INTRODUCTION TO RANDOM WALKS ON HOMOGENEOUS SPACES

YVES BENOIST AND JEAN-FRANÇOIS QUINT

Abstract. Let a_0 and a_1 be two matrices in $\mathrm{SL}(2,\mathbb{Z})$ which span a non-solvable group. Let x_0 be an irrational point on the torus $\mathbb{T}^2$. We toss a_0 or a_1, apply it to x_0, get another irrational point x_1, do it again to x_1, get a point x_2, and again. This random trajectory is equidistributed on the torus. This phenomenon is quite general on any finite volume homogeneous space.

10^{th} Takagi Lectures. Kyoto, May 2012

Contents

1. Introduction

1.1. **Empirical measures.** Let A be a finite set of continuous transformations of a locally compact metric space X and Γ be the semi-group generated by A, i.e. the set of products $g_n \cdots g_1$ with g_i in A. For x_0 in X, we want to understand the behavior of the Γ-orbit $\Gamma x_0 := \{gx_0 \mid g \in \Gamma\}$ and to decide wether this orbit is dense or not. More precisely, we ask :

(1.1) *Can one describe all the orbit closures $\overline{\Gamma x_0}$ in X?*

We want also to get more quantitative information on the way these orbits densify in their closure. One very intuitive way to express quantitatively this densification is by using the empirical measures: let μ be a probability measure on Γ whose support is equal to A, for instance one can choose $\mu := |A|^{-1} \sum_{g\in A} \delta_g$ to be the probability measure on A which gives same weight to each element of A. We start with a point x_0 in X and we consider a trajectory

$$x_1 = g_1 x_0 \ , \ \ x_2 = g_2 x_1 \ , \ \ \dots \ , \ \ x_n = g_n x_{n-1} \ , \ \ \dots$$

where the elements g_i are chosen independently in A with law μ. The empirical measures are the probability measures

$$\nu_n := \tfrac{1}{n}(\delta_{x_0} + \delta_{x_1} + \cdots + \delta_{x_{n-1}}),$$

i.e. for every continuous function $\varphi \in \mathcal{C}(X)$, $\nu_n(\varphi)$ is the orbital average

$$\nu_n(\varphi) = \tfrac{1}{n}(\varphi(x_0) + \cdots + \varphi(x_n)).$$

We want to know, for almost every trajectory starting at x_0 :

(1.2) *Do the empirical measures ν_n converge? What is the limit?*

All the measures we will consider in this paper will be Borel measures i.e. measures on the σ-algebra of Borel subsets. We endow the set $\mathcal{M}(X)$ of finite measures on X with the weak topology: A sequence of probability measures ν_n on X converges toward a measure ν if, for any continuous compactly supported function φ on X, $\nu_n(\varphi)$ converges towards $\nu(\varphi)$.

1.2. **Stationary measures.** For every measure ν on X we define the convolution $\mu * \nu$ to be the average of translates

$$\mu * \nu = \int_A g_* \nu \, \mathrm{d}\mu(g).$$

In other terms, for every compactly supported function φ on X, one has

$$\mu * \nu(\varphi) = |A|^{-1} \sum_{g \in A} \nu(\varphi \circ g).$$

The measure ν is said to be μ-stationary if $\mu * \nu = \nu$. Intuitively, when ν is a probability measure, if you choose a point x on X with law ν and apply one step of the random walk whose jumps have law μ then the law of the new point is $\mu * \nu$. Hence the μ-stationary probability measures are the laws which are invariant under the random walk.

According to Breiman law of large numbers Proposition 3.8, the empirical measures ν_n are asymptotically stationary. More presisely Breiman law says that every weak limit ν_∞ of a subsequence of ν_n is a μ-stationary measure.

Hence Question (1.2) splits into two parts. The first part of the question is :

(1.3) *Prove there is no escape of mass for the empirical measures ν_n?*

More precisely Question (1.3) asks : Does any weak limit ν_∞ have total mass $\nu_\infty(X) = 1$? Or, equivalently, for every $\varepsilon > 0$, does there exist a compact set $K_\varepsilon \subset X$ such that, for all $n \geq 1$, one has $\nu_n(K_\varepsilon) \geq 1 - \varepsilon$. This condition is a strong recurrence property for the random walk. In many of our examples, the space X will be compact and the answer to Question (1.3) will be automatically "Yes".

The second part of the question is

(1.4) *Describe all the μ-stationary probability measures ν on X?*

A μ-stationary probability measure ν is said to be μ-ergodic if it is extremal among the μ-stationary measures. This means that the only way to write ν as an average $\nu = \frac{1}{2}(\nu' + \nu'')$ of two μ-stationary probability measures ν' and ν'' is with $\nu' = \nu'' = \nu$. Every μ-stationary measure can be decomposed as an integral average of μ-ergodic μ-stationary measure. Hence, in order to answer to Question (1.3) we may assume ν to be μ-ergodic.

The last question we would like to understand is :

(1.5) *Describe the topology of the set of μ-ergodic μ-stationary probability measures on X.*

1.3. **Two examples.** In general, one can not expect to be able to answer to these five questions. We will explain why in this survey: even when A is a single transformation, i.e. even when the dynamics is deterministic, one can not expect to get a full answer to these five questions because of the chaotic behavior of many dynamical systems.

However, even when A contains more than one transformation, we will see that in some cases one can fully answer to these five questions. In most of our examples the space X will be a homogeneous space for the action of a locally compact group G and Γ will be included in G.

We describe in this section two concrete examples for which a complete answer to these five questions has been obtained recently. These examples are special cases of a general phenomenon that we will describe in Chapter 5.

First example: X is the d-dimensional torus

$$X = \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d,$$

Γ is a subsemigroup of $\mathrm{SL}(d, \mathbb{Z})$ whose action on $\mathbb{R}^d$ is strongly irreducible i.e. such that no finite union of proper vector subspaces of $\mathbb{R}^d$ is Γ-invariant, and μ is a probability measure on Γ whose support A is finite and spans Γ. For instance one can choose $d = 2$ and

$$\mu = \tfrac{1}{2}(\delta_{a_0} + \delta_{a_1}) \text{ where } a_0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } a_1 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

A point x_0 in X is said to be rational if it belongs to $\mathbb{Q}^d/\mathbb{Z}^d$ and irrational if not. We denote by $\nu_X := \mathrm{d}x_1 \dots \mathrm{d}x_d$ the translation invariant probability on $\mathbb{T}^d$. It is called the Lebesgue probability or the Haar probability.

For this example the answer to our five questions is positive.

Theorem 1.1. *Let x_0 be an irrational point on X.*
a) The Γ-orbit Γx_0 is dense.
b) For $\mu^{\otimes \mathbb{N}}$-almost every sequence $(g_1, \dots, g_n, \dots)$ in Γ, the trajectory $x_n := g_n \cdots g_1 x_0$ equidistributes towards ν_X.
c) The sequence $\frac{1}{n}\sum_{k=0}^{n-1} \mu^{*k} * \delta_{x_0}$ converges to ν_X.
d) The only atom-free μ-stationary probability measure ν on X is ν_X.
e) Any sequence of distinct finite Γ-orbits equidistributes towards ν_X.

Point *a*) is due to Guivarc'h and Starkov [20] and Muchnik [27]. Point *c*), *d*) and *e*) are due to Bourgain, Furman, Lindenstrauss and Mozes [7], in case Γ is proximal i.e. contains matrices with a leading real eigen value of multiplicity one, and is due to [2] in general. Point *b*) is in [4].

In Point a), we note that the Γ-orbits of rational points are finite. Point b) means that, for almost every independent choices of matrices g_n with law μ, the empirical measures converge towards ν_X. In Point d), "atom-free" means "$\nu(\{x\}) = 0$ for all x in X". We note that the atomic μ-ergodic μ-stationary probability measures are supported by the Γ-orbits of rational points. Point e) means that a sequence of Γ-invariant probability measures ν_{Y_n} on distinct finite Γ-orbits Y_n always converges towards ν_X.

In this example, the semi-direct product $G := \mathrm{SL}(d,\mathbb{Z}) \ltimes \mathbb{T}^d$ acts transitively on X and the stabilizer of 0 is the group $\Lambda := \mathrm{SL}(d,\mathbb{Z})$. One has then the identification $X = G/\Lambda$.

Second example: X is the set of covolume one lattices Δ in $\mathbb{R}^d$, i.e. the set of discrete subgroups Δ of $\mathbb{R}^d$ with a $\mathbb{Z}$-basis $e_1, \ldots, e_d$ such that $\det(e_1, \ldots, e_d) = 1$. The group $G := \mathrm{SL}(d,\mathbb{R})$ of real unimodular matrices acts transitively on X and the stabilizer of the point $\mathbb{Z}^d \in X$ is the group $\Lambda := \mathrm{SL}(d,\mathbb{Z})$. Hence one has the identification

$$X = G/\Lambda = \mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z}).$$

Γ is a subsemigroup of $\mathrm{SL}(d,\mathbb{Z})$ which is Zariski dense in $\mathrm{SL}(d,\mathbb{R})$ i.e. such that the adjoint action of Γ on the Lie algebra $\mathfrak{g}$ of G is irreducible. μ is a probability measure on Γ whose support A is finite and spans Γ. For instance, as above, one can choose $d = 2$ and

$$\mu = \tfrac{1}{2}(\delta_{a_0} + \delta_{a_1}) \ \text{ where } \ a_0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \ \text{ and } \ a_1 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

A point x_0 in X is said to be rational if it is included in $\lambda\mathbb{Q}^d$ for some $\lambda > 0$, and irrational if not.

With these notation the answer to our five questions can be stated exactly in the same way as in Theorem 1.1.

Theorem 1.2. *Let x_0 be an irrational point on X.*
a) The Γ-orbit Γx_0 is dense.
b) For $\mu^{\otimes\mathbb{N}}$-almost every sequence $(g_1, \ldots, g_n, \ldots)$ in Γ, the trajectory $x_n := g_n \cdots g_1 x_0$ equidistributes towards ν_X.
c) The sequence $\frac{1}{n}\sum_{k=0}^{n-1} \mu^{*k} * \delta_{x_0}$ converges to ν_X.
d) The only atom-free μ-stationary probability measure ν on X is ν_X.
e) Any sequence of distinct finite Γ-orbits equidistributes towards ν_X.

Theorem 1.2 is a special case of more general results in [2], [3] and [4] that we will describe in Chapter 5. It is very surprising that, already in this example with $d = 2$, the proof of the topological statement Theorem 1.2.a) relies on the random walk approach. In more concrete words, the only way we are able to prove that the irrational Γ-orbits

Γx_0 are dense in X is to walk at random on these orbits and to prove that these random trajectories equidistributes towards ν_X and hence are dense in X.

The main aim of this paper is to sketch a proof of Theorem 1.1 based on [2], [3], [4] and [5]. While the method in [7] to prove Theorem 1.1.d) relies on a deep analysis of the Fourier coefficients of the stationary measure ν, our method relies on more ergodic theoretic tools like the martingale theorem. That is why it gives also a proof of Theorem 1.2.

But more importantly the aim of this paper is to recall, for a wider audience, much simpler examples for which the answers to these five questions are well-known. The intuitions and the tools behind these classical examples will be also useful to a more advanced reader who wants to understand our proof of Theorem 1.1.

In chapter 2 we recall the behavior of a few simple deterministic dynamical systems.

In chapter 3 we recall a few properties of stationary measures for a few simple non-deterministic dynamical systems.

In chapter 4 we give a short proof of the equidistribution of random trajectories on the torus based on the classification of the stationary measures. We also sketch a proof for this classification.

In chapter 5 we explain how these two examples are instances of a much more general phenomenon. This phenomenon is even true for p-adic Lie groups. We end this survey by a nice application, Proposition 5.7, to an equidistribution property for a Markov chain in the space of lattices choosing randomly at each step a lattice of index p.

2. Deterministic dynamical system

Let X be a locally compact metric space. When a random walk on X is deterministic, i.e. when the probability measure μ is a Dirac mass $\mu = \delta_g$ for some continuous transformation g of X, a μ-stationary measure ν on X is nothing but a g-invariant measure, i.e. it satisfies $\nu = g_*\nu$.

In this section we would like to recall a few basic examples of deterministic dynamical systems and we would like to describe on these examples the deep relationship between the closed invariant subsets, the statistical behavior of trajectories and the invariant probability measures.

2.1. **The doubling map.** A very simple example is the doubling map

$$m_2 : x \mapsto 2x \text{ on the circle } X = \mathbb{T} = \mathbb{R}/\mathbb{Z}.$$

2.1.1. *Invariant subsets. There are many closed m_2-invariant subsets in $\mathbb{T}$.* Let us explain why.

To construct these invariant subsets, we introduce a coding of $(\mathbb{T}, m_2)$ with the one-sided Bernoulli dynamical system (B, T) on the alphabet $A = \{0, 1\}$. This means that B is the compact space $B = A^{\mathbb{N}}$ whose elements are sequences $b = (b_1, \ldots, b_n, \ldots)$ with $b_i \in A$, and $T : B \to B$ is the shift transformation $T(b) = (b_2, \ldots, b_{n+1}, \ldots)$. Since the coding map

$$\xi : B \to X \ ; \ b \mapsto \xi(b) = \sum_{i \geq 1} b_i 2^{-i}$$

intertwines T and m_2, i.e. since one has

$$\xi \circ T = m_2 \circ \xi,$$

the image $Y = \xi(C)$ of a closed T-invariant subset is a closed m_2-invariant subset. Here are a few examples obtained that way:
$*$ $Y = \{2^n x_0 \mid n \geq 1\}$ where x_0 is rational.
$*$ $Y = \{2^n x_0 \mid n \geq 1\} \cup \{2^{-n} \mid n \geq 0\}$ where $x_0 = \sum_{i \geq 1} 2^{-i^2}$.
$*$ $Y = \{\xi(b) \mid b_i b_{i+1} = 0 \ , \ \forall i \geq 1\}$.
$*$ $Y = \xi(C)$ where C is the closure of the orbit $T^{\mathbb{N}} b_0$ of the one-sided Thue-Morse sequence $b_0 := 0110100110010110...$ for which $b_{0,n}$ is the parity of the number of 1 in the dyadic expansion of $n-1$. This last example is important because it is "minimal with zero entropy".

2.1.2. *Invariant measures. There are uncountably many different m_2-ergodic m_2-invariant probability measure μ_p on $\mathbb{T}$ with $p \in (0, 1)$.* Let us explain why.

To construct intuitively these probability measures μ_p, just write the asymptotic dyadic expansion of x as $x = \sum_{i \geq 1} b_i 2^{-i}$ with b_i in $\{0, 1\}$ and choose the coefficients b_i of this expansion independently so that $\mu_p(\{x \mid b_i = 1\}) = p$.

In a more formal way, one uses the one-sided Bernoulli dynamical system (B, β_p, T) on the alphabet (A, α_p) with $A = \{0, 1\}$ and $\alpha_p = (1-p)\delta_0 + p\delta_1$. This means that $B = A^{\mathbb{N}}$, that β_p is the product probability measure $\beta_p = \alpha_p^{\otimes \mathbb{N}}$ and that T is the shift transformation. These probability measures β_p are T-invariant i.e. $T_*(\beta_p) = \beta_p$.

Since the coding map intertwines T and m_2, the image probability measures $\mu_p := \xi_*(\beta_p)$ are m_2-invariant.

2.1.3. *Empirical measures. All the probability measures μ_p may occur as limit of empirical measures.* Let us explain why.

It is easy to check that these probability measures μ_p are m_2-ergodic. According to Birkhoff ergodic theorem, for μ_p-almost every x, the empirical measures $\frac{1}{n} \sum_{k=0}^{n-1} \delta_{2^k x}$ converge to μ_p.

Note that the orbit $\{2^n x \mid n \geq 1\}$ of such a point x is dense and that the statistical behavior of its orbits depends heavily on p. For instance the proportion of 1 in the dyadic expansion (b_i) of x tends to p. In particular these measures μ_p are supported by disjoint Borel subsets $X_p \subset \mathbb{T}$.

One can also construct points x in $\mathbb{T}$ for which the empirical measures do not have any limit and more precisely such that all the μ_p are limit of subsequences of empirical measures.

Let us end this section with an informal comment. Notice that the Bernoulli dynamical system (B, β_p, T) is also the space of trajectories for a head and tail games, a fair game when $p = \frac{1}{2}$ and an unfair game when $p \neq \frac{1}{2}$. The fact that a dynamical system can be described by this pure probabilistic game is refered as a *chaotic behavior* : there is no more order in the doubling map than in the head and tail game!

2.2. **The cat map.** Another very simple example is Arnold cat map

$$a_0 : (x_1, x_2) \mapsto (2x_1 + x_2, x_1 + x_2) \text{ on the 2-torus } X = \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2.$$

This map is invertible. We will just explain in this section that the dynamics of this example is as chaotic as the dynamics of m_2.

2.2.1. *Invariant subsets. There are many closed a_0-invariant subsets in* $\mathbb{T}^2$. Let us explain why.

To construct these invariant subsets, we will first construct a nice invariant subset Y_0 of X by Smale horseshoe construction we will then introduce a coding of (Y_0, a_0).

The matrix a_0 has two real eigenvalues $k_0^{\pm 2}$ where k_0 is the golden ratio $k_0 = \frac{1+\sqrt{5}}{2}$ and the corresponding eigenspaces $E_\pm = \mathbb{R}e_\pm \subset \mathbb{R}^2$ are given by $e_+ = (k_0, 1)$ and $e_- = (-1, k_0)$. Let $R \subset \mathbb{T}^2$ be the rectangle defined, thanks to the euclidean scalar product $\langle ., . \rangle$, by

$$R = \{v + \mathbb{Z}^2 \in \mathbb{T}^2 \mid 0 \leq \langle v, e_+ \rangle \leq k_0^{-1} \quad \text{and} \quad 0 \leq \langle v, e_- \rangle \leq k_0\}.$$

This rectangle has been chosen so that the intersection $R \cap a_0 R$ is the union of two rectangles $R_0 \cup R_1$ where

$$\begin{aligned} R_0 &= \{v + \mathbb{Z}^2 \in \mathbb{T}^2 \mid 0 \leq \langle v, e_+ \rangle \leq k_0^{-1} \text{ and } 0 \leq \langle v, e_- \rangle \leq k_0^{-1}\} \\ R_1 &= \{v + \mathbb{Z}^2 \in \mathbb{T}^2 \mid 0 \leq \langle v, e_+ \rangle \leq k_0^{-1} \quad \text{and} \quad 1 \leq \langle v, e_- \rangle \leq k_0\} \end{aligned}$$

and that both rectangles R_0 and R_1 are sitting at the extremities of both R and $a_0 R$. We choose Y_0 to be the invariant subset

$$Y_0 = \cap_{n \in \mathbb{Z}} a_0^{-n} R.$$

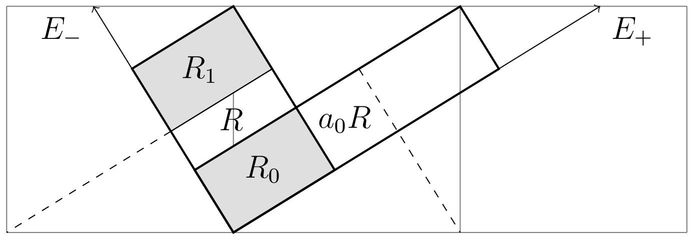

FIGURE 1. The cat map

The coding of (Y_0, a_0) uses the two-sided Bernoulli dynamical system $(\widetilde{B}, T)$ on the alphabet $A = \{0, 1\}$. This means that $\widetilde{B}$ is the compact space $\widetilde{B} = A^{\mathbb{Z}}$ whose elements are biinfinite sequences $b = (\ldots, b_n, \ldots)$ with $b_n \in A$, and $T : \widetilde{B} \to \widetilde{B}$ is the shift transformation given by $T(b) = (\ldots, b_{n+1}, \ldots)$. The coding map is the bijection ξ given by

$$\xi : \widetilde{B} \to Y_0 \ ; \ b \mapsto \xi(b) = \cap_{n\in\mathbb{Z}} a_0^{-n} R_{b_n}.$$

It intertwines T and a_0, i.e. one has

$$\xi \circ T = a_0 \circ \xi.$$

Hence the image $Y = \xi(C)$ of any closed T-invariant subset of $\widetilde{B}$ is a closed a_0-invariant subset. Here are a few examples obtained that way:
* $Y = \{\xi(b) \mid b_n = b_{n+\ell} \ , \ \forall n \in \mathbb{Z}\}$ where $\ell \geq 1$.
* $Y = \{a_0^n x_0 \mid n \geq 1\} \cup \{0\}$ where x_0 is a vertex of R.
* $Y = \{\xi(b) \mid b_n b_{n+1} = 0 \ , \ \forall n \in \mathbb{Z}\}$.
* $Y = \xi(C)$ where C is the closure of the orbit $T^{\mathbb{Z}} b_0$ of the two-sided Thue-Morse sequence $b_0 :=$ ...1001011001101001... for which $b_{0,n}$ is the parity of the number of 1 in the dyadic expansion of $\max(-n, n-1)$.

2.2.2. *Invariant measures. There are uncountably many different a_0-ergodic a_0-invariant probability measure μ_p on $\mathbb{T}^2$ with $p \in (0, 1)$.* Let us explain why.

The construction is the same as in the previous section. One can choose for μ_p the image $\mu_p := \xi_*(\beta_p)$ by the coding ξ of the product probability measure $\beta_p = \alpha_p^{\otimes\mathbb{Z}}$ where α_p is the probability $(1-p)\delta_0 + p\delta_1$ on the alphabet A. Since the probability measures β_p are T-invariant and since the coding map ξ intertwines T and a_0, the image probability measures $\mu_p := \xi_*(\beta_p)$ are a_0-invariant.

2.2.3. *Empirical measures.* As in the previous section these probability measures μ_p are a_0-ergodic. They allow us to construct points x in $\mathbb{T}^2$

with different statistical behaviors; one can construct also points x in $\mathbb{T}^2$ for which all the μ_p are limit of subsequences of empirical measures.

2.3. **Linear maps on the torus.** We recall in this section the dynamical behavior of more general linear transformations on the d-dimensional torus $X = \mathbb{T}^d$ which preserve the Haar measure ν_X.

Proposition 2.1. (Auslander) *Let $g \in \mathrm{SL}(d, \mathbb{Z})$ be a matrix with no eigenvalues being a root of unity. Then for ν_X-almost any x in X, the sequence $(g^n x)_{n\geq 1}$ is dense in X. More precisely this sequence equidistributes towards ν_X.*

For this deterministic dynamical system we do not expect equidistribution for the orbits of all the irrational points but only for almost all of them. Indeed, in the previous section, we have constructed many exceptional points i.e. irrational points whose orbit is not dense.

Proof. This proposition is a consequence of Birkhoff ergodic theorem. We only have to check that ν_X is g-ergodic i.e. the fact that,

for any Borel subset $Y \subset X$ with $g^{-1}Y = Y$, one has $\nu_X(Y) = 0$ or 1.

As an analog of the two approaches that we mentioned can be used to prove Theorem 1.1.d), we will give two proofs of this fact. The first one relying on harmonic analysis is very short. The second one relying on ergodic theory is more flexible for generalization. □

First proof. We follow [1]. Look at the Fourier coefficients of the function $\mathbf{1}_Y$,

$$c_n = \int_Y e^{-2i\pi n x}\, \mathrm{d}\nu_X(x) \ , \quad \text{where} \ \ n \in \mathbb{Z}^d.$$

Since Y is g invariant, those coefficients are constant on the orbits of g in $\mathbb{Z}^d$. Since no eigenvalues of g are roots of unity, all the orbits of g in $\mathbb{Z}^d \smallsetminus \{0\}$ are infinite. By Riemann-Lebesgue lemma these coefficients go to 0 when $|n| \to \infty$. Hence $c_n = 0$ for all $n \neq 0$, and by injectivity of the Fourier transform, the function $\mathbf{1}_Y$ is ν_X-almost everywhere constant, i.e. $\nu_X(Y) = 0$ or 1. □

Second proof. According to Birkhoff ergodic theorem, for any ν_X-integrable function $\varphi \in L^1(X, \nu_X)$, the limit

$$\varphi^*(x) = \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{k=1}^{n} \varphi(g^k x) \ \text{ exists for } \nu_X\text{-almost all } x \in X,$$

and the map $\varphi \to \varphi^*$ has norm 1 in $L^1(X, \nu_X)$: one has $\|\varphi^*\|_{L^1} \leq \|\varphi\|_{L^1}$. As a consequence one also has

$$\varphi^*(x) = \lim_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{k=n+1}^{2n} \varphi(g^k x) \ \text{ for } \nu_X\text{-almost all } x \in X.$$

The stabilizer $S_\varphi := \{y \in \mathbb{T}^d \mid \varphi(y + .) = \varphi(.)\ \ \nu_X\text{-almost surely}\}$ is a closed subgroup of $\mathbb{T}^d$. The assumption on the eigenvalues of g implies, by induction on d, that the group

$$E_- := \{v + \mathbb{Z}^d \in \mathbb{T}^d \mid \lim_{n\to\infty} g^n v = 0\} \text{ is dense in } \mathbb{T}^d.$$

By construction, when φ is continuous, the stabilizer of φ^* contains E_-, hence equals $\mathbb{T}^d$, and the function φ^* is ν_X-almost surely equal to $\nu_X(\varphi)$. Since the continuous functions are dense in $L^1(X, \nu_X)$, for any φ in $L^1(X, \nu_X)$, one also has $\varphi^* = \nu_X(\varphi)$. In particular, for our g-invariant set Y, one has $\mathbf{1}_Y = \mathbf{1}_Y^* = \nu_X(Y)$, i.e. $\nu_X(Y) = 0$ or 1. □

2.4. **Affine maps on the torus.** In contrast to the previous section we describe now two simple deterministic dynamical systems for which one can classify the invariant measures.

The first one is a translation

$$\tau_\alpha : x \mapsto x + \alpha \text{ on the circle } X = \mathbb{T} = \mathbb{R}/\mathbb{Z}.$$

When $\alpha \in \mathbb{T}$ is irrational, *there exists only one τ_α-invariant probability measure ν :* the Haar measure $\nu_X = \mathrm{d}x$.

Proof. For $n \in \mathbb{Z} \smallsetminus \{0\}$, the Fourier coefficient $c_n := \int_{\mathbb{T}} e^{-2i\pi n x}\, \mathrm{d}\nu(x)$ satisfy $c_n = e^{2i\pi n\alpha} c_n$ and hence $c_n = 0$, and $\nu = \nu_X$. □

One says that this dynamical system is uniquely ergodic. In this case all the orbits are dense and all the empirical measures converge to ν.

The second example is the transformation

$$\tau'_\alpha : (x_1, x_2) \mapsto (x_1 + \alpha, x_1 + x_2) \text{ on the 2-torus } X = \mathbb{T}^2.$$

When $\alpha \in \mathbb{T}$ is irrational, this dynamical system is also uniquely ergodic :

The only τ'_α-invariant probability ν on X is the Haar measure ν_X.

Proof. The transformations $r_t : (x_1, x_2) \mapsto (x_1, x_2 + t)$ commute with τ_α. After regularization, one can assume that ν is smooth along the orbits of the group $\{r_t \mid t \in \mathbb{T}\}$. The image of ν by the first projection $(x_1, x_2) \mapsto x_1$, which is τ_α-invariant, is the Haar measure on $\mathbb{T}$. Hence the probability ν is absolutely continuous with respect to ν_X, and its Fourier coefficients decay to zero at infinity. One can then use the same argument as in the first proof of Proposition 2.1. □

2.5. **Unipotent flow.** Here is another famous, and much more sophisticated, deterministic dynamical system for which one can classify the invariant measures.

In this example, G is a real Lie group, Λ is a lattice in G i.e. a discrete subgroup of finite covolume, $X = G/\Lambda$ and Γ is a subgroup generated by Ad-unipotent one-parameter subgroups i.e. one-parameter subgroups $u_t = e^{tN}$ where N is an element of the Lie algebra $\mathfrak{g}$ of G whose adjoint matrix ad$N$ is nilpotent. For instance

$$G = \mathrm{SL}(2, \mathbb{R}),\ \Lambda = \mathrm{SL}(2, \mathbb{Z}) \text{ and } \Gamma = \{u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mid t \in \mathbb{R}\}.$$

The following Theorem generalizes Margulis answer to the Oppenheim conjecture.

Theorem 2.2. (Ratner) *a) For every x in X, there exists a closed subgroup H of G with $\Gamma \subset H$ such that the orbit closure $Y := \overline{\Gamma x}$ is a H-orbit $Y = Hx$ and this orbit carries a H-invariant probability measure ν_Y.*
b) Every Γ-ergodic Γ-invariant probability measure ν on X is equal to one of those probability measures ν_Y.
c) Assume Γ is a Ad*-unipotent one-parameter group $\Gamma = \{u_t \mid t \in \mathbb{R}\}$. Then, for any x in X, and any bounded continuous function φ on X, the orbital averages*

$$\nu_T(\varphi) := \tfrac{1}{T} \int_0^T \varphi(u_t x)\, \mathrm{d}t$$

converge when $T \to \infty$ towards $\nu_Y(\varphi)$ with $Y = \overline{\Gamma x}$.

We omit the proof. One important point in the proof, is the fact that no mass of an orbit of a Ad-unipotent flow escape to infinity i.e. the answer to question (1.3). This point is called "Dani-Margulis recurrence phenomenon". See [11].

Another important point in the proof of b) when Γ is a Ad-unipotent flow, is called the "polynomial drift". Roughly one applies Birkhoff ergodic theorem for the u_t-invariant measure ν, one gets a set of full measure for ν, one compares how two orbits $u_t x$ and $u_t y$ of nearby points x and y on this set diverge from one another, and one deduce that ν is also invariant under another one-parameter group a_t normalizing u_t. See [29], [30], [31] and [25].

3. Stationary measures

Let X be a locally compact metric space, G be a locally compact group acting continuously on X, and μ be a probability measure on G. Let A be the support of μ and Γ be the closure of the semigroup generated by A. We recall that a probability measure ν on X is μ-stationary if one has $\mu * \nu = \nu$. We gather in this chapter a few classical properties of μ-stationary probability measures.

3.1. **Existence of stationary measures.**

Proposition 3.1. (Kakutani) *When X is compact, there always exists a μ-stationary probability measure ν on X.*

Proof. Since X is compact, the set $\mathcal{P}(X)$ of probability measures on X is also compact for the weak convergence. For any $x \in X$, we introduce the sequence of probability measures

$$\nu'_n := \tfrac{1}{n} \textstyle\sum_{k=0}^{n-1} \mu^{*k} * \delta_x .$$

By weak compactnes of $\mathcal{P}(X)$, there exists a subsequence of ν'_n which converges weakly to some probability $\nu'_\infty \in \mathcal{P}(X)$. Since

$$\mu * \nu'_n - \nu'_n = \tfrac{1}{n}(\mu^{*n} * \delta_x - \delta_x) \xrightarrow[n\to\infty]{} 0 \ ,$$

the limit ν'_∞ is μ-stationary. □

3.2. **Stationary measures on countable sets.**

Lemma 3.2. *Any μ-ergodic μ-stationary probability measure ν supported by a countable set is Γ-invariant and is supported by a finite set.*

Proof. This follows from the maximum principle : Choose a point x of maximum mass $\nu(\{x\}) = p$. The formula $\nu(\{x\}) = \int_\Gamma \nu(\{g^{-1}x\})\, \mathrm{d}\mu(g)$ implies that, for μ-almost all g, $g^{-1}x$ also has maximal mass. Hence the set of points of maximal mass which is finite is also Γ-invariant. □

3.3. **The limit probability measures.** In order to study stationary measures, we introduce the one-sided Bernoulli dynamical system (B, β, T) with alphabet (A, μ). This means that $B = A^{\mathbb{N}}$ is the space of trajectories $b = (b_1, \ldots, b_n, \ldots)$ with $b_n \in A$, that β is the product probability measure $\beta = \mu^{\otimes\mathbb{N}}$ and that T is the shift $T : b \mapsto (b_2, \ldots, b_{n+1}, \ldots)$. The following proposition tells us that the data of a stationary measure ν is equivalent to the data of an equivariant family $b \mapsto \nu_b$ of probability measures.

Proposition 3.3. (Furstenberg) *Let ν be a μ-stationary probability measure on X. Then, the limit $\nu_b := \lim\limits_{n\to\infty} (b_1 \cdots b_n)_* \nu$ exists, for β-almost any b in B,and satisfies the equivariance property :*

$$\nu_b = (b_1)_* \nu_{Tb},$$

and one can recover ν as the average

$$\nu = \int_B \nu_b \, \mathrm{d}\beta(b).$$

Proof. For the existence of the limit, apply Doob martingale theorem to the sequence $F_n : b \mapsto (b_1 \cdots b_n)_* \nu$ of $\mathcal{P}(X)$-valued functions on B. Since ν is μ-stationary, this sequence is a bounded martingale with respect to the σ-algebras $\mathcal{B}_n = \langle b_1, \cdots, b_n \rangle$. □

3.4. **Abelian group actions.** It is clear from the definition that every Γ-invariant probability measure ν on X is μ-stationary.

When the acting semigroup Γ is abelian, the converse is true :

Corollary 3.4. (Choquet, Deny) *When Γ is abelian, every μ-stationary probability measure ν on X is Γ-invariant*

Proof. There are many proofs of Choquet Deny theorem. We just explain how to deduce this theorem from Hewitt-Savage zero-one law.

Let Σ be the group of permutations σ of $\mathbb{N}$ with finite support i.e. such that $\sigma(n) = n$ for n outside a finite set. This group Σ acts on B by the formula

$$\sigma(b) = (b_{\sigma^{-1}(1)}, \ldots, b_{\sigma^{-1}(n)}, \ldots).$$

According to Hewitt-Savage Theorem, the action of Σ on (B, β) is ergodic. Since Γ is abelian, the function $b \mapsto \nu_b$ is constant on the Σ-orbits and hence ν_b is β-almost surely constant equal to its average ν. Hence for μ-almost every b_1 in Γ one has $(b_1)_*\nu = \nu$. Since the stabilizer of ν in G is closed, ν is Γ-invariant. □

However, even when $\Gamma = \mathbb{N}$ or $\mathbb{N}^2$, the classification of the Γ-invariant probability measures on X may be quite difficult. We have already seen examples for deterministic dynamical systems.

Another famous example is Furstenberg's conjecture for the doubling and tripling maps

$$m_2 : x \mapsto 2x \quad , \quad m_3 : x \mapsto 3x \qquad \text{on the circle } X = \mathbb{T}.$$

This conjecture, still open, says that the only atom-free probability measure ν on $\mathbb{T}$ which is invariant by both m_2 and m_3 is the Haar probability.

3.5. **Solvable group actions.** According to a result by Guivarc'h and Raugi, when Γ is discrete nilpotent all μ-stationary probability measures are still Γ-invariant. However this is not always the case when Γ is solvable.

Example of a μ-stationary measure which is not Γ-invariant.

We play head and tail with the two contractions

$$c_0 : x \mapsto \tfrac{x}{3} \quad , \quad c_1 : x \mapsto \tfrac{x+2}{3} \qquad \text{on the interval } X = [0,1]$$

i.e. we choose $\mu = \frac{1}{2}(\delta_{c_0} + \delta_{c_1})$ and Γ is the semigroup spanned by c_0 and c_1. In this case the Cantor set

$$K := \{x = \textstyle\sum_{i\geq 1} 2b_i 3^{-i} \mid b_i = 0 \text{ or } 1\}$$

is a closed Γ-invariant subset. Let $\beta = \mu^{\otimes \mathbb{N}}$ be the Bernoulli measure on the space $B := \{0,1\}^{\mathbb{N}}$ and $\xi : B \to K$ be the coding map given by $\xi(b) = \sum_{i\geq 1} 2b_i 3^{-i}$. The probability measure $\nu_X := \xi_*(\beta)$ is μ-stationary but is not Γ-invariant.

Eventhough the stationary probability measure ν_X is not Γ-invariant, one can easily answer to the five questions in this example :

Lemma 3.5. *a) For all x in X, the orbit closure $\overline{\Gamma x}$ is equal to $\Gamma x \cup K$.*
b) The only μ-stationary probability measure ν on X is ν_X.
c) For all x_0 in X, for β-almost all b in B, the empirical measures $\nu_n = \frac{1}{n}(\delta_{x_0} + \cdots + \delta_{x_{n-1}})$ converge to ν_X, where $x_n = b_n \cdots b_1 x_0$.

Proof. *a)* The 2^n intervals I_w, image of $[0,1]$ by a word w of length $\ell(w) = n$ in c_0, c_1, are disjoint of size 3^{-n} and the Cantor set K is equal to the intersection $K = \cap_{n\geq 1} \cup_{\ell(w)=n} I_w$.
b) The equation $\mu^{*n} * \nu = \nu$ means that each of the intervals I_w with $\ell(w) = n$ have weight $\nu(I_w) = 2^{-n}$. This proves that $\nu = \nu_X$.

c) One may use the uniqueness of ν and Breiman law of large numbers Proposition 3.8.

One may also notice that, because of the contraction property, one has only to check our statement for one point x_0 in X. The coding map $\xi : B \to K$ intertwins the two contractions

$$C_0 : b \mapsto 0b \quad , \quad C_1 : b \mapsto 1b \qquad \text{of the space } B = \{0,1\}^{\mathbb{N}},$$

with c_0 and c_1, i.e. one has

$$\xi \circ C_0 = c_0 \circ \xi \text{ and } \xi \circ C_1 = c_1 \circ \xi$$

Then Birkhoff ergodic theorem applied to $(\widetilde{B}, \widetilde{\beta}, T^{-1})$ tells us : for β-almost every b_0 and b in B, the trajectory

$$n \mapsto (b_n, \ldots, b_1, b_{0,1}, \ldots, b_{0,n}, \ldots)$$

equidistributes towards β. Our result follows. □

By construction in this example the limit measures ν_b are Dirac masses $\nu_b = \delta_{\xi(b)}$. We say that ν is μ-proximal.

3.6. **Stationary measures on projective spaces.** The last example we would like to discuss is the linear action on the projective space. Its behavior is very similar to the solvable example in section 3.5.

In this example we set $V = \mathbb{R}^d$ and $G = \mathrm{SL}(V)$, X is the real projective space $X = \mathbb{P}(V)$, $\Gamma \subset G$ is a subsemigroup, μ is a probability measure on G whose support A spans Γ. We make two assumptions on this semigroup. First, the action of Γ on V is strongly irreducible i.e. no finite union of proper subspaces of V is Γ-invariant. Second, the action of Γ on V is proximal i.e. one can find a sequence $\gamma_n \in \Gamma$ such that $\frac{\gamma_n}{\|\gamma_n\|}$ converges to a rank one matrix π. We denote by $\Lambda_\Gamma \subset X$ the set of lines which are images $\mathrm{Im}\pi$ of such a rank one matrix.

Proposition 3.6. (Furstenberg)
a) For all x in X, the orbit closure is $\overline{\Gamma x} = \Gamma x \cup K$.
b) There exist only one μ-stationary probability measure ν_X on X.
c) For β-almost every b in B, there exists a line $V_b \in \mathbb{P}(V)$ such that any cluster point $\pi \in \mathrm{End}(V)$ of the sequence $\frac{b_1\cdots b_n}{\|b_1\cdots b_n\|}$ has image $V_b = \mathrm{Im}(\pi)$; it satisfies the equivariance property $V_b = b_1 V_{Tb}$ and the limit measures are Dirac masses given by $(\nu_X)_b = \delta_{V_b}$.
d) For all x_0 in X, for β-almost all b in B, the empirical measures $\nu_n = \frac{1}{n}(\delta_{x_0} + \cdots + \delta_{x_{n-1}})$ converge to ν_X, where $x_n = b_n \cdots b_1 x_0$.

The statement of Proposition 3.6 is motivated by Lemma 3.5. We omit the proof which is a tricky application of Proposition 3.3 combined with the proximality assumption. See [17] or [6].

In the following corollary we keep the same group G and measure μ but consider its action on the vector space $X = V$.

Corollary 3.7. *The only μ-stationary probability measure ν on V is the Dirac mass δ_0.*

Proof. Let $\check{\mu} \in \mathcal{P}(G)$ be the image of μ by the map $g \mapsto g^{-1}$. According to Proposition 3.6.c applied to the action of $\check{\mu}$ on V^*, for every v in $V \smallsetminus \{0\}$, and β-almost all b in B, the sequence $b_n \cdots b_1 v$ goes to ∞. On the other hand, the dynamical system $(b, v) \mapsto (Tb, b_1 v)$ on $B \times V$ preserves the probability measure $\beta \otimes \nu$. Hence, by Poincaré Recurrence Lemma, for ν-almost all v in V, and β-almost all b in B, this sequence $b_1 \cdots b_n v$ is recurrent. This implies $\nu(V \smallsetminus \{0\}) = 0$ as required. □

3.7. **Breiman law of large number.** We say that the action of μ on X is uniquely ergodic if there exists only one μ-stationary probability measure on X.

Proposition 3.8. (Breiman) *Assume X is compact and there exists only one μ-stationary probability measure ν on X. Then, for all x_0 in*

X, for β-almost every b in B, one has the convergence of the empirical measures $\frac{1}{n}(\delta_{x_0} + \cdots + \delta_{x_{n-1}}) \xrightarrow[n\to\infty]{} \nu$, *where* $x_n = b_n \cdots b_1 x_0$.

We recall that (B, β) is the one-sided Bernoulli space and that $\mathcal{B}_n$ is the σ-subalgebra $\mathcal{B}_n = \langle b_1, \ldots, b_n \rangle$. We will use Kolmogorov law of large numbers which says

Lemma 3.9. (Kolmogorov) *For $n \geq 1$, let $\varphi_n : B \to \mathbb{R}$ be a $\mathcal{B}_n$-measurable function such that $\mathbb{E}(\varphi_n \mid \mathcal{B}_{n-1}) = 0$, and $\sup\limits_{n\geq 1} \|\varphi_n\|_{L^2} < \infty$. Then the average $\frac{1}{n}(\varphi_1 + \cdots + \varphi_n)$ converges β-almost surely to* 0.

The classical strong law of large numbers is the special case where $\varphi_n(b) = \varphi(b_n)$ with φ a centered (square-)integrable function on A.

Proof of Lemma 3.9. The sequence $\psi_n = \sum_{k=1}^n \varphi_k / k$ is a martingale which is bounded in L^2 :

$$\mathbb{E}(\psi_n^2) = \textstyle\sum_{k=1}^n \mathbb{E}(\varphi_k^2)/k^2 \leq \left(\sum_{k\geq 1} k^{-2}\right) \sup\limits_{n\geq 1} \mathbb{E}(\varphi_n^2) < \infty.$$

Hence by Doob martingale theorem, it converges almost surely. Apply Kronecker Lemma, i.e. Abel summation method, to conclude. $\square$

Proof of Proposition 3.8. See [8]. Choose

$$\varphi_n(b) = \varphi(x_n) - \mu * \varphi(x_{n-1})$$

where φ is a continuous function on X. According to Lemma 3.9 with a shift in the indices, the average

$$\tfrac{1}{n} \textstyle\sum_{k=1}^n (\varphi(x_k) - \mu * \varphi(x_k))$$

converges β-almost surely to 0. Hence any weak limit ν_∞ of a subsequence of empirical measures $\nu_n := \frac{1}{n}(\delta_{x_0} + \cdots + \delta_{x_{n-1}})$ is μ-stationary. Since the action of μ on X is uniquely ergodic, $\mu_\infty = \nu$ is the only possibility. Since $\mathcal{P}(X)$ is weakly compact, ν_n converges weakly towards ν. $\square$

This argument is very general and is very useful even when the action of μ on X is not uniquely ergodic.

4. Random walk on the torus

In this chapter we describe the main ideas in the proof of Theorem 1.1. The space X is the d-dimensional torus $X = \mathbb{T}^d$, Γ is a subsemigroup of $\mathrm{SL}(d, \mathbb{Z})$ whose action on $\mathbb{R}^d$ is strongly irreducible, and μ is a probability measure on Γ whose support A is finite and spans Γ. If needed, we might replace μ by a convolution power μ^{*n_0}.

4.1. **Empirical measures on the torus.** We will first explain in this section how Theorem 1.1.d) implies Theorem 1.1.a) and 1.1.b).

We know that the Haar probability ν_X is the only atom-free μ-stationary probability measure on X and we want to deduce that the empirical measures ν_n of irrational points converge towards ν_X (and hence the irrational Γ-orbits are dense).

Since we know, by the proof of Breiman law of large number, that any weak limit $\nu_\infty \in \mathcal{P}(X)$ of a subsequence of ν_n is μ-stationary, we only have to check that such a measure ν_∞ has no atom. By Lemma 3.2 such an atom is on a finite orbit hence is rational with denominator say $k \geq 1$. Replacing ν by $k_*\nu$, we only have to prove that $\nu_\infty(\{0\}) = 0$. This follows from the following proposition which controls the proportion of time for the excursions near 0. We denote by d the euclidean distance on $\mathbb{T}^d$.

Proposition 4.1. *For any $\varepsilon_0 > 0$, one can find $r > 0$ such that, for all x_0 in $\mathbb{T}^d \smallsetminus \{0\}$, for β-almost every b in B,*

$$\limsup_{n\to\infty} \tfrac{1}{n}|\{k \leq n \mid d(x_k, 0) \leq r\}| \leq \varepsilon_0.$$

The trajectories of the random walk on X are parametrized by couples $(b, x) \in B \times X$. We recall that $x_k = b_k \cdots b_1 x_0$.

Proof. We fix a small radius $r_0 > 0$ and introduce the compact subset Y of $\mathbb{T}^d \smallsetminus \{0\}$

$$Y = \{x \in \mathbb{T}^d \mid d(x, 0) \geq r_0\}.$$

We introduce various times: The first return time in Y is

$$\tau_Y = \tau_Y(b, x_0) = \inf\{k > 0 \mid x_k \in Y\} \in \mathbb{N} \cup \{\infty\},$$

the n^{th} return time in Y is

$$\tau_{Y,n} = \tau_{Y,n}(b, x_0) = \inf\{k > \tau_{Y,n-1}(b, x_0) \mid x_k \in Y\},$$

the n^{th} excursion time outside Y is, when $\tau_{Y,n-1} < \infty$,

$$\sigma_{Y,n} = \sigma_{Y,n}(b, x_0) = \tau_{Y,n} - \tau_{Y,n-1},$$

so that one has the equality $\tau_{Y,n} = \sum_{p=1}^{n} \sigma_{Y,p}$.

Since A is compact, the quantity $M := \max\limits_{g\in A} \log \|g\|$ is finite. At each step of the walk the distance to 0 can not decay faster than by a factor e^{-M}. Hence if one reaches the ball $B(0, r)$ starting from Y, the excursion time must be larger than $T := M^{-1} \log \frac{r_0}{r}$. That is why the time

$$\sigma^T_{Y,n} := \sigma_{Y,n} \mathbf{1}_{\{\sigma_{Y,n} \geq T\}}$$

satisfies the upper bound for the cardinality :

$$|\{k \leq n \mid d(x_k, 0) \leq r\}| \leq \sum_{p=1}^{n} \sigma_{Y,p}^{T}$$

Now Proposition 4.1 follows from the following two lemmas □

For a function F on $B \times X$ and x in X we write $\mathbb{E}_x(F)$ for its expectation $\mathbb{E}_x(F) = \int_B F(b,x)\,\mathrm{d}\beta(b)$. We denote by $\mathbb{P}_x := \beta \otimes \delta_x$ the corresponding probability measure.

The first lemma tells us that the first return time has uniformly a finite exponential moment.

Lemma 4.2. *There exists* $\alpha > 0$ *such that* $\sup_{x\in Y} \mathbb{E}_x(e^{\alpha\tau_Y}) < \infty$.

The second lemma tells us that, if the expectations of the large first return time have a uniform bound then the orbit averages of this large first return time have asymptotically the same uniform bound.

Lemma 4.3. *If one chooses* $r > 0$ *such that the number* $T := M^{-1}\log\frac{r_0}{r}$ *satisfies* $\sup_{x\in Y} \mathbb{E}_x(\sigma_Y^T) \leq \varepsilon_0$ *then, for all* x_0 *in* $\mathbb{T}^d \smallsetminus \{0\}$, *for* β*-almost all* b *in* B, *one has* $\limsup_{n\to\infty} \frac{1}{n}\sum_{p=1}^{n} \sigma_{Y,p}^{T} \leq \varepsilon_0$.

We will be able to choose such an $r > 0$ thanks to Lemma 4.2.

Proof of Lemma 4.2. During each excursion that it spends outside Y, the random walk behaves like the linear random walk on the vector space $\mathbb{R}^d$. We will take for granted Furstenberg positivity of the first Lyapounov : after replacing μ by some power μ^{*n_0} and choosing r_0 small enough, one can find $\delta > 0$ and $a < 1$ such that the function $u(x) := d(x,0)^{-\delta}$ on $\mathbb{T}^d \smallsetminus \{0\}$ satisfies

$$\int_G u(gx)\,\mathrm{d}\mu(g) \leq a\,u(x) \quad \text{for all } \ x \notin Y.$$

By induction, using the Markov property and the fact that the sets $\{\tau_Y > n-1\}$ are $\mathcal{B}_{n-1}$-measurable, one gets the bound, for $x_0 \notin Y$ and $n \geq 1$

$$\begin{aligned} \mathbb{E}_{x_0}(u(x_n)\mathbf{1}_{\{\tau_Y > n\}}) &\leq a\,\mathbb{E}_{x_0}(u(x_{n-1})\mathbf{1}_{\{\tau_Y > n-1\}}) \leq \cdots \\ &\leq a^{n-1}\mathbb{E}_{x_0}(u(x_1)) \leq a^{n-1}e^{\delta M}u(x_0), \end{aligned}$$

and hence one has

$$\mathbb{P}_{x_0}(\{\tau_Y > n\}) \leq r_0^{\delta}\,\mathbb{E}_{x_0}(u(x_n)\mathbf{1}_{\{\tau_Y > n\}}) \leq a^{n-1}e^{\delta M}.$$

This proves that the function τ_Y has an exponential moment. □

Proof of Lemma 4.3. For $\varepsilon > 0$ we estimate, choosing $\alpha > 0$ small enough, using the Markov property, and using the bound $e^t \leq 1+t+t^2e^t$ for $t > 0$,

$$\begin{aligned} \mathbb{P}_{x_0}(\{\tau^T_{Y,n} \geq n(\varepsilon_0+\varepsilon)\}) &\leq e^{-\alpha n(\varepsilon_0+\varepsilon)}\, \mathbb{E}_{x_0}(\textstyle\prod_{p=1}^n e^{\alpha\sigma^T_{Y,p}}) \\ &\leq e^{-\alpha n(\varepsilon_0+\varepsilon)}\, (\sup_{x\in X} \mathbb{E}_x(e^{\alpha\tau^T_Y}))^n \\ &\leq e^{-\alpha n(\varepsilon_0+\varepsilon)}\, (1+\alpha\varepsilon_0+O(\alpha^2))^n \leq e^{-\alpha n\varepsilon/2}. \end{aligned}$$

Since this series converges, we can apply Borel-Cantelli Lemma. □

4.2. **Equidistribution of finite orbits.** We explain in this section how Theorem 1.1.d) implies Theorem 1.1.b) and 1.1.e).

We know that the Haar probability ν_X is the only atom-free μ-stationary probability measure on X and we want to deduce that any sequence ν_{Y_n} of invariant probability on distinct finite orbits Y_n converge towards ν_X (the proof of 1.1.b) is similar).

Since any weak limit $\nu_\infty \in \mathcal{P}(X)$ of a subsequence of ν_{Y_n} is μ-stationary, we only have to check that such a measure ν_∞ has no atom. As in the previous section, we only have to prove that $\nu_\infty(\{0\}) = 0$. This follows directly from the following corollary.

Corollary 4.4. *For any $\varepsilon_0 > 0$, one can find $r > 0$ such that, for all finite Γ-orbit $Y \subset \mathbb{T}^d \smallsetminus \{0\}$, one has $\nu_Y(B(0,r)) \leq \varepsilon_0$*

Proof. Integrating over B, the conclusion of Proposition 4.1 one gets the following statement: *For any $\varepsilon_0 > 0$, one can find $r > 0$ such that, for all x_0 in $\mathbb{T}^d \smallsetminus \{0\}$, for β-almost every b in B,*

$$\limsup_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{k=0}^{n-1} \mathbb{P}_{x_0}(\{d(x_k,0) \leq r\}) \leq \varepsilon_0.$$

Averaging this statement, for x_0 in Y, one gets $\nu_Y(B(0,r)) \leq \varepsilon_0$. □

4.3. **Stationary measures on the torus.** In this section, following the general strategy of [3], we sketch the proof of Theorem 1.1.d), i.e. the fact that the Haar probability ν_X is the only atom-free μ-stationary probability measure on X. For a rigorous proof see [3].

4.3.1. *Reduction to the Key Step.* We start with an atom-free μ-stationary μ-ergodic probability measure ν on $X = \mathbb{T}^d$, and we want to prove that $\nu = \nu_X$. In order to lighten the notations *we assume that the action of Γ on $V = \mathbb{R}^d$ is proximal,* (this is certainly the case when $d = 2$). From Furstenberg propositions 3.3 and 3.6, we know that, for β-almost every b in B, we have a limit probability measure $\nu_b \in \mathcal{P}(\mathbb{T}^d)$ and a limit line $V_b \subset \mathbb{R}^d$. The key step of the proof is

Key Step. *For β-almost all b in B, the probability ν_b is V_b-invariant.*

We mean "ν_b is invariant by translations $x \mapsto x + v$ with $v \in V_b$".

Proof of (Key Step $\Longrightarrow \nu = \nu_X$). The stabilizer of ν_b in $\mathbb{T}^d$ is a closed subgroup. According to the Key Step, its connected component S_b is a non-zero subtorus of $\mathbb{T}^d$, and one has $S_b = b_1(S_{Tb})$. Let $\mathcal{F}$ be the set of non-zero subtori of $\mathbb{T}^d$. By construction, the image η of β by the map $b \mapsto S_b$ is a μ-ergodic μ-stationary probability measure on the countable set $\mathcal{F}$. Hence by Lemma 3.2 the support of η is finite and Γ-invariant. This contradicts the strong irreducibility of Γ unless $S_b = \mathbb{T}^d$. Hence one has $\nu_b = \nu_X$. Since by Proposition 3.3, the probability ν is the average of the ν_b, one gets $\nu = \nu_X$. $\square$

In order to prove the Key Step, we need more notations. We choose a vector

$$v_b \in V_b \ , \ \|v_b\| = 1 \text{ and set } \theta(b) = \log \|b_0 v_{Tb}\|.$$

Since V_b is a line, one has then

$$b_0 v_{Tb} = \pm e^{\theta(b)} v_b \text{ for } \beta\text{-almost all } b \in B.$$

In order to lighten the notations, *we assume that we can choose these signs $\pm$ to be positive.*

4.3.2. *Non degeneracy of the ν_b.* Before checking the Key Step, we need to know

Lemma 4.5. *For β-almost all b in B, for all x in $\mathbb{T}^d$, one has*

$$\nu_b(x + V_b) = 0.$$

Proof. The main point is to prove that ν_b is atom-free. For that we check a recurrence property for the action of μ on $X \times X \smallsetminus \Delta_X$ where Δ_X is the diagonal. This property is very similar to the recurrence property for the action of μ on $\mathbb{T}^d \smallsetminus \{0\}$ in Proposition 4.1. $\square$

4.3.3. *The fibered dynamical system.* We introduce the following fibered dynamical system (B^X, β^X, T_X) which is well adapted to the study of the random walk on X

$$B^X = B \times X \ , \ \beta^X = \int_B \delta_b \otimes \nu_b \, \mathrm{d}\beta(b) \ , \ T_X(b, x) = (Tb, b_1^{-1}x).$$

We note that the probability measure β^X is T_X-invariant. Inspired by [12], we desintegrate the probability measure ν_b along the leaves $x + V_b \subset \mathbb{T}^d$ using the parametrization of the leaves

$$\mathbb{R} \to x + V_b \ ; \ t \mapsto x + t v_b.$$

For β^X-almost all (b, x) in $B \times X$, the conditional measures along these leaves give Radon measures $\sigma(b, x)$ on $\mathbb{R}$ which are well-defined modulo normalization. The Key Step can be restated as an invariance by translation of these conditional measures. We denote by τ_t the translation on $\mathbb{R}$ i.e. $\tau_t(s) = s+t$.

Key Step bis *For β^X-almost every (b, x) in $B \times X$, and $\varepsilon_0 > 0$, one can find $t \in (0, \varepsilon_0)$ such that*

$$(\tau_t)_*\sigma(b, x) = \sigma(b, x).$$

Our aim now is to sketch the proof of the Key Step bis. By construction the map σ satisfies the following two crucial properties that we plan to use :

Fact 4.6. *There exists a Borel set $E \subset B^X$, $\beta^X(E) = 1$ such that*

$$((b, x) \in E \;\; , \;\; (b, x + tv_b) \in E)) \Longrightarrow \sigma(b, x) = (\tau_t)_*\sigma(b, x + tv_b).$$

Fact 4.7. *For β^X-almost every (b, x) in $B \times X$, one has*

$$\sigma(b, x) = (e^{\theta(b)})_*\sigma(T_X(b, x)).$$

Thanks to Fact 4.6 we have roughly to find many pairs of points (b', x'), $(b', x' + tv_{b'})$ in E sitting on the same leaf and on which σ takes the same value.

As a consequence of Fact 4.7, one also has, for $n \geq 1$

$$\sigma(b, x) = (e^{\theta_n(b)})_*\sigma(T_X^n(b, x)) \tag{4.1}$$

where $\theta_n(b)$ is the Birkhoff sum

$$\theta_n(b) := \theta(b) + \cdots + \theta(T^{n-1}b).$$

4.3.4. *Piece of fibers.* Another crucial point in our approach is the fact that our fibered dynamical system is not invertible. In order to lighten the notations, *we assume that* $|A| = 2$. The fibers of T_X^n contain then 2^n elements and are parametrized by

$$\begin{array}{ccc} A^n & \longrightarrow & T_X^{-n}(T_X^n(b, x)) \\ a & \mapsto & h_{n,b,x}(a) := (aT^n b, a_1 \cdots a_n b_n^{-1} \cdots b_1^{-1} x) \end{array}$$

where $aT^n b := (a_1, \ldots, a_n, b_{n+1}, b_{n+2}, \ldots) \in B$. According to Formula (4.1), one will be able to control the function σ on the "piece of fibers"

$$A_{n,b} := \{a \in A^n \mid \;\; |\theta_n(aT^n b) - \theta_n(b)| < 1\}.$$

Eventhough this piece of fiber is very small, one has

Lemma 4.8. (Equidistribution of pieces of fibers) *Let $K \subset B^X$ be a Borel subset. Then for β^X-almost all $(b,x) \in B^X$, the limit*

$$\psi_{b,x} = \lim_{n\to\infty} \frac{1}{|A_{n,b}|} \sum_{a \in A_{n,b}} \mathbf{1}_K(h_{n,b,x}(a))$$

exists and satisfies $\int_{B^X} \psi_{b,x}\, \mathrm{d}\beta(b,x) = \beta^X(K)$.

The proof of Lemma 4.8 relies on an interpretation of similar averages as conditional expectations against an increasing sequence of σ-subalgebras and on Doob martingale theorem.

Using Lusin theorem, for $\varepsilon > 0$ small, we find a compact subset $K \subset E$ with $\beta^X(K^c) < \varepsilon^2$ on which the functions θ, σ, and $(b,x) \mapsto V_b$ are continuous. Lemma 4.8 tells us that a large proportion of points of the pieces of fibers are sitting in K. Using Egoroff theorem, for $\varepsilon > 0$ small, we find a compact subset $L \subset E$ with $\beta^X(L^c) < \varepsilon$ on which the average in Lemma 4.8 is larger than $1-\varepsilon$, uniformly for $n \geq n_0$.

4.3.5. *Exponential drift.* This argument is analogous to the polynomial drift in Section 2.5, but we replace the orbit of the unipotent flow by the pieces of fibers of T_X^n and we replace Birkhoff ergodic theorem by Doob martingale theorem.

More precisely, by Lemma 4.5, for a set of full measure of points (b,x) in L, one can find a sequence $y_p = x + v_p \in X$ with $(b, y_p) \in L$, with $v_p \in V$ converging to 0, and $v_p \notin V_b$. For all p, we can wait up to some time $n = n_p$ so that

$$\theta_n(b) \| b_n^{-1} \cdots b_1^{-1} v_p \| \simeq \varepsilon_0,$$

where $\simeq$ stands for "equal up to a uniform multiplicative constant". We know that a proportion at least $1 - 2\varepsilon$ of the words $a \in A_{n,b}$ parametrize points of the two fibers which belong to K:

$$(b', x') = h_{n,b,x}(a) \in K \;\text{ and }\; (b', y') = h_{n,b,y_p}(a) \in K.$$

In order to lighten the notations, we have not explicitely writen all the dependances in p. We can write

$$y' = x' + v' \;\text{ with }\; v' = a_1 \cdots a_n b_n^{-1} \cdots b_1^{-1} v_p$$

Using twice the law of the angles below, one can control the size of this drift vector v' by the product of two norms up to a fixed multiplicative constant, i.e. one has

$$\begin{aligned} \|v'\| &\simeq \|a_1 \cdots a_n\| \, \|b_n^{-1} \cdots b_1^{-1} v_p\| \\ &\simeq e^{\theta_n(aT^n b)} \, \|b_n^{-1} \cdots b_1^{-1} v_p\| \\ &\simeq e^{\theta_n(b)} \, \|b_n^{-1} \cdots b_1^{-1} v_p\| \simeq \varepsilon_0 \end{aligned}$$

Using similar computations one can also check that the direction of v' is very near $V_{b'}$. Extracting, we can assume that $\theta_n(aT^nb)-\theta_n(b)$ converges to θ_0. Say $\theta_0=0$ to simplify notations. Passing to the limit we get points

$$(b'_\infty, x'_\infty)\in K \ , \ (b'_\infty, x'_\infty+v'_\infty)\in K \quad \text{with } \ v'_\infty = t_\infty v_{b'_\infty} \ , \ t_\infty \simeq \varepsilon_0.$$

and the equality

$$\sigma(b,x)=\sigma(b'_\infty,x'_\infty)=(\tau_{t_\infty})_*\sigma(b'_\infty,x'_\infty+v'_\infty)=(\tau_{t_\infty})_*\sigma(b,x),$$

This concludes the Key Step bis.

4.3.6. *Law of the angle.* In general when g is a matrix and w a vector one has

$$\|gw\|\simeq\|g\|\,\|w\|$$

except when w is near some "repulsing hyperplane" that we denote by $X_g^< \in \mathbb{P}(V^*)$. A key fact used in the previous exponential drift argument was a control of the norm of the drift vector $\|v'\|$ for a large proportion of the words a, up to a uniform multiplicative constant. This control follows from the following asymptotic law which allows us, for any vector w in V, to bound below, for a large proportion of the words a, the angle between $X^<_{a_1\cdots a_n}$ and w.

We denote by $\check\mu\in\mathcal{P}(G)$, the image of μ by $g\mapsto g^{-1}$ and by ν^* the unique $\check\mu$-stationary probability measure on $\mathbb{P}(V^*)$.

Lemma 4.9. (Law of the angles) *For β-almost all b in B, for all $\varphi\in C_c(\mathbb{P}(V^*))$ one has*

$$\lim_{n\to\infty}\frac{1}{|A_{n,b}|}\textstyle\sum_{a\in A_{n,b}}\varphi(X^<_{a_1\cdots a_n}) \ = \ \displaystyle\int_{\mathbb{P}(V^*)}\varphi(y)\,\mathrm{d}\nu^*(y).$$

We note that the space $X=G/\Lambda$ does not occur in the statement of the law of the angles.

If the left hand-side were an average over the whole fiber A^n, then the rough meaning of this law would just be that ν^* can be obtained by the image of $\check\mu^{\otimes\mathbb{N}}$ by Furstenberg coding map $B\to\mathbb{P}(V^*); b\mapsto\xi^*(b)$. But the left-hand side is an average over a piece $A_{n,b}$ of the fiber A^n whose relative size goes to 0.

The proof of the law of the angle requires then a precise understanding of the Birkhoff sum $\theta_n(b)$. We need to know that they satisfy, exactly like the random walk on $\mathbb{R}$, the law of the iterated logarithm, the large deviations principle, and the local limit theorem. Hence, the pioneering papers of LePage [22] and Guivarc'h-Raugi [19] on random products of matrices, see also [6], give us the main tools in order to prove our law of the angles in [3].

4.4. **Random walk on the space of lattices.** The proof of Theorem 1.1 that we have just described works also for Theorem 1.2. There is almost nothing to change. We just have to replace the torus $\mathbb{T}^d$ by the space $X := \mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z}) = G/\Lambda$ of unimodular lattices in $\mathbb{R}^d$, and the action of Γ on $\mathbb{R}^d$ by the adjoint action on the Lie algebra $V := \mathfrak{g}$ of G. In this section, we would just like to explain how to deal with two new difficulties.

4.4.1. *Empirical measures on the space of lattices.* The first new difficulty comes from the non-compactness of $\mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z})$. We have to be sure that there is no escape of mass for the empirical measures i.e. we have to answer to Question (1.3). To overcome this difficulty, we check that $\nu_\infty(X) = 1$, using a similar argument as in Proposition 4.1 with a proper function $u : X \to [0,\infty)$. This proper function u is exactly the one used by Eskin and Margulis in [14] in order to prove recurrence properties for the random walk on X. For more general spaces $X = G/\Lambda$ in the next Chapter, we will have to use a proper function u that we constructed in [5].

4.4.2. *Reduction to the Key Step.* We start with an atom-free μ-stationary μ-ergodic probability measure ν on $X = \mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z})$, and we want to prove that $\nu = \nu_X$.

One still has, for β-almost all b in B, limit probabilities $\nu_b \in \mathcal{P}(X)$ and limit lines $V_b \subset \mathfrak{g}$. These limit lines are generated by nilpotent matrices v_b. The Key Step is then exactly the same as in section 4.3.1.

Key Step. *For β-almost all b in B, the probability ν_b is V_b-invariant.*

We mean "ν_b is invariant by translations $x \mapsto \exp(v)x$ with $v \in V_b$".

The second new difficulty occurs when we want to prove the implication (*Key Step* $\Longrightarrow \nu = \nu_X$) since there are now uncountably many probability measures which are invariant and ergodic under a unipotent one-parameter subgroup.

We explain how to overcome this difficulty when $d = 2$ i.e. for $X = \mathrm{SL}(2,\mathbb{R})/\mathrm{SL}(2,\mathbb{Z})$. According to the Key Step and to Theorem 2.2, there are two cases.

In the first case, the ν_b are G-invariant β-almost surely. Since by Proposition 3.3, the probability ν is the average of the ν_b, one gets $\nu = \nu_X$ as we wanted.

In the second case, β-almost surely, the ν_b are averages of probability measures ν_Y for closed unipotent orbits $Y \subset X$. We want to show that this case does not occur. The group $G = \mathrm{SL}(2,\mathbb{R})$ acts on the set $\mathcal{F}$ of closed unipotent orbits in X. In our example, this action is transitive, hence one can write

$$\mathcal{F} = G/U_0 \text{ with } U_0 = \{u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \mid t \in \mathbb{R}\}.$$

We write the decomposition of ν_b in V_b-ergodic components :

$$\nu_b = \int_{\mathcal{F}} \nu_Y \, \mathrm{d}\eta_b(Y),$$

where $\eta_b \in \mathcal{P}(\mathcal{F})$. Because of the equivariance property in Proposition 3.3, one has the equivariance property

$$\eta_b = (b_1)_* \eta_{Tb} \text{ for } \beta\text{-almost all } b \text{ in } B.$$

Hence the average

$$\eta := \int_B \eta_b \, \mathrm{d}\beta(b)$$

is a μ-stationary probability measure on $\mathcal{F} \simeq \mathbb{R}^2 - \{0\}$, contradicting Corollary 3.7.

5. Finite volume homogeneous spaces

In this chapter, we describe with no proof a general situation in which one can compute all the orbit closures, all the limits of empirical measures, and all the stationary measures. The proofs are in [2], [3], [4] and [5].

5.1. **General Lie groups.**

Theorem 5.1. (Orbit closures) *Let G be a real Lie group, Λ be a lattice in G, $X = G/\Lambda$ and Γ be a compactly generated closed sub-semigroup of G. We assume that the Zariski closure of the adjoint semigroup $\mathrm{Ad}(\Gamma) \subset \mathrm{GL}(\mathfrak{g})$ is semisimple with no compact factor.*

For every x in X, there exists a closed subgroup H of G with $\Gamma \subset H$ such that the orbit closure $Y := \overline{\Gamma x}$ is a H-orbit $Y = Hx$ and this orbit carries a H-invariant probability measure.

We will denote by ν_Y the H-invariant probability measure on the orbit $Y = Hx$.

This result on orbit closures answers a question by Shah [32] and Margulis [24]. In case $\mathrm{Ad}\Gamma$ itself is a semisimple subgroup of $\mathrm{GL}(\mathfrak{g})$ with no compact factor, this result follows from Ratner's Theorem.

Theorem 5.1 can be strengthen in the following equidistribution result.

Theorem 5.2. (Equidistribution of trajectories) *Let G, Λ and Γ be as in Theorem 5.1. Let μ be a compactly supported Borel probability measure on Γ whose support is compact and spans a dense subsemigroup of Γ. Let $g_1, \dots, g_n, \dots$ be a sequence of independent identically distributed random elements of Γ with law μ. Then, for every x in X, almost surely,*

$$\tfrac{1}{n} \sum_{k=0}^{n-1} \delta_{g_k \cdots g_1 x} \xrightarrow[n\to\infty]{} \nu_Y \quad \text{with} \quad Y = \overline{\Gamma x}.$$

In Theorem 5.2, "almost surely" means for $\mu^{\otimes \mathbb{N}}$-almost every choice of the sequence $g_1, \dots, g_n, \dots$

This result may be seen as a random analogue of the equidistribution properties of unipotent flows on homogeneous spaces, due to Ratner [30] and Dani-Margulis [11].

As a consequence, we get the following equidistribution in law which suffices to prove Theorem 5.1.

Theorem 5.3. (Equidistribution in law) *Let G, Λ, Γ and μ be as in Theorem 5.2. Then, for every x in X, one has*

$$\tfrac{1}{n} \sum_{k=0}^{n-1} \mu^{*k} * \delta_x \xrightarrow[n\to\infty]{} \nu_Y \quad \text{with} \quad Y = \overline{\Gamma x}.$$

Intuitively $\mu^{*k} * \delta_x$ is the law at time k of a random walk starting from x with independent jumps of law μ. The left-hand side is the average of these laws before time n.

Theorem 5.4. (Stationary measures) *Let G, Λ, Γ and μ be as in Theorem 5.2. Every μ-ergodic μ-stationary probability measure ν on X is Γ-invariant and is equal to one of those probability measures ν_Y.*

The Γ-invariance of the μ-stationary measure ν was conjectured by Furstenberg who called this property "stiffness". See [18].

Our methods also allow us to describe the topology on the set $\mathcal{S}_X(\Gamma)$ of those Γ-invariant and Γ-ergodic probability measures ν_Y. As we will see in Corollary 5.6, the following Theorem 5.5 is very efficient to compute the limit of a sequence in $\mathcal{S}_X(\Gamma)$.

In order to give a simpler statement we assume that Γ has discrete centralizer.

Theorem 5.5. (Limit of stationary measures) *Let G, Λ and Γ be as in Theorem 5.1. Assume the centralizer L of Γ in G is discrete.*
a) *The set $\mathcal{S}_X(\Gamma)$ is countable and is compact.*
b) *If $(\nu_{Y_n}) \subset \mathcal{S}_X(\Gamma)$ converges to $\nu_{Y_\infty} \in \mathcal{S}_X(\Gamma)$, then, for n large, one has $Y_n \subset Y_\infty$.*
c) Every closed Γ-invariant subset of X is a finite union of orbit closures $Y_i = \overline{\Gamma x_i}$.

In particular, if (Y_n) is a sequence in $\mathcal{S}_X(\Gamma)$ such that, for any $Y \in \mathcal{S}_X(\Gamma)$ with $Y \neq X$, for all but finitely many n, one has $Y_n \not\subset Y$, then $\nu_{Y_n} \xrightarrow[n\to\infty]{} \nu_X$, that is the orbits Y_n become equidistributed in X when n is large.

Theorem 5.5 is an analogue of the main theorem of Mozes and Shah in [26] (see also [15]) which asserts, in case G is a real Lie group, if $\mathcal{E}$ is the space of finite volume homogeneous subsets of X which are invariant and ergodic under some Ad-unipotent one-parameter subgroup of G, then the set $\mathcal{E} \cup \{\delta_\infty\}$ is compact.

5.2. **Semisimple Lie groups.** Let us state a particular case of these theorems such that the answer to our five questions can be stated exactly in the same way as in Theorem 1.1.

Corollary 5.6. *We assume that G is a connected semisimple real Lie group with no compact factor, Λ is an irreducible lattice in G and Γ is a Zariski dense subgroup of G. Let x_0 be a point of X whose Γ-orbit is infinite.*
a) The Γ-orbit Γx_0 is dense.
b) For $\mu^{\otimes\mathbb{N}}$-almost every sequence $(g_1, \ldots, g_n, \ldots)$ in Γ, the trajectory $x_n := g_n \cdots g_1 x_0$ equidistributes towards ν_X.
c) The sequence $\frac{1}{n}\sum_{k=0}^{n-1} \mu^{*k} * \delta_{x_0}$ converges to ν_X.
d) The only atom-free μ-stationary probability measure ν on X is ν_X.
e) Any sequence of distinct finite Γ-orbits equidistributes towards ν_X.

The assumption "Λ irreducible" means that the image of Λ in any proper quotient G/G' by a non-compact normal subgroup G' is dense.

Corollary 5.6.e) extends previous results by Clozel, Oh and Ullmo in [10] about equidistribution of Hecke orbits (see also [16]).

5.3. **p-adic Lie groups.** These results can also be extended to products of real and p-adic Lie groups, see [3] and [4]. In this section we give a basic example in order to explain the significance of these p-adic extensions.

Last Example: X is again the set of covolume one lattices Δ in $\mathbb{R}^d$. Hence one has the identification

$$X = G/\Lambda = \mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z}).$$

We consider the following Markov chain. Let $p \geq 2$ be an integer. We start with any lattice $x_0 \in X$. This lattice contains finitely many lattices of index p. We choose uniformly at random one of them that we renormalize by an homothety of ratio $p^{-1/d}$ in order to get a new lattice of covolume 1 called $x_1 \in X$. We do it again with x_1, get a lattice x_2, and again.

Proposition 5.7. *For every x_0 in X, almost every random trajectory, x_0, x_1, x_2, ... in X equidistributes towards the Haar probability ν_X*

Proof. Since this Markov chain commutes with G, it is enough to prove this equidistribution for ν_X almost every $x_0 \in X$. Since the probability ν_X is P-invariant where P is the corresponding Markov operator. Our statement follows from Birkhoff ergodic theorem applied in the space of trajectories of this Markov chain.

We only have to check that ν_X is P-ergodic, i.e. the fact that

> *for any Borel subset $Y \subset X$ invariant by passing to any subgroup of index p, one has $\nu_X(Y) = 0$ or 1.*

In order to lighten the notations we assume that p is prime. We introduce the group $G_p = \mathrm{SL}(d, \mathbb{R}) \times \mathrm{SL}(d, \mathbb{Q}_p)$, its lattice $\Lambda_p = \mathrm{SL}(d, \mathbb{Z}[\frac{1}{p}])$ and the quotient space

$$X_p := G_p/\Lambda_p = \mathrm{SL}(d, \mathbb{R}) \times \mathrm{SL}(d, \mathbb{Z}_p)/\mathrm{SL}(d, \mathbb{Z}).$$

One has a natural G-equivariant fibration with fiber $\mathrm{SL}(d, \mathbb{Z}_p)$

$$\pi_p : X_p \to X.$$

By assumption, the set $Y_p := \pi_p^{-1}(Y)$ is invariant by translation by the diagonal matrix $g = \mathrm{diag}(p, \cdots, p, p^{1-d}) \in \mathrm{SL}(d, \mathbb{Q}_p)$. According to Moore ergodic theorem, see [23] or [33], the G_p-invariant probability measure ν_{X_p} on X_p is g-ergodic, hence $\nu_{X_p}(Y_p) = 0$ or 1. And one has also $\nu_X(Y) = 0$ or 1. □

5.4. **Conclusion.** To conclude this survey we would like to point out that improvements of the main results might be possible. Here are five questions related to the five theorems of section 5.1.

Question 1 The first question deals with Theorem 5.1. Is the description of orbit closures in Theorem 5.1 still true when the Zariski closure of $\mathrm{Ad}(\Gamma)$ is only supposed to be spanned by its one-parameter Ad-unipotent subgroups?

Question 2 The second question deals with Theorem 5.2. Is the description of empirical measures in Theorem 5.2 still true when μ is not supposed to have compact support?

Question 3 The third question deals with Theorem 5.3. Can the Cesaro average been removed from the conclusion of Theorem 5.3, i.e. does one also have $\lim\limits_{n\to\infty} \mu^{*n} * \delta_x = \nu_Y$?

Question 4 The fourth question deals with Theorem 5.4. Assume $X = \mathbb{T}^d$ and $\Gamma \subset \mathrm{SL}(d, \mathbb{Z})$. Is is true that all μ-stationary probability measures ν on X are Γ-invariant, when the Zariski closure of Γ in $\mathrm{SL}(d, \mathbb{R})$ is only supposed to be reductive?

Question 5 The fifth question deals with Theorem 5.5. Assume $X = G/\Lambda$ to be compact. Let $k \geq 1$ large enough. For all $\varphi \in \mathcal{C}^k(X)$, is there a speed, as in [13], in the convergence, $\lim_{n\to\infty} \nu_{Y_n}(\varphi) = \nu_Y(\varphi)$ i.e. is there a bound for the error term of the form $C\,\|\varphi\|_{\mathcal{C}^k}\, Vol(Y_n)^{-\alpha}$ where C and α are positive constants and where the volume $\mathrm{Vol}(Y_n)$ is computed with respect to a fixed Riemannian metric on X?

References

[1] L. Auslander, Ergodic Automorphisms, *Amer. Math. Monthly* **77** (1970), 1-19.

[2] Y. Benoist, J.-F. Quint, Mesures stationnaires et fermés invariants des espaces homogènes (I), *CRAS* (2009) and *Ann. of Math.* **174** (2011), 1111-1162.

[3] Y. Benoist, J.-F. Quint, Stationary measures and invariant subsets of homogeneous spaces (II), preprint (2011).

[4] Y. Benoist, J.-F. Quint, Stationary measures and invariant subsets of homogeneous spaces (III), preprint (2011).

[5] Y. Benoist, J.-F. Quint, Random walks on finite volume homogeneous spaces, *Invent. Math.* **187** (2012), 37-59.

[6] P. Bougerol et J. Lacroix, Products of Random Matrices with Applications to Schrödinger Operators, *PM Birkhaüser* (1985).

[7] J. Bourgain, A. Furman, E. Lindenstrauss, S. Mozes, Invariant measures and stiffness for non-abelian groups of toral automorphisms, *Journ. Am. Math. Soc.* **24** (2011), 231-280.

[8] L. Breiman, The strong law of large numbers for a class of Markov chains, *Ann. Math. Statist.* **31** (1960), 801-803.

[9] A. Bufetov, Convergence of spherical averages for actions of free groups, *Annals of Math.* **155** (2002), 929-944.

[10] L. Clozel, H. Oh, E. Ullmo, Hecke operators and equidistribution of Hecke points, *Invent. Math.* **144** (2003), 327-351.

[11] S. G. Dani, G. Margulis, Limit distributions of orbits of unipotent flows and values of quadratic forms, *Adv. Soviet Math.* **16** (1993), 91-137.

[12] M. Einsiedler, A. Katok, E. Lindenstrauss, Invariant measures and the set of exceptions to Littlewood's conjecture, *Ann. of Math.* **164** (2006), 513-560.

[13] M. Einsiedler, G. Margulis, A. Venkatesh, Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces, *Inv. Math.* **177** (2009) 137-212.

[14] A. Eskin, G. Margulis, Recurrence properties of random walks on finite volume homogeneous manifolds, *in Random walks and geometry*, W. de Gruiter (2004), 431-444.

[15] A. Eskin, S. Mozes, N. Shah, Unipotent flows and counting lattice points on homogeneous varieties, *Ann. of Math.***143** (1996), 253-299.

[16] A. Eskin, H. Oh, Ergodic theoretic proof of equidistribution of Hecke points, *Ergodic Theory Dynam. Systems* **26** (2006), 163-167.
[17] H. Furstenberg, Noncommuting random products, *Trans. Amer. Math. Soc* **108** (1963), 377-428.
[18] H. Furstenberg, Stiffness of group actions, *Lie groups and ergodic theory*, Tata Inst. Fund. Res. Stud. Math. **14** (1998), 105-117.
[19] Y. Guivarc'h, A. Raugi, Frontière de Furstenberg, propriété de contraction et théorèmes de convergence, *Zeit. Wahrsch. verw. Gebiete* **69** (1985), 187-242.
[20] Y. Guivarc'h, A. Starkov, Orbits of linear group actions, random walks on homogeneous spaces and toral automorphisms, *Ergodic Theory Dynam. Systems* **24** (2004), 767-802.
[21] F. Ledrappier, Mesures Stationnaires sur les espaces homogènes d'après Y.Benoist et JF. Quint, *Seminaire Bourbaki* **1058** (2012), 1-17.
[22] E. Le Page, Théorèmes limites pour les produits de matrices aléatoires, *LN in Math.* **928** (1982), 258-303.
[23] G. Margulis, *Discrete subgroups of semisimple Lie groups*, Springer (1991).
[24] G. Margulis, Problems and conjectures in rigidity theory, *in Mathematics: frontiers and perspectives,* Amer. Math. Soc. (2000), 161-174.
[25] G. Margulis, G. Tomanov, Invariant measures for actions of unipotent groups over local fields on homogeneous spaces, *Invent. Math.* **116** (1994), 347-392.
[26] S. Mozes, N. Shah On the space of ergodic invariant measures of unipotent flows. *Ergodic Theory Dynam. Systems* **15** (1995), 149-159.
[27] R. Muchnik, Semigroup actions on $\mathbb{T}^n$, *Geom. Dedicata* **110** (2005), 1-47.
[28] M. Raghunathan, *Discrete subgroups of Lie groups*, Springer (1972).
[29] M. Ratner, On Raghunathan's measure conjecture, *Ann. of Math.* **134** (1991), 545-607.
[30] M. Ratner, Raghunathan's topological conjecture and distributions of unipotent flows, *Duke Math. J.* **63** (1991), 235-280.
[31] M. Ratner, Raghunathan's conjectures for cartesian products of real and p-adic Lie groups, *Duke Math. J.* **72** (1995), 275-382.
[32] N. Shah, Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements, in Lie groups and ergodic theory, *Tata Inst. Fund. Res. Stud. Math.* **14** (1998), 229-271.
[33] R. Zimmer, Ergodic theory and semisimple groups, Birkhäuser (1984).

CNRS – Université Paris-Sud Bat.425, 91405 Orsay
E-mail address: yves.benoist@math.u-psud.fr

CNRS – Université Paris-Nord, LAGA, 93430 Villetaneuse
E-mail address: quint@math.univ-paris13.fr