JOININGS CLASSIFICATION AND APPLICATIONS
[after Einsiedler and Lindenstrauss]

by **Menny Aka**

INTRODUCTION

Fix a group A and consider a set of measure-preserving actions of A on $\mathrm{X_i} = (X_i, \mathcal{B}_i, \mu_i)$, $i = 1, \ldots, r$, where X_i is a Borel probability space with a measure μ_i and a σ-algebra $\mathcal{B}_i$. Consider the joint action (also called the diagonal action) of A on

$$\mathrm{X} = (X_1 \times \cdots \times X_r, \mathcal{B}_1 \times \cdots \times \mathcal{B}_r)$$

given by $a.(x_1, \ldots, x_r) = (a.x_1, \ldots, a.x_r)$. A ($r$-fold) *joining* of the systems $\{\mathrm{X_i}\}_{i=1}^r$ is an A-invariant probability measure μ on X with $(\pi_i)_* \mu = \mu_i$ for $i = 1, \ldots, r$ where $\pi_i \colon X \to X_i$ is the natural projection map.

There always exists at least one joining, namely the *trivial joining*, which is the product measure $\mu_1 \otimes \cdots \otimes \mu_r$. When this is the only possible joining of the systems $\{\mathrm{X_i}\}_{i=1}^r$ one says that these systems are *disjoint*. The systematic study of joinings stems from Furstenberg's seminal paper (FURSTENBERG, 1967). Furstenberg marked an analogy between joinings and the arithmetic of integers: saying that two measure-preserving systems are disjoint is analogous to saying that their least common multiple is their product. The analogy works in one direction; measure-preserving systems admitting a non-trivial common factor are never disjoint: recall that a *factor* of a measure-preserving system $\mathrm{X} = (X, \mathcal{B}, \mu, A)$ is a measure-preserving system $\mathrm{Y} = (Y, \mathcal{C}, \nu, A)$ and a measure-preserving map $\phi \colon X \to Y$ which intertwines the action of A, that is, for all $a \in A$ we have $a.\phi(x) = \phi(a.x)$ for μ-almost everywhere. Like integers, any measure-preserving system has itself and the trivial system (one-point system) as factors. Moreover, like integers, as stated above, if two measure-preserving systems have a common factor, they have a non-trivial joining, called the relatively independent joining over a common factor (see e.g., EINSIEDLER and WARD, 2011, §6.5). FURSTENBERG (1967) asked if this analogy also works in the other direction: if two systems do not have any common factor, must they be disjoint? RUDOLPH (1979) answered negatively, providing the first counterexamples. Joinings are nonetheless a strong tool in ergodic theory, as exemplified by GLASNER (2003) which gives a complete treatment of ergodic theory via joinings. The broad applicability of the classification of possible joinings of certain systems was

already visible in the work of FURSTENBERG (1967), where he solves a question in Diophantine approximation using joinings. We refer the reader also to the recent survey of DE LA RUE (2020) about the broad use of joinings in ergodic theory.

Roughly said, the study of joinings is the study of all possible ways two systems (or r systems) can be embedded as factors of another system, which is in turn spanned by them. When two systems are *not* disjoint, this is a sign that there is strong relation between them. The main topic of this survey is a very good example of this principle. This is a survey of the work of Einsiedler and Lindenstrauss on joinings of higher-rank torus actions on S-arithmetic homogeneous spaces (EINSIEDLER and LINDENSTRAUSS, 2019), which extends their previous paper (EINSIEDLER and LINDENSTRAUSS, 2007). They consider torus actions on two (or r) homogeneous spaces which are quotients of S-arithmetic points of perfect algebraic groups, equipped with the uniform Haar probability measure on each quotient. They show in particular that if such systems are not disjoint, there must be a strong *algebraic* relation between the corresponding perfect algebraic groups, exemplifying the principle stated above. This may remind the reader of the folklore Goursat's Lemma from group theory; while the latter is a natural structural theorem about subgroups of a product, the joining theorem of Einsiedler and Lindenstrauss is a striking instance of measure rigidity, where the existence of non-trivial joinings in this setting can only be due to a strong algebraic relation.

The main result of EINSIEDLER and LINDENSTRAUSS (2019, Theorem 1.7) classifies joinings on higher-rank torus actions on a product of two (or r) homogenous spaces of the form

$$\Gamma_1 \backslash \mathbb{G}_1(\mathbb{Q}_S) \times \Gamma_2 \backslash \mathbb{G}_2(\mathbb{Q}_S)$$

as we now state after recalling the necessary definitions. The measure spaces we consider are S-arithmetic homogeneous quotients of perfect groups. More precisely, let $\mathbb{G}$ be a perfect Zariski-connected linear algebraic group defined over $\mathbb{Q}$ and let S be a finite set of places of $\mathbb{Q}$. Let $\mathbb{Q}_S$ denote $\prod_{s \in S} \mathbb{Q}_s$ (with $\mathbb{Q}_\infty = \mathbb{R}$). An S-arithmetic quotient is a quotient space of the form $\Gamma \backslash G$ with G being a finite-index subgroup of $\mathbb{G}(\mathbb{Q}_S)$ and Γ is an irreducible arithmetic lattice commensurable to $\mathbb{G}(\mathcal{O}_S)$. Here, $\mathcal{O}_S$ denotes the ring of S-adic integers. Such an S-arithmetic quotient is said to be *saturated by unipotent* if the group generated by all unipotent elements of G acts ergodically on $\Gamma \backslash G$. For example, for $\mathbb{G} = \mathrm{SL}_n$ (or more generally simply-connected algebraic groups) the quotient $\Gamma \backslash \mathbb{G}(\mathbb{Q}_S)$ is saturated by unipotents.

A probability measure μ on an S-arithmetic quotient $\Gamma \backslash G$ is called *algebraic over* $\mathbb{Q}$ if there exists an algebraic group $\mathbb{H}$ defined over $\mathbb{Q}$ and a finite-index subgroup $H < \mathbb{H}(\mathbb{Q}_S)$ such that $\mu = m_{\Gamma H g}$ where $g \in G$ and $m_{\Gamma H g}$ denotes the normalized Haar measure on a single (necessarily closed, by the finiteness of μ) orbit - see §2.1 for a detailed definition.

The joinings we aim to classify are joinings of S-arithmetic quotients $X_i = \Gamma_i \backslash G_i$ which are saturated by unipotents, equipped with Haar probability measures $m_{X_i} = m_{\Gamma \backslash G}$, and a torus action which we now define. Following the notation of EINSIEDLER and LINDENSTRAUSS (2019) we say that a subgroup $A < G$ is of class-$\mathcal{A}'$ if it is simultaneously

diagonalizable and the projection of $a \in A$ to $\mathbb{G}(\mathbb{Q}_s)$ for any $s \in S$ satisfies the following: for $s = \infty$ it has only positive real eigenvalues, and for s equal to a finite prime p, we assume that all the eigenvalues are powers of θ_p for some $\theta_p \in \mathbb{Q}_p^\times$ with $|\theta_p|_p \neq 1$ chosen independently of $a \in A$. A homomorphism $\phi \colon \mathbb{Z}^d \to G$ is said to be of class $\mathcal{A}'$ if it is proper and $\phi(\mathbb{Z}^d)$ is of class-$\mathcal{A}'$. The term *higher-rank torus action* refers to such a homomorphism with $d \geq 2$. We are ready to state the main theorem of EINSIEDLER and LINDENSTRAUSS (2019, Theorem 1.7):

THEOREM 1.1 (Einsiedler–Lindenstrauss, 2019). — *Let $r, d \geq 2$ and let $\mathbb{G}_1, \ldots, \mathbb{G}_r$ be perfect algebraic groups defined over $\mathbb{Q}$, $\mathbb{G} = \prod \mathbb{G}_i$, and S be a finite set of places of $\mathbb{Q}$. Let $X_i = \Gamma_i \backslash G_i$ be S-arithmetic quotients for $G_i < \mathbb{G}_i(\mathbb{Q}_S)$ which are saturated by unipotents and set $G = \prod_{i=1}^r G_i$ and $X = \prod_{i=1}^r X_i$. Let $\phi_i \colon \mathbb{Z}^d \to G_i$ be homomorphisms such that $\phi = (\phi_1, \ldots, \phi_r) \colon \mathbb{Z}^d \to G$ is of class-$\mathcal{A}'$, and such that the projection of ϕ_i to every $\mathbb{Q}$-almost simple factor of $\mathbb{G}_i(\mathbb{Q}_S)$ is proper. Let $A = \phi(\mathbb{Z}^d)$ and suppose μ is an A-invariant and ergodic joining of the actions of $A_i = \phi_i(\mathbb{Z}^d)$ on X_i equipped with the Haar measure m_{X_i}. Then, μ is an algebraic measure defined over $\mathbb{Q}$.*

This theorem exemplifies the above principle concerning disjointness: let $\mathbb{H} < \mathbb{G}$ be the group showing the algebraicity of μ. If $\mathbb{H} = \mathbb{G}$ then μ is the trivial joining. Otherwise, $\mathbb{H}$ arises from a very strong relation between the algebraic groups $\mathbb{G}_i$. Indeed, certain of their $\mathbb{Q}$-simple factors need to be isogenous *over* $\mathbb{Q}$. In particular, if $\mathbb{G}_i$ are pairwise non-$\mathbb{Q}$-isogenous almost simple groups, any joining must be the trivial one. This situation strongly echoes Goursat's Lemma from group theory.

Taking again the broader viewpoint of measure rigidity for torus action (or $\mathbb{Z}^d$-actions) on homogeneous spaces, Theorem 1.1 is the most complete result in this context. Such rigidity results are currently only possible under a positive entropy assumption. In our context, the positive entropy assumption is hidden in the assumption that we join homogeneous spaces equipped with the Haar probably measure on each quotient (we give more details below). Moreover, the assumption that the groups are perfect is essential: considering more general groups in both factors would allow to recast the classification of $\mathbb{Z}^k$-actions on solenoids (including the zero entropy case - a notoriously difficult problem), as a classification problem of joinings.

Theorem 1.1 is already interesting when $r = 2$ and $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_n$ for $n \geq 3$ and $d \geq 2$, or for $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_2 \times \mathrm{SL}_2$ and $d = 2$. While reading this survey, the reader is advised to concentrate on these cases. Indeed, the techniques used and the main steps of the proofs of EINSIEDLER and LINDENSTRAUSS (2007) and EINSIEDLER and LINDENSTRAUSS (2019) are already visible when one considers the case where $\mathbb{G}_1$ and $\mathbb{G}_2$ are equal to SL_n for $n \geq 3$ or to $\mathrm{SL}_2 \times \mathrm{SL}_2$, and where $S = \{\infty\}$, that is, where we consider real Lie groups. Therefore, apart from describing the main result of EINSIEDLER and LINDENSTRAUSS (2019) in this introduction, we will reduce this survey to these cases.

To end this introduction we present a few images of the following arithmetic application (AKA, EINSIEDLER, and SHAPIRA, 2016) which appeared at the same time as (EINSIEDLER and LINDENSTRAUSS, 2019). We discuss further applications in §6.

For $D \in \mathbb{N}$ write

$$\mathbb{S}^2(D) = \left\{ (x, y, z) \in \mathbb{Z}^3 : x^2 + y^2 + z^2 = D, \gcd(x, y, z) = 1 \right\}.$$

By Legendre and Gauss we have $\mathbb{S}^2(D) \neq \emptyset$ if and only if $D \neq 0, 4, 7 \mod 8$. Consider

$$P_D := \frac{1}{\sqrt{D}} \cdot \mathbb{S}^2(D) \subset \mathbb{S}^2 := \left\{ (x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1 \right\}. \tag{1}$$

By a celebrated theorem of DUKE (1988), based on a breakthrough of IWANIEC (1987), P_D equidistribute on $\mathbb{S}^2$ when $D \to \infty$ along $D \neq 0, 4, 7 \mod 8$. That is, the following weak-* convergence

$$\mu_D := \frac{1}{|\mathbb{S}^2(D)|} \sum_{v \in \mathbb{S}^2(D)} \delta_{\frac{v}{\sqrt{D}}} \longrightarrow m_{\mathbb{S}^2}$$

holds, where $m_{\mathbb{S}^2}$ is the uniform (cone) measure on $\mathbb{S}^2$.

We wish to join this equidistribution problem with another equidistribution problem in a natural way. For each $v \in \mathbb{S}^2(D)$ we consider the two-dimensional lattice $\Lambda_v := v^\perp \cap \mathbb{Z}^3$ which we can consider up to rotation as lying in a fixed plane of $\mathbb{Q}^3$. We denote it by $[\Lambda_v]$ and call it *the (shape of the) orthogonal lattice of v*. The set $Q_D := \{ [\Lambda_v] : v \in \mathbb{S}^2(D) \}$ can be considered as a subset of the modular surface $X_2 := \mathrm{SL}_2(\mathbb{Z}) \backslash \mathbb{H}$ which parametrizes the space of two-dimensional lattices up to rotation, and carries a natural invariant probability measure m_{X_2}. A careful analysis (see e.g., ELLENBERG, MICHEL, and VENKATESH, 2013, §5.2) shows that the normalized counting measure on Q_D also equidistributes as $D \to \infty$ to m_{X_2}, by a variant of Duke's Theorem. This construction yields the following natural problem: does the normalized counting measure on

$$J_D := \left\{ (v, [\Lambda_v]) : v \in \mathbb{S}^2(D) \right\} \tag{2}$$

equidistribute to the product measure $m_{\mathbb{S}^2} \otimes m_{X_2}$ when $D \to \infty$ with $D \neq 0, 4, 7 \mod 8$?

We conjecture that it does (mainly because we don't see a reason why it shouldn't). Here is some "visible" evidence: for the D's below, we divide the modular surface X_2 using the height function into two (resp. three) equal m_{X_2}-measure regions and call lattices in each region non-stretched/stretched (resp. non-stretched/mildly stretched/super-stretched) and color each point on $\frac{1}{\sqrt{D}} \cdot \mathbb{S}^2(D)$ with a different color according to the type of its orthogonal lattice. In figures 1 and 2 below, one can see the distribution of the corresponding points together with the number of points of each type for $D = 101, 8011, 104851, 14500001$.

Both equidistribution problems in $\mathbb{S}^2$ and in X_2 may be individually phrased as two individual equidistribution problems on an S-arithmetic (or adelic) quotient as defined above (see § 6.1 for more details). Linnik could prove these results under a congruence condition on D modulo a fixed arbitrary prime (see § 6.1 for more details). It turns out that the coupling of $v \in \mathbb{S}^2(D)$ with its orthogonal lattice $[\Lambda_D]$ gives rise to a joining of
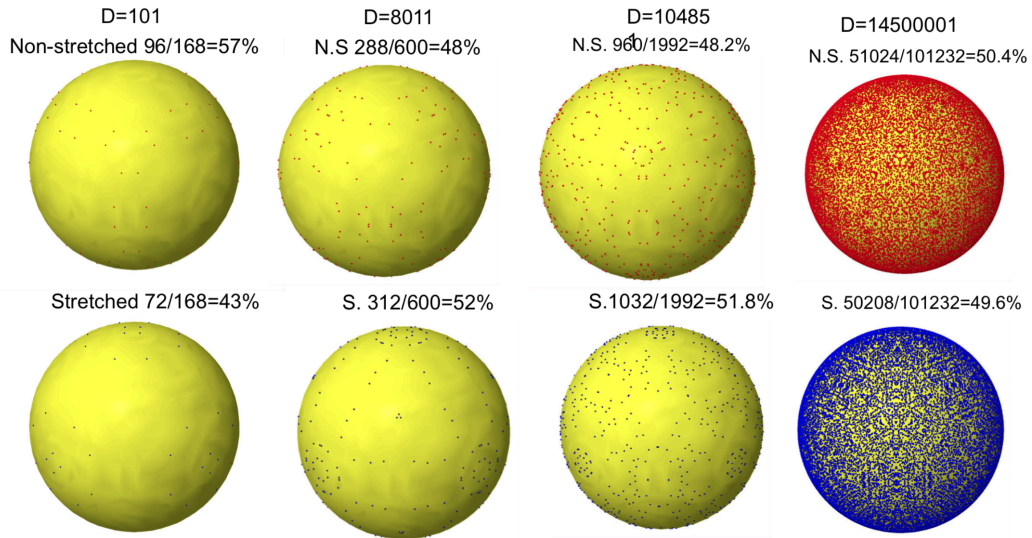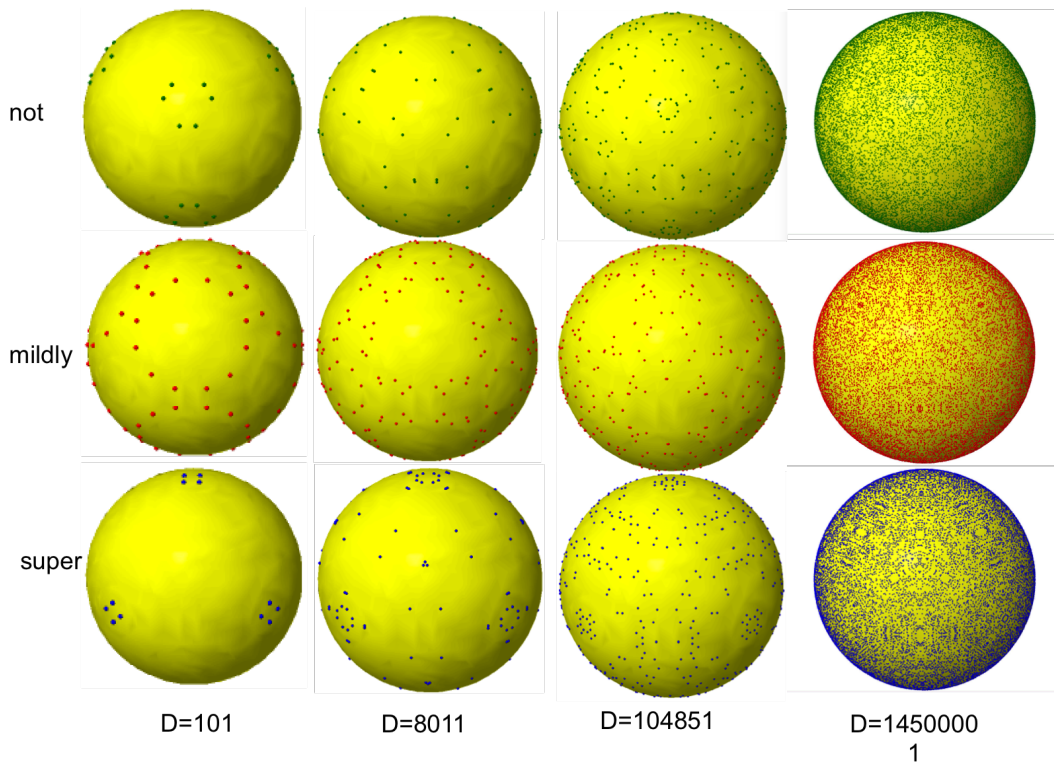
FIGURE 1. non-stretched vs. stretched

FIGURE 2. non-stretched vs. mildly stretched vs. non-stretched

the above S-arithmetic quotient. Under congruence conditions at *two* fixed arbitrary primes, one could apply Theorem 1.1 to deduce the equidistribution of the normalized counting measure on J_D to $m_{\mathbb{S}^2} \otimes X_2$ when $D \to \infty$ along $D \neq 0, 4, 7 \mod 8$ and the congruence conditions modulo the above two fixed primes. Recently BLOMER and

BRUMLEY (2020) showed that under the Generalized Riemann Hypothesis, the above equidistribution holds along $D \neq 0, 4, 7 \mod 8$ without any congruence conditions.

1.1. Bird's-eye view and the organization of this survey

Let's give a (very subjective and entropy-centred) bird's-eye view of the main ingredients used in the proofs of the main theorems of the works EINSIEDLER and LINDENSTRAUSS (2007, 2019). There are four main ingredients, all related to entropy:

– **Basic ingredient:** Leafwise measures, Lyapunov weights, entropy contribution, and in particular the relation between maximal entropy contribution and invariance.
– **Second ingredient:** Product structure for coarse Lyapunov weights, Abramov–Rokhlin Formula for coarse Lyapunov weights as a corollary.
– **Third ingredient:** The high-entropy method.
– **Fourth ingredient:** The low-entropy method.

As entropy considerations underlie all the above ingredients and every aspect of the work we survey, the title of §2 is entropy. In §2.1 we discuss a general setting and notation for the entire survey. In Sections 2.2-2.4 we review the basic ingredient, giving a bit of intuition and introducing several representative examples that will be discussed throughout the survey. In Sections 2.5-2.8 we discuss the second ingredient. Besides the product structure, which is discussed in §2.5, Lemma 2.7 also plays a central role. Its proof is based on a construction of a special partition (a partition which is subordinate to a given subgroup of a given stable horospherical subgroup). Such constructions play a major role also in the proofs of the results, which constitute the basic ingredient.

Using only the basic and the second ingredients, one can already draw several interesting corollaries, which we discuss in Section 3. For example, Corollary 3.1, which we call the "two ingredients joinings theorem", already gives the ability to reach strong conclusions on joinings in specific situations (which actually arise in applications), and Corollary 3.7 gives an important first step toward proving Theorem 1.1, and is used again and again throughout the proof.

In section 4 we introduce the third ingredient, the high-entropy method, which is summarized in theorem 4.1. The main Theorem of EINSIEDLER and LINDENSTRAUSS (2007) uses exactly these three ingredients (and could be referred to as the "three ingredients joinings theorem") and the method of its proof is exemplified by classifying all joinings in a specific representative example, see Theorem 4.2.

The main theorem of EINSIEDLER and LINDENSTRAUSS (2019) (Theorem 1.1) generalizes the main theorem of EINSIEDLER and LINDENSTRAUSS (2007) in many senses, but the main difference is that algebraic groups $\mathbb{G}_i$ which are products (with at least two factors) of forms of SL_2 had been previously excluded. Indeed, the third ingredient, the high-entropy method, implies nothing when the algebraic groups $\mathbb{G}_i$ are products of rank one groups. To also treat these cases, Einsiedler and Lindenstrauss used the fourth ingredient, the low-entropy method, yielding the "four ingredients joinings theorem" - Theorem 1.1. The general formulation of the low-entropy method may seem intimidating,

but the proof of it, although highly intricate, is very concrete. We therefore choose to give a rough sketch of a complete proof of the step involving the low-entropy method for another representative example, see §5. This sketch is based on a corresponding step in the proof of arithmetic quantum unique ergodicity by LINDENSTRAUSS (2006), where the low-entropy method was introduced (compare also to EINSIEDLER and LINDENSTRAUSS, 2010, Ch.10), and to the corresponding step in EINSIEDLER, KATOK, and LINDENSTRAUSS (2006).

The second part of the survey deals with applications of Theorem 1.1, which we discuss in Section 6. In §6.1 we give a (rather long) survey of how arithmetic problems relate to adelic torus orbits. We hope that the novice reader could gain some intuition about the use of adeles and p-adic numbers in dynamics from this subsection. In §6.2 we finally present several problems which are related (directly or retrospectively) to a coupling of the problems discussed in §6.1. We then explain how and under which conditions, Theorem 1.1 can give key input towards the solutions of these coupled problems. We finally present a new application for groups with high rank in §6.3.

Acknowledgements

2. ENTROPY

2.1. Setting

We start with a few general definitions. We recall that the action of G or its subgroups on a homogeneous space of the form $\Gamma\backslash G$ is given by $g.(\Gamma h) := \Gamma h g^{-1}$. Given a closed orbit of the form $\Gamma H g$ for $H < G$, the group preserving it is $g^{-1}Hg$. If the restriction of a Haar measure to a fundamental domain of $\Gamma \cap (g^{-1}Hg)$ in $g^{-1}Hg$ has a finite measure, we can normalize it to give a fundamental domain measure one. Pushing it forward via the above action to $\Gamma\backslash G$, we get the *normalized uniform/Haar measure* on the orbit $\Gamma H g$.

To ease the notation, we will fix a slightly simplified setting and refer to it below. We use the same notation as in Theorem 1.1 but fix $r = 2$. For clarity purposes, we repeat some notations: we let $\mathbb{G}_1$ and $\mathbb{G}_2$ be two semisimple algebraic groups defined over $\mathbb{Q}$. One may keep in mind the following "baby-cases": $\mathbb{G}_1 = \mathrm{SL}_n, \mathbb{G}_2 = \mathrm{SL}_{n'}$

(with $n = n'$ being an interesting case), and the case $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_2 \times \mathrm{SL}_2$. As in Theorem 1.1, we let $G_i < \mathbb{G}_i(\mathbb{Q}_S)$ be subgroups, Γ_i be two *irreducible* lattices in G_i and denote $X_i = \Gamma_i \backslash G_i$ and $X = X_1 \times X_2$. We assume that the action of G_i on X_i is saturated by unipotents. In the above "baby-cases" one can think about $S = \{\infty\}$ and $G_i = \mathrm{SL}_n(\mathbb{R})$ or $G_i = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ (or $G_i = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{Q}_p)$) where the saturated by unipotents assumption is satisfied. In the "baby-cases" we consider below, all the images of the class $\mathcal{A}'$-homomorphisms we consider, will be embedded in the (product of the) respective diagonal subgroup of SL_n and it will be easy to verify the class-$\mathcal{A}'$ assumption for them. We denote further by $a = (a_1, a_2)$ a diagonalizable element. The spaces X_i are equipped with the Haar measure m_{X_i} and we denote by μ an unknown measure on X, which is normally the unknown joining that we are trying to classify.

In very rough terms, given a joining μ as in the main theorem, Einsiedler and Lindenstrauss utilize and develop methods concerning entropy in order to find that μ is invariant under an unipotent element. Measure rigidity results (MARGULIS and TOMANOV, 1994; RATNER, 1991, 1995) are then employed to conclude that μ is algebraic. It is therefore essential for us to introduce and discuss a few of these entropy methods. These methods were predominantly developed by Einsiedler, Katok, Lindenstrauss and Spatzier.

We will assume that the reader knows the basic definitions and properties of entropy (see e.g., EINSIEDLER and LINDENSTRAUSS, 2010, §3) and of conditional measures (see e.g., EINSIEDLER and WARD, 2011, §5.3).

2.2. Weights and Lyapunov weights

For a diagonalizable regular element $a \in G$ the subgroup

$$G_a^- = \left\{ g \in G : a^n g a^{-n} \xrightarrow{n \to \infty} e \right\},$$

where e denotes the identity element in G, is called the *stable horospherical subgroup of a*. Its counterpart $G_a^+ := G_{a^{-1}}^-$ is called the *unstable horospherical subgroup of a*. These groups are central to the study of the action of a on homogeneous spaces of the form $\Gamma \backslash G$. For instance, we will explain later that the entropy $h_\mu(a)$ of the action of a is equal to a quantity that may be calculated solely through G_a^- — the entropy contribution $h_\mu(a, G_a^-)$. For now, let's just say that G_a^- is "defined using the dynamics of a" and that it contains "dynamical information". It turns out that G_a^- is too crude for us in order to extract the information we need for the joinings classification (like invariance under one of its elements). It is therefore interesting to know which subgroups of G_a^- can also be defined "dynamically". To this end, we will need to consider the action of the whole diagonalizable subgroup $A = \phi(\mathbb{Z}^d)$ (as in the notation of Theorem 1.1) and define weights for the action of A.

Recall that the adjoint action of an element $g \in G$ on the Lie algebra $\mathfrak{g} = \mathrm{Lie}\, G$ describes locally the conjugation action on G as exp and log (at least in characteristic zero) are local isomorphisms. This in turn describes the local dynamics on a homogeneous

space of the form $\Gamma\backslash G$, as each point in $\Gamma\backslash G$ (for a discrete group Γ) is locally isomorphic to G.

Consider now $A = \phi(\mathbb{Z}^d)$ for a class-$\mathcal{A}'$ homomorphism ϕ. A character $\lambda\colon A \to k^\times$ (here $k = \mathbb{Q}_s$ for some $s \in S$) is called a *weight* or a *Lyapunov weight* if there exists a non-zero $\mathrm{x} \in \mathfrak{g}$ which is a common eigenvector for the adjoint action of A, that is, for every $a \in A$ we have $\mathrm{Ad}_a(\mathrm{x}) = \lambda(a)\mathrm{x}$, where Ad denotes the adjoint representation $\mathrm{Ad}_a(\mathrm{x}) = a\mathrm{x}a^{-1}$. Once we consider a character via precomposition with ϕ as a map from $\mathbb{Z}^d$ to $k^\times$, it follows from the class-$\mathcal{A}'$ assumption, that characters are of the form $\lambda(\mathrm{n}) = e^{\mathrm{n}\cdot w_\lambda}$ for some $w_\lambda \in \mathbb{R}^d$ when $k = \mathbb{R}$, and of the form $\lambda(\mathrm{n}) = (\theta_p)^{\mathrm{n}\cdot w_\lambda}$ for some $w_\lambda \in \mathbb{Z}^d$ when $k = \mathbb{Q}_p$ (where θ_p was defined just before Theorem 1.1). For a fixed character λ the set of such common eigenvectors forms the weight spaces $\mathfrak{g}^\lambda$. As A (or rather $\mathrm{Ad}(A)$) is simultaneously diagonalizable, we can decompose $\mathfrak{g}$ as

$$\mathfrak{g} = \sum_{\lambda \in \Phi} \mathfrak{g}^\lambda$$

where Φ is the set of all weights for the action of A. As G_a^- is invariant under conjugation, it follows that

$$\mathrm{Lie}(G_a^-) = \sum_{\lambda \in \Phi, |\lambda(a)| < 1} \mathfrak{g}^\lambda \tag{3}$$

and that $\exp$ is a global homomorphism from this nilpotent Lie algebra to G_a^-. We may then ask ourselves if $\exp(\mathfrak{g}^\lambda)$ is a subgroup and if it is "dynamically defined". First note that $\mathfrak{g}^\lambda$ is not necessarily a subalgebra when λ^2 is also a weight (in general we have $[\mathfrak{g}^\gamma, \mathfrak{g}^\eta] \subset \mathfrak{g}^{\gamma+\eta}$). But more importantly, in order to find "dynamically-defined" subgroups, we may carry out one of the following two equivalent constructions. We can check which subalgebras of the form $\mathfrak{g}^\lambda$ are indistinguishable from one another using the whole dynamics of A. More precisely, we may say (temporarily) that two characters are equivalent and write $\lambda \sim \lambda'$ if for any $a \in A$ we have

$$\mathfrak{g}^\lambda \subset \mathrm{Lie}(G_a^-) \iff \mathfrak{g}^{\lambda'} \subset \mathrm{Lie}(G_a^-).$$

The union of an equivalence class of weights under this relation forms a subalgebra, and its image under $\exp$ is a subgroup of G_a^- for some $a \in A$, which is "defined dynamically". Alternatively, we could have asked: which subgroups are the smallest non-trivial intersection of subgroups of the form G_a^- for various $a \in A$? Both questions lead to the same answer; the resulting subgroups are called the coarse Lyapunov weights. To define these only in terms of the weights, we say that two weights λ and η are equivalent and write $\lambda \sim \eta$, if there exist positive integers m, ℓ with $\lambda^m = \eta^\ell$. The equivalence classes are called *coarse Lyapunov weights*. Given a coarse Lyapunov weight $[\lambda]$, the sum $\sum_{\eta \sim \lambda} \mathfrak{g}^\eta$ is called the *coarse Lyapunov weight space* and is denoted by $\mathfrak{g}^{[\lambda]}$. This is a nilpotent subalgebra and $\exp$ defines a global homomorphism from $\mathfrak{g}^{[\lambda]}$ to a nilpotent subgroup denoted by $G^{[\lambda]}$, which is called a *coarse Lyapunov subgroup*.

In order to visualize the weight structure in given examples, it is convenient to consider the logarithm of the (real or p-adic) absolute value of a given character: $\log|\lambda(\mathrm{n})|$. Viewing this as a map from $\mathbb{Z}^d$ to k we get a linear map that can be considered as an

element of the dual of $\mathbb{Z}^d \otimes k \cong k^d$. In the real case, $\log \circ \lambda$ is just the inner product with the vector w_λ defined above, so one may identify the weight λ with w_λ. Under this identification, the coarse Lyapunov weight $[\lambda]$ is the union of all weights η such that $w_\eta \in \mathbb{R}^+ \cdot w_\lambda$.

Recall now the setting in §2.1. Theorem 1.1 tells us that the set of possible joinings is strongly connected to the relation between G_1 and G_2, equipped with the corresponding torus action $\phi_i : \mathbb{Z}^2 \to G_i$, $i = 1, 2$. As we will explain later, the first instance where this reflects in the proof is the comparison between the weights structure of G_1 (with respect to ϕ_1) and the weights structure of G_2 (with respect to ϕ_2). We will consider four representing examples (Examples 2.1(1)-2.1(3), and Example 2.2 below). The first three are all with $G_1 = G_2 = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ but equipped with different torus action. In Example 2.1(1), both weight structures will be identical, which will give rise to the presence of possible "diagonal" joinings, e.g. the one supported on the graph of the identity isomorphism $\mathrm{Id} : G_1 \to G_2$. In Example 2.1(2) the weights structures will be "unrelated" to each other, which will allow us to show that there is no non-trivial joining. For this case, we will just need the basic and the second ingredients mentioned in §1.1, and in particular the high or the low entropy method are not needed, see §3.1. In Example 2.1(3), the weights will be related but will have "different speeds". Also in this example, only the trivial joining may occur, but in order to show this one need to use the low entropy method, see §5).

Example 2.1 (Three examples on $(\mathrm{SL}_2 \times \mathrm{SL}_2)^2$). — Let $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_2 \times \mathrm{SL}_2$, $S = \{\infty\}$, $G_i = \mathbb{G}_i(\mathbb{R})$, $i = 1, 2$, $G = G_1 \times G_2$. Let D be the diagonal group in SL_2,

$$U = \left\{ \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \right\}, \ V = \left\{ \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} \right\}, \ a_t = \begin{pmatrix} e^{-\frac{t}{2}} & \\ & e^{\frac{t}{2}} \end{pmatrix}$$

and $\mathfrak{u}, \mathfrak{v}$ the corresponding Lie algebras. We consider three different homomorphisms from $\mathbb{Z}^2$ to $D \times D$ (the latter is a subgroup of both G_1 and G_2):

$$\phi(t, s) = (a_t, a_s), \quad \psi(t, s) = (a_{t+s}, a_{t-s}), \quad \tau(t, s) = (a_{2t}, a_{2s}). \tag{4}$$

Each one of them is of class $\mathcal{A}'$-homomorphism and it has a related weight structure: let $\mathfrak{h} = \mathrm{Lie}(\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}))$ and note that with respect to each of the maps in (4), the weight spaces are

$$\mathfrak{h}^{\lambda_1} = \mathfrak{u} \times \{0\}, \ \mathfrak{h}^{-\lambda_1} = \mathfrak{v} \times \{0\}, \ \mathfrak{h}^{\lambda_2} = \{0\} \times \mathfrak{u}, \ \mathfrak{h}^{-\lambda_2} = \{0\} \times \mathfrak{v}$$

but for different characters λ_1 and λ_2:

- For $\phi(t, s)$: $w_{\lambda_1} = (-1, 0)$ (since $\lambda_1(\phi(t, s)) = e^{(t,s) \cdot (-1,0)} = e^{-t}$), and $w_{\lambda_2} = (0, -1)$.
- For $\psi(t, s)$: $w_{\lambda_1} = (-1, -1)$ (since $\lambda_1(\psi(t, s)) = e^{(t,s) \cdot (-1,-1)} = e^{-(t+s)}$), and $w_{\lambda_2} = (-1, 1)$.
- For $\tau(t, s)$: $w_{\lambda_1} = (-2, 0)$ (since $\lambda_1(\tau(t, s)) = e^{(t,s) \cdot (-2,0)} = e^{-2t}$), $w_{\lambda_2} = (0, -2)$.

In general, one has $w_{-\eta} = -w_\eta$, so we omit the weight $-\lambda_i$ above. These roots systems (only the non-trivial weights) are drawn in Figure 3.
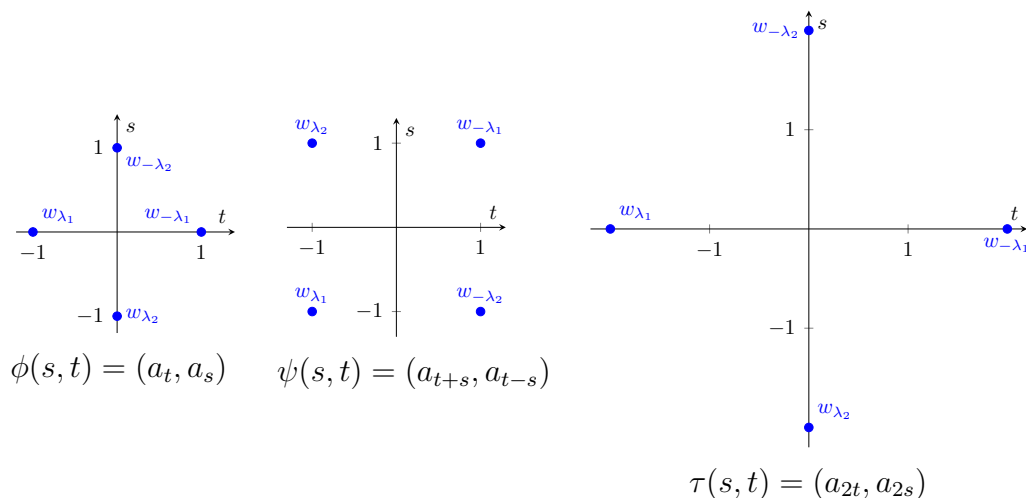
$$\phi(s,t) = (a_t, a_s) \qquad \psi(s,t) = (a_{t+s}, a_{t-s})$$

$$\tau(s,t) = (a_{2t}, a_{2s})$$

FIGURE 3. Root systems for $\mathrm{SL}_2 \times \mathrm{SL}_2$

We consider now three different torus actions on $\Gamma_1 \backslash \mathbb{G}_1(\mathbb{R}) \times \Gamma_2 \backslash \mathbb{G}_2(\mathbb{R})$ where Γ_1 and Γ_2 are irreducible arithmetic lattices, and classify joining with respect for each of these three actions. These actions are given by the following class-$\mathcal{A}'$ homomorphisms:

1. $(\phi, \phi)\colon \mathbb{Z}^2 \to (D \times D)^2 < G_1 \times G_2$. ("Identical weights")
2. $(\phi, \psi)\colon \mathbb{Z}^2 \to (D \times D)^2 < G_1 \times G_2$. ("45° rotation")
3. $(\phi, \tau)\colon \mathbb{Z}^2 \to (D \times D)^2 < G_1 \times G_2$. ("Different speeds")

The action of (ϕ, ϕ) has four (non-trivial) weight spaces and four coarse Lyapunov weights, the action of (ϕ, ψ) has eight (non-trivial) weight spaces and eight coarse Lyapunov weights, and the action of (ϕ, τ) has eight (non-trivial) weight spaces but only four coarse Lyapunov weights. The weight spaces may be visualized as in figure 4

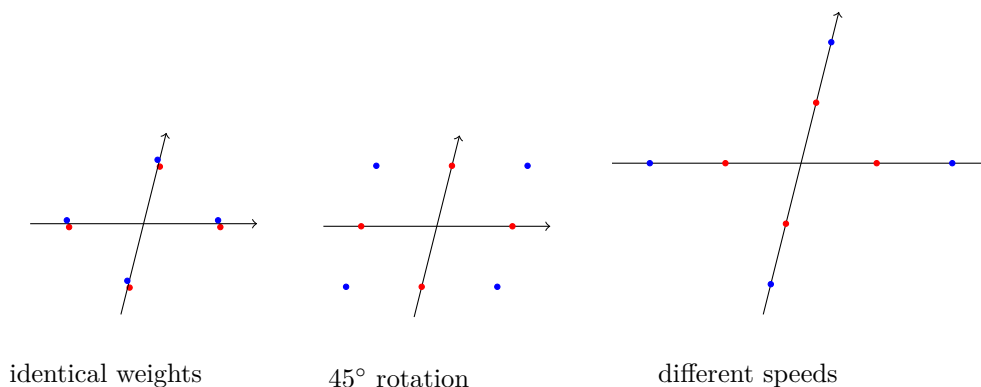identical weights   45° rotation   different speeds

FIGURE 4. Decomposition of $\mathfrak{sl}_2^4 = (\mathfrak{sl}_2 \times \mathfrak{sl}_2)^2$ into weight spaces

In the examples above, the (non-opposite) weight spaces always commute with each other. As we will see later, this means that the high entropy method (see §4) does not help at all in these cases. This is why these cases were excluded in EINSIEDLER

and LINDENSTRAUSS (2007) and dealt with later in EINSIEDLER and LINDENSTRAUSS (2019) using the low entropy method (see §5). To give a representing example where there are non-commuting Lyapunov weights, we consider the following example.

Example 2.2. — Let $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_3$, $S = \{\infty\}$, $G_i = \mathbb{G}_i(\mathbb{R})$, $i = 1, 2$, $G = G_1 \times G_2$. Let $D < G_i$ be the diagonal group and for $i = 1, 2$, let

$$\phi_i : \mathbb{Z}^2 \to D, \; \phi_i(t, s) = \begin{pmatrix} e^{-\frac{t}{2}} & & \\ & e^{-\frac{s}{2}} & \\ & & e^{\frac{t+s}{2}} \end{pmatrix}.$$

The map $\phi = (\phi_1, \phi_2)$ is of class $\mathcal{A}'$ and we set $A = \phi(\mathbb{Z}^2) < G \times G$.

Let's concentrate for a moment on one of the factors, say on $G_1 = \mathrm{SL}_3(\mathbb{R})$ and $A_1 := \phi_1(\mathbb{Z}^2)$: for $1 \le i \ne j \le 3$ let $U_{ij} < \mathrm{SL}_3(\mathbb{R})$ be the unipotent one-parameter group with 1's on the diagonal and zeroes everywhere except from the ij-entry, and $\mathfrak{u}_{ij}$ the corresponding Lie algebra. Then, the eight-dimensional Lie algebra $\mathfrak{sl}_3$ decomposes into weight spaces as $\mathrm{Lie}(D) \oplus (\oplus_{1 \le i \ne j \le 3} \mathfrak{u}_{ij})$. Furthermore, each of the spaces $\mathfrak{u}_{ij}$ forms a coarse Lyapunov weight with respect to ϕ_1, and the reader may calculate the corresponding weights to get the weight diagram in Figure 5, where we denote the weight vector corresponding to $\mathfrak{u}_{ij}$ with w_{ij}.
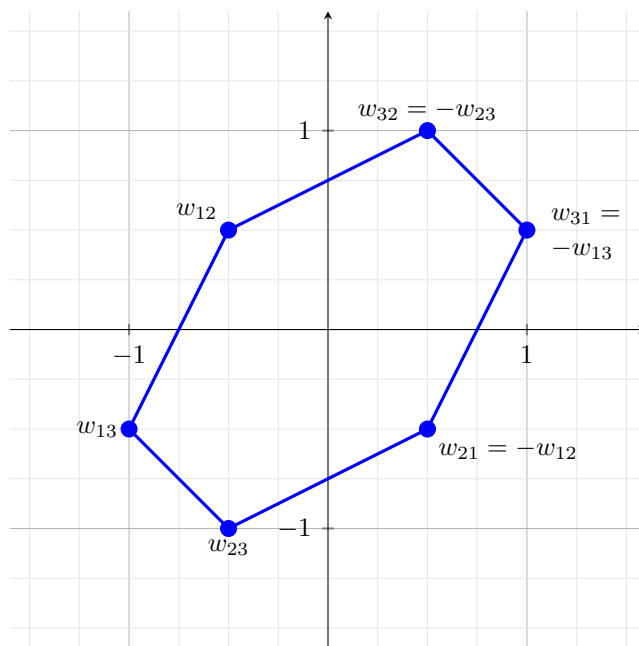
FIGURE 5. Decomposition of $\mathfrak{sl}_3$ into weight spaces

There are commutation relations between the weight spaces, which are also visible via addition of the vectors w_{ij}: for example $[\mathfrak{u}_{12}, \mathfrak{u}_{23}] = \mathfrak{u}_{13}$. Moreover, one can read from the weight diagram in Figure 5 the decomposition of the Lie algebra of any stable horospherical: for $a = \phi_1(t, s)$, it follows from the definition of the weight spaces and

their corresponding weight vectors that $(\mathfrak{sl}_3)_a^-$ is the direct sum of all the weight spaces whose weight vector has a negative inner product with (t,s). That is, the corresponding weight vectors must belong to a half-plane defined by the line $(t,s)^\perp$. Further, note that for each weight space, the corresponding weight vector can be written as a sum of two other weight vectors, such that all three weight vectors belong to a half-plane related to some $a \in A_1$. It follows that each weight space can be written as the commutator of two other weight spaces, all belonging to the same stable horospherical of some $a \in A_1$. This fact will be crucial for us in §4.

Fix now two irreducible lattices Γ_1 and Γ_2 in $\mathrm{SL}_3(\mathbb{R})$ and let's return to consider joinings for the action of ϕ on $\Gamma_1\backslash G_1 \times \Gamma_2\backslash G_2$. We have at least two types of joinings in this setting: the trivial joining, which is the product measure $m_{\Gamma_1\backslash G_1} \otimes m_{\Gamma_2\backslash G_2}$, and a "diagonal" joining, for example, the Haar measure supported on an orbit of $\{(g,g) : g \in \mathrm{SL}_3(\mathbb{R})\} < \mathrm{SL}_3^2(\mathbb{R})$. In Theorem 4.2 we give a more precise statement and see that these are the only possible types of *ergodic* joinings in this case (In general, an arbitrary joining is a convex combination of ergodic joinings).

2.3. Leafwise measures

The joining classification is a part of the classification of measures with torus actions on homogeneous spaces, under the assumptions that some elements act with positive entropy. In the joinings classification, the positive entropy assumption is "hidden" in the fact that each factor is equipped with the Haar measure and therefore the action on each factor has positive entropy, which imply that any non-trivial $a \in A$ has positive entropy (see Remark 2.6). Positive entropy is related to growth/decay properties of the measure. For example, for a hyperbolic toral automorphism positive entropy is equivalent to positive Hausdorff dimension of the measure which, in turn, is equivalent to an almost surely decay property of the measure of r-balls when $r \to 0$. In our context, the action of $a \in A$ on $\Gamma\backslash G$ is not hyperbolic as a commutes with all elements of A. Nevertheless, one could characterize again positive entropy as an almost sure decay property by considering the so-called Bowen-balls. Moreover, as we will state precisely in Proposition 2.4, one can relate the entropy $h_\mu(a)$ of a diagonalizable element a on a homogenous space $\Gamma\backslash G$ to decay of the measure along the stable horospherical subgroup G_a^-. Given the decomposition above of G_a^- to coarse Lyapunov subgroups, one may ask which subgroups of G_a^- contribute entropy. To be able to answer such questions and to be able to state the above statements in a choices-free way (e.g. independently of a choice of a specific generating partition for the entropy calculation) one is led to the notion of leafwise measures.

The aim of leafwise measures in our context is to describe a given measure μ on $X := \Gamma\backslash G$, as in Theorem 1.1, along orbits of an a-normalized subgroup $U < G_a^-$ with $h_\mu(a) > 0$. For example, along U being a coarse Lyapunov subgroup as above. Unfortunately, normally there exists no countably-generated σ-algebra whose atoms are U-orbits (at least for ergodic systems when individual orbits have measure zero).

Therefore, the natural candidates that will describe μ along orbits, conditional measures, cannot be constructed. Leafwise measures are constructed to remedy this problem. Roughly said, if one restricts to a bounded portion of a U-orbit, the orbits of this bounded portion can be the atoms of countably-generated σ-algebra. One considers the conditional measures which respect to such a σ-algebra. In order to get a measure that describes the measure μ along the full orbit, the idea is to consider larger and larger parts of the U-orbits and to patch the resulting conditional measures together using the group U itself as a reference object. This construction results in a family of locally-finite measures $\left\{\mu_x^U\right\}_{x\in X}$ on U, called the *leafwise measures* of U. They are defined almost everywhere and up to proportionality (as the "patching" can only be done up to an arbitrary constant). Defining these measures on U rather on the space itself has the advantage that we can directly exploit the group structure on U when we extract information from $\left\{\mu_x^U\right\}_{x\in X}$.

In this next subsection, we give a precise definition of leafwise measures and their characterizing property. The best way to understand their essence is via EINSIEDLER and LINDENSTRAUSS (2010, §6) where the above patchwork is carried out carefully.

2.3.1. *Definition of leafwise measures.* — The setting that we have in mind is a homogeneous space $X = \Gamma\backslash G$, an acting diagonalizable element $a \in G$ which has positive entropy with respect to an a-invariant measure μ and a closed subgroup $U < G_a^-$. Note that a priori μ has no known invariance under U. To help the reader to see what is essential for the construction, we use a slightly generalized setting in the following proposition, which defines the leafwise measures. The subgroup H that appears in the proposition will be later G_a^- or an a-normalized subgroup of it.

PROPOSITION 2.3. — *Let $X = \Gamma\backslash G$ be an S-adic homogeneous space and μ a locally-finite measure on X. Let $H < G$ be a closed subgroup which acts freely on X almost everywhere, that is, for a.e. $x \in X$ the map*

$$h \in H \mapsto h.x := xh^{-1} \tag{5}$$

is injective. Then, there exists a set of full measure $X' \subset X$ and family $\left\{\mu_x^H\right\}_{x\in X'}$ of locally-finite measures on H, unique up-to proportionality, having the following properties.

1. *(Characterizing property - describing μ along H) Let $Y \subset X$ be a measurable set with $\mu(Y) < \infty$ and $\mathcal{A}$ be a countably generated σ-algebra on Y which is H-subordinate (that is, its atoms $[y]_\mathcal{A}$ are bounded pieces of an H-orbit, or more precisely, there exists a bounded open subset $V_y \subset H$ with $[y]_\mathcal{A} = V_y.y$. The subset V_y is called the shape of the atom $[y]_\mathcal{A}$). Then, the leafwise measure μ_y^H describes the conditional measure $\mu_y^\mathcal{A}$ in the following sense:*

$$\mu_y^\mathcal{A} = \frac{1}{\mu_y^H(V_y)}\left(\mu_y^H|_{V_y}.y\right)$$

where $\mu_y^H|_{V_y}$ stands for the restriction of μ_y^H to V_y and $\mu_y^H|_{V_y}.y$ for pushing this measure under the map (5).

2. *(Compatibility/Shifting formula) The measure μ_x^H describes the orbit $H.x$ with $e \in H$ corresponding to $x \in X$. In particular, the following* shifting formula *holds:*

$$(R_h)_* \, \mu_{h.x}^H \propto \mu_x^H \qquad (6)$$

where $R_h \colon H \to H$ is defined by $h' \mapsto h'h$ and $\propto$ denote proportionality of measures.

We normalize[(1)] the measures μ_x^H to have $\mu_x^H(B_1^H) = 1$ where B_1^H is the ball of radius 1 in H around e. We further have

3. *(Invariance) The measure μ is H-invariant if and only if μ_x^H is equal to the Haar measure on H for almost every $x \in X$.*

4. *(Measure preservation implies equivariance) Assume further that there exists $a \in G$ that normalizes H and preserves the measure μ. Then*

$$\mu_{a.x}^H \propto (\theta_a)_* \, \mu_x^H \qquad (7)$$

where $\theta_a \colon H \to H$ is defined by $h \mapsto aHa^{-1}$. If particular, if θ_a acts trivially on H, we have $\mu_{a.x}^H = (\theta_a)_* \, \mu_x^H$ (by the above normalization).

Properties 2 and 4 can be visualized as follows: Imagine that the colour intensity in Figure 6 represents the distribution of the given measure μ. An embedding of an H-orbit $H.x = H.(h.x)$ gives rise to compatible distributions $\mu_x^H, \mu_{h.x}^H$ on H in the sense that wrapping them on the orbit with $e \in H$ corresponding to the point $x \in X$ (resp. to the point $h.x \in X$) will be compatible to the given distribution of the measure. This compatibility give rise to the shifting property stated in equation (6).

Similarly, imagine that the given measure μ gives rise to some distribution along a given H-orbit $H.x$ as in Figure 7. Imagine further that the a action contracts the neighbourhood around x. Then, the distribution along the orbit $H.(a.x)$ is just a contracted version of the distribution along $H.x$. This gives rise to the equivariance property stated in equation (7).
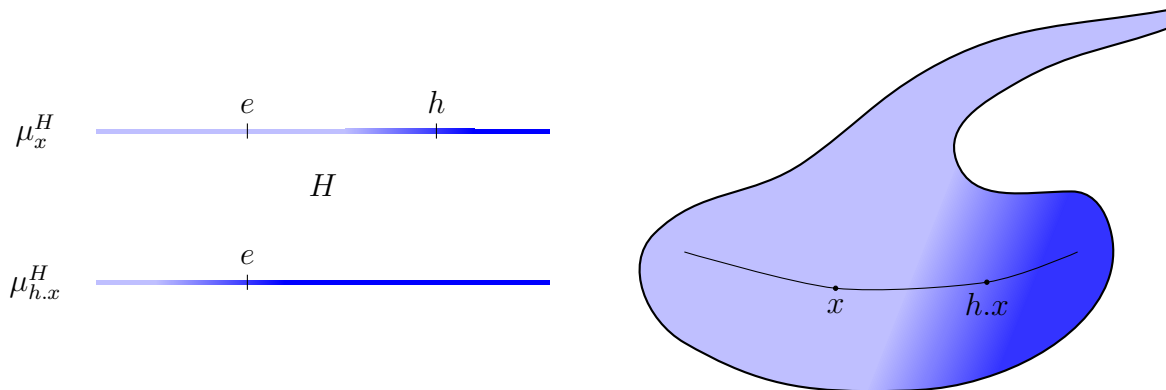

FIGURE 6. Shifting property

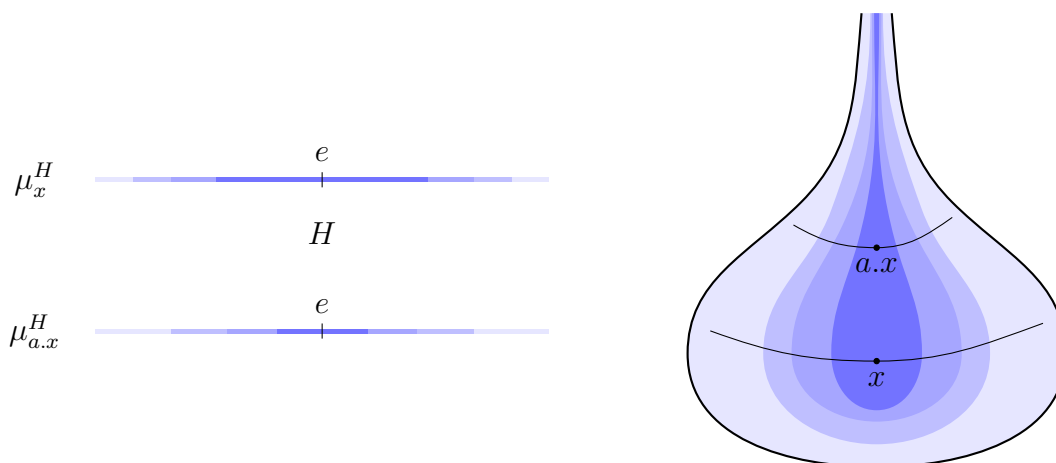[(1)] See EINSIEDLER and LINDENSTRAUSS (2010, §6.29) for a discussion of possible normalizations.

FIGURE 7. Equivariance property

We note that the assumption (5) in Proposition 2.3 holds when $H = U$ for an a-normalized subgroup $U < G_a^-$: if $u.x = x$ for some $u \in U$ then $aua^{-1}.(a.x) = a.x$. Since $U < G_a^-$ this shows that the injectivity radius of $a^n.x$ goes to zero. So either $u.x = x$ cannot hold (e.g., when X is compact) or x fails Poincaré recurrence (with respect to the measure that a preserves) and therefore x belongs to a null set.

2.4. Entropy contribution and invariance

We stated above that the entropy $h_\mu(a)$ is related to a volume decay property for the stable horospherical subgroup. We want now to use the notion of leafwise measures to define this more precisely, and use the same definition to quantify the entropy contribution of (a-normalized) subgroups of the stable horospherical subgroup of a.

Consider the setting of Proposition 2.3 with the additional assumptions in the 4-th item. We denote by $\theta_a\colon G \to G$ the conjugation map by a and recall that B_1^H is the ball of radius one in H around e. The *volume decay entropy at x* is defined by

$$\mathrm{vol}_\mu(a, H, x) = -\lim_{n\to\infty} \frac{\log \mu_x^H\left(\theta_a^n\left(B_1^H\right)\right)}{n}. \tag{8}$$

This limit exists almost everywhere and should be thought of as a pointwise entropy contribution: we define its integral

$$h_\mu(a, H) = \int_X \mathrm{vol}_\mu(a, H, x)\mathrm{d}\mu \tag{9}$$

to be the *entropy contribution of H*.

PROPOSITION 2.4 (Entropy as volume decay). — *Let $X = \Gamma\backslash G$ be an S-adic homogeneous space with a probability measure μ which is invariant under a diagonalizable element $a \in G$. Then $h_\mu(a) = h_\mu(a, G_a^-)$.*

The volume decay with respect to the Haar measure m_X is a local quantity that can be directly calculated using the adjoint map on the Lie algebra. For example,

$$\mathrm{vol}_{m_X}(a, G_a^-, x) = -\log \left| \det \mathrm{Ad}_a \left|_{\mathrm{Lie}(G_a^-)}\right| \right|, \quad \text{for every } x \in X.$$

There are many measure-preserving systems where the uniform (e.g. Haar) measure is the unique measure of maximal entropy. For many S-arithmetic quotients this fact can be deduced from the following theorem:

THEOREM 2.5. — *With the assumptions as in Proposition 2.4, let $U < G_a^-$ be an a-normalized subgroup (e.g. a coarse Lyapunov subgroup). We have*

$$h_\mu(a, U) \leq -\log \left| \det \mathrm{Ad}_a \left|_{\mathrm{Lie}(U)}\right| \right| = h_{m_X}(a, U) \tag{10}$$

with equality if and only if μ is U-invariant.

In other words, μ is U-invariant if and only if the entropy contribution of U with respect to μ agrees with the entropy contribution with respect to the Haar measure. One should keep this in mind together with the previous characterization we had for invariance in Proposition 2.3(3).

This theorem together with the decomposition of $\mathrm{Lie}(G_a^-)$ in (3) and the ability to define the entropy contribution for any coarse Lyapunov subgroup raises the question of understanding the relations between $\mu_x^{G_a^-}$ and $\mu_x^{G^{[\lambda]}}$ for the various coarse Lyapunov subgroups $G^{[\lambda]} < G_a^-$. EINSIEDLER and KATOK (2005) show that this relation is quite simple, as we explain in the next subsection. We remark that Proposition 2.4 and the results of the next subsection are variations of preceding results of LEDRAPPIER and YOUNG (1985a,b) for C^2-diffeomorphism of compact manifolds.

2.5. The product structure

The following was proven by EINSIEDLER and KATOK (2005, Theorem 8.4): let $U = G_a^-$ for some diagonalizable element a that preserves a probability measure μ on an S-arithmetic homogeneous space X. Let $G^{[\lambda_1]}, \ldots, G^{[\lambda_k]}$ denote all coarse Lyapunov subgroups and for a weight λ, and $\mu_x^{[\lambda]}$ denote the leafwise measure $\mu_x^{G^{[\lambda]}}$. Then, the map

$$\iota \colon \prod_{i=1}^{k} G^{[\lambda_i]} \to U, \ \iota(u_1, \ldots, u_k) = u_1 \ldots u_k$$

is a homeomorphism and for almost all $x \in X$ we have the so-called "product structure":

$$\mu_x^U \propto \iota_* \left(\mu_x^{[\lambda_1]} \otimes \cdots \otimes \mu_x^{[\lambda_k]} \right). \tag{11}$$

This readily implies the following addition property for entropy contribution:

$$h_\mu(a, G_a^-) = \sum_{i=1}^{k} h_\mu(a, G^{[\lambda_i]}). \tag{12}$$

2.6. Abramov–Rokhlin entropy addition formula

The joinings we consider (as in Theorem 1.1) have natural factors, both of which are equipped with the corresponding Haar measure. This should imply that we have $h_a(\mu) > 0$ for any non-trivial $a \in A$. The Abramov–Rokhlin enables to show this and to calculate the entropy in a step-wise manner through the factors. To state the formula, we shortly recall the definition of conditional entropy.

Given a Borel probability space $(X, \mathcal{B}, \mu)$ and two σ-algebras $\mathcal{A}, \mathcal{C} \subset \mathcal{B}$ with $\mathcal{C}$ countably generated, one defines the *information function on $\mathcal{C}$ given $\mathcal{A}$* as

$$I_\mu(\mathcal{C}|\mathcal{A})(x) = -\log \mu_x^{\mathcal{A}}\left([x]_{\mathcal{C}}\right),$$

and *the conditional entropy of $\mathcal{C}$ given $\mathcal{A}$*, $H_\mu(\mathcal{C}|\mathcal{A})$, as the integral of the information function

$$H_\mu(\mathcal{C}|\mathcal{A}) = \int_X I_\mu(\mathcal{C}|\mathcal{A})(x)\mathrm{d}\mu.$$

One should think about this conditional entropy as measuring how much new information $\mathcal{C}$ has, given the information that we already have from $\mathcal{A}$. To define the conditional entropy of a measure-preserving transformation $T\colon X \to X$, one defines

$$h_\mu(T|\mathcal{A}) = \sup_{\eta: H_\mu(\eta)<\infty} h_\mu(T, \eta|\mathcal{A})$$

with

$$h_\mu(T, \eta|\mathcal{A}) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\eta_0^{n-1}|\mathcal{A}).$$

With these definitions, using the setting and the notation from subsection 2.1, the Abramov–Rokhlin entropy addition formula states for a diagonalizable element $a = (a_1, a_2)$ that

$$h_\mu(a) = h_{m_{X_1}}(a_1) + h_\mu\left(a|\pi_1^{-1}\mathcal{B}_{X_1}\right), \tag{13}$$

where $\pi_1 : X \to X_1$ is the natural projection map and $\mathcal{B}_{X_1}$ is the Borel σ-algebra on X_1. By symmetry of the assumptions, the same holds with 1 and 2 interchanged.

Remark 2.6. — It follows directly from (13) that any non-trivial $a \in A$ has positive entropy.

2.7. An entropy inequality

One of the main ideas of EINSIEDLER and LINDENSTRAUSS (2007) is that Abramov–Rokhlin formula (13) may be coupled with (12) to get an "Abramov–Rokhlin type formula/inequality for the entropy contribution". In other words, a main step of the proof of the joining classification is the following inequality, which shows that the entropy contribution with respect to a joining can be bounded in a stepwise manner using the natural factors:

LEMMA 2.7. — *With the setting of §2.1 let $a = (a_1, a_2)$ be a diagonalizable element with a_i, $i = 1, 2$ acting ergodically on $X_i = \Gamma_i \backslash G_i$ equipped with the Haar measure m_{X_i}, let $U_1 < (G_1)_{a_1}^-$ and $U_2 < (G_2)_{a_2}^-$ be two subgroups, and set $U = U_1 \times U_2 < G$. Assume that U is a-normalized and let μ be a joining on X (that is, an a-invariant measure with $\pi_i(\mu) = m_{X_i}$, $i = 1, 2$). Then, we have*

$$h_\mu(a, U) \le h_{m_{X_1}}(a_1, U_1) + h_\mu(a, \{e\} \times U_2). \tag{14}$$

Moreover, for the stable horospherical subgroup $U = G_a^-$ (that is, for $U_i = (G_i)_{a_i}^-$, $i = 1, 2$) we have equality in (14).

The same results of course hold with 1 and 2 interchanged, by the symmetry of the assumptions of this lemma. The proof of this lemma uses tailor-made partitions which are subordinate to the stable horospherical subgroups $(G_i)_{a_i}^-$ and to the subgroups U_i (see EINSIEDLER and LINDENSTRAUSS, 2007, Prop. 3.1).

2.8. Equality for coarse Lyapunov weights

Lemma 2.7 together with the product structure (11) implies that we also have equality in (14) for a coarse Lyapunov subgroup U:

PROPOSITION 2.8 (Abramov–Rohklin for Coarse Lyapunov subgroups)
     With the setting of Lemma 2.7 the following equality holds for any coarse Lyapunov subgroup $G^{[\lambda]} = G_1^{[\lambda]} \times G_2^{[\lambda]}$

$$h_\mu(a, G^{[\lambda]}) = h_{m_{X_1}}(a_1, G_1^{[\lambda]}) + h_\mu\left(a, \{e\} \times G_2^{[\lambda]}\right). \tag{15}$$

By symmetry, the same statement holds with 1 and 2 interchanged.

Proof. — This is just simple arithmetic, as the quantities in (15) are additive with respect to a product of measures, and since we have equality for the stable horospherical and an inequality for each coarse Lyapunov weight. More precisely, let $G^{[\lambda]}$ be a Lyapunov weight and choose $a \in G$ such that $G^{[\lambda]} < G_a^-$. Let $G_a^- = G^{[\lambda_1]} \ldots G^{[\lambda_k]}$ be the corresponding decomposition of the stable horospherical to coarse Lyapunov subgroups (with, say, $\lambda = \lambda_1$). For each $1 \le \ell \le k$ we have by Lemma 2.7

$$h_\mu(a, G^{[\lambda_\ell]}) \le h_{m_{X_1}}(a_1, G_1^{[\lambda_\ell]}) + h_\mu\left(a, \{e\} \times G_2^{[\lambda_\ell]}\right).$$

By the product structure (11) (applied to μ and also to m_{X_1}) and the definition of entropy contribution the sum of each term over $\ell = 1, \ldots k$, sums up to the corresponding term for G_a^-. It follows that

$$h_\mu(a, G_a^-) = \sum_{\ell=1}^k h_\mu(a, G^{[\lambda_\ell]}) \tag{16}$$

$$\le \sum_{\ell=1}^k h_{m_{X_1}}(a_1, G_1^{[\lambda_\ell]}) + \sum_{\ell=1}^k h_\mu\left(a, \{e\} \times G_2^{[\lambda_\ell]}\right) \tag{17}$$

$$= h_{m_{X_1}}(a_1, (G_1)_{a_1}^-) + h_\mu\left(a, \{e\} \times (G_2)_{a_2}^-\right). \tag{18}$$

But we know from Lemma 2.7 that the left-hand side of (16) is equal to the right-hand side of (18), so all the above inequalities must be equalities. □

3. TWO COROLLARIES

We can prove now two corollaries of the equality (15). This shows that this equality already contains substantial content – that is the reason we called (15) and the following corollaries the "second ingredient" in the overview in §1.1.

The first corollary is in particular interesting and can be referred to as "the two ingredients joinings theorem": it allows for a complete classification of joinings in some situations, e.g., for Example 2.1(2)). Also, it (or more precisely, Remark 3.2) will be used again and again in various stages of the proof of Theorem 4.2 and in §5. Moreover, this corollary and its proof explain what we are looking for (utilizing entropy methods to gain an unipotent invariance) and also what we cannot hope for (finding unipotent invariance which is trivial in one of the factors). We give more details below.

3.1. Disjointness due to different Lyapunov weights

It may happen that the root structure of G_1 and G_2 are different enough, so that for some Lyapunov weight λ, one of the coarse Lyapunov subgroups $G_1^{[\lambda]}$ or $G_2^{[\lambda]}$ is trivial. This immediately "decouples" the joining:

COROLLARY 3.1 (Two ingredients joinings theorem). — *With the setting of §2.1 let μ be a joining and assume that there is a Lyapunov weight λ such that (say) $G^{[\lambda]} = G_1^{[\lambda]} \times \{e\}$. Then μ is the trivial joining.*

Proof. — For the coarse Lyapunov weight $[\lambda]$ we have $G_2^{[\lambda]} = \{e\}$ and since the entropy contribution of the trivial group is zero, (15) reads as follows

$$h_\mu(a, G^{[\lambda]}) = h_{m_{X_1}}(a_1, G_1^{[\lambda]}) + h_\mu\left(a, \{e\} \times G_2^{[\lambda]}\right) = h_{m_{X_1}}(a_1, G_1^{[\lambda]}). \qquad (19)$$

But

$$h_{m_{X_1}}\left(a_1, G_1^{[\lambda]}\right) = -\log\left|\det \mathrm{Ad}_{a_1}\big|_{\mathrm{Lie}\left(G_1^{[\lambda]}\right)}\right|$$

$$= -\log\left|\det \mathrm{Ad}_{a}\big|_{\mathrm{Lie}\left(G_1^{[\lambda])}\times\{e\}\right)}\right| \qquad = h_{m_X}\left(a, G_1^{[\lambda]} \times \{e\})\right),$$

so Theorem 2.5 then implies that μ is invariant under the unipotent group $G_1^{[\lambda]} \times \{e\}$. Recall that an unbounded closed subgroup such as G_1 acts ergodically on X_1 (see e.g. ZIMMER, 1984, Theorem 2.2.6). Therefore, we can find an element of the form (u, e) with u acting ergodically on X_1 that preserves μ. This immediately implies disjointness, that is, μ must be the trivial joining (see e.g. AKA, MUSSO, and WIESER, 2021, Lemma 7.2). □

Remark 3.2. — We record the following useful fact from the above proof: having invariance under an element of the form (u, e) with $\langle u \rangle$ unbounded, immediately implies disjointness.

Example 3.3. — Consider Example 2.1(2), that is, the "45°-rotation" torus action given by (ϕ, ψ). As we explained there, with this torus action, all the eight coarse Lyapunov weights have the form for which Corollary 3.1 is applicable. Therefore, in this setting, the only possible joining is the trivial joining. We remark that a *p*-adic variant of exactly this situation arises in the arithmetic application considered in Aka, Einsiedler, and Wieser (2021). See Example 6.7 for more details. It is interesting to remark that the statement of Corollary 3.1 is not explicitly visible in the statement of Theorem 1.1; it is nevertheless there: any non-trivial joining will be related to a local isomorphism between $\mathbb{G}_1$ and $\mathbb{G}_2$. Existence of such an isomorphism implies that the non-trivial weights for G_1 and G_2 must be identical, and therefore rules out the possibility of finding a coarse Lyapunov subgroup for which Corollary 3.1 is applicable.

The proof of Corollary 3.1 hints on how we aim to proceed in general: we aim to utilize entropy considerations to establish that a certain coarse Lyapunov subgroup (which is automatically an unipotent subgroup) has maximal entropy contribution and therefore leaves the joining invariant. Once we establish an unipotent invariance, we can utilize measure rigidity results for unipotent flows as shortly described in §2.1. But in general, this unipotent need not be trivial in one of the factors. here are two examples that the reader should keep in mind:

Example 3.4. — When we consider possible joinings for with $G_1 = G_2 = G$ and $\Gamma_1 = \Gamma_2 = \Gamma$, and with identical action, we always have the so-called "diagonal-joining": the push-forward of $m_{\Gamma \backslash G}$ under $\iota \colon G \to G \times G$, $\iota(g) = (g, g)$. This fits with the fact that Corollary 3.1 is not applicable, as the weights for G_1 and G_2 are identical so all the (coarse) Lyapunov subgroups a product of *non-trivial* weight spaces. More generally, if $\Psi \colon \mathbb{G}_1 \to \mathbb{G}_2$ is a $\mathbb{Q}$-algebraic isomorphism, then the push-forward m_{X_1} under $\iota \colon G_1 \to G_1 \times G_2$, $g \mapsto (g, \Psi(g))$ is also a non-trivial joining which is supported on the graph of the isomorphism Ψ. In the case where G_1 and G_2 are simple $\mathbb{Q}$-groups, Theorem 1.1 implies that such joinings are the only possible non-trivial joinings. For a concrete example, see Example 2.2 with $\Gamma_1 = \Gamma_2$. We will prove a classification of all (ergodic) joinings for Example 2.2 in §4.

On the other hand, it may well happen that all the coarse Lyapunov subgroups are product of non-trivial weight spaces, but nevertheless the only possible joining is the trivial one:

Example 3.5. — Consider Example 2.1(3), i.e. the "different speeds" action (ϕ, τ). Here, there are also only four weights and all the *coarse* Lyapunov subgroups are a product of non-trivial weight spaces. Nevertheless, it follows from Theorem 1.1 that ergodic (ϕ, τ)-invariant joinings must be trivial: any non-trivial joining must arise from an algebraic

automorphism Ψ of $\mathbb{G}_1 = \mathrm{SL}_2 \times \mathrm{SL}_2$ such that the image $\mathbb{Z}^2$ under (ϕ, τ) belongs to $\{(g, \Psi(g) : g \in G)\}$. But such an automorphism does not exist. It is very interesting to see how this arises in the proof, especially when one compares it to Example 2.1(1): we discuss this in §5.1.2 and §5.1.3.

3.2. Support of the leafwise measures

Our ultimate goal is to show that a joining μ admits some unipotent invariance, for example, under one of the coarse Lyapunov subgroups $G^{[\Lambda]}$, or a non-trivial subgroup of it. By Theorem 2.5 this is equivalent to showing that $G^{[\Lambda]}$ (or a non-trivial subgroup of it) has maximal entropy contribution, and by Proposition 2.3(3) this is also equivalent to showing that $\mu_x^{[\Lambda]}$ equals the Haar measure $m_{G^{[\Lambda]}}$ almost everywhere (or the Haar measure on a non-trivial subgroup of it). In this subsection, we show that a much weaker statement follows from (15), namely that the support of $\mu_x^{[\Lambda]}$ is large (at least once projected, see below) almost everywhere. In the next sections, we will try to use further methods/ingredients (the low and high entropy methods) to upgrade this partial information about the support to $\mu_x^{[\Lambda]}$ to a statement that will imply invariance under $U^{[\Lambda]}$ (or one of its non-trivial subgroups).

In general, we cannot expect $\mu_x^{[\Lambda]}$ to have full support. For instance, in the "identical weights" Example 2.1(1), if μ is the diagonal joining induced by $\iota(g) = (g, g)$, then $\mu_x^{[\Lambda]}$ is supported on

$$\left\{ \iota(u) : u \in G_1^{[\Lambda]} \right\} \subset G_1^{[\Lambda]} \times G_2^{[\Lambda]}.$$

But in general, we might hope (since having a joining should imply this fact) that the support will be large once projected to each of the factors. This is indeed the case: recall that by Proposition 2.8 we have equality in (14) for any coarse Lyapunov subgroup. This has the following consequence.

COROLLARY 3.6. — *In the setting of Lemma 2.7 assume that equality holds in* (14) *for* $U = U_1 \times U_2 < G$. *Then,* U_1 *is the smallest connected* a_1-*normalized subgroup of* G_1 *containing* $\pi_1 \left(\mathrm{supp}\, \mu_x^U \right)$ *for almost every* $x \in X$.

Proof. — Assume for contradiction that the smallest connected a_1-normalized subgroup of G_1 containing $\pi_1 \left(\mathrm{supp}\, \mu_x^U \right)$ for almost every $x \in X$ is $V_1 \lneq U_1$. We claim that the following strict inequality holds:

$$\begin{aligned} h_\mu \left(a, U_1 \times U_2 \right) &= h_\mu \left(a, V_1 \times U_2 \right) \\ &\leq h_{m_{X_1}} \left(a_1, V_1 \right) + h_\mu(a, \{e\} \times U_2) \\ &< h_{m_{X_1}} \left(a_1, U_1 \right) + h_\mu(a, \{e\} \times U_2). \end{aligned}$$

Indeed, the first equality follows from the definition of the entropy contribution (8) and (9) (strictly speaking, one needs to verify that for $U' < U$ we have $\left(\mu_x^U \right)|_{U'} = \mu_x^{U'}$, i.e. a leafwise measure variant of the double conditioning formula for conditional measures (see e.g. EINSIEDLER and WARD, 2011, Prop. 5.20). Now, the first inequality follows from (14) applied to $(V_1 \times \{e\}) \, (\{e\} \times U_2)$. For the last strict inequality, note first that

$V_1 \lneq U_1 < (G_1)_{a_1}^-$, so all the directions in $G_{a_1}^-$ are being contracted at least by some factor < 1. Therefore, we have the following strict inequality for the contribution with respect to the Haar measures (see (10)):

$$h_{m_{X_1}}(a_1, V_1) = -\log\left|\det \mathrm{Ad}_{a_1}|_{\mathrm{Lie}(V_1)}\right| < -\log\left|\det \mathrm{Ad}_{a_1}|_{\mathrm{Lie}(U_1)}\right| = h_{m_{X_1}}(a_1, U_1).$$

This implies the strict inequality above. But having a strict inequality is a contradiction to our assumption that (14) holds for U. $\qquad\square$

COROLLARY 3.7. — *In the setting of §2.1 let μ be an ergodic $A = \phi(\mathbb{Z}^d)$-invariant joining, and let $U = U_1 \times U_2$ be a coarse Lyapunov weight. Let P_x be the smallest connected A-normalized subgroup of U containing $\operatorname{supp}\mu_x^U$. Then, there is set $X' \subset X$ with $\mu(X) = \mu(X')$ and a subgroup $P < U$ with $\pi_i(P) = U_i$ for $i = 1, 2$ such that $P = P_x$ for every $x \in X'$.*

Proof. — According to Corollary 3.6 we have $\pi_i(P_x) = U_i$ for $i = 1, 2$ for almost all $x \in X$, so we just have to show that P_x is almost everywhere constant. To this end, first we note that $x \mapsto P_x$ is a measurable map (which is not at all obvious: one checks this using the characterizing property in Prop 2.3, where the set of (connected) subgroups of U is equipped with the topology on the Grassmannian of $\mathrm{Lie}(U)$). Moreover, since P_x is A-normalized, it follows from (7) that for every $a \in A$ and almost every x we have $P_{a \cdot x} = P_x$. By the ergodicity assumption, the A-invariant measurable map $x \mapsto P_x$ must be then almost everywhere constant, as we wanted to show. $\qquad\square$

The last corollary could be loosely interpreted, by saying that the leafwise measures of coarse Lyapunov weights of a joining behave like the joining: they project onto each of their natural factors. In particular, we see that leafwise measures of coarse Lyapunov weights have "interesting" support. In the next two sections (depending on the context, and on if different coarse Lyapunov weights commutes or not) we will see how one could use the non-trivial support of these measures to generate an unipotent invariance.

4. THE HIGH RANK CASE

When we have two non-commuting coarse Lyapunov weights, the following theorem (EINSIEDLER and KATOK, 2005), coupled with the input we get from Corollary 3.7, finds an unipotent invariance for us.

THEOREM 4.1 (The high entropy method). — *In the setting of Section 2.1, let μ be an A-invariant ergodic probability measure and fix $a \in A$. Then, there exist two connected and A-normalized subgroups $H \trianglelefteq P \leq G_a^-$ such that the following holds*

1. *For almost every $x \in X$, P is the minimal A-normalized subgroup supporting $\mu_x^{G_a^-}$.*
2. *For almost every $x \in X$, $\mu_x^{G_a^-}$ is left- and right-invariant under H.*

3. *If $[\lambda] \neq [\eta]$ are two different coarse Lyapunov weights with $G^{[\lambda]}, G^{[\eta]} \leq G_a^-$ and $g_1 \in P \cap G^{[\lambda]}$, $g_2 \in P \cap G^{[\eta]}$, then $[g_1, g_2] \in H$, that is, $g_1 H$ and $g_2 H$ commute with each other in P/H.*

4. *For every coarse Lyapunov weight $[\lambda]$ with $G^{[\lambda]} < G_a^-$ and for almost every $x \in X$, $\mu_x^{G^{[\lambda]}}$ is left- and right-invariant under $H \cap G^{[\lambda]}$.*

We referred to this theorem as the "third ingredient" in §1.1. We will explain now how the three ingredients we named there allow us to classify joinings when $\mathbb{G}_1$ and $\mathbb{G}_2$ are semisimple without rank one factors. As the main steps of the proof are already visible when one considers the case $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_3$, we will restrict ourselves to this case for simplicity and prove:

THEOREM 4.2. — *In the setting of Example 2.2 all ergodic joinings are algebraic. More precisely, the only possible ergodic joinings are the trivial one and joinings which are the Haar measure on an orbit of a diagonal embedding of SL_3 into SL_3^2, that is, an orbit of a group of the form $\{(g, \Psi(g)) : g \in \mathrm{SL}_3(\mathbb{R})\} < \mathrm{SL}_3^2(\mathbb{R})$, for an automorphism $\Psi \colon \mathrm{SL}_3 \to \mathrm{SL}_3$.*

4.1. Proof of Theorem 4.2

Recall that $G_i = \mathbb{G}_i(\mathbb{R})$, $i = 1, 2$ which is $\mathrm{SL}_3(\mathbb{R})$ in our case and that $G = G_1 \times G_2$. Recall that μ denotes the joining we are trying to identify, and we begin by considering the subgroup $I < G$ which is generated by all one-parameter unipotent subgroups that leave the measure μ invariant. If we knew that μ was algebraic (in the sense of Theorem 4.2), I would have projected onto each factor. Reversing the order, we begin by establishing the latter as a first step towards the algebraicity of μ:

LEMMA 4.3. — *Let $\mathfrak{i}$ be the Lie algebra of I. Then, for $i = 1, 2$ we have $\pi_i(\mathfrak{i}) = \mathfrak{g}_i$ (as above $\mathfrak{g}_i = \mathfrak{sl}_3$ denotes the Lie algebra of G_i).*

Proof. — As we explain in Example 2.2, we have six non-trivial coarse Lyapunov groups $\mathfrak{g}^{ij} = \mathfrak{u}_{ij} \times \mathfrak{u}_{ij}$. Consider, for example, $\mathfrak{g}^{13}$. We would like to write $\mathfrak{g}^{13}$ as a commutator of two other coarse Lyapunov groups, which appear together with $\mathfrak{g}^{13}$ in some stable horospherical subgroup. More precisely, recall that $A = \phi(\mathbb{Z}^2)$ is a subgroup of the diagonal group and pick $a \in A$ with $\mathfrak{g}_a^- = \mathrm{Lie}(G_a^-) = \mathfrak{g}^{12} \oplus \mathfrak{g}^{13} \oplus \mathfrak{g}^{23}$ (the Lie algebra of the Heisenberg group). For example, one can pick $a = \phi(1, 2)$ with the notation of Example 2.2. Note that $[\mathfrak{g}^{12}, \mathfrak{g}^{23}] = \mathfrak{g}^{13}$. We apply the high entropy Theorem (Theorem 4.1) with these choices and let P and H the subgroups appearing in that theorem. We know that $H \cap G^{13}$ contains all commutators $[g_1, g_2]$ with $g_1 \in P \cap G^{12}$ and $g_2 \in P \cap G^{23}$. Moreover, we know from Corollary 3.7 that $P \cap G^{12}$ must project onto G_i^{12}, $i = 1, 2$, using the notation $G^{ij} = G_1^{ij} \times G_2^{ij}$. The same holds for the projection of $P \cap G^{23}$ on the corresponding two factors. As $[G_i^{12}, G_i^{23}] = G_i^{13}$ we see that H must project onto G_i^{13} for $i = 1, 2$. Going back to the Lie algebra, this means, that for $i = 1, 2$ and $w \in \mathfrak{g}_i^{13} = \mathrm{Lie}(G_i^{13}) = \mathfrak{u}_{13}$, $\mathfrak{i}$ must contain an element whose i-component is w.

As we explain in Example 2.2, each coarse Lyapunov space can be written as commutators of other Lyapunov spaces appearing in the decomposition of some stable horospherical subgroup. Therefore, the above argument can be applied to each of the coarse Lyapunov weight spaces. As $\mathfrak{g}_i = \mathfrak{sl}_3$, $i = 1, 2$ is generated by $\mathfrak{u}_{ij}$, $i \neq j$, the lemma follows. $\qquad\square$

Loosely speaking, we already see the possible joinings only by looking at the possibilities for $H \cap G^{ij}$: it is either two-dimensional and equals G^{ij} itself, or it is one-dimensional and equal to a graph of an isomorphism from U^{ij} to itself. We note just for fun, that if $H \cap G^{ij}$ is two-dimensional for some $i \neq j$, μ will be invariant under an element of the form (u, e) or (e, u) with $u \in G_i$ acting ergodically on the corresponding quotient G_i; in other words, we will be again in the situation of the proof of Corollary 3.1 which shows that in this case, μ must be the trivial joining. For the general case, we will need the following Lemma.

LEMMA 4.4. — *The Lie Algebra* $\mathfrak{i} \cap (\mathfrak{g}_1 \times \{0\})$ *is a Lie ideal in* $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2$ *(and therefore also in* $\mathfrak{g}_1 \times \{0\} \cong \mathfrak{g}_1$*). The same holds for* $\mathfrak{i} \cap (\{0\} \times \mathfrak{g}_2)$.

Proof. — Consider $(w, 0) \in \mathfrak{i} \cap (\mathfrak{g}_1 \times \{0\})$ and $(v_1, v_2) \in \mathfrak{g}$. Using Lemma 4.3 pick an element in $\mathfrak{i}$ having the same first component as (v_1, v_2), say $(v_1, v_2') \in \mathfrak{i}$. Then,

$$\mathfrak{g}_1 \times \{0\} \ni ([w, v_1], 0) = [(w, 0), (v_1, v_2)] = [(w, 0), (v_1, v_2')] \in \mathfrak{i},$$

as we wanted to show. $\qquad\square$

Proof of Theorem 4.2. — In our case $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{sl}_3$ is a simple Lie algebra, so the Lie ideals in Lemma 4.4 are either trivial or equal to $\mathfrak{sl}_3$. If one of them is equal to $\mathfrak{sl}_3$, we can find a non-trivial element of the form (u, e) or (e, u) with u unipotent, as in the proof of Corollary 3.1, which shows that μ must be the trivial joining. So we may assume now that both ideals are trivial. By the folklore Goursat's Lemma, $\mathfrak{i}$ must be the graph of an isomorphism, that is,

$$\mathfrak{i} = \{(w, \Phi(w) : w \in \mathfrak{g}_1\}$$

for an *isomorphism* $\Phi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$. It follows that

$$I = \{(g, \Psi(g) : g \in G_1\}$$

for an *isomorphism* $\Psi\colon G_1 \to G_2$. This isomorphism must intertwine the diagonal action $\phi = (\phi_1, \phi_2)$, that is, $\phi(s, t) := (\phi_1(s, t), \phi_2(s, t))$ must agree with $(\phi_1(s, t), \Psi(\phi_1(s, t)))$. Indeed, if not, μ would be invariant under two elements having the same first coordinate but a different second coordinate, so we could again find an element of the form (e, a) with a acting ergodically on X_1, preserving μ. This puts us again in the situation of the proof of Corollary 3.1. It follows that $A \subset I$, and therefore μ must be ergodic with respect to the action of I, since μ was ergodic with respect to the action of A. Ratner's measure classification can be then applied with the group I showing that μ is the Haar measure on an orbit of I, which concludes the proof. $\qquad\square$

Remark 4.5. — We actually get more information about the possible joinings from the above proof. First, the automorphism Ψ must intertwine the action ϕ. Second, if there exists a non-trivial joining, Ratner's Theorem will imply that the lattices Γ_1 and Γ_2 must be commensurable.

The proof of the general theorem, say when $\mathbb{G}_1$ and $\mathbb{G}_2$ are both semisimple of rank ≥ 2, is very similar. One essentially replaces the ad-hoc argument for $\mathfrak{sl}_3$ in the beginning of Lemma 4.3 with a general argument that works for any semisimple Lie algebra without rank one factors; see EINSIEDLER and LINDENSTRAUSS (2007, Lemma 4.2).

5. JOININGS WITH FORMS OF SL_2

In this section we wish to classify the possible ergodic joinings in Example 2.1(1), the "identical weights" example, and Example 2.1(3), the "different speeds" example. Note that we already used Corollary 3.1 to show that all the possible joinings in Example 2.1(2) are trivial. These sub-examples of Example 2.1 may seem very specific, but they are actually very representative examples. Conceptually, classifying joinings when $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_2 \times \mathrm{SL}_2$ is the main new case, which is included in Theorem 1.1, but not included in the results of EINSIEDLER and LINDENSTRAUSS (2007). Moreover, the above three structures of the torus action cover all possible types of joined actions. Furthermore, an analogous full classification of joinings in a general S-adic version of Example 2.1, for products of SL_2-forms, will suffice for most of the applications discussed in section 6.

As we explained in §1.1, the new ingredient used in EINSIEDLER and LINDENSTRAUSS (2019) in comparison to EINSIEDLER and LINDENSTRAUSS (2007) is the low-entropy method. This method was originally developed by LINDENSTRAUSS (2006). We wish to follow the main steps of §7 of this paper (which are also nicely explained in EINSIEDLER and LINDENSTRAUSS, 2010, §10) in order to sketch a proof of the crucial step, Proposition 5.1, in classifying the possible joinings in the two examples listed above. In some sense, we "merely" run the same argument, that Lindentrauss ran in one factor of the form $\Gamma \backslash (\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}))$, in two such factors simultaneously. Our torus action is embedded "diagonally", that is, it acts on both factors simultaneously. This gives rise to several complications (and sometimes to cumbersome notation) that must be taken into account.

Let us recall and set up some notation: we let $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_2 \times \mathrm{SL}_2$, $G_i = \mathbb{G}_i(\mathbb{R})$, $i = 1, 2$ and for simplicity let $\Gamma_1 = \Gamma_2$ be an irreducible lattice in $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ (e.g., $\mathrm{SL}_2(\mathbb{Z}[\sqrt{2}])$ diagonally embedded). The letter d denotes a left-invariant metric on G, or on G_i, or on $\mathrm{SL}_2(\mathbb{R})$, depending on the context. Let $X_i = \Gamma_i \backslash G_i$ and set $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$, $G = G_1 \times G_2$, and $X = X_1 \times X_2$. Note that G (or $\mathbb{G}$) have four factors, first two in G_1 and last two in G_2. We refer to them as the first, second, third, and fourth factor accordingly. Recall the notation of $\phi, \tau \colon \mathbb{Z}^2 \to G_i$ given by $\phi(t, s) = (a_t, a_s)$

and $\tau(t, s) = (a_{2t}, a_{2s})$. We will concentrate on the map $\Phi_1 = (\phi, \phi)\colon \mathbb{Z}^2 \to G$, i.e, on Example 2.1(1), remarking, where needed, what would have changed if we had considered the map $\Phi_2 = (\phi, \tau)\colon \mathbb{Z}^2 \to G$ corresponding to Example 2.1(3). By abuse of notation, we denote $A = \Phi_1(\mathbb{Z}^2)$ or $A = \Phi_2(\mathbb{Z}^2)$ in both cases.

We recall from Example 2.1 that for Φ_1 there are four weight spaces, each giving rise to a different coarse Lyapunov subgroup, and for Φ_2 there are 8 weight spaces, coupled through coarse equivalence into pairs, which give rise to the same four Lyapunov subgroups. We denote by $[\alpha]$ the coarse Lyapunov weight with $w_\alpha = (-1, 0)$ and by $[\beta]$ be the coarse Lyapunov weight with $w_\beta = (0, -1)$. With U^+ (resp. U^-) denoting the upper (resp. lower) triangular unipotent subgroups in $\mathrm{SL}_2(\mathbb{R})$, the corresponding coarse Lyapunov subgroups in both cases are

$$U^{[\alpha]} = \left(U^+ \times \{e\}\right) \times \left(U^+ \times \{e\}\right)$$
$$U^{[\beta]} = \left(\{e\} \times U^+\right) \times \left(\{e\} \times U^+\right)$$
$$U^{[-\alpha]} = \left(U^- \times \{e\}\right) \times \left(U^- \times \{e\}\right)$$
$$U^{[-\beta]} = \left(\{e\} \times U^-\right) \times \left(\{e\} \times U^-\right).$$

We denote by μ the A-invariant joining on X that we wish to identify. Recall that $\mu_x^{[\alpha]}, \mu_x^{[\beta]}, \mu_x^{[-\alpha]}, \mu_x^{[-\beta]}$ denote the leafwise measures corresponding to the coarse Lyapunov subgroup as above. For each of the coarse Lyapunov subgroups (i.e., for $\lambda \in \{\pm\alpha, \pm\beta\}$) let

$$I_x^{[\lambda]} = \left\{u \in U^{[\lambda]} : u \text{ preserves the measure } \mu_x^{[\lambda]}\right\}. \tag{20}$$

The main goal of this section is to outline a proof of the following proposition:

PROPOSITION 5.1 (Main step). — *Let μ be an ergodic joining on X, that is, let μ be an A-invariant ergodic measure on X which projects to the Haar measure on X_i in each factor. For any coarse Lyapunov weight $[\lambda]$, and for μ-almost every x, $I_x^{[\lambda]}$ is not the trivial subgroup.*

We will sketch a proof of this proposition in §5.2.

5.1. Classification under the assumption of Proposition 5.1

5.1.1. *Preparation: coordinates on $U^{[\alpha]}$.* — Throughout this section we will concentrate on $\lambda = \alpha$. First note, that via

$$\mathrm{Lie}(U^{[\alpha]}) \cong \mathbb{R}^2, \ (n_x, 0, n_y, 0) \mapsto (x, y), \ n_x := \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \tag{21}$$

and the fact that $\exp\colon \mathrm{Lie}(U^{[\alpha]}) \to U^{[\alpha]}$ is bijective, we can identify $\mathrm{Lie}(U^{[\alpha]})$ or $U^{[\alpha]}$ with $\mathbb{R}^2$. For $u \in U^{[\alpha]}$ we define $u(s_1)$ and $u(s_2)$ as the real numbers satisfying

$$u = \left(\begin{pmatrix} 1 & u(s_1) \\ 0 & 1 \end{pmatrix}, e, \begin{pmatrix} 1 & u(s_2) \\ 0 & 1 \end{pmatrix}, e\right). \tag{22}$$

So we can define an element $u \in U^{[\alpha]}$ by specifying $u(s_1)$ and $u(s_2)$ and refer to these as the coordinates of u. Any non-trivial $u \in U^{[\alpha]}$ is contained in a unique one-parameter unipotent subgroup $L_u \subset U^{[\alpha]}$, defined by the line in $\mathrm{Lie}(U^{[\alpha]})$ through $\log(u)$. The x-axis corresponds to $U^+ \times \{e\} \times \{e\} \times \{e\}$, and the y-axis to $\{e\} \times \{e\} \times U^+ \times \{e\}$. By Remark 3.2, showing invariance of a joining μ under a non-trivial element lying on an axis, immediately implies that μ is the trivial joining.

Identifying $U^{[\alpha]}$ with $\mathbb{R}^2$ as above, let's study the action of $A = \Phi_i(\mathbb{Z}^2)$ on $U^{[\alpha]}$ for $i = 1, 2$ by conjugation. For Φ_1, the element $\Phi_1(1, 0)$ acts as scalar multiplication by $e^{-1} = (2.718\ldots)^{-1}$ and $\Phi_1(0, 1)$ acts trivially. For Φ_2, the element $\Phi_2(1, 0)$ acts on $U^{[\alpha]}$ by multiplication with $\begin{pmatrix} e^{-1} & 0 \\ 0 & e^{-2} \end{pmatrix}$, and $\Phi_2(0, 1)$ acts trivially.

5.1.2. *Classification in Example 2.1(1).* —

THEOREM 5.2. — *Any ergodic $\Phi_1(\mathbb{Z}^2)$-invariant joining μ is either the trivial joining or a diagonal joining, that is, a joining supported on a graph of an isomorphism of G_1 with G_2 (as is Example 3.4).*

Before proving this theorem, let's first analyze $I_x^{[\lambda]}$ for $\lambda \in \{\pm\alpha, \pm\beta\}$. One can show that the map $x \mapsto I_x^{[\lambda]}$ is measurable, and by (7) we also have that

$$\forall a \in A, \quad I_{a.x}^{[\lambda]} = a I_x^{[\lambda]} a^{-1}. \tag{23}$$

From these two facts we have:

LEMMA 5.3. — *For any $\lambda \in \{\pm\alpha, \pm\beta\}$, the group $I_x^{[\lambda]}$ must be almost everywhere constant.*

Proof. — For concreteness, consider $U^{[\alpha]}$. Recall that the element $\Phi_1(1, 0)$ acts as scalar multiplication by $e^{-1} = (2.718\ldots)^{-1}$ and $\Phi_1(0, 1)$ act trivially. Using Poincaré recurrence as in LINDENSTRAUSS (2006, Lemma 7.3), or with an argument similar to the proof of Lemma 5.6 below, one can show that for almost any $x \in X$, if $e \neq u \in I_x^{[\lambda]}$ then $L_u \subset I_x^{[\lambda]}$. Therefore, any subgroup of $U^{[\alpha]}$ is A-normalized. In particular, $I_x^{[\alpha]}$ is A-normalized μ-almost everywhere. Since $A = \Phi_1(\mathbb{Z}^2)$ is assumed to act ergodically on X with respect to μ, this means that $I_x^{[\alpha]}$ is almost everywhere constant. Such an argument holds for any other $\lambda \in \{\pm\alpha, \pm\beta\}$. $\qquad \square$

We denote this constant group by $I^{[\lambda]}$ and call it the invariance group for $[\lambda]$. From Item 3 of Proposition 2.3, we have:

COROLLARY 5.4. — *For any $\lambda \in \{\pm\alpha, \pm\beta\}$, μ is invariant under $I^{[\lambda]}$.*

This gives us enough information to analyse the ergodic joinings in Example 2.1(1).

Proof of Theorem 5.2. — First observe, that each of the coarse Lyapunov subgroups is two-dimensional. By Proposition 5.1 each of the invariance groups is non-trivial. If one of them is two-dimensional, we have found an element for which Remark 3.2 applies, so μ must be the trivial joining.

As we explained above, we know that if $e \neq u \in I^{[\alpha]}$, $L_u \subset I^{[\alpha]}$. Since any one-parameter subgroup of $U^{[\alpha]}$ is $\Phi(1,0)$-normalized (equivalently A-normalized), it might as well be that $I^{[\alpha]}$, and similarly all the other invariance subgroups $I^{[\lambda]}$, are one-dimensional. If one of them is supported on the axes, we will know that μ is trivial, again by applying Remark 3.2. When none of them are, then each of them projects onto both of its factors, so the group generated by them must project onto $\mathrm{SL}_2(\mathbb{R})$ in each of the four factors (as $\mathrm{SL}_2(\mathbb{R})$ is generated U^+ and U^-). We implicitly apply Goursat's Lemma: these invariance subgroups might be so compatible with each other, so the group generated by them will be the graph of an automorphism of G_1 into G_2, giving rise to a diagonal joining. Otherwise, there will be an element for which Remark 3.2 applies, resulting in the trivial joining. So this proves (modulo Proposition 5.1) everything we wanted to know about Example 2.1(1). □

5.1.3. *Classification in Example 2.1(3).* —

THEOREM 5.5. — *Any $\Phi_2(\mathbb{Z}^2)$-invariant joining μ must be the trivial joining.*

To prove this theorem, it is enough to show the following:

LEMMA 5.6. — *The group $I^{[\alpha]}$ must contain the group corresponding to the x-axis in the identification (21), that is the group $U^+ \times \{e\} \times \{e\} \times \{e\}$.*

A similar statement (and proof) will hold for any $\lambda \in \{\pm\alpha, \pm\beta\}$; it is nonetheless enough to concentrate on $\lambda = \alpha$.

Proof of Lemma 5.6. — The main difference to Example 2.1(1), is that not every subgroup of $I_x^{[\alpha]}$ is $\Phi_2(\mathbb{Z}^2)$-normalized. Recall that the element $\Phi_2(1,0)$ acts on $U^{[\alpha]}$ by multiplication with $\begin{pmatrix} e^{-1} & 0 \\ 0 & e^{-2} \end{pmatrix}$, and $\Phi_2(0,1)$ acts trivially. Dynamically, we see that $\Phi_2(1,0)^n$ pushes elements not belonging to the axes, towards the x-axis. We claim that this, coupled with Poincaré recurrence, proves the Lemma, by an argument inspired by the work of EINSIEDLER and KATOK (2005, p. 206-207): let $\epsilon > 0$ and use Lusin's Theorem (FEDERER, 1969, p. 76) to find a set K_ϵ of measure $1 - \epsilon$ on which $x \mapsto I_x^{[\alpha]}$ is continuous. Denote $a = \Phi_2(1,0)$ and let $x \in K_\epsilon$, and $e \neq u \in I_x^{[\alpha]}$ denote the element we get from Proposition 5.1. Poincaré recurrence finds for us, for almost any $x \in K_\epsilon$, a subsequence

$$a^{n_k} x \to x \text{ with } a^{n_k}.x \in K_\epsilon. \tag{24}$$

It follows from (23) that the group $I_{a^{n_k}.x}^{[\alpha]}$ contains the image of $\langle u \rangle$, under the action of a^{n_k}, that is $\langle a^{n_k} u a^{-n_k} \rangle$, which is a "squashing" of $\langle u \rangle$ towards the x-axis (here, $\langle u \rangle$ is the group generated by u). By (24), $I_x^{[\alpha]}$ contains all elements in $U^{[\alpha]}$ which are limits

of $\langle a^{n_k} u a^{-n_k} \rangle$; that is, the entire group corresponding to the x-axis. Taking $\epsilon \to 0$ we establish the above for x in a conull set of X, as we wanted to show. $\qquad\square$

Proof of Theorem 5.5 . — As in §5.1.2, we will have that μ will be invariant under $I^{[\alpha]}$. Lemma 5.6 says that $I^{[\alpha]}$ contains elements for which Remark 3.2 applies, showing that μ must be trivial. $\qquad\square$

5.2. Proof of Proposition 5.1

5.2.1. *A small reformulation.* —

LEMMA 5.7. — *Proposition 5.1 is equivalent to establishing that for any $\lambda \in \{\pm\alpha, \pm\beta\}$ and almost any $x \in X$,*

$$\exists e \neq u \in U^{[\lambda]}, \quad (R_u)_* \mu_x^{[\lambda]} \propto \mu_x^{[\lambda]}. \tag{25}$$

Proof. — This follows again from Poincaré recurrence and can be proven very similarly to LINDENSTRAUSS (2006, Lemma 7.3) or using a similar argument to the one in the proof of Lemma 5.6. $\qquad\square$

Note that if we find two points $x, y \in X$ with $\mu_x^{[\lambda]} = \mu_y^{[\lambda]}$ and $x = u.y$ with $e \neq u \in U^{[\alpha]}$, that is, two points with identical leafwise measure on the same $U^{[\lambda]}$-leaf, we get from (6) that

$$(R_u)_* \, \mu_x^{[\lambda]} = (R_u)_* \, \mu_{u.y}^{[\lambda]} \propto \mu_y^{[\lambda]} = \mu_x^{[\lambda]} \tag{26}$$

So x will satisfy (25) with $e \neq u \in U^{[\alpha]}$. This led Lindenstrauss to the following idea.

5.2.2. *An optimistic idea: using Ratner's H-principle.* — For concreteness, assume that $\lambda = \alpha$ and consider Example 2.1(1)(that is, with $A = \Phi_1(\mathbb{Z}^2)$). Using Poincaré recurrence for the action of $\Phi_1(0, 1)$, which commutes with $U^{[\alpha]}$, we can find many pairs of nearby points $x, y \in X$, with $\mu_x^{[\alpha]} = \mu_y^{[\alpha]}$, and arbitrarily small displacement g, but not necessarily belonging to the same $U^{[\alpha]}$-leaf. That is, the displacement g does not necessarily belong to $U^{[\alpha]}$ (see Lemma 5.8 below for a more precise statement). Then, in order to achieve a pair of points with the same leafwise measure and on the same $U^{[\alpha]}$-leaf, Lindenstrauss had the following, a priori extremely optimistic, idea to use the H-principle of Ratner as follows: First note, that by (6), shearing two points x and y with $\mu_x^{[\alpha]} = \mu_y^{[\alpha]}$ with the *same* $u \in U^{[\alpha]}$ preserves the equality of the leafwise measures. That is, for any $u \in U^{[\alpha]}$

$$\mu_x^{[\alpha]} = \mu_y^{[\alpha]} \implies \mu_{u.x}^{[\alpha]} = \mu_{u.y}^{[\alpha]}. \tag{27}$$

To see where this is going, think for a moment about $\mathrm{SL}_2(\mathbb{R})$ and U^+, the upper unipotent subgroup. Ratner's H-principle tells us that shearing along U^+, a pair of nearby points in general position with respect to each other, for the right amount of time, will give us two sheared points that differ, non-trivially, only in the U^+ direction, up to a small error, which tends to zero with the initial displacement (We will formulate this more precisely for the case at hand in Lemma 5.9). Doing this with pairs having displacement tending to e, and taking a limit of the sheared pairs, we find a pair of points, differing non-trivially, exactly in the U^+-direction, i.e., belonging to the same U^+-leaf. This

seminal observation of Ratner essentially amounts to a careful analysis of the matrix multiplication in (31).

So if we pretend that the map $x \mapsto \mu_x^{[\alpha]}$ is continuous, and find many nearby pairs of points (x_n, y_n) in general position with $\mu_{x_n}^{[\alpha]} = \mu_{y_n}^{[\alpha]}$, the H-principle will give us two points with identical leafwise measure on the same $U^{[\alpha]}$-leaf.

But there is one caveat: the measurable map $x \mapsto \mu_x^{[\alpha]}$ is a priori not continuous. We can try to use Lusin's Theorem to find a compact set K of arbitrarily large measure on which $x \mapsto \mu_x^{[\alpha]}$ is continuous. But then, trying to restrict the argument to K stirs up many serious complications. Seeing them, the author of this survey would have turned back and given up. As we will outline below, LINDENSTRAUSS (2006) chose to face them and dealt with each of them with an astonishing mastery.

5.2.3. *Preparations for the H-principle.* — Before we follow Lindenstrauss' footsteps to face all the difficulties that arise, let us write, in our "joined" setting, the information we have for running the H-principle of Ratner.

LEMMA 5.8 (Input to the H-machine). — *Let X' be a set of positive measure and consider a sequence $\delta_n \to 0$. We can find two sequences $(x_n)_{n=1}^{\infty}, (y_n)_{n=1}^{\infty} \subset X'$ with*

1. $g_n.x_n = y_n$ *and* $d(g_n, e) < \delta_n$,
2. $g_n = (g_n^{(1)}, g_n^{(2)}, g_n^{(3)}, g_n^{(4)})$ *with both* $g_n^{(1)}, g_n^{(3)} \notin U^+$,
3. $\mu_{x_n}^{[\alpha]} = \mu_{y_n}^{[\alpha]}$ *for all* $n \in \mathbb{N}$.

Proof. — This essentially follows from Poincaré recurrence for the map $b := \Phi_1(0,1) = (e, b_1, e, b_2)$ and irreducibility of Γ_1 and Γ_2. Indeed, we can assume first without loss of generality that $\mu_x^{[\alpha]}$ is defined for any $x \in X'$. The element b preserves μ so by Poincaré recurrence we can find a sequence $n_k \to \infty$ with $b^{n_k}.x \to x$ and $b^{n_k}.x \in X'$. As b commutes with $U^{[\alpha]}$, we also have by (7), that for every $x \in X'$ and $k \in \mathbb{N}$, $\mu_x^{[\alpha]} = \mu_{b^{n_k}.x}^{[\alpha]}$. This generates for us pairs satisfying Items 1 and 3. To show Item 2, assume for contradiction that we found a pair of nearby points x, y with $g.x = y = b^r.x$ for some large $r \in \mathbb{N}$ and with $g^{(1)} \in U^+$. Note that the first component of $U^{[\alpha]}$, U^+, is contracted by $a := \Phi_1(1,0)$. By Poincaré recurrence, we can find a sequence $m_k \to \infty$ with $a^{m_k}.x \to x$. Acting on $g.x = b^r.x$ with a subsequence of a^{m_k}, we contract $g^{(1)}$ to the identity and find a pair of nearby points x, y' with

$$\left(e, g^{(2)}, *, * \right).x = (e, b_1^r, e, b_2^r).x$$

with $g^{(2)}$ small and b_1^r large, so Γ_1 must contain a non-trivial element of the form (e, h), which is a contradiction to its irreducibility. The same argument works for $g^{(3)}$, and also for Φ_2 instead of Φ_1. □

5.2.4. *Polynomial divergence - how much should we shear?* — Still ignoring the fact that $x \mapsto \mu_x^{[\alpha]}$ is not continuous, we fix some notation and explain how much we need to

shear the pairs of nearby points we get from Lemma 5.8: for $g_n = (g_n^{(1)}, g_n^{(2)}, g_n^{(3)}, g_n^{(4)})$, the displacements between x_n and y_n from Lemma 5.8, we write

$$g_n^{(1)} = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}, \; g_n^{(3)} = \begin{pmatrix} \tilde{a}_n & \tilde{b}_n \\ \tilde{c}_n & \tilde{d}_n \end{pmatrix}. \tag{28}$$

For δ_n small enough, once we shear x_n and y_n with $u_n \in U^{[\alpha]}$ we get that the displacement between $u_n.x_n$ and $u_n.y_n$ is $u_n g_n u_n^{-1}$. We would like to choose $u_n \in U^{[\alpha]}$ such that $\{u_n g_n u_n^{-1}\}_{n \in \mathbb{N}}$ will contain a non-trivial element of $U^{[\alpha]}$ as an accumulation point.

LEMMA 5.9. — *Recall the notation* (22) *and the setting and notation of Lemma 5.8. Assuming $\delta_n < \delta_0$ for δ_0 small enough, we have the following: for any $\rho \in (0,1)$ there are $C > 0$ and $(S_n)_{n=1}^{\infty}$ with $S_n \xrightarrow{\delta_n \to 0} \infty$ such that for any*

$$s_1^{(n)}, s_2^{(n)} \in T(\rho, S_n) := [\rho S_n, S_n] \cup [-S_n, -\rho S_n],$$

there exist σ_1, σ_2 with at least one of them belonging to $[\frac{1}{C}, C]$ satisfying

$$u_n.y_n = g_n' u_\sigma u_n.x_n, \quad \text{with } d(g_n', e) < \delta_n^{\frac{1}{2}}. \tag{29}$$

where $u_n, u_\sigma \in U^{[\alpha]}$ are the elements with $u_n(s_i) = s_i^{(n)}$ and $u_\sigma(s_i) = \sigma_i$ for $i = 1, 2$.

In words: for a given $\rho \in (0,1)$ we can find a "timescale" S_n such that shearing x_n and y_n with any element of $U^{[\alpha]}$ with coordinates (s_1, s_2) in the "time window" $T(\rho, S_n)$ results in two points which differ, up to a small correction g_n', by an element $u_\sigma \in U^{[\alpha]}$ belonging, at least in one of its components, to a fixed (that is, independent of n) compact set of non-trivial unipotent elements.

Proof. — We choose $C = \frac{1}{\rho}$ and choose S_n as a *minimum* as follows:

$$S_n = \min \left(S_n^{(1)}, S_n^{(2)} \right) \text{ with } S_n^{(i)} = \min \left(\frac{C}{|d_i - a_i|}, \frac{\sqrt{C}}{\sqrt{|c_i|}} \right), \; i = 1, 2. \tag{30}$$

The fact the lemma holds with these choices, goes back to Ratner and is based on the following calculation in $\mathrm{SL}_2(\mathbb{R})$: let $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right)$ and assume that g is close to the identity, that is, that $|a - 1|, |c|, |b|, |d - 1| < \delta$. We think of g as the displacement between two δ-close points x and y in general position, for a small δ. We calculate now the displacement g_s between $u.x$ and $u.y$:

$$g_s := \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a + cs & b + (d-a)s - cs^2 \\ c & d - cs \end{pmatrix}. \tag{31}$$

For the given $\rho \in (0,1)$ we define $C = \frac{1}{\rho}$ as above and $S = \min \left(\frac{C}{|d-a|}, \frac{\sqrt{C}}{\sqrt{|c|}} \right)$. We then have for any $s \in T(\rho, S)$, that $g_s \approx \left(\begin{smallmatrix} 1 & \sigma \\ 0 & 1 \end{smallmatrix} \right)$, for $\sigma \in [\frac{1}{C}, C]$, where $\approx$ stands here for an error smaller then $\delta^{1/2}$. See also LINDENSTRAUSS (2006, Lemmata 7.4&7.5).

The above calculation should be done now for both $g_n^{(1)}$ and $g_n^{(3)}$. As we choose S_n in (30) as a minimum, we can only guarantee that one of the resulting σ's will belong to the interval $[\frac{1}{C}, C]$. See also Remark 5.10 below. $\qquad \square$

We refer below to S_n, as the "*shearing timescale*" (for x_n and y_n), and to $T(\rho, S_n)$ as the "*shearing time-window*".

Remark 5.10. — We are following the footsteps of the work of LINDENSTRAUSS (2006). But this is where we start to see a slight difference: due to the fact that in our situation we shear according to *two* displacements $g_n^{(1)}, g_n^{(3)}$ and the shearing group $U^{[\alpha]}$ is two-dimensional, we are forced to choose the shearing timescale S_n as a minimum between $S_n^{(1)}$ and $S_n^{(2)}$. Note that $S_n^{(1)}$ (resp. $S_n^{(2)}$) is the time needed for the displacement $g_n^{(1)}$ (resp. $g_n^{(3)}$) to grow in the U^+-direction. It might as well be (and we cannot control this at all since the displacements are given to us from Poincaré recurrence) that the timescale S_n was not large enough for *both* g_n and g_n' to grow. Indeed, for all we know, it may be that $|c_n|$ and $|d_n - a_n|$ might be tiny in comparison to $|\tilde{c}_n|$ and $\left|\tilde{d}_n - \tilde{a}_n\right|$. This must be taken into account once we admit that $x \mapsto \mu_x^{[\alpha]}$ is not continuous.

5.2.5. *Maximal ergodic theorems.* — Let $\epsilon > 0$. As said above, by Lusin's Theorem we can find a compact set $K = K_\epsilon$ of measure $> 1 - \epsilon$, with $x \mapsto \mu_x^{[\alpha]}$ being continuous when restricted to K. Using Lemmata 5.8 and 5.9 and their notation, we see that we "just" need to make sure that for every $n \in \mathbb{N}$ we find $u_n \in U^{[\alpha]}$ with

1. $u_n(s_1)$ and $u_n(s_2)$ belonging the time window $T(\rho, S_n)$, and
2. the sheared points $u_n.x_n, u_n.y_n$ belonging to K.

If we could ensure this, we would choose accumulation points x and y of $\{u_n.x_n\}, \{u_n.y_n\}$ to find a $u \neq e$ and x and y as in (26) leading to $u \neq e$ and x satisfying (25). But how do we actually know if such u_n exist?

To discuss this further, let's denote by

$$B_S(U^{[\alpha]}) = \left\{ u \in U^{[\alpha]} : u_n(s_1), u_n(s_2) \in [-S, S] \right\}.$$

A maximal ergodic theorem will help us "measure" the subset of $B_S(U^{[\alpha]})$ whose elements satisfy Item 2. More precisely, first notice that if μ was $U^{[\alpha]}$-invariant (which is a ridiculous assumption, since if this invariance was known to us, we wouldn't be having this discussion), then we could have used a maximal ergodic theorem with respect to the action of $U^{[\alpha]} \cong \mathbb{R}^2$. Such a theorem will tell us that we can find a set $X' = X'(K_\epsilon)$ of measure $1 - C\epsilon^{1/2}$ (with C being a universal constant), such that for every $x \in X'$ and for all timescales S, we have

$$m_{U^{[\alpha]}} \left(\left\{ u \in B_S(U^{[\alpha]}) : u.x \in K \right\} \right) \geq \left(1 - \epsilon^{\frac{1}{2}} \right) m_{U^{[\alpha]}} \left(B_S(U^{[\alpha]}) \right). \qquad (32)$$

where $m_{U^{[\alpha]}}$ denotes, under the identification (21), the Borel measure on $\mathbb{R}^2$. Having this for x_n and y_n as above would definitely help to choose u_n as above.

But μ cannot be assumed to be $U^{[\alpha]}$-invariant. The only thing we know is that μ is A-invariant. As $U^{[\alpha]}$ is A-invariant, the only information connecting μ and $U^{[\alpha]}$ is that there is a $U^{[\alpha]}$-invariant foliation of the space. Lindenstrauss (together with Rudolph) proves in the Appendix of LINDENSTRAUSS (2006), that this information is enough in order to deduce the existence of $X' = X'(K_\epsilon)$ exactly as above, satisfying

the exact statement as above, but with $m_{U^{[\alpha]}}$ interchanged with $\mu_x^{[\alpha]}$: there exists a set $X' = X'(K_\epsilon)$ of measure $1 - C\epsilon^{1/2}$ (with C a universal constant), such that for every $x \in X'$

$$\forall S > 0 \; \mu_x^{[\alpha]} \left(\left\{ u \in B_S(U^{[\alpha]}) : u.x \in K \right\} \right) \geq \left(1 - \epsilon^{\frac{1}{2}} \right) \mu_x^{[\alpha]} \left(B_S(U^{[\alpha]}) \right). \qquad (33)$$

This gives us great information with respect to satisfying Item 2, but the minute we try to couple this with satisfying Item 1 we face another problem: let

$$F(\rho, S) = \left\{ u \in U^{[\alpha]} : u_n(s_1), u_n(s_2) \in T(\rho, S) \right\} \subset B_S(U^{[\alpha]})$$

be the points that satisfy Item 1. Note that $F(\rho, S_n)$ is the set of all elements belonging to the timescale S_n and keeping distance ρS_n from the axes, i.e. the complement in $B_{S_n}(U^{[\alpha]})$ of a ρS_n-thickening of the axes. So in order to satisfy Items 1 and 2 simultaneously, we need to know, for the x_n's, y_n's and all timescales S_n given to us by Lemma 5.9, that $\mu_x^{[\alpha]}(F(\rho, S_n))$ is large in comparison to $\mu_x^{[\alpha]}(B_S(U))$. In other words, that $\mu_x^{[\alpha]}$ is not concentrated near the axes for a typical $x \in X$. How can this be guaranteed? The short answer is, that it can't. We, or more precisely Lindenstrauss, needed to find a workaround. Before describing it, let's see what we can guarantee.

5.2.6. *Where the joining assumption is used.* — As a first step, let's show that $\mu_x^{[\alpha]}$ is not supported *on* the axes. This is based on the following fact about leafwise measures:

LEMMA 5.11. — *Let U be one of the coarse Lyapunov subgroups above and $U' < U$ an A-normalized connected subgroup. We have*

$$\mu_x^U(U') > 0 \iff \mu_x^U = \mu_x^{U'}.$$

Proof. — See EINSIEDLER and LINDENSTRAUSS (2019, Lemma 5.2). The proof goes via a third equivalence, namely, that the entropy contributions are the same, $h_\mu(a, U) = h_\mu(a, U')$ for every $a \in A$. □

This lemma with $U = U^{[\alpha]}$ and U' denoting one of the subgroups corresponding to the axes, implies, that for the typical points $x \in X'$, where X' is the conull set from Corollary 3.7, we have that $\mu_x^{[\alpha]}$ gives measure zero to both axes. Otherwise, the support of $\mu_x^{[\alpha]}$ will be contained in one of the axes, and won't project surjectively on both factors.

Now, fixing $S = 1$ and taking ρ small enough, it follows that ρ-thickening of the axes inside $B_1(U^{[\alpha]})$ gets arbitrarily small $\mu_x^{[\alpha]}$-mass. In other words, given $\epsilon > 0$ there exists a set $\tilde{X} \subset X'$ with $\mu(\tilde{X}) > 1 - \epsilon$, and a $\rho > 0$ such that

$$\mu_x^{[\alpha]}(F(\rho, 1)) > \frac{1}{2} \mu_x^{[\alpha]}(B_1(U^{[\alpha]}))$$

for all $x \in \tilde{X}$.

Now, in the setting of Example 2.1(1), we can use the usual maximal ergodic theorem with the element $a := \Phi_1(1, 0)$ that preserves μ, to conclude the following: there exists a subset $Y \subset X$ with $\mu(Y) > 1 - \epsilon^{1/2}$ such that for every $x \in Y$ and every $R > 0$, most

$(1 - \epsilon^{\frac{1}{2}}$ of the times) $r \in [0, R]$ have $a^r.x \in \tilde{X}$. Together with (7), this implies that for $x \in Y$ and most $r \in [0, R]$ we have

$$\mu_x^{[\alpha]}(F(\rho, e^r)) > \frac{1}{2}\mu_x^{[\alpha]}(B_{e^r}(U^{[\alpha]})). \tag{34}$$

In other words, we can find a large set of $x \in X$ and many timescales for which we know that $\mu_x^{[\alpha]}$ is not concentrated near the axes. This is a self-similarity property for $\mu_x^{[\alpha]}$; one can expect such properties for measures invariant under a diagonalizable element with positive entropy.

Remark 5.12. — A similar statement holds for Example 2.1(3). The only difference is that we need to replace the $B_{e^r}(U^{[\alpha]})$, which is defined as "the square in $U^{[\alpha]}$ with lengths e^r" with a "rectangle in $U^{[\alpha]}$ with lengths e^r and e^{2r}" (and change similarly all the other sets appearing above).

Having (34) is unfortunately, not enough for running the H-principle argument we are after. We still cannot guarantee that the timescales given to us by Lemma 5.9 are not exactly the problematic scales we don't cover in (34). Lindenstrauss had another rabbit in the hat: he noticed that tweaking the initial pair x_n and y_n from Lemma 5.8 by $a^t := (\Phi_1(1,0))^t$ for $t \in [0, T_n]$ still preserves the equality of the corresponding leafwise measures. Note that $T_n = T_n(\delta_n)$ are needed to be chosen carefully so the displacement between $a^t.x_n$ and $a^t.y_n$ won't grow too much. Having this extra freedom in choosing the initial pairs might give more possible timescales, for which we might be able to guarantee regularity properties as in (34). An intricate calculation shows that this is indeed the case, enabling Lindenstruass to run the H-principle against all odds —see LINDENSTRAUSS (2006, §7.2). Essentially the same calculation (but done in the first and the third component simultaneously), enables us to run it also in our setting. We avoid discussing it here and finish our outline of the proof of Proposition 5.1 under the simplifying assumption that a regularity property as in (34) holds for any scale.

5.2.7. *Proof of Proposition 5.1 under a simplifying self-similarity assumption.* — For concreteness consider $\lambda = \alpha$ and to simplify notation denote $B_S = B_S\left(U^{[\alpha]}\right)$. We assume that there is a $\rho \in (0,1)$ such that for μ-almost every x we have

$$\forall S > 1 \quad \mu_x^{[\alpha]}(F(\rho, S)) > \frac{1}{2}\mu_x^{[\alpha]}(B_S), \tag{35}$$

and outline a proof of Proposition 5.1. The reader may compare this outline to LINDENSTRAUSS (2006, §7.1). Assume, for contradiction, that Proposition 5.1 does not hold. Then, (25) does not hold with $\lambda = \alpha$ on a subset of positive measure which is A-invariant by (23). Ergodicity then implies that for a typical point $x \in X$, (25) does not hold with $\lambda = \alpha$. Now, let $\epsilon > 0$ and choose a *compact* subset $X_1 \subset X$ with $\mu(X_1) > 1 - \epsilon$ and with:

1. $\mu_x^{[\alpha]}$ is defined for $x \in X_1$, and the map $x \mapsto \mu_x^{[\alpha]}$ is continuous on X_1 ,
2. (25) does not hold for $x \in X_1$ (with $\lambda = \alpha$),

By §5.2.5 there exists a compact set X_2 with $\mu(X_2) > 1 - \epsilon^{\frac{1}{2}}$ such that any $x \in X_2$ satisfies (33) with $K = X_1$.

Now, use Lemma 5.8 with $X' = X_2$ to find pairs of nearby points as stated there; let x and y be such a pair with distance $\delta = \delta(\epsilon) > 0$ small enough so that the corresponding timescale S from Lemma 5.9 will be large enough, so that (35) holds. Let $G_x = \left\{ u \in U^{[\alpha]} : u.x \in X_1 \right\}$ and $G_y = \left\{ u \in U^{[\alpha]} : u.y \in X_1 \right\}$. From (33) we know that

$$\mu_x^{[\alpha]}(G_x \cap B_S) \geq \left(1 - \epsilon^{\frac{1}{2}}\right)\mu_x^{[\alpha]}(B_S)$$

and similarly for y. So,

$$\mu_x^{[\alpha]}(G_x \cap G_y \cap B_S) \geq \mu_x^{[\alpha]}(B_S) - \mu_x^{[\alpha]}(B_S \setminus G_x) - \mu_x^{[\alpha]}(B_S \setminus G_x) \geq (1 - 2\epsilon^{\frac{1}{2}})\mu_x^{[\alpha]}(B_S).$$

In particular for ϵ small enough, we can find $u \in B_S$ enabling us to use the H-principle to find $x' = u.x, y' = u.y \in X_1$ satisfying

- $\mu_{x'}^{[\alpha]} = \mu_{y'}^{[\alpha]}$
- $y' = (g'u').x'$ with $u' \in U^{[\alpha]}$ having at least one of its coordinates, $u'(s_1)$ or $u'(s_2)$, in $[\frac{1}{C}, C]$, and with $g' \in G$ of size at most $\delta^{\frac{1}{2}}$.

Applying this to all the pairs of nearby points coming from Lemma 5.8, we get pairs $x'_n, y'_n \in X_1$ as above with $\delta_n \to 0$. We choose a subsequence where the resulting u'_n as above all belongs to a compact set in at least one of their coordinates, say their first one. That is, we assume that u'_n satisfy $u'_n(s_1) \in [\frac{1}{C}, C]$ along this subsequence. As X_1 is compact, and $x \mapsto \mu_x^{[\alpha]}$ is continuous on X_1, we explained in §5.2.2 that limit points of $\{x_n\}$ and of $\{y_n\}$ will give us points $x, y = u.x \in X_1$ with $u \neq e$ with equal leafwise measure. That is, we will find $x \in X_1$ satisfying (25), in contradiction to the definition of X_1.

6. APPLICATIONS OF THE JOININGS THEOREM

In §6.1 we survey many arithmetic problems and their relation to adelic/p-adic torus orbits; In §6.2 we couple these problems to get problems related to the classification of joinings. We then explain what Theorem 1.1 can say about these coupled problems. Finally, in §6.3 we give a new arithmetic application of Theorem 1.1 for simple groups of high rank.

We recall that when (X, m_X) is a probability measure space we say that a sequence of finite set $A_i \subset X$ *equidistribute*, when the normalized counting measure on A_i converges in the weak-* topology to m_X.

6.1. Torus orbits and arithmetic

6.1.1. *Three classical arithmetic distribution problems.* — Let us first revisit the arithmetic distribution problem we considered in the introduction (see (1) for the notation): the distribution of $\frac{1}{\sqrt{D}}\mathbb{S}^2(D)$ inside $\mathbb{S}^2$ for large $D > 0$ with $D \neq 0, 4, 7 \mod 8$. Note

that the set $\mathbb{S}^2(D)$ can be also considered as the set of integer points on the level set $Q_1 = D$ for the quadratic form $Q_1(x, y, z) = x^2 + y^2 + z^2$.

The second distribution problem we consider is very similar. For $D < 0$, one can consider the set of primitive binary quadratic forms with discriminant D

$$\mathrm{Bin}_D = \left\{ q_{(a,b,c)}(x, y) = ax^2 + bxy + cy^2 : b^2 - 4ac = D, \gcd(a, b, c) = 1 \right\}. \quad (36)$$

By identifying Bin_D with the set of the corresponding coefficient vectors (a, b, c), one can realize Bin_D as the set of integer points on the level set for the quadratic form $Q_2(X, Y, Z) = Y^2 - 4XZ$:

$$V_D := \left\{ (X, Y, Z) \in \mathbb{R}^3 : Y^2 - 4XZ = D \right\} \quad (37)$$

and ask about the distribution of $\frac{1}{\sqrt{|D|}} \mathrm{Bin}_D \subset V_{-1}$ when $D \to -\infty$ (along negative *discriminants*, which are by definition the set on D's with $\mathrm{Bin}_D \neq \emptyset$). Note that the natural "cone" measure on $V_{-1} \subset \mathbb{R}^3$ is infinite. This is not really an issue; one just measures the distribution with respect to all possible quotients of two test functions. But for other reasons, one can reformulate this distribution problem in an equivalent "modular" way, which yields a distribution problem in a finite-measure space. To this end, we recall that GAUSS (1986) defined an action of $\mathrm{GL}_2(\mathbb{Z})$ on Bin_D by

$$g.q(x, y) := \frac{1}{\det(g)} q\left((x, y)g\right), \quad q(x, y) \in \mathrm{Bin}_D, g \in \mathrm{GL}_2(\mathbb{Z})$$

and showed that this action has a finite number of orbits (forming an abelian group, see below). For $q(x, y) \in \mathrm{Bin}_D$ let z_q denote the unique root of $q(z, 1)$ belonging to the upper half plane $\mathbb{H} = \{z \in \mathbb{C} : \Im(x) \geq 0\}$. The set of such roots are called *complex multiplication points*[2] or *Heegner points* of discriminant D. Furthermore, let $[z_q]$ denote the orbit of z_q under the action of $\mathrm{SL}_2(\mathbb{Z})$ by Möbius transformations on the $\mathbb{H}$. One can then verify that the $\mathrm{GL}_2(\mathbb{Z})$-orbit of $q \in \mathrm{Bin}_D$ corresponds to the $\mathrm{SL}_2(\mathbb{Z})$-orbit $[z_q]$. So the set $\mathcal{H}_D := \{[z_q] : q \in \mathrm{Bin}_D\} \subset \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ is finite. A dual reformulation of the above distribution problem on V_{-1} is the distribution of the normalized counting measures of the finite sets $\mathcal{H}_D$ on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ equipped with the finite, normalized uniform measure $\frac{3}{\pi}(ds)^2 = \frac{3}{\pi} \frac{(dx)^2 + (dy)^2}{y^2}$ when $D \to -\infty$. Figures 8-10 were made by Andreas Wieser to show the distribution of $\mathcal{H}_D$ with growing $|D|$'s.

We shortly mention a third, related distribution problem. In the last distribution problem, one can also consider positive discriminants $D > 0$ and the corresponding distribution of $\frac{1}{\sqrt{D}} \mathrm{Bin}_D \subset V_1$ as $D \to \infty$ along positive discriminants. This problem also admits a modular formulation as above: one sees that $\mathrm{GL}_2(\mathbb{Z})$-orbits in Bin_D correspond to a finite set of closed geodesics on the modular surface $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. See the work of EINSIEDLER, LINDENSTRAUSS, MICHEL, and VENKATESH (2012) for more

---

[2]Equivalently, a point $z \in \mathbb{H}$ is called a *complex multiplication point of discriminant D* if the automorphisms group of the corresponding lattice $\mathrm{span}_{\mathbb{Z}}(1, z) \subset \mathbb{C}$ is the quadratic order of discriminant D.
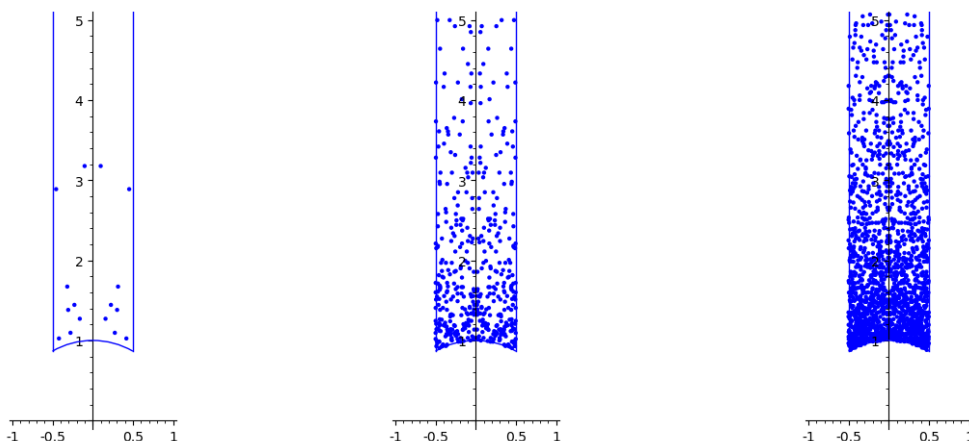
FIGURE 8. $D = -1009$ FIGURE 9. $D = -105509$ FIGURE 10. $D = -1299821$

details and nice images of this distribution for some large D's. As we partially explain in 6.1.2, roughly speaking, the sets $\mathbb{S}^2(D)$ for $D > 0$, and the sets $\mathcal{H}_D$ or Bin_D for $D < 0$ with $D = 3 \mod 4$, and the sets $\mathcal{H}_{4D}$ or Bin_{4D} for $D < 0$ with $D = 1, 2 \mod 4$, are all torsors[3] for the class group $\mathrm{Pic}(\mathcal{O}_D)$ of the quadratic order of discriminant D, and therefore have size $h_D = |\mathrm{Pic}(\mathcal{O}_D)|$. For $D > 0$, the finite set of closed geodesics corresponding to the $\mathrm{GL}_2(\mathbb{Z})$-orbits in Bin_D has size h_D. Gauss conjectured by Gauss that for $D > 0$, h_D is infinitely often equal to 1. Figure 11 shows images made by Alex Kontorovich showing the distribution of these geodesics for two comparable discriminants, one with a trivial class group and one with a class group of order 4 (The notation $[a, b, c]$ stands for the binary form $ax^2 + bxy + cy^2$).

We avoid giving historical details for these problems. We just mention that, as the images might suggest, in these three problems the sets $\mathbb{S}^2(D)$ (resp. Bin_D or equivalently $\mathcal{H}_D$) equidistribute with the respect to the natural measures on $\mathbb{S}^2$ (resp. V_{-1}, V_1 or equivalently $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$). Under an *auxiliary* congruence condition on D these results were proven by Linnik and Skubenko in the late 50s (see LINNIK, 1968). DUKE (1988) proved it in full generality, building on a work of IWANIEC (1987). We refer below to these equidistribution results (and some other variants of them mentioned below) collectively as *Duke's Theorem.*

There are several surveys of these results which might interest the reader: for a modern approach to the methods of Linnik and his school, we refer to the work of EINSIEDLER, LINDENSTRAUSS, MICHEL, and VENKATESH (2012), ELLENBERG, MICHEL, and VENKATESH (2013) and to the work of WIESER (2019). Some methods used in these works are partially related to the methods used in the proof of Theorem 1.1. For a broader overview, also on the works of Duke and Iwaniec mentioned above, the reader can consult the survey by DUKE (2007) and by MICHEL and VENKATESH (2006).

---

[3] That is, they are acted upon freely and transitively by $\mathrm{Pic}(\mathcal{O}_D)$

(a) $D = 1337$, $h_D = 1$

$\blacksquare = \{[19, 27, -8]\}$

(b) $D = 1365$, $h_D = 4$

$\blacksquare = \{[35, 35, -1]\}$, $\blacksquare = \{[7, 35, -5]\}$,
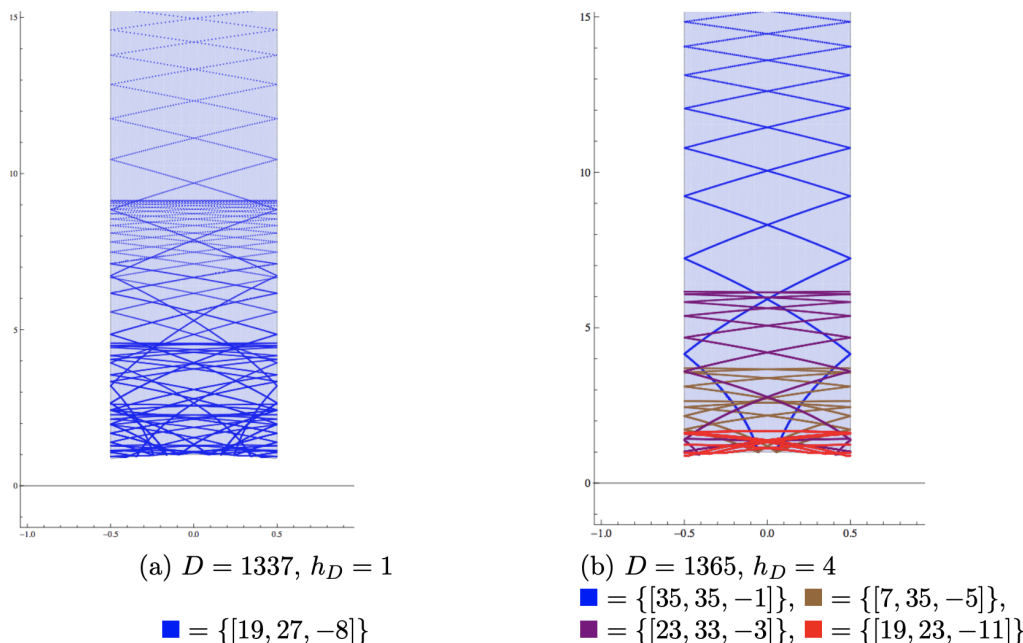$\blacksquare = \{[23, 33, -3]\}$, $\blacksquare = \{[19, 23, -11]\}$

FIGURE 11. Distribution of closed geodesics (by Alex Kontorovich)

Finally, these results (and most of the applications mentioned in this survey) are also considered in the recent survey by LINDENSTRAUSS (2021).

6.1.2. *Relation to class groups and adelic torus orbits.* — It is not a priori clear how the above distribution questions relate to questions about the distribution of torus orbits in (S-arithmetic/ adelic) homogeneous spaces. Let us concentrate on the first problem above, the distribution of $\frac{1}{\sqrt{D}}\mathbb{S}^2(D) \subset \mathbb{S}^2$, $D > 0$. We claim that the distribution properties of $\mathbb{S}^2(D)$ can be studied through the distribution properties of an orbit of the group $\mathbb{H}_v(\mathbb{A})$ in the quotient $\mathrm{SO}_3(\mathbb{Q}) \backslash \mathrm{SO}_3(\mathbb{A})$. Here, v denotes a vector in $\mathbb{S}^2(D)$, $\mathbb{H}_v$ denotes the stabilizer group of v with respect to the standard action of SO_3 on the three-dimensional space, and $\mathbb{A}$ denotes the ring of adèles of $\mathbb{Q}$. Note that, in this case, the group $\mathbb{H}_v(\mathbb{A})$ is the adelic points of a rank one, $\mathbb{Q}$-anisotropic algebraic torus. We remark that $\mathrm{SO}_3 = \mathrm{SO}_{Q_1}$ where Q_1 was defined in §6.1.1. Everything we will say below can be done analogously with Q_2 instead of Q_1, when one considers $\mathcal{H}_D$ instead of $\mathbb{S}^2(D)$.

To convince the reader that this relation is plausible, and relate it also to the class group $\mathrm{Pic}(\mathcal{O}_{-D})$, let us define a map that seems contradictory: we fix a vector $v \in \mathbb{S}^2(D)$ and we wish to define an (essentially bijective) map from a set of double-cosets related to $\mathbb{H}_v$ (see (43)) to the set $\mathbb{S}^2(D)$, by using elements of $\mathbb{H}_v(\mathbb{A}_f)$ to move v *non-trivially* to other vectors in $\mathbb{S}^2(D)$. Here, $\mathbb{A}_f$ denotes the finite adèles. To define this map, we

first note that SO_3 has class number one, that is,

$$\mathrm{SO}_3(\mathbb{A}_f) = \mathrm{SO}_3(\mathbb{Q})\,\mathrm{SO}_3(\widehat{\mathbb{Z}}) \tag{38}$$

where $\mathrm{SO}_3(\mathbb{Q})$ is embedded diagonally in $\mathrm{SO}_3(\mathbb{A}_f)$ and $\widehat{\mathbb{Z}}$ denotes the (compact) ring $\prod_{p\text{ prime}} \mathbb{Z}_p$ (For a proof of (38), one may consult ELLENBERG, MICHEL, and VENKATESH, 2013, §5.1). Note that this fact is true for this specific ternary form Q_1, although general forms can be treated with the same techniques too. One can therefore write any element of $\mathrm{SO}_3(\mathbb{A}_f)$, and in particular any $h = (h_p)_p \in \mathbb{H}_v(\mathbb{A}_f)$ as

$$h = \gamma^{-1} k, \ \text{ with } \gamma \in \mathrm{SO}_3(\mathbb{Q}), \ k = (k_p)_p \in \mathrm{SO}_3(\widehat{\mathbb{Z}}). \tag{39}$$

Using this decomposition we can take $h \in \mathbb{H}_v(\mathbb{A}_f)$ and use it to generate a vector in $\mathbb{S}^2(D)$: for $h = (h_p)_p \in \mathbb{H}_v(\mathbb{A}_f)$ let γ and $k = (k_p)_p$ be as in equation (39) and define

$$h * v := \gamma.v \in \mathbb{S}^2(D). \tag{40}$$

To show that the vector $h * v$ is indeed in $\mathbb{S}^2(D)$ and that this process is well-defined, first note that for any prime p we have (as γ is diagonally-embedded)

$$\mathbb{Q}^3 \ni \gamma.v = \gamma.h_p.v \stackrel{(39)}{=} (\gamma\gamma^{-1}k_p).v = k_p.v \in \mathbb{Z}_p^3. \tag{41}$$

As $\mathbb{Q} \cap \widehat{\mathbb{Z}} = \mathbb{Z}$ we have that $\gamma.v \in \mathbb{Z}^3$ and since we act by the orthogonal group, it has the same length and is therefore indeed an element of $\mathbb{S}^2(D)$. We also have $\gamma.v = k_p.v$ for every prime p. However, $h * v$ does depend slightly on the choices made in (39): choosing a different decomposition in (39) will yield the same vector up-to $\mathrm{SO}_3(\mathbb{Z})$, so this process gives a well-defined map to $\mathrm{SO}_3(\mathbb{Z})\backslash \mathbb{S}^2(D)$. The finite group $\mathrm{SO}_3(\mathbb{Z})$ leaves all the measures we are interested in invariant, so this small detail does not matter too much for us (and we will keep saying, imprecisely, that we got a map to $\mathbb{S}^2(D)$). A priori, there is no reason to believe that we get new vectors from this process, that is, that the image of this map is interesting. Purely algebraically (see PLATONOV and RAPINCHUK, 1994, Theorem 8.2) one can show that the set of new vectors we generate from v out of $\mathbb{H}(\mathbb{A}_f)$ via the above process, is the so-called *genus* of v:

$$\mathrm{gen}(v) = \mathrm{SO}_3(\mathbb{Z})\backslash \left\{ w \in \mathbb{Q}^3 : \forall p \ w \sim_{\mathbb{Z}_p} v, \ w \sim_{\mathbb{Q}} v \right\} \tag{42}$$

where with $w \sim_R v$ for a ring R we mean that there exists $g \in \mathrm{SO}_3(R)$ with $g.w = v$. Moreover, it follows from purely algebraic considerations, that this process defines a bijection between $\mathrm{gen}(v)$ and the following double quotient:

$$\mathbb{H}_v(\mathbb{Q})\backslash \mathbb{H}_v(\mathbb{A}_f)/\mathbb{H}_v(\widehat{\mathbb{Z}}). \tag{43}$$

Note that this double quotient is a (abelian) group since $\mathbb{H}_v$ is abelian. Assume for simplicity that D is fundamental. As an algebraic torus, $\mathbb{H}_v$ is isomorphic to the rank one $\mathbb{Q}$-anisotropic torus $\mathrm{Res}_{K/\mathbb{Q}}\mathbb{G}_m/\mathbb{G}_m$ for $K = \mathbb{Q}(\sqrt{-D})$. This identifies (with small bounded index depending on this isomorphism of the algebraic tori) the double quotient in (43) with the similar double quotient for $\mathrm{Res}_{K/\mathbb{Q}}\mathbb{G}_m/\mathbb{G}_m$ (see § 6.1.4 for more details). The latter is by definition the idèle class group $K^*\backslash \mathbb{A}_{K,f}^\times / \widehat{\mathcal{O}_{-D}}^\times$, which is isomorphic

to the class group $\operatorname{Pic}(\mathcal{O}_{-D})$. In this specific case, one can show that $\operatorname{gen}(v)$ is equal to $\operatorname{SO}_3(\mathbb{Z})\backslash\mathbb{S}^2(D)$ (see for example ELLENBERG, MICHEL, and VENKATESH, 2013, §6). In summary, the above describes a (essentially bijective) map from the class group $\operatorname{Pic}(\mathcal{O}_{-D})$ onto $\mathbb{S}^2(D)$ as promised (ignoring again the division by $\operatorname{SO}_3(\mathbb{Z})$ and the above-mentioned small bounded index issues). This map depends of course on the choice of the "base point" v. In ELLENBERG, MICHEL, and VENKATESH (2013, Proposition 3.5) they explain how to construct from these base-dependent maps a base-free way to give $\mathbb{S}^2(D)$ a torsor structure for $\operatorname{Pic}(\mathcal{O}_{-D})$, that is, how to construct a free and transitive action of $\operatorname{Pic}(\mathcal{O}_{-D})$ on $\mathbb{S}^2(D)$.

We remark that the above algebraic construction does not tell us *anything* about how $\mathbb{S}^2(D)$ distributes, and in particular does not give us any reason to believe that $\frac{1}{\sqrt{D}}\mathbb{S}^2(D)$ equidistributes. In section 6.1.5 we will give a dynamical interpretation of this construction, from which the above construction arises, and through which one can see why $\frac{1}{\sqrt{D}}\mathbb{S}^2(D)$ should equidistribute.

6.1.3. *A hidden distribution problem.* — We shortly describe another arithmetic problem because it relates in a slightly different way to the distribution of adelic torus orbits and because it gives rise naturally to a problem about joinings. We avoid giving many details, skip some definitions, and instead refer the interested reader to the introduction of AKA, LUETHI, MICHEL, and WIESER (2020).

Let $\mathcal{O}_D$ be the quadratic imaginary order with discriminant $D < 0$. We say that an elliptic curve E over $\mathbb{C}$ has complex multiplication (CM) by $\mathcal{O}_D$ if its endomorphism ring is isomorphic to $\mathcal{O}_D$. We let CM_D denote the set of isomorphism classes of elliptic curves having complex multiplication by $\mathcal{O}_D$ (Roughly speaking, one may realize this set as the isomorphism classes of $\{\mathbb{C}/\Lambda_z : [z] \in \mathcal{H}_D\}$ where Λ_z denotes the lattice $\operatorname{span}_{\mathbb{Z}}(1, z) \subset \mathbb{C}$). This set is finite, of order $h_D \asymp |D|^{\frac{1}{2}+\epsilon}$ and consists of curves which are defined over a finite algebraic extension of $\mathbb{Q}$.

Fix an odd prime p and consider now only negative D's for which p is inert in $\mathcal{O}_D$. Fixing an embedding of $\overline{\mathbb{Q}}$ in $\overline{\mathbb{Q}_p}$, one can naturally define the reduction map $\operatorname{red}_p \colon \operatorname{CM}_D \to \mathcal{S}_p$ where $\mathcal{S}_p$ is the set of supersingular elliptic curves over $\overline{\mathbb{F}_p}$. The set $\mathcal{S}_p$ is finite, of order $\asymp \frac{p-1}{12}$. DEURING (1941) showed that any element of $\mathcal{S}_p$ is the reduction of an element of CM_D for some negative D. It raised the question if there is a discriminant D (with p being inert in $\mathcal{O}_D$) for which this reduction map is surjective. Proving a variant of the works of Duke and Iwaniec mentioned above, and based on a work of GROSS (1987), MICHEL (2004) proved that surjectivity holds for large enough $|D|$. In fact, he explained that this question is a slight variant of the problems considered above: the distribution of the adelic points of a rank one, $\mathbb{Q}$-anisotropic tori in a homogeneous spaces of the form $\operatorname{SO}_Q(\mathbb{Q})\backslash \operatorname{SO}_Q(\mathbb{A})$ for some specific quadratic form Q (related only to p). In §6.1.4 we shortly review a generalized setting, in which one can consider all the above problems in a uniform way. Before we do this, let us just hint how this surjectivity problem could lead to problems involving joinings: fix two odd primes p, q and consider now only negative D's with both p and q being inert in $\mathcal{O}_D$. Is the map $(\operatorname{red}_p, \operatorname{red}_q)$

surjective for $|D|$ large enough? If so, can this be generalized to any finite number of primes?

6.1.4. *A common generalization.* — All the above problems fit in the following setting. Let B be a quaternion algebra over $\mathbb{Q}$. Recall that quaternion algebras over $\mathbb{Q}$ are classified by a finite, even-sized set of places S, where they ramify, that is, where $B \otimes \mathbb{Q}_p$ is a division-algebra. Let Nr be the associated reduced norm and $Q := \mathrm{Nr}\,|_{B^0}$ the quadratic form induced by the restriction of Nr to the traceless quaternions B^0. Recall that

$$
\begin{aligned}
B^\times \xrightarrow{\pi} PB^\times := B^\times/Z(B^\times) &\cong \ \mathrm{SO}_Q \\
g \mapsto \ &\left(x \in B^0 \mapsto gxg^{-1} \in B^0\right)
\end{aligned}
\tag{44}
$$

where $Z(B^\times)$ denotes the center of $B^\times$. For $v \in B^0$, the preimage of the stabilizer group $\mathbb{H}_v < \mathrm{SO}_Q$ under this map is the group of the invertible elements of the ring

$$
\mathbb{Q}[v] \cong \mathbb{Q}[X]/\left(X^2 + \mathrm{Nr}(v)\right) =: K,
$$

since for $v \in B^0$ we have $v^2 = -v\bar{v} = -\mathrm{Nr}(v) \in \mathbb{Q}$. Let $\mathbb{G}$ denote the algebraic group $PB^\times$ and $\mathbb{T}_v = \mathrm{Res}_{K/\mathbb{Q}}\,\mathbb{G}_m/\mathbb{G}_m < \mathbb{G}$ denote the image of $\mathbb{Q}[v]$ under π in $\mathbb{G}$. Note that $\mathbb{T}_v$ is a rank one $\mathbb{Q}$-anisotropic algebraic torus.

All the above arithmetic problems, and many others, can be recast naturally, as follows. We consider the adelic homogeneous space $\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})$, and for a *primitive* vector $v \in B^0(\mathbb{Z})$ and call an orbit of the form

$$
\mathbb{G}(\mathbb{Q})\mathbb{T}_v(\mathbb{A})\,(g_\infty, g_f)\,, \quad (g_\infty, g_f) \in \mathbb{G}(\mathbb{R}) \times \mathbb{G}(\mathbb{A}_f)
\tag{45}
$$

a *toral packet* related to v. We equip a toral packet with the pushforward of the natural Haar probability measure on the quotient $\mathbb{T}_v(\mathbb{Q})\backslash\mathbb{T}_v(\mathbb{A})$ under the map $t \mapsto t\,(g_\infty, g_f)$. One can attach to such an orbit a *discriminant* which is, roughly said, related to $-\mathrm{Nr}(v)$. More precisely, it can be defined as the product of local discriminants; for a non-Archimedean place, say at p, the local discriminant is the discriminant of the of order $\mathbb{Q}_p[v] \cap g_f(B^0(\mathbb{Z}_p))g_f^{-1}$, and for the Archimedean place, it is a measure of the distortion of the torus $g_\infty^{-1}\mathbb{T}(\mathbb{R})g_\infty$.

For example, for $B = B_{\infty,2}$, which is Hamilton's quaternion algebra, $\mathrm{Nr}\,|_{B^0}$ is the sum of three squares. So for any $v \in B^0(\mathbb{Z})$, we have $\mathbb{S}^2(D) \subset B^0(\mathbb{Z})$ with $D = \|v\|_2^2$. Recall that in this case $\mathbb{G} = PB^\times$ is SO_3 and that $\mathbb{H}_v$ denotes the stabilizer of v. We will explain in §6.1.5, how the toral packet $\mathbb{G}(\mathbb{Q})\mathbb{H}_v(\mathbb{A})\,(k_v, e_f)$ with $k_v \in \mathrm{SO}_3(\mathbb{R})$ moving $(0,0,1)^t$ to $\frac{1}{\sqrt{D}}v$ and $e_f \in \mathbb{G}(\mathbb{A}_f)$ denoting the identity element, relates naturally to the set $\mathrm{gen}(v)$ defined in (42). Note that in this case the local discriminant in the real place is constant, as $k_v^{-1}\mathbb{H}_v(\mathbb{R})k_v$ is the fixed torus $\mathrm{SO}_2(\mathbb{R})$ (see §6.1.5 for more details). This explains a bit the above terminology: the orbit (45) "codes" a packet of vectors which are in the genus of v.

In this generalized context, we refer to the following statement as Duke's Theorem: toral packets equidistribute, when their discriminants tend to $+\infty$ or $-\infty$, to a $\mathbb{G}(\mathbb{A})^+$-invariant probability measure[4] on $\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})$. Note that in many cases, but not in general, this probability measure could be shown to be $\mathbb{G}(\mathbb{A})$-invariant, i.e. the Haar measure on $\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})$. This formulation of Duke's Theorem is the conglomeration of several results by many authors, achieved before and after the seminal result of DUKE (1988). We refer the reader to EINSIEDLER, LINDENSTRAUSS, MICHEL, and VENKATESH (2011, Theorem 4.6) for a more precise formulation and for a list of references. Until now, there were three main approaches for proving such a statement: via ergodic theory (essentially Linnik's original method, under a congruence condition assumption), via modular form (as DUKE, 1988) and via subconvexity estimates. The last two approaches yield an error estimate. The subconvexity approach works naturally in the generality stated above, and can be further generalized to obtain stronger variants, such as the work of HARCOS and MICHEL (2006) (see also the discussion in Example 6.4 below).

The problems mentioned above fit in the above setup as follows:

– The distribution of $\mathbb{S}^2(D)$, as already mentioned above, relates to the choice $B = B_{\infty,2}$ for which the norm of $v \in B^0$ calculates as $Q(v) = \mathrm{Nr}(v) = \mathrm{Nr}(xi+yj+zk) = x^2 + y^2 + z^2$.

– The distribution of $\mathcal{H}_D$ (resp. closed geodesics) in $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ relates to the choice $B = B_\emptyset = M_{2\times 2}$ for which the norm of $v \in B^0$ calculates as $Q(v) = \mathrm{Nr}(v) = \mathrm{Nr}\left(\begin{smallmatrix} x & y \\ z & -x \end{smallmatrix}\right) = -(x^2 + yz)$, where the scalar matrices are identified with the underlying field. This corresponds, when $Q(v) \to \infty$ to the distribution of Heegner points, and when $Q(v) \to -\infty$ to the distribution of closed geodesics.

– The problem considered in 6.1.3, that is, the reduction of elliptic curves with CM by $\mathcal{O}_D$ to $\mathcal{S}_p$, the set of supersingular elliptic curves over $\overline{\mathbb{F}_p}$, corresponds to the case $B = B_{\infty,p}$, but in an interesting way: one can show that with this choice the double cosets in

$$\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})/\mathbb{G}(\mathbb{R} \times \widehat{\mathbb{Z}}), \tag{46}$$

or said differently, the clopen (closed and open) orbits of $\mathbb{G}(\mathbb{R} \times \widehat{\mathbb{Z}})$ on $\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})$, are in bijection with $\mathcal{S}_p$. As p is inert in $\mathcal{O}_D$ there is an embedding of $\mathrm{Res}_{\mathbb{Q}(\sqrt{D})/\mathbb{Q}}\mathbb{G}_m/\mathbb{G}_m$ into $\mathbb{G}$ and the CM curves are related to a toral packet as above. Surjectivity of red_p translates then to the question if the toral packet visits all the above clopen set, or equivalently, if the projection of toral packet to (46) is surjective. In this respect, this application is closer in nature to the work of ELLENBERG and VENKATESH (2008), which is the first instance known to the author where the algebraic ideas we explain here, were coupled with homogeneous dynamics considerations. Duke's theorem implies in this context that $\mathrm{red}_p\colon \mathrm{CM}_D \to \mathcal{S}_p$ is surjective and actually it gives a strengthening of surjectivity:

---

[4]The group $\mathbb{G}(\mathbb{A})^+$ is the image of the adelic points of the algebraic simply-connected cover of $\mathbb{G}$ under the canonical projection map (which is the Spin group of Q in our case).

the asymptotic size of the fibres for $E \in \mathcal{S}_p$, $\lim_{D \to -\infty} \frac{\left|\mathrm{red}_p^{-1}(E)\right|}{|\mathrm{CM}_D|}$ is equal to the $m_{\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})}$-measure of the $\mathbb{G}(\mathbb{R} \times \widehat{\mathbb{Z}})$-orbit corresponding to E. Here $m_{\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})}$ denotes the natural $\mathbb{G}(\mathbb{A})$-invariant probability measure on $\mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A})$. In this sense, one may say that the projection map red_p equidistributes.

6.1.5. *Why and how this works?* — In section 6.1.2 we defined a map from $\mathbb{H}_v(\mathbb{Q})\backslash\mathbb{H}_v(\mathbb{A}_f)/\mathbb{H}_v(\widehat{\mathbb{Z}})$ to $\frac{1}{\sqrt{D}}\mathbb{S}^2(D) \subset \mathbb{S}^2$. In this subsection, we would like to reinterpret this map via a projection of a toral packet between two homogeneous spaces. Consider the natural projection

$$\pi \colon \mathrm{SO}_3(\mathbb{Q})\backslash \mathrm{SO}_3(\mathbb{A}) \to \mathrm{SO}_3(\mathbb{Q})\backslash \mathrm{SO}_3(\mathbb{A})/ \mathrm{SO}_3(\widehat{\mathbb{Z}}) \tag{47}$$

where as above, $\mathrm{SO}_3(\mathbb{Q})$ is diagonally embedded in $\mathrm{SO}_3(\mathbb{A})$. We will show below the identification

$$\mathrm{SO}_3(\mathbb{Q})\backslash \mathrm{SO}_3(\mathbb{A})/ \mathrm{SO}_3(\widehat{\mathbb{Z}}) \cong \mathrm{SO}_3(\mathbb{Z})\backslash \mathrm{SO}_3(\mathbb{R}). \tag{48}$$

With this identification assumed, we may interpret π as a covering map

$$\pi \colon \mathrm{SO}_3(\mathbb{Q})\backslash \mathrm{SO}_3(\mathbb{A}) \to \mathrm{SO}_3(\mathbb{Z})\backslash \mathrm{SO}_3(\mathbb{R}) \tag{49}$$

from an adelic homogeneous space to a real homogeneous space (with compact fibres isomorphic to $\mathrm{SO}_3(\widehat{\mathbb{Z}})$). The identification (48) follows "tautologically" from (38):

$$\mathrm{SO}_3(\mathbb{Q})\backslash \mathrm{SO}_3(\mathbb{A}) \overset{(38)}{=} \mathrm{SO}_3(\mathbb{Q})\backslash \left(\mathrm{SO}_3(\mathbb{Q})\, \mathrm{SO}_3(\mathbb{R} \times \widehat{\mathbb{Z}})\right) \tag{50}$$

$$\cong \left(\mathrm{SO}_3(\mathbb{Q}) \cap \mathrm{SO}_3(\mathbb{R} \times \widehat{\mathbb{Z}})\right) \backslash \mathrm{SO}_3(\mathbb{R} \times \widehat{\mathbb{Z}}) \tag{51}$$

$$= \mathrm{SO}_3(\mathbb{Z})\backslash \mathrm{SO}_3(\mathbb{R} \times \widehat{\mathbb{Z}}) \tag{52}$$

where the third equality follows from the fact that $\mathbb{Z} = \mathbb{Q} \cap \widehat{\mathbb{Z}}$ as subsets of $\mathbb{A}_f$. Dividing now everything from the right by $\mathrm{SO}_3(\widehat{\mathbb{Z}})$ gives the desired identification. It is nevertheless useful to explicate this identification to start to "see" dynamics: consider the double coset $\mathrm{SO}_3(\mathbb{Q})\,(g_\infty, g_f)\,\mathrm{SO}_3(\widehat{\mathbb{Z}})$ with $g_\infty \in \mathrm{SO}_3(\mathbb{R}), g_f \in \mathrm{SO}_3(\mathbb{A}_f)$. This coset may be identified with $\mathrm{SO}_3(\mathbb{R})g_\infty$ if g_f is "small", that is, if g_f belongs to the compact group $\mathrm{SO}_3(\widehat{\mathbb{Z}})$. But if g_f is "large", that is, if $g_f \notin \mathrm{SO}_3(\widehat{\mathbb{Z}})$, we first need to "bring it back to $\mathrm{SO}_3(\widehat{\mathbb{Z}})$" using the discrete group $\mathrm{SO}_3(\mathbb{Q})$. This means that, we first need to find $\gamma \in \mathrm{SO}_3(\mathbb{Q})$ with $\gamma g_f =: k \in \mathrm{SO}_3(\widehat{\mathbb{Z}})$ and then, recalling that $\mathrm{SO}_3(\mathbb{Q})$ is embedded diagonally, we have

$$\mathrm{SO}_3(\mathbb{Q})(g_\infty, g_f)\,\mathrm{SO}_3(\widehat{\mathbb{Z}}) = \mathrm{SO}_3(\mathbb{Q})(\gamma g_\infty, \gamma g_f)\,\mathrm{SO}_3(\widehat{\mathbb{Z}}) = \mathrm{SO}_3(\mathbb{Q})(\gamma g_\infty, k)\,\mathrm{SO}_3(\widehat{\mathbb{Z}}).$$

As $k = \gamma g_f$ is "small", the latter is identified as above with $\mathrm{SO}_3(\mathbb{Z})\gamma g_\infty$.

Although everything we said so far is purely algebraic, the adjectives "small", "large" and the process of "bringing g_f back" stem from a dynamical perspective. Before we can explain this perspective, we just need to understand how the sphere $\mathbb{S}^2$, or more precisely $\mathrm{SO}_3(\mathbb{Z})\backslash\mathbb{S}^2$, can be described as a homogeneous space. Consider the natural action of $\mathrm{SO}_3(\mathbb{R})$ on $\mathbb{S}^2$ (or equivalently on $\mathbb{R}^3$) and let $\mathrm{SO}_2(\mathbb{R}) := \mathrm{Stab}_{\mathrm{SO}_3(\mathbb{R})}(e_3) < \mathrm{SO}_3(\mathbb{R})$,

where e_3 is the unit vector $(0,0,1)^t$. This yield the identification $\mathrm{SO}_3(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \cong \mathbb{S}^2$ and therefore also

$$\mathrm{SO}_3(\mathbb{Z})\backslash \mathrm{SO}_3(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \cong \mathrm{SO}_3(\mathbb{Z})\backslash\mathbb{S}^2. \tag{53}$$

Said differently, $\mathrm{SO}_2(\mathbb{R})$-orbits in the homogeneous space $\mathrm{SO}_3(\mathbb{Z})\backslash \mathrm{SO}_3(\mathbb{R})$ correspond to points in $\mathrm{SO}_3(\mathbb{Z})\backslash\mathbb{S}^2$.

After all of these preparations, we are finally ready to state the upshot of this subsection, which we can then also interpret dynamically. For $v \in \mathbb{S}^2(D)$ choose $k_v \in \mathrm{SO}_3(\mathbb{R})$ with $k_v(e_3) = \frac{1}{\sqrt{D}} \cdot v$. With this choice, we have $\mathbb{H}_v(\mathbb{R}) = k_v \mathrm{SO}_2(\mathbb{R})k_v^{-1}$. Tracing through the definitions and identifications above, one is led to the following conclusion: the projection (via π from (49)) of the adelic torus orbit

$$\mathrm{SO}_3(\mathbb{Q})\mathbb{H}_v(\mathbb{A})(k_v, e_f) \subset \mathrm{SO}_3(\mathbb{Q})\backslash \mathrm{SO}_3(\mathbb{A}) \tag{54}$$

to the real homogeneous space $\mathrm{SO}_3(\mathbb{Z})\backslash \mathrm{SO}_3(\mathbb{R})$, where $e_f \in SO_3(\mathbb{A}_f)$ denotes the identity element, is a collection of $\mathrm{SO}_2(\mathbb{R})$-orbits which, under (53), corresponds precisely to $\mathrm{gen}(v)$ viewed as a subset of $\mathrm{SO}_3(\mathbb{Z})\backslash\mathbb{S}^2$. We urge the reader to verify this as a good way to repeat all the algebraic considerations above. In particular, this reproduces the map related to $\mathbb{H}_v$ discussed in 6.1.2, and explains how it arises naturally.

Up to now we were just chasing definition and identifications, and from time to time, were hinting at a dynamical interpretation. We finally can explain how dynamics can give us valuable input about the distribution of $\mathrm{gen}(v)$ (which is equal to $\mathbb{S}^2(D)$ in our specific case). For all we know, $\mathrm{gen}(v)$ can be quite a boring set. What we know from the above, is that elements of $\mathrm{gen}(v)$ are of the form $\gamma.v$, $\gamma \in \mathrm{SO}_3(\mathbb{Q})$ for γ's that were found in order to "bring large elements in $\mathrm{SO}_3(\mathbb{A}_f)$ back to $\mathrm{SO}_3(\widehat{\mathbb{Z}})$". Let's explain this more precisely, under the assumption that the adelic orbit (54) "goes everywhere". That is, we assume that it is dense in $\mathrm{SO}_3(\mathbb{Q})\backslash \mathrm{SO}_3(\mathbb{A})$. Then, in particular, for an arbitrary $g_\infty \in \mathrm{SO}_3(\mathbb{R})$ we can find $h \in \mathbb{H}_v(\mathbb{A}_f)$ with

$$\mathrm{SO}_3(\mathbb{Q})(k_v, h) \approx \mathrm{SO}_3(\mathbb{Q})(g_\infty, e_f). \tag{55}$$

Projecting (55) to the real homogeneous space, this approximation found for us a $\gamma \in \mathrm{SO}_3(\mathbb{Q})$ with $\gamma h \in \mathrm{SO}_3(\widehat{\mathbb{Z}})$, such that the vector $\gamma v \in \mathbb{Z}^3$ satisfy

$$\gamma.v = \gamma k_v(e_3) \approx g_\infty(e_3) = \text{An arbitrary direction in } \mathbb{S}^2! \tag{56}$$

Note that in order to have a chance that the orbit (54) "will go everywhere", $\mathbb{H}_v(\mathbb{A})$ needs to be a large group. In particular, our dynamical considerations are (unfortunately!) usually restricted to working in one particular place, say at a fixed prime p. This explains the congruence conditions that we must impose on $D = \|v\|^2$ in order to use these dynamical considerations: they are there to verify that $\mathbb{H}_v$ splits at the fixed prime p and consequently to ensure that $\mathbb{H}_v(\mathbb{Q}_p)$ contains large elements.

Remark 6.1. — One can consider the above methods in higher dimensions, say, for the quadratic form $\sum_{i=1}^n x_i^2$ with $n \geq 6$. The dynamical methods showing the equidistribution of orbits as in (54) with 3 replaced by n are much simpler (they are easy special cases of MOZES and SHAH (1995) and GORODNIK and OH (2011) since $\mathbb{H}_v < \mathrm{SO}_n$ is maximal).

Based on this, one can work out from our discussion above a proof of the equidistribution of integer points on $\mathbb{S}^{n-1}$ based on homogeneous dynamics.

6.2. Arithmetically motivated joinings problems

In §6.2.1 we consider several arithmetic problems which are mainly coupling of individual equidistribution problems of the type that we discussed in §6.1. For all the problems presented below, there is no complete solution. In §6.2.2 we discuss partial results have been achieved without using Theorem 1.1. Then, in §6.2.3 we discuss partial results that use Theorem 1.1. For some of them, Theorem 1.1 gives a key input, which allows for a solution under congruence conditions at two distinct prime numbers. We will also discuss in §6.2.3 why and how these congruence conditions are used in order to make Theorem 1.1 applicable.

6.2.1. *Several explicit and implicit joint distribution problems.* — The first application of Theorem 1.1 was already discussed in the introduction:

Example 6.2. — With notation as in equation (2), one asks about the equidistribution of

$$J_D = \left\{ \left(\tfrac{1}{\sqrt{D}} v, [\Lambda_v] \right) : v \in \mathbb{S}^2(D) \right\} \subset \mathbb{S}^2 \times (\mathrm{SL}_2(\mathbb{Z}) \backslash \mathbb{H})$$

when $D \to \infty$ along $D = 0, 4, 7 \mod 8$. Recall from §6.1 that[(5)] $\mathbb{S}^2(D)$ is a torsor of class group $\mathrm{Pic}(\mathcal{O}_{-D})$. It turns out, that the set $\{[\Lambda_v] : v \in \mathbb{S}^2(D)\}$ is also contained in another torsor of $\mathrm{Pic}(\mathcal{O}_{-D})$: the set $\{[\Lambda_v] : v \in \mathbb{S}^2(D)\}$ is a subset of $\mathcal{H}_{-D}$ when $D = 3 \mod 4$ and of $\mathcal{H}_{-4D}$ when $D = 1, 2 \mod 4$, and the containing set in both cases is a torsor of the class group $\mathrm{Pic}(\mathcal{O}_{-D})$. The novelty of AKA, EINSIEDLER, and SHAPIRA (2016) was to relate (in the spirit of §6.1) the distribution of J_D to the distribution of a *joined* adelic orbit of a stabilizer group (as in §6.1.2) in a product of two homogeneous spaces, which correspond to $\mathbb{S}^2$ and $\mathrm{SL}_2(\mathbb{Z}) \backslash \mathbb{H}$. This stabilizer group is again a $\mathbb{Q}$-anisotropic algebraic torus of rank one. The distribution of these adelic orbits can be studied via Theorem 1.1, once one assumes congruence conditions on D. We will give more details in §6.2.3.

Although Example 6.2 was motivated by a geometric construction, examining it closely through the $\mathrm{Pic}(\mathcal{O}_{-D})$-torsor structure on $\mathbb{S}^2(D)$ and on $\mathcal{H}_{-D}$ (or on $\mathcal{H}_{-4D}$), one can see that J_D corresponds to the set

$$\left\{ \left(t.v, t^2.[\Lambda_v] \right) : t \in \mathrm{Pic}(\mathcal{O}_{-D}) \right\} \tag{57}$$

where t^2 stand for the square of t in the class group $\mathrm{Pic}(\mathcal{O}_{-D})$. This leads to several interesting and arithmetically motivated, "coupled" problems. Before stating them, we want to stress our point of view in this section: in the first component of (57) we consider an orbit of the class group $\mathrm{Pic}(\mathcal{O}_{-D})$ and in the second component of (57)

[(5)] to be completely precise, we must consider $\mathrm{SO}_3(\mathbb{Z}) \backslash \mathbb{S}^2(D)$ and possibly other finite-index issues, but we ignore them for simplicity

we consider an orbit of a subgroup of $\mathrm{Pic}(\mathcal{O}_{-D})$, namely, the subgroup of squares in $\mathrm{Pic}(\mathcal{O}_{-D})$.

Let us first concentrate on equidistribution in each component, that is, on the interesting problem of *individual* equidistribution of orbits of subgroups of $\mathrm{Pic}(\mathcal{O}_{-D})$, or more generally of cosets of such subgroups. We recall that by SIEGEL (1935) we know that $|\mathrm{Pic}(\mathcal{O}_{-D})| = D^{\frac{1}{2}+o(1)}$. One can easily find (for example for the torsor $\mathcal{H}_{-D}$) subgroups of size $D^{o(1)}$ whose orbits do not equidistribute (see also the work of MCMULLEN (2009) for interesting results in this context). EINSIEDLER, LINDENSTRAUSS, MICHEL, and VENKATESH (2009, Conjecture 1.11) make a bold conjecture in this context. We phrase it in the above context (e.g., for the $\mathrm{Pic}(\mathcal{O}_{-D})$-torsors $\mathcal{H}_{-D}$ or $\mathbb{S}^2(D)$), but it can also be phrased for all the problems related to toral packets we considered in §6.1.4.

CONJECTURE 6.3. — *Let $\rho > 0$. Individual equidistribution holds for orbits of cosets of subgroups of size D^ρ.*

Note that the generalized Riemann conjecture implies this conjecture for $\rho > \frac{1}{4}$. HARCOS and MICHEL (2006) prove this conjecture for $\rho \in (\frac{1}{2} - \epsilon_0, \frac{1}{2}]$ for some $\epsilon_0 > 0$ (e.g. for $\epsilon_0 = \frac{1}{2827}$ in the context of $\mathcal{H}_{-D}$). See also AKA and EINSIEDLER (2016) for a much weaker result achieved by homogeneous dynamics techniques. It is a fascinating conjecture from the homogeneous dynamics perspective: until now there are no results that could utilize the algebraic structure of the class group $\mathrm{Pic}(\mathcal{O}_{-D})$ (as the latter is rather mysterious, this unfortunately makes sense).

Individual equidistribution plays an important role in the following two examples:

Example 6.4. — We generalize the above setting. For $1 \le i \le r$ we let X_i denote a measure space with a uniform measure m_{X_i}. The reader should think about $X_i = \mathbb{S}^2$ or about $X_i = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ as in the examples above. Let $\mathcal{G}_D^{(i)} \subset X_i$ denote (identical or different) torsors of $\mathrm{Pic}(\mathcal{O}_{-D})$, for D's in a sequence tending to ∞. Although the above is phrased generally, we actually only consider here the problems we discussed in §6.1 (more precisely, problems that fit to the general formulation given in §6.1.4 or small variants of them). For concreteness, the reader should think about $\mathcal{H}_{-D}$, or $\mathbb{S}^2(D)$, or CM_{-D} from §6.1.3.

Assume that $\mathcal{G}_D^{(i)}$ equidistribute to m_{X_i}. Fix some exponents $k_1, \dots, k_r$ and base points $g_D^{(i)} \in \mathcal{G}_D^{(i)}$, and consider

$$P_D := \left\{ \left(t^{k_1}.g_D^{(1)}, \dots, t^{k_r}.g_D^{(r)} \right) : t \in \mathrm{Pic}(\mathcal{O}_{-D}) \right\} \tag{58}$$

Does P_D equidistribute to $m_{X_1} \otimes \cdots \otimes m_{X_r}$? A necessary condition is of course that individual equidistribution holds in each component. As we explained before this example, individual equidistribution relates to the growth of the subgroup of k_i-powers in $\mathrm{Pic}(\mathcal{O}_{-D})$, or equivalently, to the growth of the k_i-torsion subgroup of $\mathrm{Pic}(\mathcal{O}_{-D})$. Assuming a folklore conjecture (see ELLENBERG, PIERCE, and WOOD (2017, (1.2))) and the references within) about the growth of k-torsion in $\mathrm{Pic}(\mathcal{O}_{-D})$, the growth of the k-powers subgroup of $\mathrm{Pic}(\mathcal{O}_{-D})$ has conjecturally the same exponent as the growth of

Pic($\mathcal{O}_{-D}$). Therefore, conjecturally, individual equidistribution follows from the work of HARCOS and MICHEL (2006) mentioned above. This folklore conjecture is holds for $k = 2$, so individual equidistribution in this case is known.

As the work of HARCOS and MICHEL (2006) handles cosets equally well (or just by changing the base point for the torsor), individual equidistribution, under the conjecture about the k-torsion, is true also when we consider a shifted version of P_D: choose arbitrary $t_D^i \in \mathrm{Pic}(\mathcal{O}_{-D})$ and consider

$$P_D^{\mathrm{shifted}} := \left\{ \left((t^{k_1} t_D^1).g_D^{(1)}, \ldots, (t^{k_r} t_D^r).g_D^{(r)} \right) : t \in \mathrm{Pic}(\mathcal{O}_{-D}) \right\}. \tag{59}$$

As we said above, the X_i's which are of interest to us in this survey, come from the general setting explained in §6.1.4, and correspond to homogeneous spaces which are related to algebraic groups of the form $\mathbb{G}_i = PB_i^\times$ for quaternion algebras B_i over $\mathbb{Q}$, i.e., to $\mathbb{Q}$-forms of SL_2. We will explain in §6.2.3, that assuming individual equidistribution (for which the shifts t_D^i are irrelevant), we could essentially show via Theorem 1.1, that equidistribution of P_D^{shifted}, for arbitrary shifts, follows in the following cases (as usual, under congruence condition at two primes):

1. All the corresponding $\mathbb{G}_i$ are non-isogeneous over $\mathbb{Q}$.
2. For any $i \neq j$ with $\mathbb{G}_i$ and $\mathbb{G}_j$ being $\mathbb{Q}$-isogeneous, we have $k_i \neq k_j$.

But when the exponents are equal, we cannot expect equidistribution of arbitrary shifts even if individual equidistribution is known. Indeed, say $r = 2$, $X = X_1 = X_2$ and $k_1 = k_2$. If one chooses $t_D^1 = t_D^2$ for every D, the collection P_D^{shifted} will converge to the diagonal joining, that is, to $\iota_*(m_X)$ where $\iota : X \to X \times X, x \mapsto (x, x)$. Also, if the quotient $t_D^1 (t_D^2)^{-1}$ does not "grow", one cannot expect equidistribution. In order to explain what we mean by "grow", and state a necessary condition for equidistribution, we define a "size"-function N on $\mathrm{Pic}(\mathcal{O}_{-D})$. Recall that $\mathrm{Pic}(\mathcal{O}_{-D})$ is the ideal class group of $\mathcal{O}_{-D}$ which, by definition, is the set of classes of invertible fractional ideals in the field of fractions of $\mathcal{O}_{-D}$ modulo principal fractional ideals. Using the Norm map Nr on ideals of $\mathcal{O}_{-D}$, we define for $t \in \mathrm{Pic}(\mathcal{O}_{-D})$

$$N(t) = \min \left\{ \mathrm{Nr}(\mathfrak{a}) : \mathfrak{a} \subset \mathcal{O}_{-D} \text{ is an invertible ideal in the class of } t \right\}.$$

With this size-function N, we can formulate a necessary condition for equidistribution of P_D^{shifted}: for any $1 \leq i \neq j \leq r$ with $k_i = k_j$ and $\mathbb{G}_i = \mathbb{G}_j$, the shifts must satisfy

$$N \left(t_D^i t_D^{j^{-1}} \right) \overset{D \to \infty}{\longrightarrow} \infty. \tag{60}$$

Otherwise, say, if one of these size-functions is bounded by M (say, for some $i \neq j$), any limit measure will be supported on a union of the graphs of Hecke-correspondences of level $\leq M$ of a diagonal joining (in the i, j-components). MICHEL and VENKATESH (2006) conjectured that this is the only obstruction:

CONJECTURE 6.5 (Mixing conjecture of Michel and Vekatesh)

In the above setting (assuming individual equidistribution or just $k_1 = \cdots = k_r = 1$), if (60) *is satisfied, then* P_D^{shifted} *equidistributes to* $m_{X_1} \otimes \cdots \otimes m_{X_r}$.

In a celebrated work, KHAYUTIN (2019) essentially proves this conjecture, under the usual assumption of congruence conditions on D at two primes, and some other (arguably) mild assumptions. More precisely, Khayutin considers the case of the torsors $\mathcal{H}_{-D}$ with $r = 2$ and $k_1 = k_2 = 1$ and proves that under the usual congruence conditions at two arbitrary primes (specifically, $-D$ must be a square modulo these two arbitrary primes) and under the assumption that the Dedekind ζ-function of the fields $\mathbb{Q}(\sqrt{-D})$ have no exceptional Landau–Siegel zeroes, equidistribution holds. His methods should work equally well for the general case we stated here (assuming individual equidistribution holds). This is the only application of Theorem 1.1 so far, where Theorem 1.1 does not rule out diagonal joinings, so other tools must be exploited. We will explain how Theorem 1.1 was used by Khayutin in §6.2.3.

Example 6.6. — We return to the problem we posed in the end of §6.1.3: is the map $(\mathrm{red}_p, \mathrm{red}_q)$ surjective for $|D|$ large enough? Can this be generalized to any finite number of primes? This arithmetic question is implicitly a joint distribution problem, and fit exactly in the setting we discuss above. It turns out that it can answered positively under congruence conditions at two distinct primes with Theorem 1.1. See §6.2.3 below for more details.

Example 6.7. — Going back to Example 6.2, one can consider analogues of it in higher dimensions. Consider $\mathbb{Q}^n$ and fix $0 < k < n$. Let $\mathcal{L}_D$ denote the set of k-dimensional subspaces in $\mathbb{Q}^n$ with discriminant D. By the discriminant of a subspace L we mean the square of the covolume of the lattice $\Lambda_L := \mathbb{Z}^n \cap L$ inside L. One can consider the joint equidistribution of the following three objects: the orientation of L, viewed via the corresponding point in the Grassmannian of k-dimensional subspaces of $\mathbb{Q}^n$, the k-dimensional lattice $[\Lambda_L]$, and the $n-k$ dimensional orthogonal lattice $[\Lambda_{L^\perp}] = [L^\perp \cap \mathbb{Z}^n]$ (both considered up-to rotation). This leads to a similar construction of adelic orbits and therefore to interesting joinings, but in the case where k or $n-k$ are greater than 2, such joinings will be invariant under a group containing unipotent elements. This enables the use of tools from unipotent dynamics to classify the possible joinings.

The remaining case, where $k = 2, n = 4$, turns out to be very much related to the classical problems we considered above: first, for $L \in \mathcal{L}_D$ the lattice in $[\Lambda_L]$ and the orthogonal lattice $[\Lambda_{L^\perp}]$ are both two-dimensional lattices and as such (essentially, up-to primitivity issues) are points in $\mathcal{H}_{-D}$ (e.g. by considering the binary quadratic forms $(x_1^2 + x_2^2 + x_3^2 + x_4^2)|_{\Lambda_L}$ and $(x_1^2 + x_2^2 + x_3^2 + x_4^2)|_{\Lambda_{L^\perp}}$). Second, the local isomorphism between SO_4 and $\mathrm{SO}_3 \times \mathrm{SO}_3$ implies that elements $L \in \mathcal{L}_D$ are essentially parameterized by $\mathbb{S}^2(D) \times \mathbb{S}^2(D)$, say by $\left(v_1^L, v_2^L\right)$. Hoping for a miracle, one can also consider the orthogonal lattices $[\Lambda_{v_1^L}] := [\left(v_1^L\right)^\perp \cap \mathbb{Z}^3]$ and $[\Lambda_{v_2^L}] := [\left(v_2^L\right)^\perp \cap \mathbb{Z}^3]$ and ask about the joint distribution of

$$J_D = \left\{ \left(v_1^L, v_2^L, [\Lambda_L], [\Lambda_{L^\perp}], [\Lambda_{v_1^L}], [\Lambda_{v_2^L}] \right) : L \in \mathcal{L}_D \right\}. \tag{61}$$

Does J_D equidistribute to $(m_{\mathbb{S}^2})^{\otimes 2} \otimes \left(m_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}}\right)^{\otimes 4}$? It turns out, that this geometric construction corresponds to a natural arithmetic problem, of the form we discussed above: with the notation from Example 6.4, consider the $\mathrm{Pic}(\mathcal{O}_{-D})$-torsors

$$X_1 = X_2 = \mathbb{S}^2, \ \mathcal{G}_D^{(1)} = \mathcal{G}_D^{(2)} = \tfrac{1}{\sqrt{D}}\mathbb{S}^2(D), \ X_i = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}, \ \mathcal{G}_D^{(i)} = \mathcal{H}_D, \ i = 3, 4, 5, 6.$$

Then, roughly speaking, and after choosing correctly base points $g_D^{(i)} \in \mathcal{G}_D^{(i)}$, the set J_D in (61) corresponds to

$$\left\{ \left(t.g_D^{(1)}, s.g_D^{(2)}, (ts).g_D^{(3)}, \left(ts^{-1}\right).g_D^{(4)}, \left(t^2\right).g_D^{(5)}, \left(s^2\right).g_D^{(6)}\right) : (t, s) \in (\mathrm{Pic}(\mathcal{O}_{-D}))^2 \right\}. \quad (62)$$

This implies that understanding the distribution of this set, as we will explain in §6.2.3, is closely related to classifying joinings in S-adic analogues of examples 2.1(2) and 2.1(3). We refer the reader to AKA, EINSIEDLER, and WIESER (2021, §8) for a further discussion of arithmetic distribution problems in the spirit of (62).

Generally, and in particular in homogeneous dynamics, one expects equidistribution to hold, unless there is a "visible" obstacle. In particular, we expect the Mixing conjecture (Conjecture 6.5) to hold, the equidistribution of J_D from Example 6.2 (or equivalently of the set in (57)) to hold, the equidistribution of J_D from Example 6.7 to hold, and after reformulating it as an equidistribution problem, we expect the reduction maps from Example 6.6 to "equidistribute". As we said above, none of these question admits a complete, unconditional solution.

6.2.2. *Partial results without using Theorem 1.1.* — Direct generalization of the methods used in the various analytic proofs of Duke's Theorem can yield averaged results (that is, the equidistribution of the union over $\{D : D \leq M\}$ with $M \to \infty$), see the appendix by Ruixiang Zhang in the ArXiv version of AKA, EINSIEDLER, and SHAPIRA (2015). This is related to the work of MAASS (1956), who considered special cases of Example 6.7 also on the average. Similar averaged results of special cases of Example 6.7 were proven by SCHMIDT (1998) and recently sharpened by HORESH and KARASIK (2020).

BLOMER and BRUMLEY (2020) found recently an analytic approach to study joined equidistribution problems involving only two individual equidistribution problems as above, for forms of SL_2 when $r = 2$. They can prove the conjectured equidistribution under the assumption of the generalized Riemann hypothesis, but without the need for auxiliary congruence conditions. This gives a conditional solution to many of the problems considered above, and in particular to the problems mentioned in Example 6.2, and in 6.6 for the reduction with respect to two distinct primes.

6.2.3. *Results toward equidistibution using Theorem 1.1.* — We begin by explaining, in the spirit of §6.1.4, how the problems mentioned in §6.2.1 correspond to "joined" adelic orbits in adelic homogeneous spaces. We then explain the need for auxiliary congruence conditions and how they give rise to a class-$\mathcal{A}'$ homomorphism as in Theorem 1.1,

enabling the use of this theorem. Finally, we describe how Theorem 1.1 was used so far to obtain partial solutions of the problems we mentioned in §6.2.1.

Coupling toral packets together. — Going back to the setting of §6.1.4, we consider r quaternion algebras $B_1, \ldots, B_r$ over $\mathbb{Q}$ and let $\mathbb{G}_i = PB_i^\times$ be the corresponding algebraic groups. Assume that there exists a rank one $\mathbb{Q}$-anisotropic algebraic torus $\mathbb{T}_n$ such that for any $1 \le i \le r$, there is an embedding $\iota_{i,n} \colon \mathbb{T}_n \to \mathbb{G}_i$ (e.g., via finding corresponding vectors in B_i^0, see §6.1.4) giving rise to a toral packet

$$\mathbb{G}_i(\mathbb{Q})\iota_{i,n}(\mathbb{T}_n(\mathbb{A}))g_{i,n} \subset \mathbb{G}_i(\mathbb{Q})\backslash\mathbb{G}_i(\mathbb{A}), \quad g_{i,n} \in \mathbb{G}_i(\mathbb{A}) \tag{63}$$

with discriminant D_n (as in (45)). Assume that $D_n \overset{n\to\infty}{\Longrightarrow} \infty$ or $D_n \overset{n\to\infty}{\Longrightarrow} -\infty$ and that these toral orbits converge to a measure μ_i on $\mathbb{G}_i(\mathbb{Q})\backslash\mathbb{G}_i(\mathbb{A})$ when $n \to \infty$. For example, we know by Duke's Theorem (as stated in §6.1.4) that each μ_i must be a $\mathbb{G}_i(\mathbb{A})^+$-invariant measure. Now, let $\mathbb{G} = \prod_{i=1}^r \mathbb{G}_i$, define $\iota_n \colon \mathbb{T}_n \to \mathbb{G}$ as $\iota_n = (i_{1,n}, \ldots, i_{r,n})$, and consider the following joined orbit:

$$\mathcal{J}_n := \mathbb{G}(\mathbb{Q})\iota_n(\mathbb{T}_n(\mathbb{A})) \underbrace{(g_{1,n}, \ldots, g_{r,n})}_{g_n :=} \subset \mathbb{G}(\mathbb{Q})\backslash\mathbb{G}(\mathbb{A}). \tag{64}$$

As in (45), we equip $\mathcal{J}_n$ with $m_{\mathcal{J}_n}$, the pushforward of the uniform probability measure on $\mathbb{T}_n(\mathbb{Q})\backslash\mathbb{T}_n(\mathbb{A})$ under $t \mapsto \iota_n(t)g_n$. It follows that any weak-* limit η of $\{m_{\mathcal{J}_n}\}$ is a probability measure projecting via the natural projections $\pi_i \colon \mathbb{G} \to \mathbb{G}_i$ to the measures μ_i, $i = 1, \ldots r$.

To be able to use homogeneous dynamics methods, we must ask ourselves, under which elements of $\mathbb{G}(\mathbb{A})$ is η invariant. Generally, the measure η is invariant under limits in the Chabauty-topology of the subgroups $g_n^{-1}\iota_n(\mathbb{T}_n(\mathbb{A}))g_n$. We already explained in the paragraph after (56), that (most) of the methods in homogeneous dynamics, essentially work only when we restrict to a *fixed* place (or finitely many fixed places). We also gave intuition there to why having "large" elements in the toral packet (or equivalently, having a $\mathbb{Q}_p$-*split* torus $\mathbb{T}_n(\mathbb{Q}_p)$) is needed in order to expect equidistribution and in order to use dynamical methods. Going back to the situation at hand, we see, that in order to obtain elements at the place p, leaving the measure η invariant, we must look at limits of subgroups of the form

$$g_{n,p}^{-1}\iota_{n,p}(\mathbb{T}_n(\mathbb{Q}_p))g_{n,p} \tag{65}$$

where $g_{n,p}$ denotes the p-adic component of g_n, and similarly for $\iota_{n,p}$. Our only chance to pinpoint a non-trivial element in the limit of these $\mathbb{Q}_p$-tori is to assume that all of them split over $\mathbb{Q}_p$. Asking for this splitting, is equivalent to imposing a congruence condition on the corresponding discriminant D_n modulo p (normally, depending slightly on the exact definition of discriminant, $-D_n$ should be congruent to a square modulo p). Then, in the limit we have two options: either the groups in (65) "degenerate" to a one-parameter unipotent subgroup, in which case, we can use methods from unipotent dynamics mentioned in §2.1 (compare also to the work of ESKIN, MOZES, and SHAH, 1996), or we can find a $\mathbb{Q}_p$-split torus in the limit (for the problems considered above,

the latter is more plausible: if the $\mathbb{Q}_p$-split tori $\mathbb{T}_n$ are the stabilizers of integral vectors, say in $\mathbb{Z}^3$, their Chabauty limit will contain the stabilizer of a vector in $\mathbb{Z}_p^3$). Therefore, fixing two distinct places p and q, and assuming that we don't get any unipotent invariance in the limit, we get invariance under a split torus at p and at q. That is, we can find elements a_p and a_q in the corresponding tori, such that the homomorphism $\phi \colon \mathbb{Z}^2 \to \mathbb{G}(\mathbb{A})$ mapping two generators of $\mathbb{Z}^2$ to the elements of $\mathbb{G}(\mathbb{A})$ corresponding to a_p and a_q, will be of class-$\mathcal{A}'$ and the limit measure η will be invariant under its image. In summary, under the assumption of congruence conditions at two arbitrary fixed primes, η is a joining for which Theorem 1.1 applies.

Under these two congruence conditions, Theorem 1.1 reduces the classification of each of the *ergodic component* of η to an algebraic problem concerning the algebraic nature of $\mathbb{G}_i$ and $\iota_{i,n}$. We remark, that the latter is true, at least up to the so-called character spectrum, that is up-to the difference between $\mathbb{G}(\mathbb{A})^+$-invariance and $\mathbb{G}(\mathbb{A})$-invariance. Indeed, note that the measures μ_i above are known to be in general only $\mathbb{G}_i(\mathbb{A})^+$-invariant. Any input we will get from Theorem 1.1, will give us information on η, up-to the so-called character-spectrum $\mathbb{G}(\mathbb{A})/\mathbb{G}(\mathbb{A})^+$. We will avoid giving more details here about this issue (the interested reader may consult AKA, LUETHI, MICHEL, and WIESER (2020, §9.1) or KHAYUTIN, 2019, §3.3).

We discuss now each of the possible algebraic situations that may occur separately, explain what Theorem 1.1 can tell us in each situation, and say which of the problems we considered in 6.2 fit in each situation. Before starting, note that we formulated Theorem 1.1 in an S-adic setting (see EINSIEDLER and LINDENSTRAUSS, 2019, Therom 1.8 for an adelic statement), so let's restrict ourselves for simplicity to an S-adic situation where the primes p and q above are contained in S.

Case I: when the groups are non-isogenous over $\mathbb{Q}$. — When the groups $\mathbb{G}_i$ are non-isogenous over $\mathbb{Q}$, Theorem 1.1 tells us that each ergodic component of η must be invariant under a finite-index subgroup of $\mathbb{G}(\mathbb{Q}_S)$ where $\mathbb{G} = \prod_i \mathbb{G}_i$. Indeed, any stricly-contained algebraic $\mathbb{Q}$-subgroup of $\mathbb{G}$ projecting onto $\mathbb{G}_i$, must give rise to a $\mathbb{Q}$-isogeny between two factors. In other words, essentially (depending on small details as the character spectrum, for example), each ergodic component of η must be the trivial joining, and therefore also η itself. Theorem 1.1 has been applied so far in such situations in the following cases:

– The desired equidistribution in Example 6.2 follows under congruence conditions at two primes, since the non-$\mathbb{Q}$-isogenous groups SO_3 and SL_2 are considered there. See AKA, EINSIEDLER, and SHAPIRA (2016) for more details.
– The equidistribution of P_D or P_D^{shifted} from Example 6.4 follows, under the following three assumptions: congruence conditions at two primes, that individual equidistribution is assumed/known, and that the corresponding $\mathbb{G}_i$ are pairwise non-$\mathbb{Q}$-isogenous. One can find a more detailed treatment in the recent survey of LINDENSTRAUSS (2021, §4).

– Under congruence conditions at two primes p and q, the desired surjectivity of the reduction map we described in Example 6.6 follows from Theorem 1.1, with respect to finitely many primes $p_1, \ldots, p_r$, different from p and q, as in Example 6.6 . Indeed, the $\mathbb{G}_i$'s here correspond to the non-$\mathbb{Q}$-ismorphic quaternion algebras $B_{\infty,p_1}, \ldots, B_{\infty,p_r}$. See AKA, LUETHI, MICHEL, and WIESER (2020) for more details. This example shows how the fact that ergodic theory handles arbitrary products relatively well, enabling to achieve strong arithmetic applications, related to arbitrary products. As far as the author knows, the surjectivity of the above reduction maps for $r > 1$ was not known before for a *single* discriminant.

– For the six-fold product considered in (61), the algebraic groups corresponding to the first two factors are identical and non-$\mathbb{Q}$-isogeneous to ones corresponding to the last four identical factors. So, currently, without considering the corresponding embedding $\iota_{i,n}$, as in the general notation above, we can only conclude from Theorem 1.1 the disjointness of the first two factors from the last four factors. In this case, the structure of $\iota_{i,n}$ saves the day; we will discuss it further below.

Case II: when the groups are isogenous over $\mathbb{Q}$, with a non-compatible torus action. — To simplify notation, let ν denote an ergodic component of the limit measure η that we are trying to classify. In all the problems we consider here, let's assume that we have congruence conditions at p and q giving rise to $\phi \colon \mathbb{Z}^2 \to \mathbb{G}(\mathbb{Q}_S)$, a class-$\mathcal{A}'$ homomorphism, with ν being invariant under $\phi(\mathbb{Z}^2)$. Assuming that the groups are isogenous, ν might be supported on a graph of an isogeny, which implies that the invariance group of ν will be supported on a graph of an isogeny too. This invariance group must also contain $\phi(\mathbb{Z}^2)$. But, the exact structure of ϕ depends on $\iota_{i,n}$. In some cases, this rules out the possibility of being supported on a graph of isogeny, implying that ν must be (essentially) the trivial joining and therefore also η. This is the case in the following examples:

– Going back to the six-fold product considered in (61), the structure of ϕ in the third and fourth factor is the exact S-adic analogue of Example 2.1.(2), the "45°-rotation" example (strictly speaking, as (62) hints at, we actually consider here a torsor for $(\mathrm{Pic}(\mathcal{O}_d))^2$ which, together with the two congruence conditions, results in having an embedding of $\mathbb{Z}^4$, rather than $\mathbb{Z}^2$). This "45°-rotation" structure of ϕ enables to rule out rather simply[6] any joinings supported on graph of $\mathbb{Q}$-isogenies, as we explained in Example 3.3. So ν, and therefore also η must be (essentially) the trivial joining. See AKA, EINSIEDLER, and WIESER (2021) for more details.

– The equidistribution of P_D or P_D^{shifted} from Example 6.4 when $\mathbb{G}_i$ are isogenous to each other is trickier. For concreteness, consider $r = 2$ with $\mathbb{G}_1 = \mathbb{G}_2$ and with exponents $k_1 = 1$ and $k_2 = 2$. In this case individual equidistribution is known, and examining the corresponding to $\iota_{1,n}$ and $\iota_{2,n}$, the resulting class-$\mathcal{A}'$ homomorphism ϕ is the exact S-adic analogue of Example 2.1.(3), the "different

---

[6] A similar argument, as in the "two ingredients joinings theorem"- Corollary 3.1, would suffice here.

speeds" example. Therefore, also in this case, the only possibility for ν and therefore for η is (essentially) the trivial joining. See LINDENSTRAUSS (2021, §4) for more details.

Case III: when the groups are isogenous over $\mathbb{Q}$, with a compatible torus action. — The above-mentioned results were "easy" applications of Theorem 1.1. The Mixing Conjecture (Conjecture 6.5) does not fall in the above cases: for concreteness consider again P_D^{shifted} from Example 6.4 with $r = 2$, $\mathbb{G}_1 = \mathbb{G}_2 = \mathrm{SL}_2$ and with exponents $k_1 = 1$ and $k_2 = 2$, and $X := X_1 = X_2 = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, which is a case treated in the work of KHAYUTIN (2019). Let m_X denote the uniform measure on X. As we explained above, the structure of ϕ in this case, does not rule out joinings which are supported on a diagonal embedding of SL_2 into $\mathrm{SL}_2 \times \mathrm{SL}_2$. So, as far as we know, assuming congruence conditions at two primes and applying Theorem 1.1, we only get that ν, an ergodic component of a limit measure η, can either be the trivial joining or a diagonal joining. This is nonetheless very strong information! It implies the following corollary, which is utilize and generalized by Khayutin in his work:

COROLLARY 6.8. — *Let $U_n \subset X$ be positive measure sets with $m_X(U_n) \xrightarrow{n\to\infty} 0$ (e.g., a sequence of open neighbourhood of a point/of the cusp). If we can show that*

$$\eta(U_n \times U_n) \ll m_X(U_n)^{1+\epsilon_0} \tag{66}$$

for some $\epsilon_0 > 0$, then almost every ergodic component of η is the trivial joining, and so is η.

Proof. — We explained above that Theorem 1.1 implies that almost every ergodic component ν of η is either trivial or diagonal. If ν is the trivial joining, then $\nu(U_n \times U_n) \asymp m_X(U_n)^2$, and if ν is a diagonal joining, then $\nu(U_n \times U_n) \asymp m_X(U_n)$. Assume, for contradiction, that a positive proportion of the ergodic components are diagonal joinings. Then it follows that $\eta(U_n \times U_n) \gg m_X(U_n)^{1+o(1)}$ contradicting (66). □

Said differently, we believe and wish to show that $\eta(U_n \times U_n)$ decays with exponent 2. Using the joining Theorem, it is enough to establish a decay with any exponent > 1.

We explained above that in the Mixing conjecture (Conjecture 6.5) diagonal joinings correspond to measures supported on Hecke-correspondences. Khayutin strengthens Corollary 6.2.3 to consider what he calls the cross-correlation between two measures, allowing to check a condition similar to (66) for measures ν_D corresponding to P_D^{shifted}, correlated against possible measures which are supported on Hecke-correspondences. He then uses many tools (from analytic number theory and geometric invariant theory) to analyse these cross-correlations and obtain a bound similar to (66).

6.3. An application of Theorem 1.1 for simple groups with high rank

There are no published applications of Theorem 1.1 which are not related to forms of SL_2. There might be several reasons for that: as we've seen above, many classical arithmetic objects are related to toral packets of forms of SL_2. The generalization of these problems to higher dimensions normally involve invariance under a group containing unipotent elements, making Theorem 1.1 either inapplicable or superfluous.

We end this survey by sketching a new application of Theorem 1.1 with $\mathbb{G}_1$ and $\mathbb{G}_2$ being two distinct $\mathbb{Q}$-forms of PGL_3.

Let $\mathbb{G}_1 = PGL_3$ and set $\Gamma_1 = PGL_3(\mathbb{Z})$. Let $\mathbb{G}_2$ be the $\mathbb{Q}$-algebraic group associated to the group of invertible elements of a degree three (i.e., nine-dimensional) division algebra D, modulo its center. Assume that D splits over $\mathbb{R}$, that is $D \otimes \mathbb{R} = M_{3\times3}(\mathbb{R})$. Let $\mathcal{O}$ be a full order contained in $D(\mathbb{Q})$; one can show that its image in $\mathbb{G}_2(\mathbb{R})$ is an arithmetic, cocompact lattice, which we denote by Γ_2. Set $G_i = \mathbb{G}_i(\mathbb{R})$, $X_i = \Gamma_i\backslash G_i$, $X = X_1 \times X_2$, $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$, and $\Gamma = \Gamma_1 \times \Gamma_2$. Let m_{X_i} denote the Haar probability measure on X_i.

THEOREM 6.9. — *Let K_n be a sequence of totally real cubic number fields, with $\mathrm{disc}(K_n) \to \infty$, and such that for $i = 1, 2$ there exist embeddings $\iota_{i,n}$ of $\mathbb{T}_n :=$ $\mathrm{Res}_{K_n/\mathbb{Q}}\,\mathbb{G}_m/\mathbb{G}_m$ into $\mathbb{G}_i$, with $\Gamma_i\iota_{i,n}(\mathbb{T}_n(\mathbb{R}))g_i$ for some $g_i \in G_i$ being a periodic orbit, i.e., it supports a $g_i^{-1}\iota_{i,n}(\mathbb{T}_n(\mathbb{R}))g_i$-invariant probability measure $\nu_{i,n}$. We define $\iota_n = (\iota_{1,n}, \iota_{2,n})\colon \mathbb{T}_n \to \mathbb{G}$ and consider the "joined" real orbit*

$$\Gamma\iota_n(\mathbb{T}_n(\mathbb{R}))(g_1, g_2) \tag{67}$$

which is also a periodic orbit supporting a probability measure ν_n. Then, any weak-* limit of ν_n has full support. More precisely, any weak-* limit of ν_n is the product of m_{X_1} with a weak-* limit μ_2 of $\nu_{2,n}$, and m_{X_2} appears with positive proportion in the ergodic decomposition of μ_2.

Sketch of proof. — We consider first what we know about individual equidistribution. In the first factor, it is a notable result of EINSIEDLER, LINDENSTRAUSS, MICHEL, and VENKATESH (2011, Theorem 1.4) that $\nu_{1,n}$ converges to m_{X_1}. For the second factor, with easier methods, it is shown in EINSIEDLER, LINDENSTRAUSS, MICHEL, and VENKATESH (2012), that m_{X_2} must appear with positive proportion in the ergodic decomposition of any weak-* limit of $\nu_{2,n}$. Consider now a weak-* limit of ν_n, and denote it by η. Then η projects to m_{X_1} in the first factor and to a weak-* limit η_2 of $\nu_{2,n}$ in the second factor. As we explained above, the limit measure η is either invariant under a unipotent element (in which case we declare victory), or under a split $\mathbb{R}$-torus action, that is, η is a joining (of m_{X_1} with η_2) with a $\mathbb{Z}^2$-torus action (as the split torus has dimension two). Assume we are in the second case, we let μ denote one of the ergodic components of η with respect to this $\mathbb{Z}^2$-torus action. Also here we know that μ projects to m_{X_1} in the first factor and to a weak-* limit μ_2 of $\nu_{2,n}$ in the second factor. With respect to this torus

action in the second factor, we start by considering an ergodic decomposition of μ_2:

$$\mu_2 = \int \mu_2(\xi)d\omega(\xi).$$

Since m_{X_1} is ergodic with respect to the action in the first factor, and the torus action is embedded diagonally in both factors, μ admits an ergodic decomposition of the form

$$\mu = \int \mu(\xi)d\omega(\xi)$$

with $\mu(\xi)$ being a joining of m_{X_1} and $\mu_2(\xi)$. We can then proceed as follows, according to the type of $\mu_2(\xi)$:

- If $\mu_2(\xi) = m_{X_2}$, a case occurring with positive proportion, then Theorem 1.1 can be applied to $\mu(\xi)$ to conclude that $\mu(\xi) = m_{X_1} \times m_{X_2}$, since $\mathbb{G}_1$ and $\mathbb{G}_2$ are not isogenous over $\mathbb{Q}$.
- If $\mu_2(\xi) \neq m_{X_2}$ it must have zero entropy (otherwise, by the results of EINSIEDLER, KATOK, and LINDENSTRAUSS (2006) it will be the Haar measure), and since (X_1, m_{X_1}) equipped with the torus action in the first factor is a so-called K-system, it must be disjoint to any zero entropy system (see for example a recent survey by DE LA RUE, 2020, Theorem 1).

Therefore, disjointness follows in both cases and so

$$\mu = \int \mu(\xi)d\omega(\xi) = \int m_{X_1} \otimes \mu_2(\xi)d\omega(\xi) = m_{X_1} \otimes \left(\int \mu_2(\xi)d\omega(\xi) \right) = m_{X_1} \otimes \mu_2,$$

as we wanted to show. □

REFERENCES

Menny AKA and Manfred EINSIEDLER (2016). "Duke's theorem for subcollections", *Ergodic Theory Dynam. Systems* **36** (2), pp. 335–342.

Menny AKA, Manfred EINSIEDLER, and Uri SHAPIRA (2015). "Integer points and their orthogonal lattices", *arXiv preprint arXiv:1502.04209*.

——— (2016). "Integer points on spheres and their orthogonal lattices", *Invent. Math.* **206** (2), pp. 379–396.

Menny AKA, Manfred EINSIEDLER, and Andreas WIESER (2021). "Planes in four space and four associated CM points", *Duke Math. J.* (in press).

Menny AKA, Manuel LUETHI, Philippe MICHEL, and Andreas WIESER (2020). "Simultaneous supersingular reductions of CM elliptic curves", *arXiv preprint arXiv:2005.01537*.

Menny AKA, Andrea MUSSO, and Andreas WIESER (2021). "Equidistribution of rational subspaces and their shapes", *arXiv preprint arXiv:2103.05163*.

Valentin BLOMER and Farrell BRUMLEY (2020). "Simultaneous equidistribution of toric periods and fractional moments of L-functions", *arXiv preprint arXiv:2009.07093*.

Max DEURING (1941). "Die Typen der Multiplikatorenringe elliptischer Funktionenkörper", *Abh. Math. Sem. Hansischen Univ.* **14**, pp. 197–272.

William DUKE (1988). "Hyperbolic distribution problems and half-integral weight Maass forms", *Invent. Math.* **92** (1), pp. 73–90.

——— (2007). "An introduction to the Linnik problems", in: *Equidistribution in number theory, an introduction.* Vol. 237. NATO Sci. Ser. II Math. Phys. Chem. Springer, Dordrecht, pp. 197–216.

Manfred EINSIEDLER and Anatole KATOK (2005). "Rigidity of measures—the high entropy case and non-commuting foliations", in: vol. 148. Probability in mathematics, pp. 169–238.

Manfred EINSIEDLER, Anatole KATOK, and Elon LINDENSTRAUSS (2006). "Invariant measures and the set of exceptions to Littlewood's conjecture", *Ann. of Math. (2)* **164** (2), pp. 513–560.

Manfred EINSIEDLER and Elon LINDENSTRAUSS (2007). "Joinings of higher-rank diagonalizable actions on locally homogeneous spaces", *Duke Math. J.* **138** (2), pp. 203–232.

——— (2010). "Diagonal actions on locally homogeneous spaces", in: *Homogeneous flows, moduli spaces and arithmetic.* Vol. 10. Clay Math. Proc. Amer. Math. Soc., Providence, RI, pp. 155–241.

——— (2019). "Joinings of higher rank torus actions on homogeneous spaces", *Publ. Math. Inst. Hautes Études Sci.* **129**, pp. 83–127.

Manfred EINSIEDLER, Elon LINDENSTRAUSS, Philippe MICHEL, and Akshay VENKATESH (2009). "Distribution of periodic torus orbits on homogeneous spaces", *Duke Math. J.* **148** (1), pp. 119–174.

——— (2011). "Distribution of periodic torus orbits and Duke's theorem for cubic fields", *Ann. of Math. (2)* **173** (2), pp. 815–885.

——— (2012). "The distribution of closed geodesics on the modular surface, and Duke's theorem", *Enseign. Math. (2)* **58** (3-4), pp. 249–313.

Manfred EINSIEDLER and Thomas WARD (2011). *Ergodic theory with a view towards number theory.* Vol. 259. Graduate Texts in Mathematics. Springer-Verlag London, Ltd., London, pp. xviii+481. ISBN: 978-0-85729-020-5.

Jordan S. ELLENBERG, Philippe MICHEL, and Akshay VENKATESH (2013). "Linnik's ergodic method and the distribution of integer points on spheres", in: *Automorphic representations and L-functions.* Vol. 22. Tata Inst. Fundam. Res. Stud. Math. Tata Inst. Fund. Res., Mumbai, pp. 119–185.

Jordan S. ELLENBERG, Lillian B. PIERCE, and Melanie Matchett WOOD (2017). "On ℓ-torsion in class groups of number fields", *Algebra Number Theory* **11** (8), pp. 1739–1778.

Jordan S. ELLENBERG and Akshay VENKATESH (2008). "Local-global principles for representations of quadratic forms", *Invent. Math.* **171** (2), pp. 257–279.

Alex ESKIN, Shahar MOZES, and Nimish SHAH (1996). "Unipotent flows and counting lattice points on homogeneous varieties", *Ann. of Math. (2)* **143** (2), pp. 253–299.

Herbert FEDERER (1969). *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, pp. xiv+676.

Harry FURSTENBERG (1967). "Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation", *Math. Systems Theory* **1**, pp. 1–49.

Carl Friedrich GAUSS (1986). *Disquisitiones arithmeticae.* Translated and with a preface by Arthur A. Clarke, Revised by William C. Waterhouse, Cornelius Greither and A. W. Grootendorst and with a preface by Waterhouse. Springer-Verlag, New York, pp. xx+472. ISBN: 0-387-96254-9.

Eli GLASNER (2003). *Ergodic theory via joinings.* Vol. 101. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, pp. xii+384. ISBN: 0-8218-3372-3.

Alex GORODNIK and Hee OH (2011). "Rational points on homogeneous varieties and equidistribution of adelic periods", *Geom. Funct. Anal.* **21** (2). With an appendix by Mikhail Borovoi, pp. 319–392.

Benedict H. GROSS (1987). "Heights and the special values of L-series", in: *Number theory (Montreal, Que., 1985).* Vol. 7. CMS Conf. Proc. Amer. Math. Soc., Providence, RI, pp. 115–187.

Gergely HARCOS and Philippe MICHEL (2006). "The subconvexity problem for Rankin-Selberg L-functions and equidistribution of Heegner points. II", *Invent. Math.* **163** (3), pp. 581–655.

Tal HORESH and Yakov KARASIK (2020). "Equidistribution of primitive lattices in $\mathbb{R}^n$", *arXiv preprint arXiv:2012.04508.*

Henryk IWANIEC (1987). "Fourier coefficients of modular forms of half-integral weight", *Invent. Math.* **87** (2), pp. 385–401.

Ilya KHAYUTIN (2019). "Joint equidistribution of CM points", *Ann. of Math. (2)* **189** (1), pp. 145–276.

François LEDRAPPIER and Lai-Sang YOUNG (1985a). "The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula", *Ann. of Math. (2)* **122** (3), pp. 509–539.

——— (1985b). "The metric entropy of diffeomorphisms. II. Relations between entropy, exponents and dimension", *Ann. of Math. (2)* **122** (3), pp. 540–574.

Elon LINDENSTRAUSS (2006). "Invariant measures and arithmetic quantum unique ergodicity", *Ann. of Math. (2)* **163** (1), pp. 165–219.

——— (2021). "Recent progress on rigity properties of higher rank diagonalizable actions and applications", *arXiv preprint arXiv:2101.11114.*

Yuri V. LINNIK (1968). *Ergodic properties of algebraic fields.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 45. Translated from the Russian by M. S. Keane. Springer-Verlag New York Inc., New York, pp. ix+192.

Hans MAASS (1956). "Spherical functions and quadratic forms", *J. Indian Math. Soc. (N.S.)* **20**, pp. 117–162.

Gregory A. MARGULIS and George M. TOMANOV (1994). "Invariant measures for actions of unipotent groups over local fields on homogeneous spaces", *Invent. Math.* **116** (1-3), pp. 347–392.

Curtis T. MCMULLEN (2009). "Uniformly Diophantine numbers in a fixed real quadratic field", *Compos. Math.* **145** (4), pp. 827–844.

Philippe MICHEL (2004). "The subconvexity problem for Rankin-Selberg *L*-functions and equidistribution of Heegner points", *Ann. of Math. (2)* **160** (1), pp. 185–236.

Philippe MICHEL and Akshay VENKATESH (2006). "Equidistribution, *L*-functions and ergodic theory: on some problems of Yu. Linnik", in: *International Congress of Mathematicians. Vol. II.* Eur. Math. Soc., Zürich, pp. 421–457.

Shahar MOZES and Nimish SHAH (1995). "On the space of ergodic invariant measures of unipotent flows", *Ergodic Theory Dynam. Systems* **15** (1), pp. 149–159.

Vladimir PLATONOV and Andrei RAPINCHUK (1994). *Algebraic groups and number theory.* Vol. 139. Pure and Applied Mathematics. Translated from the 1991 Russian original by Rachel Rowen. Academic Press, Inc., Boston, MA, pp. xii+614. ISBN: 0-12-558180-7.

Marina RATNER (1991). "On Raghunathan's measure conjecture", *Ann. of Math. (2)* **134** (3), pp. 545–607.

——— (1995). "Raghunathan's conjectures for Cartesian products of real and *p*-adic Lie groups", *Duke Math. J.* **77** (2), pp. 275–382.

Daniel J. RUDOLPH (1979). "An example of a measure preserving map with minimal self-joinings, and applications", *J. Analyse Math.* **35**, pp. 97–122.

Thierry DE LA RUE (2020). "Joinings in Ergodic Theory", in: *Encyclopedia of Complexity and Systems Science.* Ed. by Robert A. MEYERS. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–20. ISBN: 978-3-642-27737-5.

Wolfgang M. SCHMIDT (1998). "The distribution of sublattices of $\mathbf{Z}^m$", *Monatsh. Math.* **125** (1), pp. 37–81.

Carl SIEGEL (1935). "Über die Classenzahl quadratischer Zahlkörper", ger. *Acta Arithmetica* **1** (1), pp. 83–86.

Andreas WIESER (2019). "Linnik's problems and maximal entropy methods", *Monatsh. Math.* **190** (1), pp. 153–208.

Robert J. ZIMMER (1984). *Ergodic theory and semisimple groups.* Vol. 81. Monographs in Mathematics. Birkhäuser Verlag, Basel, pp. x+209. ISBN: 3-7643-3184-4.

Menny Aka

ETH Zürich - DMATH
Rämistrasse 101
8092 Zürich, Switzerland
E-mail : menashea@math.ethz.ch