[Esk] Alex Eskin. Unipotent flows and applications. in this volume.

[EW09a] Manfred Einsiedler and Thomas Ward. Ergodic theory: with a view towards number theory. available online at http://www.mth.uea.ac.uk/ergodic/., 2009.

[EW09b] Manfred Einsiedler and Thomas Ward. Some chapters on entropy. in preparation, available online at http://www.mth.uea.ac.uk/ergodic/, 2009.

[Fur73] Harry Furstenberg. The unique ergodicity of the horocycle flow. In *Recent advances in topological dynamics (Proc. Conf., Yale Univ., New Haven, Conn., 1972; in honor of Gustav Arnold Hedlund)*, pages 95–115. Lecture Notes in Math., Vol. 318. Springer, Berlin, 1973.

[Gla03] Eli Glasner. *Ergodic theory via joinings*, volume 101 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.

[Hu93] Hu Yi Hu. Some ergodic properties of commuting diffeomorphisms. *Ergodic Theory Dynam. Systems*, 13(1):73–100, 1993.

[Hur44] Witold Hurewicz. Ergodic theorem without invariant measure. *Ann. of Math. (2)*, 45:192–206, 1944.

[KH95] Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.

[KS96] A. Katok and R. J. Spatzier. Invariant measures for higher-rank hyperbolic abelian actions. *Ergodic Theory Dynam. Systems*, 16(4):751–778, 1996.

[Lan75] Serge Lang. $SL_2(\mathbf{R})$. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1975.

[Lin05] Elon Lindenstrauss. Rigidity of multiparameter actions. *Israel J. Math.*, 149:199–226, 2005.

[Lin06] Elon Lindenstrauss. Invariant measures and arithmetic quantum unique ergodicity. *Ann. of Math. (2)*, 163(1):165–219, 2006.

[Lin07] Elon Lindenstrauss. Some examples how to use measure classification in number theory. In *Equidistribution in number theory, an introduction*, volume 237 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 261–303. Springer, Dordrecht, 2007.

[LV07] Elon Lindenstrauss and Akshay Venkatesh. Existence and Weyl's law for spherical cusp forms. *Geom. Funct. Anal.*, 17(1):220–251, 2007.

[LW01] Elon Lindenstrauss and Barak Weiss. On sets invariant under the action of the diagonal group. *Ergodic Theory Dynam. Systems*, 21(5):1481–1500, 2001.

[Mar97] G. A. Margulis. Oppenheim conjecture. In *Fields Medallists' lectures*, volume 5 of *World Sci. Ser. 20th Century Math.*, pages 272–327. World Sci. Publishing, River Edge, NJ, 1997.

[Mar00] Gregory Margulis. Problems and conjectures in rigidity theory. In *Mathematics: frontiers and perspectives*, pages 161–174. Amer. Math. Soc., Providence, RI, 2000.

[Mor] Dave Witte Morris. Introduction to arithmetic groups. in preparation, `http://people.uleth.ca/%7Edave.morris/LectureNotes.shtml#ArithmeticGroups`.

[Moz95] Shahar Mozes. Actions of Cartan subgroups. *Israel J. Math.*, 90(1-3):253–294, 1995.

[MT94] G. A. Margulis and G. M. Tomanov. Invariant measures for actions of unipotent groups over local fields on homogeneous spaces. *Invent. Math.*, 116(1-3):347–392, 1994.

[OW87] Donald S. Ornstein and Benjamin Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.*, 48:1–141, 1987.

[PS92] R. Phillips and P. Sarnak. Automorphic spectrum and Fermi's golden rule. *J. Anal. Math.*, 59:179–187, 1992. Festschrift on the occasion of the 70th birthday of Shmuel Agmon.

[Rag72] M. S. Raghunathan. *Discrete subgroups of Lie groups*. Springer-Verlag, New York, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68.

[Rt82a] Marina Ratner. Factors of horocycle flows. *Ergodic Theory Dynam. Systems*, 2(3-4):465–489, 1982.

[Rt82b] Marina Ratner. Rigidity of horocycle flows. *Ann. of Math. (2)*, 115(3):597–614, 1982.

[Rt83] Marina Ratner. Horocycle flows, joinings and rigidity of products. *Ann. of Math. (2)*, 118(2):277–313, 1983.

[Rt90] Marina Ratner. On measure rigidity of unipotent subgroups of semisimple groups. *Acta Math.*, 165(3-4):229–309, 1990.

[Rt91] Marina Ratner. On Raghunathan's measure conjecture. *Ann. of Math. (2)*, 134(3):545–607, 1991.

[Ree82] M. Rees. Some R^2-anosov flows. unpublished, 1982.

[RS94] Zeév Rudnick and Peter Sarnak. The behaviour of eigenstates of arithmetic hyperbolic manifolds. *Comm. Math. Phys.*, 161(1):195–213, 1994.

[Sar03] Peter Sarnak. Spectra of hyperbolic surfaces. *Bull. Amer. Math. Soc. (N.S.)*, 40(4):441–478 (electronic), 2003.

[Sel56] A. Selberg. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Indian Math. Soc. (N.S.)*, 20:47–87, 1956.

[Sou09] K. Soundararajan. Quantum unique ergodicity for $SL_2(\mathbb{Z})\backslash\mathbb{H}$. preprint, `arXiv:0901.4060v1[math.NT]`, 2009.

[Sri98] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[SV04] Lior Silberman and Akshay Venkatesh. On quantum unique ergodicity for locally symmetric spaces I: a micro local lift. to appear in GAFA (37 pages), 2004.

[SV06] Lior Silberman and Akshay Venkatesh. Entropy bounds for hecke eigenfunctions on division algebras. preprint (22 pages), 2006.

[Wal82] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982.

[Wol94] Scott A. Wolpert. Disappearance of cusp forms in special families. *Ann. of Math. (2)*, 139(2):239–291, 1994.