MESURES STATIONNAIRES ET FERMÉS INVARIANTS DES ESPACES HOMOGÈNES

YVES BENOIST ET JEAN-FRANÇOIS QUINT

Résumé. Soient G un groupe de Lie réel simple, Λ un réseau de G et Γ un sous-semi-groupe Zariski dense de G. On montre que toute partie infinie Γ-invariante dans le quotient $X = G/\Lambda$ est dense.

Soit μ une probabilité sur G dont le support est compact et engendre un sous-groupe Zariski dense de G. On montre que toute probabilité μ-stationnaire sans atome sur X est G-invariante.

On montre aussi des énoncés analogues pour le tore $X = \mathbb{T}^d$.

Abstract. **Stationary measures and closed invariant subsets of homogeneous spaces.** Let G be a real simple Lie group, Λ be a lattice of G and Γ be a Zariski dense subsemigroup of G. We prove that every infinite Γ-invariant subset in the quotient $X = G/\Lambda$ is dense.

Let μ be a probability measure on G whose support is compact and spans a Zariski dense subgroup of G. We prove that every atom free μ-stationary probability measure on X is G-invariant.

We also prove similar results for the torus $X = \mathbb{T}^d$.

1. Introduction

Le but de ce texte est d'introduire un nouvel outil dans l'étude des probabilités stationnaires sur les espaces homogènes que nous appelons la « dérive exponentielle ».

1.1. Motivation et principaux résultats.

Nous l'utiliserons ici pour montrer le théorème 1.1 ci-dessous.

Théorème 1.1. *Soient G un groupe de Lie réel quasi-simple connexe, Λ un réseau de G, $X = G/\Lambda$ et μ une probabilité sur G dont le support est compact et engendre un sous-semi-groupe Zariski dense de G.*

Alors toute probabilité μ-stationnaire sans atome ν sur X est la probabilité de Haar.

Précisons tout d'abord quelques définitions bien classiques que nous avons utilisées dans cet énoncé. Un groupe de Lie réel G est dit *quasi-simple* si son algèbre de Lie est simple. Une probabilité ν sur X est dite

μ-stationnaire si on a $\mu * \nu = \nu$. Elle est dite sans atome si, pour tout x dans X, on a $\nu(\{x\}) = 0$.

Précisons aussi la définition d'un sous-semi-groupe *Zariski dense* Γ lorsque G n'est pas linéaire : cela signifie que l'image $\mathrm{Ad}(\Gamma)$ est Zariski dense dans le groupe adjoint $\mathrm{Ad}(G)$ qui est linéaire.

Il découle du théorème 1.1 que toutes ces probabilités μ sur G vérifient la propriété de « stiffness », introduite par Furstenberg dans [14], pour l'action sur X : toutes les probabilités μ-stationnaires μ-ergodiques sur X sont invariantes par le support de μ.

Les théorèmes de Ratner, dans [30], [31] ou [28], décrivent les probabilités et les fermés Γ-invariants du quotient $X = G/\Lambda$ lorsque Γ est un groupe connexe engendré par des éléments unipotents. Dans [34] p.232, N. Shah parle des extensions de ces théorèmes à des groupes non connexes comme d'un « challenging problem ». Dans [27] p.162, G. Margulis signale que de telles extensions « look extremely difficult » à cause de l'absence même de conjecture. Nous obtenons ici une extension des théorèmes de Ratner à des groupes Γ Zariski denses. Plus précisément :

Corollaire 1.2. *Soient G un groupe de Lie réel quasi-simple connexe, Λ un réseau de G, $X = G/\Lambda$ et Γ un sous-semi-groupe Zariski dense de G. Alors,*
a) Toute probabilité Γ-invariante sans atome ν sur X est la probabilité de Haar.
b) Tout fermé Γ-invariant infini F de X est égal à X.
c) Tout suite de Γ-orbites finies distinctes X_n dans X s'équirépartit vers la probabilité de Haar.

Un fermé F de X est Γ-invariant si, pour tout γ dans Γ, on a $\gamma F \subset F$.

Le point c) signifie que la suite de probabilités $\nu_n := \frac{1}{\# X_n} \sum_{x \in X_n} \delta_x$ converge faiblement vers la probabilité de Haar sur X.

L'exemple le plus simple auquel s'appliquent ce théorème et son corollaire est bien sûr pour $G = \mathrm{SL}(d, \mathbb{R})$, $\Lambda = \mathrm{SL}(d, \mathbb{Z})$ avec $d \geq 2$, $\mu = \frac{1}{2}(\delta_{g_1} + \delta_{g_2})$ et Γ le semi-groupe engendré par g_1 et g_2 lorsque celui-ci est Zariski dense. L'espace X est alors l'espace des réseaux de $\mathbb{R}^d$ de covolume 1.

Même pour $d = 2$, nos énoncés sont nouveaux dans ce cas.

Le point b) du corollaire généralise un résultat d'Eskin et Margulis dans [9] qui affirme que les fermés Γ-invariants et infinis de X ne sont pas discrets.

Le point c) du corollaire généralise le théorème d'équirépartition des orbites de Hecke dû à Clozel-Oh-Ullmo (cf. [6] et [11]).

Notre méthode s'adapte à une classe bien plus large d'espaces homogènes. Par exemple, elle permet de généraliser un résultat récent de Bourgain, Furman, Lindenstrauss et Mozes dans [3] (le résultat de [3] suppose l'existence d'éléments proximaux dans Γ).

Théorème 1.3. *Soit* Γ *un sous-semi-groupe de* $\mathrm{SL}(d,\mathbb{Z})$ *dont l'action sur* $\mathbb{R}^d$ *est fortement irréductible. Soit* μ *une probabilité sur* $\mathrm{SL}(d,\mathbb{Z})$ *dont le support est fini et engendre* Γ.

Alors toute probabilité μ*-stationnaire sans atome* ν *sur le tore* $X = \mathbb{T}^d$ *est la probabilité de Haar.*

Rappelons que l'action de Γ sur $\mathbb{R}^d$ est dite *fortement irréductible* si tous les sous-groupes d'indice fini du groupe engendré par Γ agissent de façon irréductible sur $\mathbb{R}^d$.

Remarquons que, lorsqu'une probabilité μ-stationnaire ν a des atomes, la partie non atomique de ν est encore μ-stationnaire. On peut donc lui appliquer le théorème 1.1 ou 1.3 : elle est une mesure de Haar. Quant à la partie atomique de ν, elle est, par le lemme 8.3, la somme d'une famille de mesures μ-stationnaires à support fini.

Corollaire 1.4. *Soit* Γ *un sous-semi-groupe de* $\mathrm{SL}(d,\mathbb{Z})$ *dont l'action sur* $\mathbb{R}^d$ *est fortement irréductible. Alors,*
a) Toute probabilité Γ*-invariante sans atome* ν *sur le tore* $X = \mathbb{T}^d$ *est la probabilité de Haar.*
b) Tout fermé Γ*-invariant infini* F *de* X *est égal à* X.
c) Tout suite de Γ*-orbites finies distinctes* X_n *dans* X *s'équirépartit vers la probabilité de Lebesgue.*

Le point b) de ce corollaire est dû à Muchnik dans [29] et Guivarc'h et Starkov dans [19].

L'approche de [3] est basée sur une étude subtile des coefficients de Fourier de la probabilité ν. Notre approche est purement ergodique. C'est pourquoi elle s'étend à des espaces homogènes bien plus généraux. Par exemple, le théorème 1.1, son corollaire 1.2 et leurs stratégies de démonstration sont encore valables pour des groupes de Lie quasi-simples p-adiques G sans sous-groupe ouvert distingué strict.

1.2. **Stratégie.** La preuve de ces théorèmes et de leurs corollaires occupe l'ensemble de cet article. Avant de commencer, donnons, en quelques mots, sans définitions ni formules, les grandes lignes de cette preuve.

Notre approche est basée sur une étude des marches aléatoires sur $X = G/\Lambda$ (resp. $X = \mathbb{T}^d$) induites par les marches aléatoires indépendantes de loi μ sur G (resp. sur $\mathrm{SL}(d,\mathbb{Z})$). Pour étudier ces marches aléatoires nous introduisons un système dynamique non inversible que nous notons $(B^{\tau,X}, \mathcal{B}^{\tau,X}, \beta^{\tau,X}, T_\ell^{\tau,X})$. Sans rentrer dans les détails, disons juste que notre système dynamique est fibré, de fibre X, au dessus d'une suspension $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$ du décalage de Bernoulli unilatère associé à μ et donc que l'espace $B^{\tau,X}$ est le produit $B^\tau \times X$. L'utilisation d'une telle suspension nous a été inspirée par l'article [22] de Lalley.

Ce système dynamique a deux qualités. D'une part, des formules très simples y expriment l'espérance conditionnelle $\varphi_\ell := \mathbb{E}(\varphi \mid \mathcal{Q}_\ell^{\tau,X})$, des fonctions $\mathcal{B}^{\tau,X}$-mesurables bornées φ sur $B^{\tau,X}$, relativement à la tribu $\mathcal{Q}_\ell^{\tau,X} := (T_\ell^{\tau,X})^{-1}\mathcal{B}^{\tau,X}$ des évènements postérieurs au temps ℓ. D'autre part, on a un bon contrôle de la norme des produits d'éléments de G associés aux mots qui interviennent dans ces formules d'espérance conditionnelle. Pour construire ce système dynamique, on utilise divers théorèmes classiques sur les marches aléatoires indépendantes équidistribuées de loi μ dûs en grande partie à Furstenberg : positivité du premier exposant de Lyapounov, proximalité de la marche induite sur la variété drapeau, existence de probabilités limites ν_b pour les probabilités image d'une probabilité stationnaire ν par un mot aléatoire b.

Notre argument central, que nous appelons *dérive exponentielle*, reprend l'idée de dérive de Ratner, mais, au lieu de s'appuyer sur le théorème ergodique de Birkhoff, il s'appuie sur le théorème de convergence des martingales de Doob, dont l'utilisation nous a été inspirée par l'article [4] de Bufetov. Ce théorème nous permet ici d'affirmer que la suite $\varphi_\ell(c,x)$ converge, pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$, vers $\varphi_\infty(c,x)$ où $\varphi_\infty = \mathbb{E}(\varphi \mid \mathcal{Q}_\infty^{\tau,X})$ est l'espérance conditionnelle de φ pour la tribu queue $\mathcal{Q}_\infty^{\tau,X} := \cap_{\ell\geq 0}\mathcal{Q}_\ell^{\tau,X}$. L'idée est de comparer $\varphi_\ell(c,x)$ et $\varphi_\ell(c,y)$ pour des points x et y très proches en choisissant convenablement le temps ℓ.

Pour initier notre argument de dérive, il faut montrer qu'on peut choisir, quand ν est sans atome, les points (c,x) et (c,y) hors de la même feuille stable relative au facteur $B^{\tau,X} \to B^\tau$. C'est un point crucial dans notre argumentation : il signifie, grosso modo, que l'entropie relative de notre système fibré est non nulle. Pour le vérifier, nous démontrons des phénomènes de récurrence pour la marche aléatoire dans X, analogues à ceux apparaissant dans les travaux d'Eskin et Margulis [9], et nous combinons ces phénomènes au théorème ergodique de Chacon-Ornstein.

Pour développer notre argument de dérive, il nous faut un bon contrôle sur les normes des produits de matrices aléatoires de loi μ dans l'espace vectoriel $V = \mathrm{Lie}(G)$ (resp. $V = \mathbb{R}^d$). L'existence, due à Furstenberg, d'un sous-espace attracteur limite V_b s'avère très utile.

Pour appliquer notre argument de dérive, il reste du travail : contrairement à la dérive de Ratner, notre argument de dérive ne permet d'obtenir que des propriétés d'invariance très parcellaires pour nos probabilités stationnaires. C'est pourquoi, on introduit la fonction qui, à un point (c, x), associe la mesure conditionnelle $\sigma(c, x)$ de la probabilité limite ν_c le long du sous espace limite V_c. On identifie tous ces espaces V_c construisant ainsi une action d'un seul espace vectoriel V_0, action que nous appelons *flot horocyclique* Φ_v. Ce point est important car il permet, comme dans [8], de considérer la fonction σ comme une fonction à valeurs dans un espace fixe, l'espace des mesures de Radon sur V_0 modulo normalisation. C'est à cette fonction σ que nous appliquons notre argument de dérive. Un point crucial est que cette fonction σ est $\mathcal{Q}_\infty^{\tau,X}$-mesurable. Cela résulte des relations de commutation entre Φ_v et $T_\ell^{\tau,X}$, relations analogues à celles existant entre le flot horocyclique et le flot géodésique dans le plan hyperbolique.

Cet argument de dérive prouve que la composante connexe $J(c, x)$ du stabilisateur de $\sigma(c, x)$ dans V_0 est presque sûrement non triviale. Il permet alors de voir la probabilité ν_c, et donc, ν comme une moyenne de probabilités $\nu_{c,x}$ dont chacune est invariante par un sous-espace non trivial $J(c, x)$ de V_0.

Dans le cas du tore, on en déduit que les probabilités $\nu_{c,x}$, et donc ν, sont des moyennes de probabilités portées par des sous-tores non-triviaux. Comme le support de μ agit fortement irréductiblement sur $\mathbb{R}^d$, ν est nécessairement la mesure de Lebesgue de $\mathbb{T}^d$.

Dans le cas semi-simple, on applique la théorie de Ratner pour écrire $\nu_{c,x}$ comme une moyenne de probabilités portées par des orbites de sous-groupes fermés connexes non-triviaux H de G. La G-invariance de ν s'en déduit, grâce à un phénomène d'inexistence de probabilités stationnaires sur les espaces homogènes G/H à stabilisateurs unimodulaires et non discrets.

Il est remarquable que notre argument de dérive fonctionne sans que nous ayons besoin de décrire explicitement la tribu queue $\mathcal{Q}_\infty^{\tau,X}$. Néanmoins, nous décrirons cette tribu queue dans un article à venir et utiliserons pour cela des travaux de Blanchard, Conze, Guivarch, Raugi et Rohlin : [1], [7], [18] et [33].

1.3. **Structure de l'article.** Expliquons maintenant l'organisation de cet article.

Les chapitres 2 à 5 rassemblent, parmi les constructions et les propriétés des systèmes dynamiques associés à une marche aléatoire dont nous auront besoin, celles qui sont valables dans un cadre très général.

Les chapitres 6 à 8 sont consacrés directement à l'étude des probabilités stationnaires sur les espaces $X = G/\Lambda$ et $X = \mathbb{T}^d$. Ces deux cas sont traités simultanément. Nous suggérons au lecteur de penser en première lecture que X est le tore $\mathbb{T}^2$. Presque tous les arguments que nous développons sont déjà indispensables dans ce cas !

Le but du chapitre 2 est d'obtenir des formules pour les espérances conditionnelles dans les suspensions et les fibrations au dessus de systèmes dynamiques non inversibles, dont la remarquable "loi des derniers sauts".

Le chapitre 3 est consacré à quelques propriétés des probabilités stationnaires sur un espace borélien muni d'une action borélienne : existence de probabilités limites et le très utile phénomène de récurrence hors de la diagonale.

Dans le chapitre 4 nous rappelons la construction des mesures conditionnelles d'une probabilité le long des orbites d'une action borélienne à stabilisateurs discrets.

C'est dans le chapitre 5 que nous étudions les marches aléatoires linéaires fortement irréductibles. Nous rappelons les résultats de Furstenberg et introduisons le système dynamique $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$ qui est une suspension d'un décalage de Bernoulli.

Dans le chapitre 6, nous introduisons le système dynamique fibré $(B^{\tau,X}, \mathcal{B}^{\tau,X}, \beta^{\tau,X}, T^{\tau,X})$ associé à la marche aléatoire induite sur $X = G/\Lambda$ ou $X = \mathbb{T}^d$. Nous vérifions que ces marches satisfont non seulement les phénomènes de récurrence hors de la diagonale dont nous avons besoin pour initier la dérive, mais aussi des phénomènes de récurrence hors des orbites finies qui nous seront utiles pour montrer les corollaires topologiques. Nous y décrivons aussi un phénomène d'inexistence de probabilités stationnaires sur certains espaces homogènes de groupes de Lie semi-simples qui nous servira à la conclusion de l'étude des probabilités stationnaires dans le cas $X = G/\Lambda$. Nous introduisons enfin dans ce chapitre le flot horocyclique Φ_v sur $B^{\tau,X}$ et la fonction conditionnelle horocyclique σ, et vérifions que σ est $\mathcal{Q}_\infty^{\tau,X}$-mesurable.

Dans le chapitre 7, nous présentons notre argument central de dérive et son application à la fonction $(c, x) \mapsto \sigma(c, x)$.

Nous exploitons dans le chapitre 8 les propriétés d'invariance des probabilités stationnaires qui sont fournies par l'argument de dérive, ce qui permet de terminer la démonstration des théorèmes 1.1 et 1.3. Nous en déduisons facilement leurs corollaires 1.2 et 1.4.

Remerciements. Nous remercions Y. Hu, F. Ledrappier, H. Oh, R. Spatzier et J.-P. Thouvenot pour d'intéressantes discussions sur ce sujet et le département de Mathématiques de l'Université de Brown pour son hospitalité.

2. Suspensions et Extensions

Le but de ce chapitre est d'obtenir des formules pour les espérances conditionnelles contre les tribus queues dans les suspensions et les fibrations au dessus de systèmes dynamiques non inversibles (proposition 2.3 et lemme 2.5).

2.1. **Fonctions cohomologues.**

Le lemme suivant permettra de nous restreindre à des suspensions dont la fonction « toit » est positive.

Lemme 2.1. *Soit $(B, \mathcal{B}, \beta)$ un espace de probabilité de Lebesgue, muni d'une transformation T préservant la mesure et ergodique. Soit $\theta : B \to \mathbb{R}$ une fonction intégrable d'intégrale $\int_B \theta \, \mathrm{d}\beta > 0$. Alors, il existe une fonction positive φ presque sûrement finie, et une fonction positive τ intégrable telle que*

$$\theta - \varphi \circ T + \varphi = \tau.$$

Cette fonction τ peut être choisie minorée par une contante $\varepsilon_0 > 0$. La fonction τ peut être choisie bornée dès que θ l'est.

Autrement dit, la fonction θ est cohomologue à τ via φ.

Démonstration. Notons, pour $p \geq 1$, $\theta_p = \theta + \ldots + \theta \circ T^{p-1}$ et

$$\psi = \min_{p \geq 1} \theta_p, \quad \tau = \max(\psi, 0) \ \text{ et } \ \varphi = -\min(\psi, 0).$$

D'après le théorème ergodique de Birkhoff, pour β-presque tout b dans B, $\theta_p(b) \xrightarrow[p \to \infty]{} \infty$. Donc la fonction φ est presque sûrement finie. Par ailleurs, comme $\psi \leq \theta$, on a $\tau \leq \max(\theta, 0)$ et donc, τ est intégrable. Enfin, par définition, on a

$$\tau - \varphi = \psi = \min(\theta, \theta + \psi \circ T) = \theta + \min(0, \psi \circ T) = \theta - \varphi \circ T.$$

Pour obtenir τ minoré par $\varepsilon_0 > 0$, on applique le raisonnement précédent à la fonction $\theta - \varepsilon_0$. C'est possible dès que $\varepsilon_0 < \int_B \theta \, \mathrm{d}\beta$.

La fonction τ donnée par cette construction est bornée dès que θ l'est. $\square$

2.2. **Suspension d'un système non inversible.**

Nous définissons dans cette section la suspension de systèmes dynamiques dont la fonction toit a une composante à valeurs dans un groupe compact.

Soient $(B, \mathcal{B}, \beta)$ un espace de probabilité de Lebesgue muni d'une transformation mesurable T préservant la probabilité β. Soient M un groupe topologique compact métrisable et

$$\tau = (\tau_{\mathbb{R}}, \tau_M) : B \to \mathbb{R} \times M$$

une application mesurable telle que $\tau_{\mathbb{R}} : B \to \mathbb{R}$ soit une fonction positive intégrable non nulle. Pour tout $p \geq 0$ et β-presque tout b dans B, on note

$$\tau_{\mathbb{R},p}(b) = \tau_{\mathbb{R}}(T^{p-1}b) + \cdots + \tau_{\mathbb{R}}(b) \quad \text{et}$$
$$\tau_{M,p}(b) = \tau_M(T^{p-1}b) \cdots \tau_M(b).$$

Définissons la suspension $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$: l'espace B^τ est

$$B^\tau = \{c = (b, k, m) \in B \times \mathbb{R} \times M \mid 0 \leq k < \tau_{\mathbb{R}}(b)\},$$

la probabilité β^τ s'obtient par normalisation de la restriction à B^τ de la mesure produit de β par la mesure de Haar de $\mathbb{R} \times M$, la tribu $\mathcal{B}^\tau$ est la complétée pour β^τ de la restriction à B^τ de la tribu produit, et pour tous $\ell \in \mathbb{R}_+$ et $c = (b, k, m)$ dans B^τ,

$$T^\tau_\ell(c) = (T^{p_\ell(c)}b, k + \ell - \tau_{\mathbb{R},p_\ell(c)}(b), \tau_{M,p_\ell(c)}m) \quad \text{où}$$
$$p_\ell(c) = \max(\{p \in \mathbb{N} \mid k + \ell - \tau_{\mathbb{R},p}(b) \geq 0\}).$$

Ce flot T^τ_ℓ n'est donc défini que pour les temps positifs.

Lemme 2.2. *Le semi-groupe $(T^\tau_\ell)_{\ell \geq 0}$ de transformations de B^τ préserve la probabilité β^τ.*

Démonstration. Le plus simple est d'éviter tout calcul en considérant $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$ comme un facteur de la suspension $(\widetilde{B}^\tau, \widetilde{\mathcal{B}}^\tau, \widetilde{\beta}^\tau, \widetilde{T}^\tau)$ de l'extension naturelle $(\widetilde{B}, \widetilde{\mathcal{B}}, \widetilde{\beta}, \widetilde{T})$ de $(B, \mathcal{B}, \beta, T)$ et de se ramener ainsi au cas où T est inversible.

Lorsque T est inversible, on peut identifier le système dynamique suspendu comme le quotient du produit $B \times \mathbb{R} \times M$ par la transformation $S : (b, k, m) \mapsto (Tb, k - \tau_{\mathbb{R}}(b), \tau_M(b)m)$. Le flot T^τ_ℓ est induit par le flot $\widetilde{T}^\tau_\ell$ défini sur $B \times \mathbb{R} \times M$ par $\widetilde{T}^\tau_\ell(b, k, m) = (b, k + \ell, m)$. Il est clair que le flot $\widetilde{T}^\tau_\ell$ préserve la mesure produit sur $B \times \mathbb{R} \times M$. Donc T^τ_ℓ préserve β^τ. $\square$

Remarquons enfin, en utilisant le théorème ergodique de Birkhoff que, pour β^τ-presque tout c dans B^τ,

$$\text{(2.1)} \qquad \lim_{p\to\infty} \tfrac{1}{p}\tau_{\mathbb{R},p}(b) = \textstyle\int_B \tau \,\mathrm{d}\beta, \quad \lim_{\ell\to\infty} \tfrac{1}{\ell}\, p_\ell(c) = (\textstyle\int_B \tau \,\mathrm{d}\beta)^{-1}.$$

2.3. **La loi des derniers sauts.**

Établissons à présent la loi des derniers sauts qui jouera un rôle crucial pour controler la « dérive » dans la section 7.1. Cette loi est une formule explicite pour l'espérance conditionnelle d'un évènement de B^τ sachant $(T_\ell^\tau)^{-1}(\mathcal{B}^\tau)$ lorsque le système de base est un décalage de Bernoulli.

Soient $(A, \mathcal{A}, \alpha)$ un espace de probabilité de Lebesgue et $(B, \mathcal{B}, \beta, T)$ le décalage de Bernoulli unilatère d'alphabet $(A, \mathcal{A}, \alpha)$, c'est-à-dire $B = A^{\mathbb{N}}$, $\beta = \alpha^{\otimes\mathbb{N}}$, $\mathcal{B}$ est la tribu produit $\mathcal{A}^{\otimes\mathbb{N}}$ complétée pour β et T est le décalage à droite, donné, pour tout $b = (b_0, b_1, \ldots)$ dans B, par $Tb = (b_1, b_2, \ldots)$. Soient M un groupe topologique compact métrisable, $\tau = (\tau_{\mathbb{R}}, \tau_M) : B \to \mathbb{R}\times M$ une application mesurable telle que $\tau_{\mathbb{R}} : B \to \mathbb{R}$ soit une fonction positive intégrable non nulle et $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$ la suspension définie en 2.2.

Nous aurons besoin de notations pour paramétrer les branches inverses de T_ℓ^τ. Pour $q \geq 0$ et a et b dans B, on note $a[q]$ le début du mot a écrit de droite à gauche $a[q] = (a_{q-1}, \ldots, a_0)$ et $a[q]b \in B$ le mot concaténé

$$a[q]b = (a_{q-1}, \ldots, a_1, a_0, b_0, b_1, \ldots, b_p, \ldots).$$

Pour $c = (b, k, m)$ dans B^τ et ℓ dans $\mathbb{R}_+$, notons $q_{\ell,c} : B \to \mathbb{N}$ et $h_{\ell,c} : B \to B^\tau$ les applications données, pour a dans B, par

$$q_{\ell,c} = \widetilde{q}_{\ell,c'} \ \text{ et } \ h_{\ell,c} = \widetilde{h}_{\ell,c'} \ \text{ où } \ c' = T_\ell^\tau(c) \ \text{ et}$$

$$\begin{aligned} \widetilde{q}_{\ell,c}(a) &= \min(\{q \in \mathbb{N} \mid k - \ell + \tau_{\mathbb{R},q}(a[q]b) \geq 0\}), \\ \widetilde{h}_{\ell,c}(a) &= (a[q]b, k - \ell + \tau_{\mathbb{R},q}(a[q]b), \tau_{M,q}(a[q]b)^{-1}m) \ \text{ avec } q = \widetilde{q}_{\ell,c}(a). \end{aligned}$$

Par le théorème de Birkhoff appliqué au décalage bilatère, pour β-presque tout a dans B, et β^τ-presque tout c dans B^τ, on a l'égalité

$$\lim_{q\to\infty} \tfrac{1}{q}\, \tau_{\mathbb{R},q}(a[q]b) = \textstyle\int_B \tau \,\mathrm{d}\beta > 0.$$

Donc la fonction $\widetilde{q}_{\ell,c}$ est presque sûrement finie et l'image de l'application $\widetilde{h}_{\ell,c}$ est la fibre $(T_\ell^\tau)^{-1}(c)$. La fonction $q_{\ell,c}$ est donc aussi presque sûrement finie. En outre, pour β-presque tout a dans B, pour tout

$q \geq 1$, la fonction $b \mapsto \tau_{\mathbb{R},q}(a[q]b)$ est β-intégrable, donc par le théorème de Birkhoff, pour β^τ-presque tout c dans B^τ, on a,

$$\lim_{p\to\infty} \tfrac{1}{p}\, \tau_{\mathbb{R},q}(a[q]T^p b) = 0$$

et donc, d'après (2.1),

$$\lim_{\ell\to\infty} q_{\ell,c}(a) = \infty. \tag{2.2}$$

Enfin, l'image de l'application $h_{\ell,c}$ est la fibre de T^τ_ℓ passant par c :

$$\{c'' \in B^\tau \mid T^\tau_\ell(c'') = T^\tau_\ell(c)\},$$

c'est-à-dire l'atome de c pour la partition associée à la tribu $(T^\tau_\ell)^{-1}(\mathcal{B}^\tau)$.

Proposition 2.3. *L'espérance conditionnelle selon la tribu $(T^\tau_\ell)^{-1}(\mathcal{B}^\tau)$ est donnée, pour toute fonction mesurable positive φ sur B^τ et pour β^τ-presque tout $c = (b, k)$ dans B^τ, par*

$$\mathbb{E}(\varphi \mid (T^\tau_\ell)^{-1}(\mathcal{B}^\tau))(c) = \int_B \varphi(h_{\ell,c}(a))\, \mathrm{d}\beta(a).$$

Autrement dit, si on regarde chaque élément de la fibre de T^τ_ℓ au-dessus du point $c' = (b', k', m') = T^\tau_\ell(c)$ de B^τ, comme obtenu en complétant le mot infini b' par un mot fini $a[q]$ écrit de droite à gauche, la loi de ce mot fini est obtenue par un tirage aléatoire indépendant de loi α des lettres a_i dans l'alphabet A, tirage stoppé au temps $q_{\ell,c}(a)$.

En particulier, si τ est bornée et si $\ell \geq \sup \tau_{\mathbb{R}}$ la loi du dernier saut a_0 est α. Plus généralement, si $\ell \geq q \sup \tau_{\mathbb{R}}$ la loi des q derniers sauts $(a_{q-1}, \ldots, a_0)$ est $\alpha^{\otimes q}$.

Démonstration. Pour alléger la preuve, on supposera M trivial et donc $\tau = \tau_{\mathbb{R}}$. Le cas général se démontre de la même façon.

Introduisons la fonction $\varphi_0(c) = \int_B \varphi(\widetilde{h}_{\ell,c}(a))\, \mathrm{d}\beta(a)$. Pour montrer que la fonction $\varphi_0 \circ T^\tau_\ell$ est l'espérance conditionnelle cherchée, il suffit de montrer que, pour toute fonction positive $\mathcal{B}^\tau$-mesurable ψ, on a l'égalité

$$\int_{B^\tau} \psi(T^\tau_\ell c)\, \varphi(c)\, \mathrm{d}\beta^\tau(c) = \int_{B^\tau} \psi(T^\tau_\ell c)\, \varphi_0(T^\tau_\ell c)\, \mathrm{d}\beta^\tau(c). \tag{2.3}$$

Pour cela, on remarque que le membre de gauche G est égal à $G =$

$$\sum_{p=0}^{\infty} \int_{B^\tau} \mathbf{1}_{\{p_\ell(c)=p\}} \psi(T^p b, k + \ell - \tau_p(b))\, \varphi(b, k)\, \mathrm{d}\beta(b)\, \mathrm{d}k.$$

Introduisons les variables $c' = (b', k') = (T^p b, k + \ell - \tau_p(b)) \in B^\tau$ et $a \in B$ tel que $a[p] = (b_0, \ldots, b_{p-1})$. On obtient, en notant $B(c', p) =$

$\{a \in B \mid \widetilde{q}_{\ell,c'}(a) = p\}$, l'égalité $G =$

$$\int_{B^\tau} \psi(b', k') \sum_{p=0}^{\infty} \int_{B(c',p)} \varphi(a[p]b', k' - \ell + \tau_p(a[p]b)) \,\mathrm{d}\beta(a)\,\mathrm{d}\beta(b')\,\mathrm{d}k',$$

d'où $G =$

$$\int_{B^\tau} \psi(c') \int_B \varphi(\widetilde{h}_{\ell,c'}(a)) \,\mathrm{d}\beta(a)\,\mathrm{d}\beta^\tau(c') = \int_{B^\tau} \psi(c')\,\varphi_0(c')\,\mathrm{d}\beta^\tau(c').$$

L'égalité (2.3) s'en déduit car T_ℓ^τ préserve la mesure β^τ. □

2.4. **Espérance conditionnelle pour un système fibré.**

Terminons ce chapitre par un lemme abstrait général de construction d'une probabilité invariante sur un système dynamique fibré et de calcul d'espérance conditionnelle.

Soient $(B, \mathcal{B})$ un espace borélien standard, i.e. isomorphe à espace métrique complet séparable muni de sa tribu borélienne, β une probabilité borélienne sur B et T un endomorphisme de B préservant β. Soient $(X, \mathcal{X})$ un espace borélien standard, $\pi : B \times X \to B$ la projection sur le premier facteur et $\widehat{T}$ une transformation mesurable de $B \times X$ telle que $\pi \circ \widehat{T} = T \circ \pi$. Écrivons donc, pour (b, x) dans $B \times X$,

$$\widehat{T}(b, x) = (Tb, \rho(b)x).$$

L'espace $\mathcal{P}(X)$ des probabilités sur $(X, \mathcal{X})$ possède une structure naturelle d'espace borélien : c'est la structure engendrée par les applications $\mathcal{P}(X) \to \mathbb{R}, \nu \mapsto \int_X \varphi \,\mathrm{d}\nu$, où $\varphi : X \to \mathbb{R}$ est une fonction borélienne bornée. Si on réalise X comme un espace métrique compact muni de sa tribu borélienne, cette structure est engendrée par les applications $\mathcal{P}(X) \to \mathbb{R}, \nu \mapsto \int_X \varphi \,\mathrm{d}\nu$, où $\varphi : X \to \mathbb{R}$ est une fonction continue. En particulier, muni de cette structure borélienne, l'espace $\mathcal{P}(X)$ est lui-même un espace borélien standard.

Donnons nous une famille $\mathcal{B}$-mesurable $B \to \mathcal{P}(X), b \mapsto \nu_b$ de probabilités sur X telle que, pour β presque tout b dans B, on a

$$\nu_{Tb} = \rho(b)_* \nu_b. \tag{2.4}$$

On note alors λ la probabilité borélienne sur $(B \times X, \mathcal{B} \otimes \mathcal{X})$ donnée, pour toute fonction borélienne positive $\varphi : B \times X \to \mathbb{R}_+$, par

$$\lambda(\varphi) = \int_B \int_X \varphi(b, x) \,\mathrm{d}\nu_b(x)\,\mathrm{d}\beta(b).$$

On écrira, en bref,

$$\lambda = \int_B \delta_b \otimes \nu_b \,\mathrm{d}\beta(b). \tag{2.5}$$

Lemme 2.4. *a) Cette probabilité λ est $\widehat{T}$-invariante et on a l'égalité $\pi_*\lambda = \beta$.*
b) Réciproquement, lorsque T est inversible, toute probabilité $\widehat{T}$-invariante sur $B \times X$ telle que $\pi_*\lambda = \beta$ est donnée par (2.5) *pour une famille mesurable de probabilités $b \mapsto \nu_b$ vérifiant* (2.4)*.*

Démonstration. a) La $\widehat{T}$-invariance de λ est le résultat d'un simple calcul. Pour une fonction $(\mathcal{B} \otimes \mathcal{X})$-mesurable $\varphi : B \times X \to \mathbb{R}_+$, on a

$$\begin{aligned}\int_{B\times X} \varphi(\widehat{T}(b,x))\,\mathrm{d}\lambda(b,x) &= \int_B \int_X \varphi(Tb, \rho(b)x)\,\mathrm{d}\nu_b(x)\,\mathrm{d}\beta(b)\\ &= \int_B \int_X \varphi(Tb, x)\,\mathrm{d}\nu_{Tb}(x)\,\mathrm{d}\beta(b)\\ &= \int_B \int_X \varphi(b, x)\,\mathrm{d}\nu_b(x)\,\mathrm{d}\beta(b)\\ &= \int_{B\times X} \varphi(b,x)\,\mathrm{d}\lambda(b,x).\end{aligned}$$

Lorsque φ ne dépend pas de x, comme les mesures ν_b sont des probabilités, on a

$$\begin{aligned}\int_{B\times X} \varphi(b)\,\mathrm{d}\lambda(b,x) &= \int_B \int_X \varphi(b)\,\mathrm{d}\nu_b(x)\,\mathrm{d}\beta(b)\\ &= \int_B \varphi(b)\,\mathrm{d}\beta(b).\end{aligned}$$

Ce qui donne bien l'égalité $\pi_*\lambda = \beta$.

b) Les probabilités ν_b sont les probabilités conditionnelles de λ le long des fibres de π. Comme T est inversible, la condition (2.4) résulte de la $\widehat{T}$-invariance de λ et de l'unicité des probabilités conditionnelles. □

Rappelons rapidement le théorème de Rohlin (voir [32]) de désintégration des mesures que nous venons d'utiliser et son lien avec l'espérance conditionnelle.

Soit η une probabilité sur un espace borélien standard $(Y, \mathcal{Y})$. Pour toute sous tribu $\mathcal{Y}' = p^{-1}(\mathcal{Z})$ de $\mathcal{Y}$ correspondant à un facteur borélien $p : (Y, \mathcal{Y}) \to (Z, \mathcal{Z})$, on note $y \mapsto \eta_y^{\mathcal{Y}'} \in \mathcal{P}(Y)$ la désintégration de η relativement à $\mathcal{Y}'$: c'est une application $\mathcal{Y}'$-mesurable telle que, pour η-presque tout y dans Y, $\eta_y^{\mathcal{Y}'}$ est portée par $p^{-1}(p(y))$ et on a

$$\eta = \int_Y \eta_y^{\mathcal{Y}'}\,\mathrm{d}\eta(y). \tag{2.6}$$

Cette application $y \mapsto \eta_y^{\mathcal{Y}'}$ est η-presque sûrement unique.

En outre, pour toute fonction positive $\mathcal{Y}$-mesurable $\varphi : Y \to \mathbb{R}_+$, pour η-presque tout y dans Y, on a

$$\mathbb{E}(\varphi \mid \mathcal{Y}')(y) = \int_B \varphi(y') \, \mathrm{d}\eta_y^{\mathcal{Y}'}(y').$$

Le lemme suivant affirme que la désintégration de λ pour le facteur $\widehat{T} : B \times X \to B \times X$ se déduit facilement de la désintégration de β pour le facteur $T : B \to B$.

Lemme 2.5. *On suppose que, pour β-presque tout b dans B l'application $\rho(b) : X \to X$ est bijective. Alors, pour toute fonction $(\mathcal{B} \otimes \mathcal{X})$-mesurable et λ-intégrable $\varphi : B \times X \to \mathbb{C}$ et pour λ-presque tout (b, x) dans $B \times X$, on a l'égalité*

$$\mathbb{E}(\varphi \mid \widehat{T}^{-1}(\mathcal{B} \otimes \mathcal{X}))(b, x) = \int_B \varphi(b', \rho(b')^{-1}\rho(b)x) \, \mathrm{d}\beta_b^{T^{-1}\mathcal{B}}(b').$$

Démonstration. Comme nous venons de le rappeler ci-dessus, pour λ-presque tout (b, x) dans $B \times X$, on a l'égalité

$$\mathbb{E}(\varphi \mid \widehat{T}^{-1}(\mathcal{B} \otimes \mathcal{X}))(b, x) = \int_{B \times X} \varphi(b', x') \, \mathrm{d}\lambda_{(b,x)}^{\widehat{T}^{-1}(\mathcal{B} \otimes \mathcal{X})}(b', x')$$

Il s'agit donc d'identifier ces probabilités $\lambda_{(b,x)}^{\widehat{T}^{-1}(\mathcal{B} \otimes \mathcal{X})}$.

Remarquons tout d'abord que, comme $\rho(b)$ est bijectif, pour λ-presque tout (b, x) dans $B \times X$, la projection π induit une bijection de la fibre $\widehat{T}^{-1}(\widehat{T}(b, x))$ sur $T^{-1}(Tb)$ dont l'inverse est donnée par $b' \mapsto (b', \rho(b')^{-1}\rho(b)x)$. Notons $\mu_{(b,x)}$ la probabilité sur $B \times X$ donnée par le deuxième membre de l'égalité cherchée :

$$\int_{B \times X} \varphi(b', x') \, \mathrm{d}\mu_{(b,x)}(b', x') = \int_B \varphi(b', \rho(b')^{-1}\rho(b)x) \, \mathrm{d}\beta_b^{T^{-1}\mathcal{B}}(b').$$

On veut montrer l'égalité, pour λ-presque tout (b, x) dans $B \times X$,

$$\lambda_{(b,x)}^{\widehat{T}^{-1}(\mathcal{B} \otimes \mathcal{X})} = \mu_{(b,x)}.$$

Pour cela, d'une part, on remarque que l'application $(b, x) \mapsto \mu_{(b,x)}$ est $\widehat{T}^{-1}(\mathcal{B} \otimes \mathcal{X})$-mesurable et que la probabilité $\mu_{(b,x)}$ est portée par $\widehat{T}^{-1}(\widehat{T}(b, x))$. D'autre part, on calcule, pour toute fonction λ-intégrable $\varphi : B \times X \to \mathbb{C}$, l'intégrale I suivante :

$$I = \int_{B \times X} \int_{B \times X} \varphi(b', x') \, \mathrm{d}\mu_{(b,x)}(b', x') \, \mathrm{d}\lambda(b, x).$$

En appliquant, pour β-presque tout b, le théorème de Fubini dans l'espace $(B\times X, \mathcal{B}\times\mathcal{X}, \beta_b^{T^{-1}\mathcal{B}}\otimes\nu_b)$, on obtient

$$I = \int_B\int_B\int_X \varphi(b', \rho(b')^{-1}\rho(b)x)\,\mathrm{d}\nu_b(x)\,\mathrm{d}\beta_b^{T^{-1}\mathcal{B}}(b')\,\mathrm{d}\beta(b).$$

D'après l'égalité (2.4), on a

$$\begin{aligned} I &= \int_B\int_B\int_X \varphi(b', \rho(b')^{-1}x)\,\mathrm{d}\nu_{Tb}(x)\,\mathrm{d}\beta_b^{T^{-1}\mathcal{B}}(b')\,\mathrm{d}\beta(b)\\ &= \int_B\int_B\int_X \varphi(b', \rho(b')^{-1}x)\,\mathrm{d}\nu_{Tb'}(x)\,\mathrm{d}\beta_b^{T^{-1}\mathcal{B}}(b')\,\mathrm{d}\beta(b).\end{aligned}$$

Enfin, en utilisant de nouveau l'égalité (2.4) puis (2.6), il vient

$$\begin{aligned} I &= \int_B\int_B\int_X \varphi(b', x)\,\mathrm{d}\nu_{b'}(x)\,\mathrm{d}\beta_b^{T^{-1}\mathcal{B}}(b')\,\mathrm{d}\beta(b)\\ &= \int_B\int_X \varphi(b, x)\,\mathrm{d}\nu_{b}(x)\,\mathrm{d}\beta(b) \;=\; \int_{B\times X}\varphi(b,x)\,\mathrm{d}\lambda(b,x).\end{aligned}$$

Par unicité de la désintégration, on a donc l'égalité $\lambda_{(b,x)}^{\widehat{T}^{-1}(\mathcal{B}\otimes\mathcal{X})} = \mu_{(b,x)}$, pour λ-presque tout (b,x) dans $B\times X$. $\square$

3. Marches aléatoires sur les G-espaces

Nous rassemblons dans cette partie quelques propriétés fondamentales des probabilités stationnaires qui sont valides dans un cadre très général.

3.1. **Probabilités stationnaires et mesure de Furstenberg.**

Nous associons, à toute probabilité stationnaire ν, un système dynamique probabilisé $(B^X, \mathcal{B}^X, \beta^X, T^X)$.

Soient G un groupe localement compact métrisable séparable, $\mathcal{G}$ sa tribu borélienne, μ une probabilité borélienne sur G et $(B, \mathcal{B}, \beta, T)$ le décalage de Bernoulli unilatère d'alphabet $(G, \mathcal{G}, \mu)$.

Soit $(X, \mathcal{X})$ un espace borélien standard muni d'une action borélienne de G. Soit ν une probabilité borélienne sur X qui est μ-stationnaire i.e. telle que $\mu * \nu = \nu$.

On note T^X la transformation de $B^X := B\times X$ donnée par, pour tout (b,x) dans B^X,

$$T^X(b,x) = (Tb, b_0^{-1}x). \tag{3.1}$$

Notons, pour $n\geq 0$, $\mathcal{B}_n$ la sous-tribu de $\mathcal{B}$ engendrée par les fonctions coordonnées b_i pour $0\leq i<n$ et $\pi : B\times X\to B$ la projection sur le premier facteur.

Lemme 3.1. *Soit ν une probabilité μ-stationnaire sur X.*
a) Il existe une unique probabilité β^X sur $(B \times X, \mathcal{B} \otimes \mathcal{X})$ telle que, pour toute fonction $\mathcal{B}_n \otimes \mathcal{X}$-mesurable bornée φ,

$$\int_{B\times X} \varphi(b,x)\,\mathrm{d}\beta^X(b,x) = \int_{B\times X} \varphi(b, b_0 \dots b_{n-1}y)\,\mathrm{d}\beta(b)\,\mathrm{d}\nu(y). \tag{3.2}$$

b) Cette probabilité β^X est T^X-invariante et on a l'égalité $\pi_*\beta^X = \beta$.

Démonstration. a) Introduisons, pour $n \geq 0$, la probabilité sur $\mathcal{B}_n \otimes \mathcal{X}$ $\beta_n^X = \int_B \delta_b \otimes (b_0 \cdots b_{n-1})_*\nu\,\mathrm{d}\beta(b)$. Comme ν est μ-stationnaire, pour tout $n \geq 0$, la probabilité β_{n+1}^X coïncide sur $\mathcal{B}_n \otimes \mathcal{X}$ avec β_n^X. D'après le théorème de Carathéodory, il existe donc une unique probabilité β^X sur $\mathcal{B} \otimes \mathcal{X}$ qui, pour tout $n \geq 0$, coïncide sur $\mathcal{B}_n \otimes \mathcal{X}$ avec β_n^X.

b) Pour tout $n \geq 0$, on a $(T^X)^{-1}(\mathcal{B}_n \otimes \mathcal{X}) \subset (\mathcal{B}_{n+1} \otimes \mathcal{X})$ et, pour toute fonction $\mathcal{B}_n \otimes \mathcal{X}$-mesurable bornée φ, par définition,

$$\begin{aligned}\int_{B\times X} \varphi(T^X(b,x))\,\mathrm{d}\beta_{n+1}^X(b,x) &= \int_{B\times X} \varphi(Tb, b_0^{-1}b_0 \dots b_n y)\,\mathrm{d}\beta(b)\,\mathrm{d}\nu(y)\\ &= \int_{B\times X} \varphi(b,x)\,\mathrm{d}\beta_n^X(b,x)\end{aligned}$$

si bien que $T_*^X\beta^X = \beta^X$. En outre, la formule (3.2) avec $n = 0$ donne l'égalité $\pi_*\beta^X = \beta$. □

On note $\mathcal{B}^X$ la tribu complétée de $\mathcal{B} \otimes \mathcal{X}$ pour la probabilité β^X.

3.2. **Martingales et probabilités conditionnelles.**

Dans cette section, nous associons à toute probabilité stationnaire ν sur X une famille mesurable et T-équivariante $(\nu_b)_{b\in B}$ de probabilités sur X.

La désintégration de β^X le long de π, prouve qu'il existe une application $\mathcal{B}$-mesurable $B \to \mathcal{P}(X), b \mapsto \nu_b$ telle que

$$\beta^X = \int_B \delta_b \otimes \nu_b\,\mathrm{d}\beta(b). \tag{3.3}$$

Autrement dit, pour toute fonction $\mathcal{B}^X$-mesurable bornée φ sur $B \times X$, on a

$$\beta^X(\varphi) = \int_B \int_X \varphi(b,y)\,\mathrm{d}\nu_b(y)\,\mathrm{d}\beta(b). \tag{3.4}$$

On a aussi l'égalité, pour β^X-presque tout (b,x) dans $B \times X$,

$$\mathbb{E}(\varphi \mid \pi^{-1}\mathcal{B})(b,x) = \int_X \varphi(b,y)\,\mathrm{d}\nu_b(y), \tag{3.5}$$

où l'espérance conditionnelle est prise relativement à la probabilité β^X.

Le lemme suivant interprète les probabilités conditionnelles ν_b comme des probabilités limites.

Lemme 3.2. *Soient ν une probabilité μ-stationnaire sur X et $b \mapsto \nu_b$ la famille $\mathcal{B}$-mesurable de probabilités sur X construite ci-dessus.*
a) Pour toute fonction borélienne bornée f sur X, pour β-presque tout b dans B, on a l'égalité,

$$\nu_b(f) = \lim_{p\to\infty} (b_{0*} \cdots b_{p*}\nu)(f). \tag{3.6}$$

b) Pour β-presque tout b dans B, on a l'égalité

$$\nu_b = b_{0*}\nu_{Tb}. \tag{3.7}$$

c) On a l'égalité

$$\nu = \int_B \nu_b \, \mathrm{d}\beta(b)\ . \tag{3.8}$$

d) Cette application $b \mapsto \nu_b$ est l'unique application $\mathcal{B}$-mesurable $B \to \mathcal{P}(X)$ vérifiant (3.7) *et* (3.8)*.*
e) Réciproquement, pour toute famille $\mathcal{B}$-mesurable $b \mapsto \nu_b \in \mathcal{P}(X)$ vérifiant (3.7)*, la probabilité ν donnée par* (3.8) *est μ-stationnaire.*

Démonstration. a) Pour β-presque tout b dans B, on note $\nu_{b,p}$ la probabilité $\nu_{b,p} = b_{0*} \cdots b_{p*}\nu \in \mathcal{P}(X)$. La démonstration est basée sur une formule explicite d'espérance conditionnelle : pour tout $p \geq 0$, pour toute fonction $\mathcal{X}$-mesurable bornée f sur X, qu'on considère comme une fonction sur $B \times X$, pour β^X-presque tout (b, x) dans B^X, on a

$$\mathbb{E}(f \mid \pi^{-1}\mathcal{B}_p)(b, x) = \int_X f(b_0 \cdots b_{p-1}x') \, \mathrm{d}\nu(x'). \tag{3.9}$$

En effet, le membre de droite de cette équation est une fonction $\pi^{-1}\mathcal{B}_p$-mesurable et, pour toute fonction $\pi^{-1}\mathcal{B}_p$-mesurable ψ, on a, d'après (3.2),

$$\int_{B\times X} f\psi \, \mathrm{d}\beta^X = \int_B \psi(b_0, \ldots, b_{p-1}) \int_X f(b_0 \cdots b_{p-1}x') \, \mathrm{d}\nu(x') \, \mathrm{d}\beta(b),$$

d'où la formule. Le résultat est alors une conséquence directe du théorème de convergence des martingales, puisque, par définition, pour β-presque b dans B, $\nu_b(f) = \mathbb{E}(f \mid \pi^{-1}\mathcal{B})(b)$.

b) Cette égalité résulte du point a) appliqué à une famille dénombrable de fonctions f qui engendrent la tribu borélienne $\mathcal{X}$.

c) D'après (3.2) et (3.3), pour toute fonction borélienne bornée f sur X, on a $\nu(f) = \int_{B\times X} f(x) \, \mathrm{d}\beta^X(b, x) = \int_B \nu_b(f) \, \mathrm{d}\beta(b)$.

d) Soit $b \mapsto \nu'_b$ une famille $\mathcal{B}$-mesurable de probabilités sur X vérifiant ces conditions. Introduisons la probabilité $\lambda = \int_B \delta_b \otimes \nu'_b \,\mathrm{d}\beta(b)$ sur B^X et montrons que $\lambda = \beta^X$. Pour cela, calculons, pour toute fonction positive $\mathcal{B}_n \otimes \mathcal{X}$-mesurable φ sur B^X, en utilisant les deux propriétés (3.7) et (3.8) pour la famille ν'_b et l'égalité (3.2),

$$\begin{aligned}\lambda(\varphi) &= \int_B \int_X \varphi(b,x)\,\mathrm{d}\nu'_b(x)\,\mathrm{d}\beta(b)\\ &= \int_B \int_B \int_X \varphi(b_0..b_{n-1}b', b_0..b_{n-1}y)\,\mathrm{d}\nu'_{b'}(y)\,\mathrm{d}\beta(b')\,\mathrm{d}\beta(b)\\ &= \int_B \int_X \varphi(b, b_0..b_{n-1}y)\,\mathrm{d}\beta(b)\,\mathrm{d}\nu(y) = \beta^X(\varphi).\end{aligned}$$

On a donc l'égalité $\lambda = \beta^X$ puis, par unicité de la désintégration, pour β-presque tout b, on a l'égalité $\nu'_b = \nu_b$.

e) On a

$$\mu * \nu = \int_G \int_B g_*\nu_b \,\mathrm{d}\beta(b)\,\mathrm{d}\mu(g) = \int_B b_{0*}\nu_{Tb}\,\mathrm{d}\beta(b) = \int_B \nu_b \,\mathrm{d}\beta(b) = \nu.$$

□

Remarque 3.3. Lorsque X est un espace métrisable séparable localement compact et que l'action de G sur X est continue (c'est toujours le cas dans nos applications) on a donc

$$\nu_b = \lim_{p\to\infty} b_{0*}\cdots b_{p*}\nu. \tag{3.10}$$

C'est là la présentation originale de cet objet par Furstenberg [13].

Remarque 3.4. On montre aisément que la probabilité ν est μ-ergodique si et seulement si la probabilité β^X est T^X-ergodique.

Signalons un joli corollaire de ces constructions.

Corollaire 3.5. *Soient μ une probabilité sur G, ν et ν' deux probabilités μ-stationnaires sur deux espaces boréliens standards $(X,\mathcal{X})$ et $(X',\mathcal{X}')$ munis d'une action borélienne de G. Alors, la probabilité $\nu'' := \int_B \nu_b \otimes \nu'_b \,\mathrm{d}\beta(b)$ est une probabilité borélienne μ-stationnaire sur l'espace produit $X \times X'$.*

Démonstration. En effet, la famille $\mathcal{B}$-mesurable $b \mapsto \nu''_b = \nu_b \otimes \nu'_b$ de probabilités sur $X \times X'$ vérifie, pour β-presque tout b dans B, $b_{0*}\nu''_{Tb} = \nu''_b$. □

3.3. **Système fibré au dessus d'une suspension.**

Le système dynamique dont nous aurons besoin pour notre problème est un produit fibré au dessus d'une suspension.

Soient M un groupe topologique compact métrisable et $\tau = (\tau_{\mathbb{R}}, \tau_M) : B \to \mathbb{R}_+ \times M$ une application $\mathcal{B}$-mesurable avec $\tau_{\mathbb{R}} \neq 0$. On note $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$ le semi-flot obtenu par suspension de $(B, \mathcal{B}, \beta, T)$ par τ défini dans la section 2.2. Nous allons construire un semi-flot $B^{\tau,X}$ fibré au dessus de B^τ.

Introduisons, pour $\ell \geq 0$ et β^τ-presque tout $c = (b, k, m)$ dans B^τ, la transformation $\rho_\ell(c)$ de X donnée par, pour tout x dans X,

$$\rho_\ell(c)x = b_{p_\ell(b,k)-1}^{-1} \cdots b_0^{-1} x,$$

et notons $\nu_c = \nu_b$. On a alors une propriété d'équivariance de ces probabilités sur X :

Lemme 3.6. *Pour β^τ-presque tout $c = (b, k, m)$ dans B^τ et tout $\ell \geq 0$, on a l'égalité*

$$\nu_{T_\ell^\tau c} = \rho_\ell(c)_* \nu_c.$$

Démonstration. D'après le lemme 3.2.b et l'égalité $\nu_{T_\ell^\tau c} = \nu_{T^{p_\ell(b,k)} b}$, on a bien $\nu_c = (b_0 \cdots b_{p_\ell(b,k)-1})_* \nu_{T_\ell^\tau c}$. □

On définit le semi-flot $(B^{\tau,X}, \mathcal{B}^{\tau,X}, \beta^{\tau,X}, T^{\tau,X})$ fibré au-dessus de $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$ de la façon suivante. On pose $B^{\tau,X} = B^\tau \times X$ et

$$\beta^{\tau,X} = \int_{B^\tau} \delta_c \otimes \nu_c \, \mathrm{d}\beta^\tau(c).$$

On note $\mathcal{B}^{\tau,X}$ la tribu complétée de la tribu produit $\mathcal{B}^\tau \times \mathcal{X}$ pour la probabilité $\beta^{\tau,X}$, et, pour (c, x) dans $B^{\tau,X}$ et $\ell \geq 0$, on pose

$$T_\ell^{\tau,X}(c, x) = (T_\ell^\tau c, \rho_\ell(c)x).$$

Lemme 3.7. *Pour tout $\ell \geq 0$, la transformation $T_\ell^{\tau,X}$ de $B^{\tau,X}$ préserve la probabilité $\beta^{\tau,X}$.*

Démonstration. Cela résulte des lemmes 2.4 et 3.6. □

Notons $\mathcal{Q}_\ell^{\tau,X} = (T_\ell^{\tau,X})^{-1}(\mathcal{B}^{\tau,X})$ et notons $\mathcal{Q}_\infty^{\tau,X}$ la tribu queue de $(B^{\tau,X}, \mathcal{B}^{\tau,X}, \beta^{\tau,X}, T^{\tau,X})$, c'est-à-dire l' intersection de cette famille décroissante de sous-tribus, $\mathcal{Q}_\infty^{\tau,X} = \cap_{\ell \geq 0} \mathcal{Q}_\ell^{\tau,X}$. De même, notons $\mathcal{Q}_\ell$ la famille décroissante de sous-tribus $\mathcal{Q}_\ell = (T_\ell^\tau)^{-1}(\mathcal{B}^\tau)$ et $c \mapsto \beta_c^\ell$ les mesures conditionnelles de β^τ relativement à $\mathcal{Q}_\ell$.

On peut résumer la discussion précédente en le corollaire suivant qui sera au coeur de l'argument de dérive.

Corollaire 3.8. *Pour toute fonction $\beta^{\tau,X}$-intégrable $\varphi : B^{\tau,X} \to \mathbb{R}$, pour tout $\ell \geq 0$, pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$, on a*

$$\mathbb{E}(\varphi \mid \mathcal{Q}_\ell^{\tau,X})(c,x) = \int_{B^\tau} \varphi(c', \rho_\ell(c')^{-1}\rho_\ell(c)x)\, \mathrm{d}\beta_c^\ell(c'). \tag{3.11}$$

Démonstration. Cela résulte du lemme 2.5. □

3.4. Mesure des feuilles stables relatives.

Pour pouvoir appliquer notre argument de dérive, nous aurons besoin de savoir que les probabilités ν_b ne massent pas les feuilles stables relatives au facteur $B^{\tau,X} \to B^\tau$. La proposition 3.9 ci-dessous donne un critère maniable qui permet de l'affirmer.

On suppose désormais que X est un espace topologique métrisable localement compact et que l'action de G sur X est continue. On note d une distance sur X induisant la topologie de X. Pour (b,x) dans $B \times X$, on note

$$W_b(x) = \{x' \in X \mid \lim_{p\to\infty} d(\rho_p(b)x, \rho_p(b)x') = 0\}$$

la feuille stable relative de (b,x). Cette feuille ne dépend pas du choix de d lorsque X est compact. Elle pourrait en dépendre en général. Néanmoins, dans tous les cas, on a la proposition ci-dessous. Rappelons qu'une application continue est *propre* si l'image inverse de tout compact est compacte. Notons A_μ l'opérateur de moyenne sur $X \times X$ donné par, pour toute fonction positive v sur $X \times X$ et tout (x,y) dans $X \times X$,

$$A_\mu(v)(x,y) = \int_G v(gx, gy)d\mu(g)$$

Cet opérateur A_μ est donc l'opérateur de convolution par la probabilité $\check{\mu}$ image de μ par l'inversion $g \mapsto g^{-1}$. On note Δ_X la diagonale de $X \times X$.

Proposition 3.9. *On suppose que*

(HC) : *il existe une fonction $v : (X \times X) \smallsetminus \Delta \to [0,\infty[$, telle que, pour tout compact K de X, la restriction de v à $K \times K \smallsetminus \Delta_K$ est propre et il existe des constantes $0 < a < 1$, $C > 0$ telles que $A_\mu(v) \leq av + C$.*

Soit ν une probabilité borélienne μ-stationnaire sans atome sur X. Alors, pour β^X-presque tout (b,x) dans $B \times X$, on a

$$\nu_b(W_b(x)) = 0.$$

L'hypothèse (HC) signifie que la moyenne par μ contracte, à une constante près, la fonction v .

La démonstration se fait en trois étapes. La première étape est la plus délicate, c'est le lemme suivant.

Lemme 3.10. *Sous l'hypothèse (HC), soit ν une probabilité borélienne μ-stationnaire telle que, pour β-presque tout b dans B, la probabilité ν_b est une masse de Dirac. Alors ν est une masse de Dirac.*

Démonstration. Notons $\kappa : B \to X$ l'application $\mathcal{B}$-mesurable telle que, pour β-presque tout b dans B, on a

$$\nu_b = \delta_{\kappa(b)}. \tag{3.12}$$

La stratégie consistera, en gros, à étudier la marche aléatoire correspondante sur $X \times X$. L'existence de κ et le théorème ergodique de Chacon-Ornstein assureront que cette marche s'approche de la diagonale Δ_X tandis que l'existence de v éloignera cette marche de la diagonale. Détaillons tout cela.

Pour g dans G et $b = (b_0, b_1, \ldots)$ dans B, on note $gb = (g, b_0, b_1, \ldots)$. On a donc, par le lemme 3.2.b, pour μ-presque tout g dans G et β-presque tout b dans B,

$$\kappa(gb) = g\kappa(b).$$

Par le lemme 3.2.c, on a aussi l'égalité

$$\nu = \kappa_*(\beta).$$

Munissons $B = G^{\mathbb{N}}$ de la topologie produit. D'après le théorème de Lusin, pour tout $\varepsilon > 0$, il existe une partie compacte K_0 de B telle que $\beta(K_0) = 1-\varepsilon$ et que la restriction de κ à K_0 est uniformément continue. Notons K le compact image $K = \kappa(K_0)$. Comme la restriction de v à $K \times K \smallsetminus \Delta_K$ est propre, on a

$$\begin{aligned} &\forall M > 0,\ \exists n_M > 0,\ \forall n \geq n_M,\ \forall b, b' \in B,\ \forall g_1, \ldots, g_n \in G \\ &\text{tels que } g_1 \cdots g_n b \in K_0 \text{ et } g_1 \cdots g_n b' \in K_0, \text{ on a} \\ &v(\kappa(g_1 \cdots g_n b), \kappa(g_1 \cdots g_n b')) \geq M. \end{aligned} \tag{3.13}$$

Introduisons maintenant l'opérateur de transfert L_μ sur B donné par, pour tout φ_0 dans $\mathrm{L}^1(B, \beta)$, pour β-presque tout b dans B

$$(L_\mu \varphi_0)(b) = \int_G \varphi_0(gb) d\mu(g).$$

Comme il est l'adjoint du décalage T, l'opérateur L_μ est ergodique. Le théorème ergodique de Chacon-Ornstein [5], appliqué à la fonction $\varphi_0 = \mathbf{1}_{K_0}$, entraine que, pour b en dehors d'une partie β-négligeable $N \subset B$, on a l'égalité

$$\lim_{p\to\infty} \frac{1}{p} \sum_{0\le n<p} (L_\mu^n \mathbf{1}_{K_0})(b) = \beta(K_0) = 1-\varepsilon. \tag{3.14}$$

Quitte à augmenter la taille de N, on peut aussi supposer que, pour tout b dans $B \smallsetminus N$, pour tout entier $n \ge 0$ et pour $\mu^{\otimes n}$-presque tout $(g_1, \dots, g_n)$ dans G^n, on a $\kappa(g_1 \dots g_n b) = g_1 \dots g_n \kappa(b)$.

Supposons par l'absurde que ν ne soit pas une masse de Dirac. Alors, l'ensemble

$$E := \{(b, b') \in B \times B \mid \kappa(b) \ne \kappa(b')\}$$

est de $\beta \otimes \beta$-mesure non nulle. On peut donc trouver deux points b_0 et b_0' hors de N tels que

$$\kappa(b_0) \ne \kappa(b_0'). \tag{3.15}$$

Exploitons maintenant la condition (HC). Elle implique que, pour tout $n \ge 0$, on a

$$A_\mu^n v \le a^n v + (1 + \cdots + a^{n-1}) C.$$

On en déduit la majoration, pour tous $x \ne x'$ dans X,

$$\frac{1}{p} \sum_{0\le n<p} (A_\mu^n v)(x, x') \le \frac{1}{p(1-a)} v(x, x') + \frac{1}{1-a} C. \tag{3.16}$$

Nous allons appliquer cette majoration avec $x = \kappa(b_0)$ et $x' = \kappa(b_0')$. Fixons $M > 0$. Remarquons que, grâce à (3.14), il existe un entier $p_0 \ge n_M$ tel que, pour tout $p \ge p_0$,

$$\frac{1}{p} \sum_{0\le n<p} (L_\mu^n \mathbf{1}_{K_0})(b_0) \ge 1 - 2\varepsilon \ \text{ et } \ \frac{1}{p} \sum_{0\le n<p} (L_\mu^n \mathbf{1}_{K_0})(b_0') \ge 1 - 2\varepsilon.$$

et donc

$$\frac{1}{p} \sum_{0\le n<p} \mu^{\otimes n}(\{(g_1, \dots, g_n) \in G^n \mid g_1 \cdots g_n b_0 \in K_0 \text{ et } g_1 \cdots g_n b_0' \in K_0\}) \ge 1 - 4\varepsilon.$$

Donc, pour $p \ge p_0$, on a, grâce à (3.13), la minoration

$$\frac{1}{p} \sum_{0\le n<p} (A_\mu^n v)(\kappa(b_0), \kappa(b_0')) \ge (1 - 4\varepsilon - \frac{p_0}{p}) M$$

En prenant la limite quand p tend vers l'infini, on obtient grâce à (3.16),

$$(1-4\varepsilon)M \le C/(1-a).$$

Comme M est arbitraire, ceci donne une contradiction dès que $\varepsilon < \frac{1}{4}$. Donc ν est une masse de Dirac. □

La deuxième étape est le lemme suivant.

Lemme 3.11. *Sous l'hypothèse* (HC)*, soit* ν *une probabilité* μ*-stationnaire sans atome sur* X*. Alors, pour* β*-presque tout* b *dans* B*, la probabilité* ν_b *n'a pas d'atome.*

Démonstration. La stratégie consiste, après quelques réductions basées sur l'ergodicité de β, à construire une probabilité stationnaire sur un espace Y à laquelle on pourra appliquer le lemme 3.10.

Supposons par l'absurde que l'ensemble $D := \{b \in B \mid \nu_b$ a des atomes$\}$ est de mesure non nulle. Comme $\nu_b = b_{0*}\nu_{Tb}$, cet ensemble D est T-invariant. Comme la probabilité β est T-ergodique, on a alors $\beta(D) = 1$. Le même argument prouve aussi que la masse maximale m_b des atomes de ν_b est une fonction β-presque sûrement constante et que le nombre N_b d'atomes de ν_b de masse m_b est aussi presque sûrement constant. On note $m_0 > 0$ cette masse et $N_0 \ge 1$ ce nombre d'atomes. Notons ν'_b la probabilité moyenne des N_0 atomes de ν_b de masse m_0. On a aussi l'égalité $\nu'_b = b_{0*}\nu'_{Tb}$ Par le lemme 3.2.e, la probabilité $\nu' := \int_B \nu'_b \,\mathrm{d}\beta(b)$ sur X est donc μ-stationnaire et on peut écrire ν comme la somme de $m_0\nu'$ et d'une mesure stationnaire de masse $(1-m_0)$. Par hypothèse, la probabilité ν' est aussi sans atome et, d'après le lemme 3.2.d, les probabilités ν'_b sont les probabilités limites de ν', si bien qu'on peut dorénavant supposer qu'on a $\nu = \nu'$.

Notons S_{N_0} le groupe des permutations de $\{1,\dots,N_0\}$, Y le quotient $Y := X^{N_0}/S_{N_0}$ et $p : X^{N_0} \to Y$ la projection. Le groupe G agit naturellement sur Y. Vérifions que Y satisfait l'hypothèse (HC). Notons v la fonction et a, C les constantes donnée par l'hypothèse (HC) sur X et introduisons l'application $w : (Y \times Y) \smallsetminus \Delta_Y \to [0,\infty[$ donnée, pour $y = p(x_1,\dots,x_{N_0})$ et $y' = p(x'_1,\dots,x'_{N_0})$ avec x_i et x'_i dans X, par

$$w(y,y') = \sum_{\sigma\in S_{N_0}} \min_{1\le i\le N_0} v(x_i, x'_{\sigma(i)}).$$

Cette application w est bien continue et propre sur $K \times K \smallsetminus \Delta_K$ pour tout compact K de Y. En outre elle vérifie la majoration

$$A_\mu(w) \le a\,w + C\,N_0!.$$

Introduisons la famille $b \mapsto \nu''_b := p_*({\nu_b}^{\otimes N_0})$ de probabilités sur Y. On a aussi l'égalité $\nu''_b = b_{0*}\nu''_{Tb}$. D'après le lemme 3.2.e, la probabilité $\nu'' := \int_B p_*({\nu_b}^{\otimes N_0})\,\mathrm{d}\beta(b)$ est μ-stationnaire. Par construction, pour β-presque tout b dans B la probabilité ν''_b est une masse de Dirac. Le lemme 3.10 prouve alors que ν'' est une masse de Dirac δ_{y_0}. Donc, pour β-presque tout b dans B, $\nu''_b = \delta_{y_0}$ et, par suite, ν est à support fini, ce qui est contradictoire. □

La dernière étape n'utilise pas l'hypothèse (HC).

Lemme 3.12. *Soit ν une probabilité μ-stationnaire sur X telle que, pour β-presque tout b dans B, la probabilité ν_b n'a pas d'atome.*

Alors, pour β^X-presque tout (b, x) dans $B \times X$, on a $\nu_b(W_b(x)) = 0$.

Démonstration. Considérons la transformation sur $B \times X \times X$ donnée par, pour (b, x, x') dans $B \times X \times X$,

$$R(b, x, x') = (Tb, b_0^{-1}x, b_0^{-1}x').$$

Le lemme 3.1 et le corollaire 3.5 prouvent que la transformation R préserve la probabilité

$$\Lambda = \int_B \delta_b \otimes \nu_b \otimes \nu_b \,\mathrm{d}\beta(b).$$

Notons

$$Z = \{(b, x, x') \in B \times X \times X \mid \lim_{p\to\infty} d(\rho_p(b)x, \rho_p(b)x') = 0\}$$

et, pour (b, x, x') dans $B \times X \times X$, $\varphi(b, x, x') = d(x, x')$. Par hypothèse, pour β-presque tout b, la probabilité ν_b n'a pas d'atomes, donc $\nu_b \otimes \nu_b$ ne masse pas la diagonale de $X \times X$, et donc cette fonction φ est Λ-presque sûrement non nulle. Par construction, pour Λ-presque tout z dans Z, on a $\lim\limits_{p\to\infty} \varphi(R^p(z)) = 0$, si bien que, d'après le théorème de récurrence de Poincaré, $\Lambda(Z) = 0$, ce qu'il fallait démontrer. □

Démonstration de la proposition 3.9. Elle résulte des lemmes 3.11 et 3.12. □

4. Mesures conditionnelles

Nous rassemblons dans ce chapitre quelques propriétés des mesures conditionnelles d'une probabilité pour une action borélienne d'un groupe localement compact.

4.1. **Mesures conditionnelles.**

On rappelle dans cette section la construction de ces mesures conditionnelles.

Soient R un groupe localement compact métrisable séparable et $(Z, \mathcal{Z})$ un espace borélien standard muni d'une action borélienne de R. Soit λ une probabilité borélienne sur Z. On suppose l'action de R à stabilisateurs discrets. Expliquons alors comment l'action de R sur Z permet en « désintégrant la probabilité λ le long des R-orbites » d'obtenir des mesures sur R et qui sont uniques modulo normalisation. Plus précisément :

Notons $\mathcal{M}(R)$ l'espace des mesures de Radon positives non nulles sur R et $\mathcal{M}_1(R) = \mathcal{M}(R)/_{\simeq}$ l'espace de ces mesures modulo normalisation : deux mesures de Radon σ_1 et σ_2 sont dites égales modulo normalisation, et on écrit

$$\sigma_1 \simeq \sigma_2 \text{ s'il existe } c > 0 \text{ tel que } \sigma_2 = c\sigma_1.$$

On peut choisir un représentant dans chaque classe d'équivalence : on fixe une suite croissante (K_n) de parties compactes de R qui recouvrent R et on choisit σ de masse 1 sur le compact K_n, où n est le plus petit entier m tel que $\sigma(K_m) \neq 0$.

On dit qu'une partie borélienne Σ de Z est une *section discrète* à l'action de R, si, pour tout z dans Z, l'ensemble $\{r \in R \mid r\, z \in \Sigma\}$ est discret et fermé dans R. D'après le théorème principal de [21], il existe une section discrète Σ à l'action de R telle que $R\Sigma = Z$.

Choisissons donc une section discrète Σ pour l'action de R et notons $a : R \times \Sigma \to Z, (r, z) \to r\, z$. La mesure $a^*\lambda$ sur $R \times \Sigma$ définie par, pour toute fonction borélienne positive f sur $R \times \Sigma$,

$$a^*\lambda(f) = \int_Z \left(\sum_{\{(r,z') \in a^{-1}(z)\}} f(r, z') \right) \mathrm{d}\lambda(z) \tag{4.1}$$

est alors une mesure borélienne σ-finie sur $R \times \Sigma$. Cela résulte de ce que, pour tout compact C de R, pour tout z dans Z, l'ensemble $(C \times \Sigma) \cap a^{-1}(z)$ est fini.

On note $\pi_\Sigma : R \times \Sigma \to \Sigma$ la deuxième projection et λ_Σ la mesure image par π_Σ d'une mesure finie sur $R \times \Sigma$ équivalente à $a^*\lambda$. On note, pour λ_Σ-presque tout z dans Σ, $\sigma_\Sigma(z) \in \mathcal{M}(R)$, les mesures conditionnelles obtenues par désintégration de $a^*\lambda$ le long de π_Σ. On a donc, pour toute fonction borélienne positive f sur $R \times Z$,

$$a^*\lambda(f) = \int_\Sigma \int_R f(r, z)\, \mathrm{d}\sigma_\Sigma(z)(r)\, \mathrm{d}\lambda_\Sigma(z) \tag{4.2}$$

Remarquons que ces mesures conditionnelles $\sigma_\Sigma(z)$ sont bien des mesures de Radon sur R. Cela résulte de nouveau de la finitude des parties $(C\times\Sigma)\cap a^{-1}(z)$.

On note t_r la translation à droite par un élément r de R.

Lemme 4.1. *Soit Σ une section discrète à l'action de R dans Z. Pour λ_Σ-presque tout z dans Σ, tout r dans R tel que $r\,z$ est dans Σ, on a*

$$\sigma_\Sigma(z) \simeq t_{r*}\sigma_\Sigma(r\,z).$$

Démonstration. La difficulté vient de ce que l'on exige que cette condition soit satisfaite pour une famille non dénombrable d'éléments r de R. Pour contourner cette difficulté, il suffit de remarquer qu'il existe une famille dénombrable, indexée par $i \in \mathbb{N}$, de parties boréliennes Σ_i de Σ et d'applications boréliennes $r_i : \Sigma_i \to R$ définies sur des parties boréliennes Σ_i de Σ, telle que

$$\{(z,r) \in \Sigma \times R \mid r\,z \in \Sigma\} = \bigcup_{i\in\mathbb{N}} \{(z, r_i(z)) \mid z \in \Sigma_i\}.$$

et que, pour λ_Σ-presque tout z dans Σ_i, $\sigma_\Sigma(z) \simeq t_{r_i(z)_*}\sigma_\Sigma(r_i(z)\,z)$.$\square$

Proposition 4.2. *Considérons une action borélienne à stabilisateurs discrets d'un groupe localement compact métrisable séparable R sur un espace borélien standard $(Z, \mathcal{Z})$.*

Alors il existe une application borélienne σ de Z dans $\mathcal{M}_1(R)$ et une partie borélienne $E \subset Z$ telle que $\lambda(E^c) = 0$ de sorte que, pour toute section discrète Σ à l'action de R, pour λ_Σ-presque tout z_0 dans Σ, pour tout r dans R tel que rz_0 est dans E,

$$\sigma(z_0) \simeq t_{r*}\sigma_\Sigma(r\,z_0).$$

Cette application σ est unique à un ensemble de λ-mesure nulle près.

Pour tout $r \in R$ et tout $z \in E$ tel que $r\,z \in E$, on a,

$$\sigma(z) \simeq t_{r*}(\sigma(r\,z)). \tag{4.3}$$

La mesure $\sigma(z)$ est appelée la mesure conditionnelle en z le long de l'action de R.

Démonstration. On choisit une section discrète Σ_0 telle que $R\Sigma_0 = Z$. D'après le lemme 4.1, pour λ presque tout z dans Z, si on écrit $z = r\,z_0$ avec r dans R et z_0 dans Σ_0, la mesure modulo normalisation $\sigma(z) := t_r^{-1}{}_*\sigma_{\Sigma_0}(z_0) \in \mathcal{M}_1(R)$ ne dépend pas du choix de cette écriture.

Cela définit l'application σ. La propriété annoncée de σ résulte du lemme appliqué à $\Sigma \cup \Sigma_0$ qui est aussi une section discrète. L'assertion (4.3) s'en déduit. L'unicité de σ est claire. $\square$

L'utilisation des mesures conditionnelles en théorie ergodique géométrique remonte, au moins, aux travaux de Ledrappier-Young [23, 24]. Leur utilisation dans des problèmes de classification de mesures invariantes sur les espaces homogènes apparait déjà dans [20] et, plus récemment, dans [8].

4.2. **Désintégration le long des stabilisateurs.**

Nous expliquons dans cette section comment exploiter des propriétés d'invariance par translation des mesures conditionnelles le long d'une action.

Notons $\mathrm{Gr}(\mathbb{R}^d)$ la variété grassmannienne de $\mathbb{R}^d$. La proposition suivante affirme que la désintégration de λ le long du stabilisateur des conditionnelles donne des probabilités invariantes par ce stabilisateur. Dans un groupe topologique S, on note S_0 la composante connexe de l'élément neutre.

Proposition 4.3. *Soit $(Z, \mathcal{Z})$ un espace borélien standard muni d'une action borélienne de $\mathbb{R}^d$ à stabilisateurs discrets, et soit λ une probabilité borélienne sur Z. Notons, pour λ-presque tout z dans Z,*

$\sigma(z)$ la mesure conditionnelle en z pour l'action de $\mathbb{R}^d$,

$$V_z := \{r \in \mathbb{R}^d \mid t_{r*}\sigma(z) = \sigma(z)\}_0,$$

$$\lambda = \textstyle\int_Z \lambda_z \, \mathrm{d}\lambda(z)$$

la désintégration de λ le long de l'application $Z \to \mathrm{Gr}(\mathbb{R}^d), z \mapsto V_z$. Alors, pour λ-presque tout z dans Z, la probabilité λ_z est V_z-invariante.

Cette proposition est une conséquence des trois lemmes ci-dessous. Le premier lemme utilise des notations différentes de celles de la proposition 4.3.

Lemme 4.4. *Soient $(Z, \mathcal{Z}, \lambda)$ un espace de Lebesgue, $(Y, \mathcal{Y})$ un espace borélien standard muni d'une action borélienne de $\mathbb{R}^d$, $f : Z \to Y$ une application mesurable et $I : Z \to \mathrm{Gr}(\mathbb{R}^d)$ une application mesurable telle que, pour λ-presque tout z dans Z, $I(z)$ stabilise $f(z)$.*

Notons $\lambda = \int_Z \lambda_z \, \mathrm{d}\lambda(z)$ la désintégration de λ le long de I.

Alors, pour λ-presque tout z dans Z, pour λ_z-presque tout z' dans Z, l'élément $f(z')$ est $I(z)$-invariant.

Démonstration. En effet, pour λ-presque tout z dans Z, pour λ_z-presque tout z' dans Z, on a, par définition des mesures conditionnelles, $I(z) = I(z')$ et, par hypothèse, $f(z')$ est $I(z')$-invariant. □

Le deuxième lemme reprend les notations de la proposition 4.3.

Lemme 4.5. *Soit* $(Z, \mathcal{Z})$ *un espace borélien standard muni d'une action borélienne de* $\mathbb{R}^d$ *à stabilisateurs discrets, et soit* λ *une probabilité borélienne sur* Z*. Soient* $(Y_0, \mathcal{Y}_0)$ *un espace borélien standard et* $\varphi : Z \to Y_0$ *une application mesurable telle qu'il existe une partie* $E \subset Z$ *avec* $\lambda(E^c) = 0$ *et, pour tous* z *dans* E *et* r *dans* $\mathbb{R}^d$ *avec* rz *dans* E*,* $\varphi(z) = \varphi(rz)$*. Notons* $z \mapsto \sigma(z) \in \mathcal{M}(\mathbb{R}^d)$ *les mesures conditionnelles en* z *de* λ *pour l'action de* $\mathbb{R}^d$*. Notons* $\lambda = \int_Z \lambda_z \,\mathrm{d}\lambda(z)$ *la désintégration de* λ *le long de* φ*. Alors, pour* λ*-presque tout* z *dans* Z*, pour* λ_z*-presque tout* z' *dans* Z*,* $\sigma(z')$ *est aussi la mesure conditionnelle en* z' *de* λ_z *pour l'action de* $\mathbb{R}^d$*.*

Démonstration. Il s'agit d'adapter l'argument de transitivité de désintégration des mesures à ce cadre.

Rappelons juste cet argument dans le cadre classique : on a un espace de Lebesgue $(A, \mathcal{A}, \alpha)$ et deux espaces boréliens standard $(B, \mathcal{B})$, $(C, \mathcal{C})$ ainsi que des applications mesurables $f : A \to B$ et $g : B \to C$. Alors, presque sûrement, les conditionnelles de α le long de f coïncident avec les conditionnelles le long de f des conditionnelles de α le long de $g \circ f$. De façon plus précise, on note $\alpha = \int_A \alpha_a \,\mathrm{d}\alpha(a)$ et $\alpha = \int_A \beta_a \,\mathrm{d}\alpha(a)$, les désintégrations de α respectivement le long de f et de $g \circ f$. On a alors, pour α-presque tout a, l'égalité $\beta_a = \int_A \alpha_{a'} \,\mathrm{d}\beta_a(a')$ qui donne la désintégration de β_a le long de f. □

Lemme 4.6. *Soient* $(Z, \mathcal{Z})$ *un espace borélien standard, muni d'une action borélienne de* $\mathbb{R}^d$ *à stabilisateurs discrets,* W *un sous-espace vectoriel de* $\mathbb{R}^d$*,* λ *une probabilité sur* $(Z, \mathcal{Z})$ *et* $z \mapsto \sigma(z) \in \mathcal{M}(\mathbb{R}^d)$ *les mesures conditionnelles en* z *de* λ *pour l'action de* $\mathbb{R}^d$*. On suppose que, pour* λ*-presque tout* z *dans* Z*, la mesure* $\sigma(z)$ *est invariante par translation par* W*. Alors, la probabilité* λ *est aussi invariante par l'action de* W*.*

Démonstration. Comme en 4.1, on note Σ une section discrète à l'action de $\mathbb{R}^d$ telle que $\mathbb{R}^d\Sigma = Z$ et a l'application $a : \mathbb{R}^d \times \Sigma \to Z, (r, z) \mapsto rz$. Par hypothèse la mesure $a^*\lambda$ est W-invariante, la mesure λ aussi. □

Démonstration de la proposition 4.3. On applique le lemme 4.4 avec $Y = \mathcal{M}(\mathbb{R}^d)$, $f = \sigma$ et $I(z) = V_z$, puis le lemme 4.5 avec $Y_0 = \mathrm{Gr}(\mathbb{R}^d)$ et $\varphi(z) = V_z$. On obtient que, pour λ-presque tout z dans Z, pour λ_z-presque tout z' dans Z, la mesure conditionnelle $\sigma_z(z')$ de λ_z pour l'action de $\mathbb{R}^d$ sur Z est V_z-invariante et donc, par le lemme 4.6, que la probabilité λ_z est V_z-invariante. □

5. Marches aléatoires sur les groupes de Lie

On introduit dans ce chapitre, pour toute marche aléatoire linéaire fortement irréductible, un système dynamique $(B^\tau, \mathcal{B}^\tau, \beta^\tau, T^\tau)$ suspension du système de Bernoulli $(B, \mathcal{B}, \beta, T)$. On étudie ensuite le comportement asymptotique de ces marches aléatoires pour pouvoir, dans la section 7.1, contrôler la dérive.

5.1. **Probabilité stationnaire sur la variété drapeau.** Soit G un groupe de Lie semi-simple réel virtuellement connexe, c'est-à-dire ayant un nombre fini de composantes connexes.

Definition 5.1. *Nous dirons qu'une probabilité borélienne sur G est Zariski dense si le semi-groupe Γ_μ engendré par le support de μ a une image Zariski dense dans le groupe adjoint* $\mathrm{Ad}G$.

Soit μ une probabilité Zariski dense sur G à support compact. On note encore $(B, \mathcal{B}, \beta, T)$ le décalage de Bernoulli unilatère d'alphabet $(G, \mathcal{G}, \mu)$ où $\mathcal{G}$ désigne la tribu borélienne de G.

Soit P un sous-groupe parabolique minimal de G. Ecrivons $P = ZU$, où U est le radical unipotent de P et Z est un sous-groupe réductif maximal de P, et notons A le sous-espace de Cartan de Z et A^+ la chambre de Weyl de A pour l'ordre associé au choix de P. Choisisons une involution de Cartan de G qui laisse stable Z et notons K le sous-groupe compact maximal de G formé des points fixes de cette involution de Cartan.

Soit V une représentation réelle de G de dimension d qui est fortement irréductible, c'est-à-dire, dont la restriction à la composante connexe de G est encore irréductible. Fixons une fois pour toutes une norme euclidienne K-invariante $\|.\|$ sur V de sorte que les éléments de A agissent de façon symétrique sur V.

On note χ le plus grand poids de A dans V, $V_0 = V_\chi$ l'espace poids associé dans V, de sorte que $PV_0 \subset V_0$ et $d_0 := \dim V_0$. On note V_0' le sous-espace de V somme des autres sous-espaces poids de sorte que $V = V_0 \oplus V_0'$

La proposition suivante est essentiellement due à Furstenberg et Kesten [15]. Notons $\mathrm{Gr}_{d_0}(V)$ la variété grassmannienne des d_0-plans dans V.

Proposition 5.2. *Il existe des applications $\mathcal{B}$-mesurables $B \to \mathrm{Gr}_{d_0}(V), b \mapsto V_b$ et $B \to \mathrm{Gr}_{d-d_0}(V), b \mapsto V_b'$ telles que :*
a) Pour β-presque tout $b \in B$, toute valeur d'adhérence m d'une suite $\frac{b_0 \cdots b_n}{\|b_0 \cdots b_n\|}$, a pour image $\mathrm{Im}(m) = V_b$ et est une isométrie sur $\mathrm{Ker}(m)^\perp$.

b) Pour β-presque tout $b \in B$, toute valeur d'adhérence m d'une suite $\frac{b_n \cdots b_0}{\|b_n \cdots b_0\|}$, a pour noyau $\mathrm{Ker}(m) = V'_b$ *et est une isométrie sur* $\mathrm{Ker}(m)^\perp$.
c) Pour tout hyperplan $W \subset V$, on a $\beta(\{b \in B \mid V_b \subset W\}) = 0$.
d) Pour tout vecteur non nul $v \in V$, on a $\beta(\{b \in B \mid v \in V'_b) = 0$.
e) Pour tout $W \in \mathrm{Gr}_{d_0}(V)$, on a $\beta(\{b \in B \mid W \cap V'_b \neq 0\}) = 0$.
f) Pour β-presque tout $b \in B$, la limite $\lambda_1 := \lim_{n\to\infty} \frac{1}{n} \log \|b_0 \cdots b_n\|$ existe et est positive : $\lambda_1 > 0$.

Démonstration. Pour a), c), f), voir [13] ou [2]. Le fait que les valeurs d'adhérence m soient de rang $d_0 = \dim V_0$ est dû à Gol'dsheid et Margulis dans [16]. On peut aussi le voir comme une conséquence de l'existence d'éléments loxodromiques dans Γ_μ. Le fait que la restriction de m à l'orthogonal de son noyau soit une similitude est vrai pour toute matrice π de rang d_0 dans l'adhérence $\overline{\mathbb{R}_* G} \subset \mathrm{End} V$. On vérifie aisément cette assertion grâce à la décomposition de Cartan $G = KA^+K$.

Les points b) et d) se déduisent des points a) et c) en passant à la représentation duale.

Le point e) se déduit du d) en passant à la sous-représentation irréductible de la représentation de G dans $\Lambda^{d_0} V$ engendrée par la droite de plus haut poids $\Lambda^{d_0} V_0$. □

En appliquant 5.2.a à une représentation convenable de G, on montre qu'il existe une unique application $\mathcal{B}$-mesurable $\xi : B \to G/P$ telle que, pour β-presque tout b dans B,

$$\xi(b) = b_0 \xi(Tb).$$

La probabilité image $\xi_* \beta$ est alors l'unique probabilité μ-stationnaire sur G/P

Remarque 5.3. Bien sûr, les espaces V_b et V'_b de la proposition 5.2 sont reliés à l'application bord ξ. Pour b dans B, on note $\check{b}$ l'élément $\check{b} = (b_0^{-1}, b_1^{-1}, \ldots)$ de B. Pour β-presque tout b dans B, on a, $V_b = \xi(b) V_0$ et $V'_b = \xi(\check{b}) V'_0$. On a aussi $V_b = b_0 V_{Tb}$ et $V'_b = b_0 V'_{Tb}$.

5.2. **Le système dynamique B^τ.**

Nous voulons construire une $\mathbb{R} \times M$-suspension (B^τ, T^τ) du décalage de Bernoulli associé à μ qui permette d'estimer le comportement asymptotique de la marche induite dans une représentation irréductible de G.

On construit tout d'abord une fonction $\theta : B \to Z$.

Soit $s : G/P \to G/U$ une section borélienne de la projection $G/U \to G/P$. En pratique, pour construire une telle section, on peut utiliser la décomposition d'Iwasawa ou la décomposition de Bruhat. Une formule explicite pour s n'est pas très importante car les constructions que nous allons faire dépendent peu du choix de cette section s. Néanmoins, pour simplifier, on supposera s construit à l'aide de la décomposition d'Iwasawa. Plus précisément, notons $M = Z \cap K$. La décomposition d'Iwasawa $G = KP$ permet de choisir la section s de sorte que, pour tout k dans K,

$$s(kP) = k\, m(k) U \quad \text{avec } m(k) \in M. \tag{5.1}$$

On dira alors que la section s est *à valeurs dans K modulo U*. Le groupe Z agit par multiplication à droite sur G/U.

On note $\sigma : G \times G/P \to Z$ le cocycle borélien donnée par, pour tout g dans G et tout x dans G/P,

$$g\, s(x) = s(gx)\, \sigma(g, x).$$

On note $\theta : B \to Z$ la fonction $\mathcal{B}$-mesurable donnée par, pour β-presque tout b dans B,

$$\theta(b) = \sigma(b_0, \xi(Tb)).$$

Introduisons la fonction bornée $\theta_{\mathbb{R}} : B \to \mathbb{R}$ donnée, pour β-presque tout b dans B, par

$$\theta_{\mathbb{R}}(b) := \log |\chi(\theta(b))| \tag{5.2}$$

Utilisons la formule de Furstenberg pour le premier exposant de Lyapounov

$$\lambda_1 = \textstyle\int_B \theta_{\mathbb{R}}(b)\, \mathrm{d}\beta(b) \tag{5.3}$$

([13], voir aussi [12] theorème 1.8), et la positivité du premier exposant de Lyapounov, $\lambda_1 > 0$ (voir proposition 5.2.f). Il existe donc, d'après le lemme 2.1, deux fonctions $\mathcal{B}$-mesurables bornées $\tau_{\mathbb{R}} : B \to \mathbb{R}_+^*$ et $\varphi : B \to \mathbb{R}$ telles que,

$$\theta_{\mathbb{R}} = \tau_{\mathbb{R}} + \varphi \circ T - \varphi. \tag{5.4}$$

Notons $\theta_M(b)$ la composante dans M de $\theta(b)$, $\tau_M(b) = \theta_M(b)^{-1}$ et

$$\tau = (\tau_{\mathbb{R}}, \tau_M) : B \to \mathbb{R} \times M. \tag{5.5}$$

C'est la suspension B^τ associée à cette fonction τ que nous utiliserons désormais.

Cette suspension permet de contrôler la norme des mots intervenant dans les formules d'espérance conditionnelle grâce au lemme suivant.

Lemme 5.4. *Pour β-presque tout b dans B, pour tout w dans V_b, on a*

$$\left\|b_0^{-1}w\right\| = e^{-\theta_{\mathbb{R}}(b)} \left\|w\right\| . \tag{5.6}$$

Démonstration. Par définition de θ, pour β-presque tout b dans B, on a

$$b_0 s(\xi(Tb)) = s(\xi(b))\theta(b).$$

Comme w est dans V_b, on peut écrire $w = s(\xi(b))v$ avec v dans V_0. Remarquons que cette expression a un sens car U agit trivialement sur V_0. Comme la norme est K-invariante, on a

$$\left\|b_0^{-1}w\right\| = \left\|b_0^{-1}s(\xi(b))v\right\| = \left\|\theta(b)^{-1}v\right\| = e^{-\theta_{\mathbb{R}}(b)} \left\|v\right\| = e^{-\theta_{\mathbb{R}}(b)} \left\|w\right\| .$$

□

5.3. Comportement des marches aléatoires.

Continuons notre étude du comportement asymptotique des marches aléatoires sur G.

Nous utiliserons la proposition 5.2 pour le contrôle de la dérive dans le lemme 7.3 sous la forme du corollaire suivant.

Corollaire 5.5. *a) Pour tout $\alpha > 0$, il existe $r_0 \geq 1$, $q_0 \geq 1$, tels que, pour tout $v \in V \smallsetminus \{0\}$, on a*

$$\beta(\{a \in B \mid \forall q \geq q_0 \ \ \|a_q \cdots a_0 v\| \geq \frac{1}{r_0}\|a_q \cdots a_0\|\|v\|\}) \geq 1 - \alpha.$$

b)Pour tout $\alpha > 0$, $\eta > 0$, il existe $q_0 \geq 1$, tels que, pour tout $v \in V \smallsetminus \{0\}$ et tout $W \in \mathrm{Gr}_{d_0}(V)$, on a

$$\beta(\{a \in B \mid \forall q \geq q_0 \ \ d(\mathbb{R}a_q \cdots a_0 v, a_q \cdots a_0 W) \leq \eta\}) \geq 1 - \alpha.$$

Pour montrer ce corollaire, nous aurons besoin du lemme suivant d'algèbre linéaire. Notons

$$O_{d_0}(V) = \{\pi \in \mathrm{End}(V) \mid \mathrm{rang}(\pi) = d_0 \ \text{ et } \pi|_{(\mathrm{Ker}\pi)^\perp} \text{ est une isométrie}\}.$$

C'est une partie compacte de $\mathrm{End}(V)$.

Lemme 5.6. *a) Pour tout $\varepsilon > 0$, il existe $r_0 \geq 1$, $\varepsilon' > 0$ tels que, pour tous $g \in \mathrm{GL}(V)$ et $\pi \in O_{d_0}(V)$ avec $\|g - \pi\| \leq \varepsilon'$, pour tout $v \in V \smallsetminus \{0\}$ avec $d(\mathbb{R}v, \mathrm{Ker}\pi) \geq \varepsilon$, on a $\|gv\| \geq \frac{1}{r_0}\|v\|$.*
b) Pour tout $\varepsilon > 0$, $\eta > 0$, il existe $\varepsilon' > 0$ tel que, pour tous $g \in \mathrm{GL}(V)$ et $\pi \in O_{d_0}(V)$ avec $\|g - \pi\| \leq \varepsilon'$, on a, pour tous $v \in V \smallsetminus \{0\}$ et $W \in \mathrm{Gr}_{d_0}(V)$, si $d(\mathbb{R}v, \mathrm{Ker}\pi) \geq \varepsilon$ et $\inf\limits_{w \in W \smallsetminus 0} d(\mathbb{R}w, \mathrm{Ker}\pi) \geq \varepsilon$, alors $d(\mathbb{R}gv, gW) \leq \eta$.

Démonstration. a) Sinon, on pourrait trouver des suites π_n dans $O_{d_0}(V)$, g_n dans $\mathrm{GL}(V)$, v_n de norme 1 dans V, avec $\|g_n - \pi_n\| \to 0$, $d(\mathbb{R}v_n, \mathrm{Ker}\pi_n) \geq \varepsilon$ et $\|g_n v_n\| \to 0$. Par compacité, on peut supposer que π_n converge vers un élément π dans $O_{d_0}(V)$ et v_n converge vers un élément de norme 1 dans V. Nos assertions impliquent que v est simultanément dans $\mathrm{Ker}\pi$ et à distance au moins ε de $\mathrm{Ker}\pi$, ce qui est contradictoire.

b) L'argument est semblable à celui employé pour démontrer le point *a*). □

Démonstration du corollaire 5.5. a) D'après la proposition 5.2.d, pour tout $\alpha > 0$, il existe $\varepsilon > 0$ tel que pour tout $v \in V \smallsetminus \{0\}$,

$$\beta(\{a \in B \mid d(\mathbb{R}v, V'_a) \leq \varepsilon\}) \geq 1 - \alpha/2.$$

D'autre part, d'après la proposition 5.2.b, pour tout $\varepsilon' > 0$, il existe $q_0 \geq 0$ tel que

$$\beta(\{a \in B \mid \forall q \geq q_0 \;\; d(\tfrac{a_q \cdots a_0}{\|a_q \cdots a_0\|}, O_{d_0}(V)) \geq \varepsilon'\}) \geq 1 - \alpha/2.$$

Il suffit alors d'appliquer le lemme 5.6.a.

b) D'après la proposition 5.2.e, pour tout $\alpha > 0$, il existe $\varepsilon > 0$ tel que pour tout $W \in \mathrm{Gr}_{d_0}(V)$,

$$\beta(\{a \in B \mid \inf_{w \in W \smallsetminus 0} d(\mathbb{R}w, V'_b) \geq \varepsilon\}) \geq 1 - \alpha/2.$$

Il suffit alors d'appliquer, comme ci-dessus, la proposition 5.2.b et le lemme 5.6.b. □

6. Espaces homogènes de groupes semi-simples

Ce chapitre rassemble diverses propriétés ergodiques des marches aléatoires sur des espaces homogènes. Ces propriétés nous permettront dans le chapitre 7 de développer l'argument de dérive exponentielle.

6.1. **Notations.** Pour montrer le théorème 1.1 et le théorème 1.3, nous allons utiliser une méthode et des notations communes.

Nous gardérons les notations suivantes jusqu'à la fin de cet article.

Dans le premier cas, i.e. celui du théorème 1.1, G est un groupe de Lie quasi-simple connexe et Λ un réseau de G. On note X le quotient G/Λ et R la représentation adjointe dans $V = \mathfrak{g}$, l'algèbre de Lie de G.

Dans le deuxième cas, i.e. celui du théorème 1.3, G est l'adhérence de Zariski de Γ_μ dans $\mathrm{SL}(d,\mathbb{R})$. On note X le tore $\mathbb{T}^d$ et R la représentation de G dans l'algèbre de Lie $V=\mathbb{R}^d$ du groupe $\mathbb{T}^d$, c'est-à-dire l'action naturelle de G sur $\mathbb{R}^d$.

Dans les deux cas, G est un groupe de Lie semi-simple (on détaillera ce point dans le lemme 8.5), la représentation R de G dans V est fortement irréductible, μ est une probabilité Zariski dense à support compact sur G, $\Gamma=\Gamma_\mu$ est le sous-semi-groupe Zariski dense de G engendré par le support de μ, ν est une probabilité μ-stationnaire sans atome sur X et τ est l'application donnée par (5.5). En outre, on supposera que G est non compact (le cas, beaucoup plus facile, où G est compact, est détaillé dans le lemme 8.4).

La démonstration, qui s'étend jusqu'à la fin de cet article, repose sur les propriétés des systèmes dynamiques

$$(B^X,\mathcal{B}^X,\beta^X,T^X) \text{ et } (B^{\tau,X},\mathcal{B}^{\tau,X},\beta^{\tau,X},T^{\tau,X})$$

que nous avons introduits dans les sections 3.1 et 3.3, pour ces valeurs de G, V, X, τ...

6.2. Récurrence hors de la diagonale.

Vérifions maintenant la condition (HC) de la section 3.4,
ce qui nous permettra d'utiliser la proposition 3.9.

Pour tout x dans X, on note r_x le rayon d'injectivité en x, c'est à dire la borne supérieure des rayons $r>0$ pour lesquels l'application $V\to X, w\mapsto e^w x$ est injective sur la boule $B(0,r)$.

Proposition 6.1. *Dans les deux cas de 6.1, l'opérateur de moyenne A_μ sur $X\times X$ vérifie la condition (HC).*

La démonstration de cette proposition reprend des idées de Eskin et Margulis dans [9]. On notera le contraste entre cette proposition 6.1 et le théorème 1 de LePage dans [26] qui montre que sur la variété drapeau, c'est une puissance positive de la distance qui est contractée par la convolution. Nous aurons besoin des deux lemmes suivants. Nous noterons de la même façon A_μ tous les opérateurs de moyenne par μ sur tout espace où agit Γ_μ. Le premier lemme, dû à Eskin et Margulis, exhibe une fonction sur laquelle A_μ agit par contraction.

Lemme 6.2. ([9]) *Soient $V=\mathbb{R}^d$ et G un sous-groupe de Lie semi-simple de $\mathrm{GL}(V)$ tel que, pour tout sous-espace vectoriel non nul G-invariant V' de V, l'image de G dans $\mathrm{GL}(V')$ n'est pas compacte. Notons φ la fonction $\varphi: V\smallsetminus 0\to\mathbb{R}^*, v\mapsto \|v\|^{-1}$. Alors, il existe $a_0<1$,*

$\delta_0 > 0$ et $n_0 \geq 1$ tels que

$$A_\mu^n(\varphi^\delta) \leq a_0^n \varphi^\delta \quad \text{pour tous } \delta \leq \delta_0 \text{ et } n \geq n_0. \tag{6.1}$$

Démonstration. C'est le lemme 4.2 de [9]. Il est basé sur un développement limité à l'ordre 2 de $e^{-\delta \log(\|gv\|/\|v\|)}$ et sur le théorème de Furstenberg et Kesten de positivité du premier exposant de Lyapounov λ_1. □

Lorsque X n'est pas compact, nous aurons besoin d'une variante d'un lemme dû à Eskin et Margulis qui affirme l'existence d'une application propre sur X contractée, à une constante près, par l'opérateur de moyenne.

Lemme 6.3. *Soient G un groupe de Lie semi-simple réel connexe sans facteur compact, Λ un réseau de G, $X = G/\Lambda$ et μ une probabilité Zariski dense sur G dont le support est compact. Alors, il existe une application propre $u : X \to [0, \infty[$ et des constantes $a < 1$, $C > 0$ et $\kappa > 0$ telles que*

$$A_\mu(u) \leq au + C \tag{6.2}$$

et, pour tout x dans X,

$$u(x) \geq r_x^{-\kappa}. \tag{6.3}$$

Démonstration. Comme le centre de G rencontre Λ en un sous-groupe d'indice fini, on peut se ramener au cas où G est adjoint, et donc, en particulier, linéaire. Dans la section (3.2) de [9] est construite explicitement une application propre u vérifiant (6.2). D'après cette construction, si on regarde G comme un groupe de matrices, il existe des constantes $C_0 > 0$ et $\kappa_0 > 0$ telles que, pour tout $x = g\Lambda$ dans X, on a la minoration

$$u(x) \geq C_0 \min_{\gamma \in \Lambda} \|g\gamma\|^{\kappa_0}. \tag{6.4}$$

Il suffit alors de remarquer qu'il existe des constantes $C_1 > 0$ et $\kappa_1 > 0$ telles que, pour tout $x = g\Lambda$ dans X, on a la minoration

$$r_x \geq C_1 \,(\min_{\gamma \in \Lambda} \|g\gamma\|)^{-\kappa_1}. \tag{6.5}$$

En effet, si $h = e^w$ est un élément non trivial de G tel que $hx = x$, alors, pour tout γ dans Λ, $\delta := \gamma^{-1} g^{-1} h g \gamma$ est dans Λ et on a

$$\|h - e\| \geq \|\delta - e\| \|\mathrm{Ad}(g\gamma)^{-1}\|^{-1}.$$

On a donc, en notant $C_2 = \min_{\delta\in\Lambda\smallsetminus e} \|\delta - e\|$,

$$\min_{hx=x\,,\,h\neq e} \|h-e\| \geq C_2\,(\min_{\gamma\in\Lambda} \|\mathrm{Ad}(g\gamma)^{-1}\|)^{-1}.$$

La minoration (6.5) s'en déduit. □

Démonstration de la proposition 6.1. Remarquons tout d'abord que, si la condition (HC) est satisfaite pour une puissance μ^{*n_0}, elle est satisfaite pour μ. On choisit $a_0 < 1$, $0 < \delta_0 < 1$ et $n_0 \geq 1$ comme dans le lemme 6.2. Quitte à remplacer μ par μ^{*n_0}, on peut donc supposer $n_0 = 1$. Soit $\delta \leq \delta_0$.

Pour tout $x \neq x'$ dans X, on note $r_{x,x'} = \frac{1}{2}\min(r_x, r_{x'})$,

$$d_0(x,x') = \begin{cases} \|w\| & \text{si } x' = e^w x \ \text{ avec } \ w \in V\ ,\ \ \|w\| \leq r_{x,x'} \\ r_{x,x'} & \text{sinon,} \end{cases}$$
$$v_0(x,x') = d_0(x,x')^{-\delta}.$$

Lorsque X est compact, la fonction $v = v_0$ convient. Dans le cas général, il faut introduire la fonction u et les constantes $a < 1$, $C > 0$ et $\kappa > 0$ données par le lemme 6.3. On peut supposer $a = a_0$. On pose $R_0 = \sup_{g\in\mathrm{supp}\mu} \max(\|R(g)\|, \|R(g)^{-1}\|)$. Si on choisit $\delta < \kappa$ et $C_0 = \frac{2R_0^{2\delta}}{1-a_0}$, la fonction v donnée par, pour tout $x \neq x'$ dans X,

$$v(x,x') = v_0(x,x') + C_0(u(x) + u(x')) \tag{6.6}$$

vérifie la condition (HC).

En effet, si $d_0(x,x') \geq R_0^{-1} r_{x,x'}$, d'après (6.3),

$$\begin{aligned}(A_\mu v_0)(x,x') &\leq R_0^{2\delta} r_{x,x'}^{-\delta} \\ &\leq 2R_0^{2\delta}(r_x^{-\delta} + r_{x'}^{-\delta}) \leq 2R_0^{2\delta}(u(x) + u(x')).\end{aligned}$$

D'autre part, si $d_0(x,x') \leq R_0^{-1} r_{x,x'}$ alors, en écrivant $x' = e^w x$ avec $w \in V$, $\|w\| \leq r_{x,x'}$, on a, pour tout g dans G de norme au plus R_0,

$$v_0(gx, gx') = \|gw\|^{-\delta},$$

et donc, d'après (6.1),

$$(A_\mu v_0)(x,x') \leq a_0 \|w\|^{-\delta} = a_0 v_0(x,x').$$

Dans tous les cas, on a donc la majoration

$$(A_\mu v_0)(x,x') \leq a_0 v_0(x,x') + R_0^{2\delta}(u(x) + u(x')).$$

L'inégalité (6.2) et la définition (6.6) de v donnent alors la majoration

$$\begin{aligned}(A_\mu v)(x,x') &\leq a_0 v_0(x,x') + (R_0^{2\delta} + a_0 C_0)(u(x)+u(x')) + 2C\,C_0 \\ &\leq \tfrac{1+a_0}{2} v(x,x') + 2C\,C_0\end{aligned}$$

qui prouve la propriété (HC). □

6.3. **Récurrence hors des orbites finies.**

Nous montrons dans cette section un phénomène de récurrence hors des orbites finies pour les marches aléatoires sur X analogue au phénomène de récurrence dans les compacts de [9].

Proposition 6.4. *Dans les deux cas de 6.1, soit $F \subset X$ une partie finie Γ-invariante. Alors, pour tout $\varepsilon > 0$, il existe un compact K_ε de F^c tel que, pour tout point $x \in X \smallsetminus F$, il existe une constante $M = M_x$ uniforme sur les compacts de $X \smallsetminus F$ telle que pour tout $n \geq M$,*

$$A_\mu^n(\mathbf{1}_{K_\varepsilon})(x) \geq 1 - \varepsilon\ .$$

Nous aurons besoin des deux lemmes suivants.

Le premier traduit un phénomène de récurrence dans les compacts, dû à Foster, et utilisé dans ce contexte par Eskin et Margulis.

Lemme 6.5. **(**[9]**)** *Soient H un groupe localement compact agissant continûment sur un espace localement compact Y et μ une probabilité borélienne sur H.*

On suppose qu'il existe une application propre $f : Y \to [0,\infty[$, et des constantes $a < 1$, $b > 0$ telles que $A_\mu(f) \leq af + b$.

Alors, pour tout $\varepsilon > 0$, il existe un compact $K \subset Y$ tel que, pour tout point $y \in Y$, il existe une constante $M = M_y$ uniforme sur les compacts de Y telle que pour tout $n \geq M$,

$$A_\mu^n(\mathbf{1}_K)(y) \geq 1 - \varepsilon\ .$$

Rappelons la courte démonstration de ce lemme.

Démonstration. D'après l'hypothèse, on a, pour tout $n \geq 1$

$$A_\mu^n(f) \leq a^n f + b(1 + \cdots + a^{n-1}) \leq a^n f + B$$

avec $B = \frac{b}{1-a}$. Comme f est propre, on peut prendre pour compact

$$K = \{z \in Y \mid f(z) \leq \frac{2B}{\varepsilon}\}$$

de sorte que $\mathbf{1}_{K^c} \leq \frac{\varepsilon}{2B} f$. On a alors les majorations

$$A_\mu^n(\mathbf{1}_{K^c})(y) \leq \frac{\varepsilon}{2B} A_\mu^n(f)(y) \leq \frac{\varepsilon a^n}{2B} f(y) + \frac{\varepsilon}{2} \leq \varepsilon$$

dès que n est suffisamment grand pour que $f(y) \leq \frac{B}{a^n}$. $\square$

Le second lemme est une variante de la proposition 6.1.

Lemme 6.6. *Dans les deux cas de 6.1, soit $F \subset X$ une partie finie Γ-invariante. Alors, il existe une application propre $u_F : X \smallsetminus F \to [0, \infty[$ et des constantes $a < 1$ et $C > 0$ telles que*

$$A_\mu(u_F) \leq a u_F + C. \tag{6.7}$$

Démonstration. On procède comme pour la démonstration de la proposition 6.1. On choisit $a_0 < 1$, $\delta_0 > 0$ et $n_0 \geq 1$ comme dans le lemme 6.2. Quitte à remplacer μ par μ^{*n_0}, on peut supposer $n_0 = 1$. Soit $\delta \leq \delta_0$.

Soit $r_0 > 0$ un réel tel que, pour tout point x_0 de F, on a $r_0 \leq \frac{1}{2} r_{x_0}$ et que, pour toute paire x_0, x_0' de points distincts de F, on a $r_0 \leq \frac{1}{2} d_0(x_0, x_0')$. Pour tout x dans X, on note

$$d_0(x) = \begin{cases} \|w\| & \text{si } x = e^w x_0 \text{ avec } x_0 \in F \text{ et } \|w\| \leq r_0 \\ r_0 & \text{sinon,} \end{cases}$$

$$u_0(x) = d_0(x)^{-\delta}$$

Lorsque X est compact, la fonction $u_F = u_0$ convient. Dans le cas général, la fonction $u_F = u_0 + u$ avec u comme dans le lemme 6.3 convient. La présence de u sert seulement à assurer la propreté de u_F. Pour vérifier que u_F convient, on pose $R_0 = \sup_{g \in \operatorname{supp} \mu} \max(\|R(g)\|, \|R(g)^{-1}\|)$.

D'une part, si $d_0(x) \geq R_0^{-1} r_0$ alors, on a

$$(A_\mu u_0)(x) \leq R_0^{2\delta} r_0^{-\delta}.$$

D'autre part, si $d_0(x) \leq R_0^{-1} r_0$ alors, en écrivant $x = e^w x_0$ avec $x_0 \in F$, on a, pour tout g dans G de norme au plus R_0,

$$d_0(gx) \leq \|gw\|^{-\delta},$$

et donc, d'après (6.1),

$$(A_\mu u_0)(x) \leq a_0 \|w\|^{-\delta} = a_0 u_0(x).$$

Dans tous les cas, on a donc la majoration

$$(A_\mu u_0)(x) \leq a_0 u_0(x) + R_0^{2\delta} r_0^{-\delta}.$$

Cette inégalité et celle du lemme 6.3 donnent l'inégalité cherchée pour u_F. $\square$

Démonstration de la proposition 6.4. Cela résulte du lemme 6.5 appliqué à $Y = X \smallsetminus F$ et à la fonction $f = u_F$ du lemme 6.6. $\square$

6.4. **Probabilités stationnaires sur G/H.**

Pour exploiter l'argument de dérive, nous aurons besoin, dans le premier cas de 6.1, de la proposition suivante qui est d'intérêt indépendant.

Proposition 6.7. *Soient G un groupe de Lie semi-simple réel connexe sans facteur compact, μ une probabilité Zariski dense à support compact sur G et $H \subset G$ un sous-groupe unimodulaire. S'il existe une probabilité μ-stationnaire sur l'espace homogène G/H, alors l'algèbre de Lie de H est un idéal de celle de G.*

Pour montrer cette proposition, nous utiliserons le lemme suivant.

Lemme 6.8. *Soient $V = \mathbb{R}^d$, G un sous-groupe semi-simple sans facteur compact de $\mathrm{GL}(V)$ et μ une probabilité Zariski dense à support compact sur G. Alors, toute probabilité μ-stationnaire ν sur l'espace vectoriel V est portée par le sous-espace V^G des points fixes de G dans V.*

Démonstration. Supposons par l'absurde qu'il existe une probabilité μ-stationnaire ν sur V non portée par V^G. Il existe donc une sous-représentation irréductible $W \subset V$ de dimension au moins 2 telle que la projection de ν sur W n'est pas la masse de Dirac en 0. Cette projection est encore μ-stationnaire. On peut donc supposer V irréductible et G non compact.

On utilise encore le système de Bernoulli $(B, \mathcal{B}, \beta, T)$ d'alphabet (G, μ) et le système dynamique fibré $B \times V$ muni de la transformation $R : (b, v) \mapsto (Tb, b_0 v)$ qui laisse invariante la probabilité $\beta \otimes \nu$.

Le théorème de Furstenberg et Kesten de positivité du premier exposant de Lyapounov ([15], voir aussi [12] chapitre 1) assure alors que, pour β-presque tout b dans B, il existe un sous-espace $W_b \subsetneq V$ tel que, pour tout v dans $V \smallsetminus W_b$, la norme $\|b_n \cdots b_0 v\|$ tend (exponentiellement vite) vers l'infini. Introduisons la partie R-invariante,

$$Z = \{(b, v) \in B \times V \mid v \notin W_b\}$$

et la fonction φ sur Z donnée par

$$\varphi(b, v) = \|v\|.$$

Comme ν est μ-stationnaire, que μ est Zariski dense dans G et que l'action de G sur V est irréductible, ν ne charge pas les sous-espaces propres de V. On a donc $(\beta \otimes \nu)(Z) = 1$. Par construction, pour $\beta \otimes \nu$-presque tout z dans Z, on a

$$\lim_{n \to \infty} \varphi(R^n z) = \infty,$$

ce qui contredit le théorème de récurrence de Poincaré. □

Démonstration de la proposition 6.7. Notons ν une probabilité μ-stationnaire sur G/H, notons $\mathfrak{g}$ l'algèbre de Lie de G, $\mathfrak{h}$ celle de H, $r = \dim \mathfrak{h}$, $V = S^2(\Lambda^r \mathfrak{g})$ et v un point non nul de la droite $S^2(\Lambda^d \mathfrak{h}) \subset V$.

Comme H est unimodulaire, H est inclus dans le stabilisateur N du point v. L'orbite $G\,v \simeq G/N$ admet donc aussi une probabilité μ-stationnaire : l'image ν' de ν par la projection $G/H \to G/N$. D'après le lemme 6.8, cette probabilité μ-stationnaire ν' est portée par le sous-espace V^G des vecteurs G-invariants. Donc $N = G$. Comme N normalise $\mathfrak{h}$, $\mathfrak{h}$ est un idéal de $\mathfrak{g}$. □

6.5. **Flot horocyclique.**

> Le but de cette section est de construire une action de V_0 qui jouera un rôle analogue au flot horocyclique des surfaces hyperboliques compactes, en ce sens que les orbites de cette action sont contenues dans les feuilles stables relatives du facteur $B^{\tau,X} \to B^\tau$ et qu'elles sont uniformément dilatées par le semi-flot T^τ.

On garde les notations de la section 6.1.

Definition 6.9. *Le flot horocyclique est l'action Φ de V_0 sur $B^{\tau,X}$ donnée par l'égalité, pour tout v dans V_0, β^τ-presque tout $c = (b, k, m)$ dans B^τ et tout x dans X,*

$$\Phi_v(c, x) = (c, \exp(D_c(v))\, x), \tag{6.8}$$

où $D_c(v)$ est l'élément de V_c donné par

$$D_c(v) = e^{k-\varphi(b)} s(\xi(b))\, m\, v. \tag{6.9}$$

Rappelons que s, ξ, φ ont été définis dans les sections 5.1 et 5.2. Géométriquement, ce flot Φ « translate chaque point (c, x) dans la direction de V_c ». Signalons qu'à ce stade du raisonnement, on ne sait pas si ce flot préserve la probabilité $\beta^{\tau,X}$: nous ne le saurons qu'après avoir démontré le théorème 1.1. Cette difficulté est bien sûr une source de complications qui est au coeur du sujet.

La propriété fondamentale du flot horocyclique est son lien avec le flot $(T_\ell^{\tau,X})_{\ell \geq 0}$ sur $B^{\tau,X}$.

Lemme 6.10. *Dans les deux cas de 6.1, pour tout v dans V_0 et tout $\ell \geq 0$, on a l'égalité, pour β^τ-presque tout c dans B^τ et pour tout x dans X,*

$$T_\ell^{\tau,X} \circ \Phi_v\,(c, x) = \Phi_{e^{-\ell} v} \circ T_\ell^{\tau,X}\,(c, x). \tag{6.10}$$

Démonstration. Notons S la transformation de $B\times\mathbb{R}\times M\times X$ donnée par

$$S(b,k,m,x)=(Tb,k-\tau_{\mathbb{R}}(b),\tau_M(b)m,b_0^{-1}x).$$

Remarquons que $B^{\tau,X}$ est l'ensemble des points de $B\times\mathbb{R}_+\times M\times X$ que S fait sortir de ce produit.

Introduisons le flot $\widetilde{T}_\ell^{\tau,X}$ défini sur $B\times\mathbb{R}\times M\times X$ par

$$\widetilde{T}_\ell^{\tau,X}(b,k,m,x)=(b,k+\ell,m,x).$$

Le flot $T_\ell^{\tau,X}$ est donné, pour $\ell\geq 0$ et (b,k,m,x) dans $B^{\tau,X}$, par

$$T_\ell^{\tau,X}(b,k,m,x)=(S^p\circ\widetilde{T}_\ell^{\tau,X})(b,k,m,x)$$

où $p\geq 0$ est l'unique entier tel que cette expression est dans $B^{\tau,X}$. Définissons alors une action $\widetilde{\Phi}$ de V_0 sur $B\times\mathbb{R}\times M\times X$ par la même formule que ci-dessus.

$$\widetilde{\Phi}_v(b,k,m,x)=(b,k,m,\exp(D_{(b,k,m)}(v))\,x)\ \text{ où}$$

$$D_{(b,k,m)}(v)=e^{k-\varphi(b)}s(\xi(b))\,m\,v. \tag{6.11}$$

Avant de poursuivre démontrons l'égalité suivante, pour β-presque tout b dans B, tout (k,m) dans $\mathbb{R}\times M$ et tout v dans V_0,

$$b_0^{-1}D_{(b,k,m)}(v)=D_{S(b,k,m)}(v)\ \text{ où} \tag{6.12}$$

$$S(b,k,m)=(Tb,k-\tau_{\mathbb{R}}(b),\tau_M(b)m). \tag{6.13}$$

Pour cela, on calcule comme dans le lemme 5.4,

$$\begin{aligned}b_0^{-1}D_{(b,k,m)}(v)&=e^{k-\varphi(b)}b_0^{-1}s(\xi(b))\,m\,v\\&=e^{k-\varphi(b)}s(\xi(Tb))\theta(b)^{-1}\,m\,v\\&=e^{k-\varphi(b)-\theta_{\mathbb{R}}(b)}s(\xi(Tb))\theta_M(b)\,m\,v,\end{aligned}$$

d'où, en utilisant (5.4)

$$\begin{aligned}b_0^{-1}D_{(b,k,m)}(v)&=e^{k-\tau_{\mathbb{R}}(b)-\varphi(Tb)}s(\xi(Tb))\tau_M(b)\,m\,v\\&=D_{S(b,k,m)}(v).\end{aligned}$$

On déduit, grâce à (6.12), les deux égalités suivantes

$$S\circ\widetilde{\Phi}_v=\widetilde{\Phi}_v\circ S\ \text{ et} \tag{6.14}$$

$$\widetilde{T}_\ell^{\tau,X}\circ\widetilde{\Phi}_v=\widetilde{\Phi}_{e^{-\ell}v}\circ\widetilde{T}_\ell^{\tau,X} \tag{6.15}$$

qui prouvent que le flot Φ_v vérifie la relation (6.10). $\square$

6.6. **Probabilités conditionnelles horocycliques.**

On introduit dans cette section la fonction « mesure conditionnelle horocyclique » et on montre que cette fonction est mesurable pour la tribu queue.

On garde les notations de la section 6.1 et on note encore t_v la translation sur V_0 par un élément v de V_0. On note $\sigma : B^{\tau,X} \to \mathcal{M}_1(V_0)$ l'application « mesure conditionnelle de la probabilité $\beta^{\tau,X}$ pour l'action horocyclique de V_0 ».

Lemme 6.11. *Dans les deux cas de 6.1, il existe une partie borélienne $E \subset B^{\tau,X}$ telle que $\beta^{\tau,X}(E^c) = 0$ et que, pour tout $v \in V_0$ et $(c,x) \in E$ tel que $\Phi_v(c,x) \in E$,*

$$t_{v*}\sigma(\Phi_v(c,x)) \simeq \sigma(c,x). \tag{6.16}$$

Démonstration. Cela résulte de la proposition 4.2. □

Rappelons que le symbole $\simeq$ signifie une égalité à normalisation près.

Géométriquement, pour $\beta^{\tau,X}$ presque tout (c,x) dans $B^{\tau,X}$, $\sigma(c,x)$ est la mesure conditionnelle de $\delta_c \otimes \nu_c$ pour l'action de V_0 sur $\{c\} \times X$.

Lemme 6.12. *Dans les deux cas de 6.1, pour tout $\ell \geq 0$, pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$, on a l'égalité*

$$\sigma(T_\ell^{\tau,X}(c,x)) \simeq (e^{-\ell})_*(\sigma(c,x)).$$

Dans cette égalité $e^{-\ell}$ désigne l'homothétie de rapport $e^{-\ell}$ dans V_0.

Démonstration. Cela résulte de l'unicité de σ, de l'égalité (6.10) et du fait que, pour β-presque tout b dans B, pour tout p dans $\mathbb{N}$, l'action de $b_{p-1}^{-1} \dots b_0^{-1}$ induit un isomorphisme entre les espaces mesurés (X, ν_b) et $(X, \nu_{T^p b})$. □

Corollaire 6.13. *Dans les deux cas de 6.1, l'application $\sigma : B^{\tau,X} \to \mathcal{M}_1(V_0)$ est $\mathcal{Q}_\infty^{\tau,X}$-mesurable.*

Démonstration. Il suffit de montrer que, pour tout $\ell \geq 0$, elle est $\mathcal{Q}_\ell^{\tau,X}$-mesurable. Cela résulte de l'égalité, pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$, $\sigma(c,x) \simeq (e^{\ell})_*(\sigma(T_\ell^{\tau,X}(c,x)))$. □

6.7. **L'approche hors des feuilles W.**

Pour pouvoir démarrer l'argument de dérive on aura besoin, dans tout compact de $\beta^{\tau,X}$-mesure positive, d'approcher presque tout point x par des points qui ne sont pas dans la même feuille que x pour un certain sous-feuilletage du feuilletage stable relatif.

Pour b dans B, introduisons sous-espace vectoriel de V

$$W_b = \left\{ v \in V \mid \sup_{p\in\mathbb{N}}(e^{\theta_{\mathbb{R},p}(b)} \left\| b_p^{-1} \dots b_0^{-1} v \right\|) < \infty \right\} \tag{6.17}$$

et, pour $c = (b, k, m)$ dans B^τ, posons $W_c = W_b$.

Lemme 6.14. *Dans les deux cas de 6.1, pour β^X-presque tout (b, x) dans B^X, on a $\nu_b(\exp(W_b)x) = 0$.*

Démonstration. Par ergodicité du système de Bernoulli $(B, \mathcal{B}, \beta, T)$ et par la formule (5.3) de Furstenberg, pour β-presque tout b dans B, on a $\lim\limits_{p\to\infty} \frac{1}{p}\theta_{\mathbb{R},p}(b) = \int_B \theta_{\mathbb{R}}(b)\, \mathrm{d}\beta(b) = \lambda_1 > 0$. Donc, d'après le lemme 5.4, pour tout v dans W_b, on a $\lim\limits_{p\to\infty} \left\| b_p^{-1} \dots b_0^{-1} v \right\| = 0$. Choisissons une distance d sur X, provenant d'une distance invariante à droite sur le groupe $\widetilde{X}$, revêtement universel de X. Pour β-presque tout b dans B, tout x dans X, et tout v dans W_b, on a

$$d(b_p^{-1} \dots b_0^{-1} \exp(v)x, b_p^{-1} \dots b_0^{-1} x) \xrightarrow[p\to\infty]{} 0.$$

Comme, d'après la proposition 6.1, la mesure μ satisfait la propriété (HC), d'après la proposition 3.9, pour β^X-presque tout (b, x) dans B^X, on a $\nu_b(\exp(W_b)x) = 0$, ce qu'il fallait démontrer.□

Corollaire 6.15. *Dans les deux cas de 6.1, notons $F \subset B^{\tau,X}$ une partie $\mathcal{B}^{\tau,X}$-mesurable telle que $\beta^{\tau,X}(F) > 0$. Alors, pour $\beta^{\tau,X}$-presque tout (c, x) dans F, il existe une suite (u_n) d'éléments de $V \smallsetminus W_c$ telle que $u_n \xrightarrow[n\to\infty]{} 0$ et que, pour tout entier n, on ait $(c, \exp(u_n)x) \in F$.*

Démonstration. Soit (U_n) une base de voisinages de 0 dans V. Pour β^τ-presque tout c dans B^τ, la tranche $F_c = \{x \in X | (c, x) \in F\}$ vérifie $\nu_c(F_c) > 0$. Pour $\beta^{\tau,X}$-presque tout (c, x) dans F, pour tout $n \geq 0$, on a alors $\nu_c(F_c \cap \exp(U_n)x) > 0$ et, donc, comme, d'après le lemme 6.14, $\nu_c(\exp(W_c)x) = 0$, on a bien $\nu_c(F_c \cap (\exp(U_n \smallsetminus W_c)x)) > 0$.□

7. Invariance des probabilités stationnaires

Le but de ce chapitre est de présenter l'argument de dérive exponentielle et d'en déduire des propriétés d'invariance pour certaines mesures conditionnelles des probabilités stationnaires (proposition 7.6).

Pour cela, nous rassemblons, les pièces du puzzle qui sont éparpillées dans les chapitres précédents.

7.1. **La dérive exponentielle.**

Le cœur de cet article est la proposition ci-dessous.

On garde toujours les notations de 6.1. En particulier, μ est une probabilité Zariski dense sur G, ν est une probabilité borélienne μ-stationnaire et μ-ergodique sur X et les symboles s, ξ, θ, $\theta_{\mathbb{R}}$, φ, $\tau_{\mathbb{R}}$, τ_M, τ, B^τ, β^τ, $\beta^{\tau,X}$, σ, R, etc... ont le même sens qu'aux chapitres 5 et 6.

Proposition 7.1. *Dans les deux cas de 6.1, soient $(Y, \mathcal{Y})$ un espace borélien standard, $f : B^{\tau,X} \to Y$ une application $\mathcal{Q}_\infty^{\tau,X}$-mesurable et $E \subset B^{\tau,X}$, un ensemble $\mathcal{B}^{\tau,X}$-mesurable avec $\beta^{\tau,X}(E^c) = 0$. Alors, pour $\beta^{\tau,X}$-presque tout (c, x) dans $B^{\tau,X}$, pour tout $\varepsilon > 0$, il existe un élément v non nul de V_0 de norme au plus ε et un élément (c', x') de E tels que $\Phi_v(c', x')$ est aussi dans E et que*

$$f(\Phi_v(c', x')) = f(c', x') = f(c, x). \tag{7.1}$$

Remarque 7.2. Comme on ne sait pas encore que le flot horocyclique préserve la mesure $\beta^{\tau,X}$ (on ne le saura que dans la section 8.1), il n'est pas clair a priori qu'*il existe un élément (c', x') dans E et un vecteur v non nul dans V_0 tel que $\Phi_v(c', x')$ est dans E*. Cette assertion sera donc une conséquence non triviale de la proposition 7.1.

Début de la démonstration de la proposition 7.1. Par définition, on peut considérer que Y est muni d'une topologie d'espace métrique séparable complet pour laquelle $\mathcal{Y}$ est la tribu borélienne. De même, choissons une topologie d'espace métrique compact sur B^τ dont la tribu borélienne est égale, aux ensembles de mesure nulle près, à $\mathcal{B}^\tau$ et telle que la projection naturelle $B^\tau \to M$ soit continue, et munissons $B^\tau \times X$ de la topologie produit de cette topologie et de la topologie usuelle de X.

Soit $\alpha > 0$ très petit. D'après le théorème de Lusin, il existe un compact $K \subset E$ de $B^{\tau,X}$ tel que $\beta^{\tau,X}(K^c) < \alpha^2$ et tel que toutes les fonctions que nous allons rencontrer comme les fonctions f, θ, $(c, x) \mapsto \varphi(b)$, $(c, x) \mapsto V_c$ ou encore $(c, x) \mapsto D_c \in \mathrm{Hom}(V_0, V_c)$ sont uniformément continues sur K.

La preuve repose alors sur une étude de la fonction $\mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\infty^{\tau,X})$.

D'une part, cette fonction est majorée par 1 et sa moyenne est minorée par $1 - \alpha^2$ car :

$$\int_{B^{\tau,X}} \mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\infty^{\tau,X})(c, x)\, \mathrm{d}\beta^{\tau,X}(c, x) = \beta^{\tau,X}(K) > 1 - \alpha^2, \tag{7.2}$$

Cette fonction $\mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\infty^{\tau,X})$ est donc minorée par $1-\alpha$ sur un ensemble de mesure au moins $1 - \alpha$. Il existe donc un compact $L \subset E$ de $B^{\tau,X}$

tel que $\beta^{\tau,X}(L^c) < \alpha$ et que, pour tout (c,x) dans L, on a

$$\mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\infty^{\tau,X})(c,x) > 1-\alpha. \tag{7.3}$$

D'après le théorème de Lusin, on peut aussi supposer que f est continue sur L.

D'autre part, d'après le théorème de convergence des martingales, pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$ on a,

$$\lim_{\ell\to\infty} \mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\ell^{\tau,X})(c,x) = \mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\infty^{\tau,X})(c,x). \tag{7.4}$$

D'après le corollaire 3.8, on peut aussi supposer que, pour tout (c,x) dans L et tout ℓ rationnel, le terme de gauche dans cette égalité (7.4) est donné par la formule (3.11). Grâce à la loi des sauts (proposition 2.3), dont on reprend les notations $h_{\ell,c}(a)$, cette formule se réécrit,

$$\mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\ell^{\tau,X})(c,x) = \int_B \mathbf{1}_K(h_{\ell,c,x}(a))\,\mathrm{d}\beta(a) \tag{7.5}$$

où

$$h_{\ell,c,x}(a) = (c',x') \ \text{ avec } \ c' = h_{\ell,c}(a) \ \text{ et } \ x' = \rho_\ell(c')^{-1}\rho_\ell(c)x.$$

Par ailleurs, comme f est $\mathcal{Q}_\infty^{\tau,X}$-mesurable, elle est $\mathcal{Q}_\ell^{\tau,X}$-mesurable pour tout $\ell \geq 0$ et, donc, toujours d'après le corollaire 3.8 et la proposition 2.3, on peut aussi supposer que, pour tout (c,x) dans K, pour β-presque tout a dans B, pour tout $\ell \geq 0$ rationnel, on a $f(h_{\ell,c,x}(a)) = f(c,x)$.

Le théorème d'Egoroff assure que, quitte à ôter à L un ensemble de $\beta^{\tau,X}$-mesure arbitrairement petite, la convergence dans (7.4) est uniforme sur L. Il existe donc $\ell_0 \geq 0$ tel que pour tout entier $\ell \geq \ell_0$, pour tout (c,x) dans L, on a

$$\mathbb{E}(\mathbf{1}_K \mid \mathcal{Q}_\ell^{\tau,X})(c,x) \geq 1-\alpha \tag{7.6}$$

Comme la $\beta^{\tau,X}$-mesure de L^c est au plus α et que α est arbitrairement petit, il suffit pour conclure de montrer l'assertion (7.1) pour $\beta^{\tau,X}$-presque tous les points (c,x) de L.

Le corollaire 6.15 permet de supposer, pour ces points (c,x) de L, qu'il existe une suite (u_n) d'éléments de $V \smallsetminus W_c$ convergeant vers 0 et telle que les points $(c,y_n) := (c,\exp(u_n)x)$ sont aussi dans L.

On applique les deux formules (7.5) et (7.6) pour l'espérance conditionnelle à chacun de ces deux points (c,x) et (c,y_n). Pour $\ell \geq \ell_0$, on a donc

$$\beta(\{a\in B \mid h_{\ell,c,x}(a)\in K\}) \geq 1-\alpha \tag{7.7}$$

et

$$\beta(\{a\in B \mid h_{\ell,c,y_n}(a)\in K\}) \geq 1-\alpha. \tag{7.8}$$

Disons maintenant quelques mots sur la suite de la stratégie de preuve. Par construction, lorsque $y = \exp(u)x$ avec $u \in \mathfrak{g}$, les paramétrisations des deux fibres de $T_\ell^{\tau,X}$ passant par (c, x) et (c, y) sont reliés par une *dérive* facile à calculer : si $(c', x') = h_{\ell,c,x}(a)$ et $(c', y') = h_{\ell,c,y}(a)$, alors on a

$$y' = \exp(F_{\ell,c}(a)u)x' \tag{7.9}$$

où la dérive est donnée par

$$F_{\ell,c}(a)u = R_\ell(c') \circ R_\ell(c)^{-1}u. \tag{7.10}$$

où, comme dans la section 3.3, en écrivant $c = (b, k, m)$, et $p = p_\ell(b, k)$ on a

$$R_\ell(c) = R(b_0) \circ \cdots \circ R(b_{p-1}). \tag{7.11}$$

Pour alléger les notations, on écrira parfois b_0 pour $R(b_0)$. On vient de voir que, pour la paramétrisation des deux fibres de $T_\ell^{\tau,X}$ passant par les points (c, x) et (c, y_n), une grosse proportion des paramètres $a \in B$ correspondent à deux points (c'_n, x'_n) et (c'_n, y'_n) qui sont tous deux dans K. Nous allons maintenant ajuster le rayon $\ell = \ell_n$ à la suite u_n de façon à contrôler la norme et la direction de la dérive séparant ces deux points.

Ce sera possible grâce au lemme suivant.

Lemme 7.3. *Dans les deux cas de 6.1, pour tous $\alpha > 0$ et $\eta > 0$, il existe $r_0 > 1$, tel que, pour β^τ-presque tout $c \in B^\tau$, pour ℓ suffisament grand, on a, pour tout $u \in V \smallsetminus 0$,*

$$\beta(\{a \in B \mid \frac{1}{r_0} \leq \frac{\|F_{\ell,c}(a)u\|}{e^{\theta_{\mathbb{R},\ell}(c)}\|R_\ell(c)^{-1}u\|} \leq r_0\}) \geq 1 - \alpha. \tag{7.12}$$

et

$$\beta(\{a \in B \mid d(\mathbb{R}\, F_{\ell,c}(a)u, \mathbb{P}(V_{h_{\ell,c}(a)})) \leq \eta\}) \geq 1 - \alpha. \tag{7.13}$$

Démonstration. Rappelons que, d'après la section 2.3, pour β^τ-presque tout $c \in B^\tau$, pour β-presque tout $a \in B$, on a $\lim_{\ell\to\infty} q_{\ell,c}(a) = \infty$.

Pour obtenir la majoration (7.12), on applique le corollaire 5.5.a aux vecteurs $v_1 = R_\ell(c)^{-1}u$ et v_2 dans $V_{T_\ell^\tau(c)}$ de sorte qu'on a les égalités, avec $q = q_{\ell,c}(a)$,

$$\frac{\|F_{\ell,c}(a)u\|}{\|R_\ell(c)^{-1}u\|} = \frac{\|a_{q-1}\cdots a_0 v_1\|}{\|v_1\|}$$

et, grâce aussi au lemme 5.4, on a l'égalité, avec $c' := h_{\ell,c}(a)$,

$$e^{\theta_{\mathbb{R},\ell}(c')} = \frac{\|a_{q-1}\cdots a_0 v_2\|}{\|v_2\|}$$

Pour montrer (7.13), appliquons le corollaire 5.5.b au même vecteur $v = R_\ell(c)^{-1}u$ et avec $W = V_{T_\ell^\tau(c)}$. Pour β^τ-presque tout $c \in B^\tau$, pour tous $\alpha, \eta > 0$, il existe donc $\ell_0 \geq 0$ tel que, pour tous $u \in V \smallsetminus 0$ et $\ell \geq \ell_0$,

$$\beta(\{a \in B \mid d(\mathbb{R}\, F_{\ell,c}(a)u, \mathbb{P}(V_{h_{\ell,c}(a)})) \leq \eta\}) \geq 1 - \alpha,$$

ce qu'il fallait démontrer. □

Fin de la démonstration de la proposition 7.1 Détaillons maintenant de façon plus précise notre stratégie. On ajuste le paramètre $\ell = \ell_n$ de la façon suivante.

Comme la mesure μ sur G est à support compact et que la section s du paragraphe 5.2 est à image bornée, il existe $C_0 > 0$ tel que, pour β-presque tout b dans B, pour tous u dans $V \smallsetminus 0$ et p dans $\mathbb{N}$, on ait

$$\frac{e^{\theta_{\mathbb{R},p+1}(b)} \left\| b_{p+1}^{-1} \dots b_0^{-1} u \right\|}{e^{\theta_{\mathbb{R},p}(b)} \left\| b_p^{-1} \dots b_0^{-1} u \right\|} \leq C_0.$$

Comme u_n n'est pas dans W_c, la suite $p \mapsto e^{\theta_{\mathbb{R},p}(b)} \left\| b_p^{-1} \dots b_0^{-1} u_n \right\|$ n'est pas majorée. Pour n assez grand, il existe donc un entier p_n tel que

$$\frac{e^{-M_0}\varepsilon}{r_0 C_0} \leq e^{\theta_{\mathbb{R},p_n}(b)} \left\| b_{p_n}^{-1} \dots b_0^{-1} u_n \right\| \leq \frac{e^{-M_0}\varepsilon}{r_0} \tag{7.14}$$

où $M_0 = \sup \tau$. On choisit un rationnel l_n tel que $p_n = p_{l_n}(c)$. C'est possible car τ est strictement positive.

On peut donc, dès que $\alpha < \frac{1}{4}$, choisir un élément $a = a_n$ dans B de sorte qu'il soit simultanément dans les ensembles donnés par (7.7) et (7.8), (7.12) et (7.13) avec $\ell = \ell_n$, $u = u_n$ et $\eta = \eta_n \to 0$ et que

$$f(h_{\ell_n,c,x}(a_n)) = f(c,x) \text{ et } f(h_{\ell_n,c,y_n}(a_n)) = f(c,y_n). \tag{7.15}$$

Quitte à extraire une sous-suite,

(1) la suite $(c'_n, x'_n) := h_{\ell_n,c,x}(a_n)$ a une limite (c', x') dans K,
(2) la suite $(c'_n, y'_n) := h_{\ell_n,c,y_n}(a_n)$ a une limite (c', y') dans K et
(3) la limite de la dérive $w = \lim_{n\to\infty} F_{\ell_n,c}(a_n)u_n$ existe, est non nulle, de norme au plus $e^{-M_0}\varepsilon$, et appartient à $V_{c'}$.

On en déduit, en passant à la limite dans (7.15), comme tous les suites considérées prennent leurs valeurs dans K et L et que f est continue sur ces ensembles,

$$f(c', x') = \lim_{n\to\infty} f(c'_n, x'_n) = \lim_{n\to\infty} f(c, x) = f(c, x),$$
$$f(c', y') = \lim_{n\to\infty} f(c'_n, y'_n) = \lim_{n\to\infty} f(c, y_n) = f(c, x) \text{ et}$$
$$y' = \exp(w)x'.$$

De plus, en notant $v \in V_0$ l'élément non nul $v = D_{c'}^{-1}(w)$, on a

$$\|v\| \leq \varepsilon \ \text{ et } \ (c', y') = \Phi_v(c', x'),$$

ce qui est l'assertion cherchée. □

7.2. **Le stabilisateur des mesures conditionnelles.**

Explicitons l'information que nous fournit la dérive sur les mesures conditionnelles horocycliques $\sigma(c, x)$, pour $\beta^{\tau,X}$-presque tout (c, x) dans $B^{\tau,X}$.

Introduisons les stabilisateurs connexes de la mesure $\sigma(c, x)$ et de sa classe $\mathbb{R}_+^*\sigma(c, x)$ modulo normalisation :

$$J(c, x) := \{v \in V_0 \mid t_{v*}\sigma(c, x) = \sigma(c, x)\}_0 \ .$$

$$J_1(c, x) := \{v \in V_0 \mid t_{v*}\sigma(c, x) \simeq \sigma(c, x)\}_0 \ ,$$

Ce sont des sous-groupes fermés connexes et donc des sous-espaces vectoriels de V_0.

Proposition 7.4. *Dans les deux cas de 6.1, pour $\beta^{\tau,X}$-presque tout (c, x) dans $B^{\tau,X}$, on a*
a) $J_1(c, x) \neq \{0\}$,
b) $J(c, x) = J_1(c, x)$.

Démonstration. a) On va montrer que, pour $\beta^{\tau,X}$-presque tout (c, x) et pour tout $\varepsilon > 0$, le stabilisateur de $\sigma(c, x)$ modulo normalisation contient un vecteur non nul de norme au plus ε.

D'après le corollaire 6.11, il existe une partie borélienne $E \subset B^{\tau,X}$ telle que $\beta^{\tau,X}(E^c) = 0$ et que, pour tous $v \in V_0$ et $(c', x') \in E$ tel que $\Phi_v(c', x') \in E$, on a

$$t_{v*}\sigma(\Phi_v(c', x')) \simeq \sigma(c', x'). \tag{7.16}$$

D'après le corollaire 6.13, la fonction σ est $\mathcal{Q}_\infty^{\tau,X}$-mesurable. La dérive (proposition 7.1) appliquée à cet ensemble E et à cette fonction $f = \sigma$ permet de trouver, pour $\beta^{\tau,X}$-presque tout (c, x) dans $B^{\tau,X}$ et pour tout $\varepsilon > 0$, un vecteur non nul v de V_0 de norme au plus ε et un élément (c', x') de E tel que $\Phi_v(c', x')$ est aussi dans E et que

$$\sigma(\Phi_v(c', x')) \simeq \sigma(c', x') \simeq \sigma(c, x).$$

En appliquant l'égalité (7.16) à cet élément (c', x'), on obtient

$$t_{v*}\sigma(\Phi_v(c', x')) \simeq \sigma(c', x')$$

et donc

$$t_{v*}\sigma(c, x) \simeq \sigma(c, x).$$

Le vecteur v est bien dans le stabilisateur de $\sigma(c,x)$ modulo normalisation. Ce stabilisateur est non discret et fermé. Il contient donc un sous-espace vectoriel non nul de V_0.

b) Pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$, il existe une forme linéaire $\alpha(c,x) \in J_1(c,x)^*$, telle que, pour tout v dans $J_1(c,x)$,

$$t_{v*}\sigma(c,x) = e^{\alpha(c,x)(v)}\sigma(c,x).$$

On veut montrer que $\alpha = 0$. Le lemme 6.12 entraine, pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$, l'égalité $J_1(T_\ell^{\tau,X}(c,x)) = J_1(c,x)$ et, pour tout $\ell \geq 0$, l'égalité de formes linéaires sur $J_1(c,x)$

$$\alpha(T_\ell^{\tau,X}(c,x)) = e^\ell \alpha(c,x), \tag{7.17}$$

si bien que, d'après le théorème de récurrence de Poincaré on a, $\beta^{\tau,X}$-presque partout, $\alpha = 0$. □

7.3. Désintégration de ν_b le long des stabilisateurs.

On désintègre dans cette section les mesures limites ν_b selon les valeurs de la composante connexe du stabilisateur des conditionnelles horocycliques. On obtient ainsi des probabilités $\nu_{b,x}$ qui sont invariantes par un sous-groupe unipotent non trivial.

Commençons par traduire le fait que les stabilisateurs des conditionnelles horocycliques ne sont pas discrets en un énoncé qui ne fait pas intervenir la suspension B^τ.

Pour β-presque tout b dans B, et ν_b-presque tout x dans X, on note $\sigma_{b,x} \in \mathcal{M}(V_b)$ la mesure conditionnelle en x de ν_b pour l'action sur X de V_b via le groupe $\exp V_b$ (voir section 4.1), et on note $V_{b,x} \subset V_b$ la composante connexe du stabilisateur dans V_b de $\sigma_{b,x}$.

Proposition 7.5. *Dans les deux cas de 6.1, pour β^X-presque tout (b,x) dans B^X, on a $\sigma_{b,x} \simeq b_{0*}\sigma_{T^X(b,x)}$, $V_{b,x} = b_0(V_{T^X(b,x)})$ et $V_{b,x} \neq 0$.*

Démonstration. La première égalité résulte des égalités, pour β-presque tout b dans B, $\nu_{Tb} = (b_0^{-1})_*\nu_b$ et, pour tout x dans X et $v \in \mathfrak{g}$,

$$T^X(b, \exp(v)x) = (Tb, \exp(b_0^{-1}v)b_0^{-1}x).$$

La deuxième égalité s'en déduit.

La non-nullité de $V_{b,x}$ résulte de la proposition 7.4 et de l'égalité, pour $\beta^{\tau,X}$-presque tout (c,x) dans $B^{\tau,X}$, $V_{b,x} = R(s(\xi(b))m)(J(c,x))$, où $c = (b,k,m)$. □

La désintégration de β^X le long de l'application $(b,x) \mapsto (b, V_{b,x})$, ou, ce qui revient au même, la désintégration pour β-presque tout b de ν_b

le long de l'application $x \mapsto V_{b,x}$, s'écrit

$$\nu_b = \int_X \nu_{b,x}\, \mathrm{d}\nu_b(x)$$

où, pour β^X-presque tout (b,x) dans B^X, la probabilité $\nu_{b,x}$ sur X est portée par la fibre $\{x' \in X \mid V_{b,x'} = V_{b,x}\}$

Proposition 7.6. *Dans les deux cas de 6.1, pour β^X-presque tout (b,x) dans B^X, la probabilité $\nu_{b,x}$ est $V_{b,x}$-invariante et on a la propriété d'équivariance : $\nu_{b,x} = b_{0*}\nu_{Tb,b_0^{-1}x}$.*

Démonstration. La première assertion résulte de la proposition 4.3.

La deuxième assertion résulte de l'égalité $\nu_b = b_{0*}\nu_{Tb}$, de la proposition 7.5 et de l'unicité de la désintégration d'une probabilité. □

8. Applications

Dans ce chapitre, nous terminons la démonstration des théorèmes 1.1 et 1.3 et de leurs corollaires.

8.1. **Invariance des probabilités stationnaires.**

Nous gardons toujours les notations de 6.1 et nous terminons dans cette section la classification des probabilités stationnaires sur X.

Proposition 8.1. *Dans les deux cas de 6.1, la probabilité ν est la probabilité de Haar sur X.*

Pour déduire cette proposition de la proposition 7.6, nous aurons besoin du lemme suivant. Soit α dans $\mathcal{P}(X)$.

Dans le premier cas de 6.1, on note S_α la composante connexe du stabilisateur de α dans G qui agit par translation sur $X = G/\Lambda$.

Dans le deuxième cas de 6.1, on note S_α la composante connexe du stabilisateur de α dans $V = \mathbb{R}^d$ qui agit par translations sur $X = \mathbb{T}^d$.

Dans les deux cas, on pose

$$\mathcal{F} := \{\alpha \in \mathcal{P}(X) \mid S_\alpha \neq \{1\} \text{ et } \alpha \text{ est portée par une } S_\alpha\text{-orbite}\}.$$

et on le munit de la topologie de la convergence vague des mesures.

Remarquons que le groupe G agit naturellement sur $\mathcal{F}$. Notons ν_0 la probabilité de Haar sur X. La mesure ν_0 est un élément de $\mathcal{F}$.

Lemme 8.2. *Dans les deux cas de 6.1, la seule probabilité borélienne μ-stationnaire η sur $\mathcal{F}$ est δ_{ν_0}*

Démonstration. On peut supposer que η est μ-ergodique. On va distinguer les deux cas :

Premier cas de 6.1 Dans ce cas, on a $X = G/\Lambda$.

D'après [30, Theorem 1.1], l'ensemble $\mathcal{G}$ des orbites de G dans $\mathcal{F}$ est un ensemble dénombrable.

L'image $\overline{\eta}$ de η dans $\mathcal{G}$ est une probabilité μ-stationnaire ergodique sur un ensemble dénombrable. D'après le lemme 8.3, la probabilité $\overline{\eta}$ est à support fini.

Comme la probabilité μ-stationnaire η est μ-ergodique, elle est portée par une seule orbite $G\alpha \simeq G/G_\alpha \subset \mathcal{F}$. Par définition de $\mathcal{F}$, le groupe G_α n'est pas discret. Comme ce groupe G_α contient un réseau, il est unimodulaire. La proposition 6.7 prouve alors que $G_\alpha = G$. La probabilité ν est donc égale à ν_0.

Deuxième cas de 6.1 Dans ce cas, on a $X = \mathbb{T}^d$.

Notons $\mathcal{G}$ l'ensemble des sous tores non-triviaux de X, et, pour Y dans $\mathcal{G}$, notons $\mathcal{F}_Y$ l'ensemble des mesures translatées de la probabilité de Haar sur Y. L'espace $\mathcal{F}$ est alors la réunion dénombrable des parties compactes $\mathcal{F}_Y$.

L'image $\overline{\eta}$ de η dans $\mathcal{G}$ est une probabilité μ-stationnaire ergodique sur un ensemble dénombrable. D'après le lemme 8.3, la probabilité $\overline{\eta}$ est à support fini $Y_1, \ldots, Y_n$ et Γ permute les sous-espaces $V_1, \ldots, V_n$, directions des tores $Y_1, \ldots, Y_n$.

Comme l'action de Γ sur V est fortement irréductible, on a nécessairement $V_1 = \ldots = V_n = V$, ce qu'il fallait démontrer. $\square$

On a utilisé le lemme classique suivant

Lemme 8.3. *Soient Γ un groupe agissant sur un espace dénombrable X et μ une mesure de probabilité sur Γ. Toute probabilité μ-stationnaire μ-ergodique ν sur X est Γ-invariante et a un support fini.*

Démonstration du lemme 8.3. Soit Y l'ensemble fini des atomes de masse maximale de ν. L'égalité $\mu * \nu = \nu$ et le principe du maximum impliquent que, pour μ-presque tout γ dans Γ, on a $\gamma^{-1}Y \subset Y$ et, donc, $\gamma^{-1}Y = Y$. Comme $\nu(Y) > 0$ et que ν et μ-ergodique, on a $\nu(Y) = 1$, d'où le résultat. $\square$

Démonstration de la proposition 8.1. D'après la proposition 7.5, fruit de tous nos efforts, pour β^X presque tout (b, x) dans B^X, les sous-groupes $V_{b,x}$ sont non triviaux.

Le principal intérêt de *l'ensemble $\mathcal{F}$ est qu'il contient toutes les probabilités invariantes et ergodiques sous l'action d'un groupe unipotent connexe non trivial.* Ce fait résulte du théorème de Ratner (voir [30] ou [28]) dans le premier cas ; il est élémentaire dans le second.

Pour β^X-presque tout (b,x) dans B^X, les décompositions de $\nu_{b,x}$ en composantes $V_{b,x}$-ergodiques peuvent donc s'écrire simultanément sous la forme

$$\nu_{b,x} = \int_X \zeta(b,x')\,\mathrm{d}\nu_{b,x}(x'). \tag{8.1}$$

où $\zeta : B^X \to \mathcal{F}$ est une application $\mathcal{B}^X$-mesurable telle que, pour β^X-presque tout (b,x) dans B^X, la restriction de ζ à la fibre $\{(b,x') \mid V_{b,x'} = V_{b,x}\}$ est constante sur les $V_{b,x}$-orbites.

L'unicité des décompositions ergodiques et les propositions 7.5 et 7.6 prouvent que, pour β^X-presque tout (b,x) dans B^X, on a

$$\zeta(b,x) = (b_0)_*\zeta(T^X(b,x)). \tag{8.2}$$

D'après le lemme 3.2.e, la probabilité image $\eta := \zeta_*\beta^X$ est donc une probabilité μ-stationnaire sur $\mathcal{F}$. D'après le lemme 8.2. Cette probabilité est la masse de Dirac en ν_0. Autrement dit $\zeta(b,x)$ est β^X-presque sûrement égale à ν_0 et on a $\nu = \nu_0$. □

Démonstration des théorèmes 1.1 *et* 1.3. Rappelons que, dans le deuxième cas, on a noté G l'adhérence de Zariski de Γ_μ dans $\mathrm{SL}(d,\mathbb{R})$. Le lemme 8.5 ci-dessous affirme que G est aussi semi-simple.

Dans les deux cas, le lemme 8.4 ci-dessous permet de supposer que G est un groupe de Lie semi-simple non compact. On peut donc appliquer la proposition 8.1 : la probabilité ν est G-invariante. □

Nous avons utilisé les deux lemmes faciles suivants.

Lemme 8.4. *Soient K un groupe compact métrisable agissant de façon borélienne sur un espace borélien X et μ une probabilité borélienne sur K. Alors toute probabilité borélienne μ-stationnaire ν sur X est invariante par le groupe Γ_μ engendré par le support de μ.*

Démonstration. Par le théorème de Varadarajan (proposition 2.1.19 de [35]) on peut supposer que X est un espace compact et que l'action est continue. On peut aussi supposer que ν est μ-ergodique. Elle est alors portée par une seule K-orbite $K\,x_0$. On donc considérer ν comme une probabilité H-invariante à droite sur K, où H est le stabilisateur de x_0. Cette probabilité relevée est encore μ-stationnaire. On se ramène ainsi au cas où $X = K$.

Quitte à convoler ν à droite par une approximation de l'identité, on peut supposer que ν est absolument continue par rapport à la mesure de Haar, avec une densité continue. On regarde alors ν comme un élément de $L^2(K)$ vérifiant $\mu * \nu = \nu$. Mais, dans un espace de Hilbert, une moyenne de vecteurs de même norme a une norme strictement plus petite à moins que ces vecteurs soient tous égaux, ce qui prouve que ν est Γ_μ-invariante. □

Lemme 8.5. *Soit* Γ *un sous-semi-groupe de* $\mathrm{SL}(d, \mathbb{Z})$ *qui agit de façon fortement irréductible sur* $\mathbb{R}^d$*. Alors son adhérence de Zariski* G *dans* $\mathrm{SL}(d, \mathbb{R})$ *est un groupe semi-simple.*

Démonstration. On peut supposer G Zariski connexe. Comme la représentation de G dans $\mathbb{R}^d$ est irréductible, G est un groupe réductif. Comme en outre G est formé de matrices de déterminant 1, le centre Z de G est un groupe compact. On veut montrer que ce centre Z est fini.

Supposons, par l'absurde que le centre Z est infini. Le commutant de G dans $\mathrm{End}(\mathbb{Q}^d)$ est alors une extension quadratique imaginaire K de $\mathbb{Q}$. On peut donc regarder $\mathbb{Q}^d$ comme un K-espace vectoriel. L'application déterminant $g \mapsto \det_K(g)$ envoie Γ dans le groupe U_K des unités de K. Comme ce groupe U_K est fini, l'application déterminant envoie aussi G dans U_K. Donc Z est aussi fini, ce qui est contradictoire. □

8.2. **Mesures invariantes.**

Pour déduire les corollaires des théorèmes, nous devons choisir convenablement la probabilité μ.

Démonstration des corollaires 1.2.*a et* 1.4.*a.* Comme G est simple, tout sous-semi-groupe Zariski dense Γ de G contient un sous-semi-groupe Γ' de type fini qui est encore Zariski dense dans G. Notons $g_1, \ldots, g_\ell$ un système de générateurs de Γ' et $\mu = \frac{1}{\ell}(\delta_{g_1} + \cdots + \delta_{g_\ell}) \in \mathcal{P}(G)$.

Soit ν une probabilité sans atome sur X qui soit invariante par Γ. La probabilité ν est donc μ-stationnaire. Le théorème 1.1 affirme que ν est G-invariante, ce qu'il fallait démontrer. □

8.3. **Fermés invariants.**

Pour montrer les corollaires 1.2.b et 1.4.b, nous aurons besoin du lemme suivant.

Lemme 8.6. *Dans les deux cas de 6.1, l'ensemble des parties finies* Γ*-invariantes de* X *est dénombrable.*

Démonstration. Comme précédemment, on peut supposer Γ de type fini. Comme Γ n'a qu'un ensemble dénombrable de sous-groupes Δ d'indice fini, il suffit de montrer que les points fixes dans X d'un tel groupe Δ sont isolés. Ce dernier point résulte du fait qu'au voisinage d'un tel point fixe, l'action de Δ se linéarise en son action dans V et que, comme l'action de Γ sur V est fortement irréductible, Δ n'a pas de vecteur invariant non nul dans V. $\square$

Démonstration des corollaires 1.2.*b et* 1.4.*b.* On peut toujours suposer Γ de type fini. On note alors, comme dans le point a), μ la probabilité $\mu = \frac{1}{\ell}(\delta_{g_1} + \cdots + \delta_{g_\ell})$, où $g_1, \ldots, g_\ell$ est un système de générateurs de Γ. Soit F un fermé infini de X qui soit Γ-invariant. Grâce au lemme 8.6, on peut construire une suite croissante $F_1 \subset \cdots \subset F_i \subset \cdots$ de parties finies Γ-invariantes (éventuellement vides) de X telle que toute partie finie Γ-invariante de X est incluse dans l'un des F_i. Comme F est infini, on peut choisir une suite de points deux à deux distincts $x_1, \ldots, x_i, \ldots$ de F de sorte que x_i ne soit pas dans F_i.

D'après la proposition 6.4 de récurrence hors des orbites finies, il existe une famille $(K_i)_{i \geq 1}$ où, pour tout $i \geq 1$, K_i est un compact de F_i^c telle que, pour tout $j \geq 1$, il existe un entier $M_j \geq 1$ avec, pour tout $n \geq M_j$ et tout $i \leq j$,

$$(\mu^{*n} * \delta_{x_j})(K_i^c) \leq \frac{1}{i}. \tag{8.3}$$

Posons $n_j = jM_j$ et introduisons la moyenne de Birkhoff-Kakutani

$$\nu_j := \frac{1}{n_j}(\mu * \delta_{x_j} + \cdots + \mu^{*n_j} * \delta_{x_j}). \tag{8.4}$$

On a, pour tout $i \leq j$,

$$\nu_j(K_i^c) \leq \frac{M_j}{n_j} + \frac{n_j - M_j}{n_j}\frac{1}{i} \leq \frac{2}{i} \tag{8.5}$$

La condition (8.5) assure que valeurs d'adhérence de la suite (ν_j) pour la convergence vague des mesures boréliennes positives finies sur X sont des mesures de probabilité et qu'elles ne chargent pas les ensembles F_i, $i \geq 1$. Si ν_∞ est une telle valeur d'adhérence, ν_∞ est donc une probabilité μ-stationnaire avec $\nu_\infty(F) = 1$ et ν_∞ est sans atome, d'après le lemme 8.3. Les théorèmes 1.1 et 1.3 montrent que ν_∞ est la probabilité de Haar. On en déduit l'égalité cherchée, $F = X$. $\square$

8.4. **Equirépartitions des orbites finies.**

Les mêmes arguments permettent de montrer l'équirépartition des orbites finies.

Démonstration des corollaires 1.2.*c et* 1.4.*c*. On peut encore supposer que Γ est engendré par le support fini d'une probabilité μ. Nous voulons montrer que la suite de probabilités Γ-invariantes

$$\nu_j := \tfrac{1}{\# X_j} \textstyle\sum_{x \in X_j} \delta_x$$

converge vaguement vers la probabilité de Haar sur X. D'après le point a), nous devons juste montrer que tout limite faible ν_∞ de ν_j est une probabilité qui ne charge pas les orbites finies. La preuve repose sur le phénomène de récurrence hors des orbites finies. Elle est donc analogue au point b) dont nous gardons les notations F_i et K_i.

Comme les Γ-orbites finies X_j sont distinctes, on peut supposer, après extraction d'une sous-suite, que pour tous $j \geq i$, on a $\nu_j(F_i) = 0$. Comme ν_j est Γ-invariante, pour tout $n \geq 0$, on a $\mu^{*n} * \nu_j = \nu_j$, et, donc, comme en b), pour tout $j \geq i$, $\nu_j(K_i^c) \leq \frac{1}{i}$. On en déduit que, pour tout $i \geq 1$, on a $\nu_\infty(K_i^c) \leq \frac{1}{i}$. Ce qui prouve bien que, d'une part, ν_∞ est une probabilité et que, d'autre part, $\nu_\infty(F_i) = 0$ pour tout i, si bien que ν_∞ est la probabilité de Haar. $\square$

Références

[1] F. Blanchard, K-flots et théorème de renouvellement, *Zeitschrift fur Wahrsch. und verw. Gebiete* **36** (1976), 345-358.

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

[3] J. Bourgain, A. Furman, E. Lindenstrauss, S. Mozes, Invariant measures and stiffness for non-abelian groups of toral automorphisms . *C. R. Math. Acad. Sci. Paris* **344** (2007), 737-742.

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

[5] R. Chacon, D. Ornstein, A general ergodic theorem. *Illinois J. Math.* **4** (1960), 153-160.

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

[7] J.-P. Conze, Extensions de systèmes dynamiques par des groupes compacts, *Annales de l'Institut Henri Poincaré* **8** (1972), 33-66.

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

[9] 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.

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

[11] A. Eskin, H. Oh, Ergodic theoretic proof of equidistribution of Hecke points, *Erg. The. Dyn. Sys.* **26** (2006), 163-167.

[12] A. Furman, Random walks on groups and random transformations, *Handbook of dynamical systems*, **1A** North-Holland (2002) 931-1014.

[13] H. Furstenberg, Noncommuting random products, *Trans. Amer. Math. Soc* **108** (1963), 377-428.

[14] H. Furstenberg, Stiffness of group actions, *Lie groups and ergodic theory*, Tata Inst. Fund. Res. Stud. Math. **14** (1998), 105-117.

[15] H. Furstenberg, H. Kesten, Products of random matrices, *Ann. Math. Statist* **31** (1960), 457-469.

[16] I. Gol'dsheid, G. Margulis, Lyapunov exponents of a product of random matrices, *Russian Math. Surveys* **44** (1989), 11-71.

[17] Y. Guivarc'h, A. Raugi, Frontière de Furstenberg, propriété de contraction et théorèmes de convergence, *Zeit. Wahrsch. verw. Gebiete* **69** (1985), 187-242.

[18] Y. Guivarc'h, A. Raugi, Actions of large semigroups and random walks on isometric extensions of boundaries, *Ann. Sci. Ecole Norm. Sup.* **40** (2007), 209-249.

[19] 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.

[20] A. Katok, R. Spatzier, Invariant measures for higher-rank hyperbolic abelian actions, *Ergodic Th. Dynam. Systems* **16** (1996), 751-778.

[21] A. Kechris, Countable sections for locally compact group actions, *Ergodic Th. Dynam. Systems* **12** (1992), 283-295.

[22] S. P. Lalley, Renewal theorems in symbolic dynamics, with application to geodesic flows, noneuclidean tessellations and their fractal limits, *Acta mathematica* **163** (1989), 1-55.

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

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

[25] F. Ledrappier, J.-M. Strelcyn, A proof of the estimation from below in Pesin's entropy formula, *Ergodic Th. Dynam. Systems* **2** (1982), 203-219.

[26] E. Le Page, Théorèmes limites pour les produits de matrices aléatoires, *LN in Math.* **928** (1982) p. 258-303.

[27] G. Margulis, Problems and conjectures in rigidity theory, *in Mathematics : frontiers and perspectives,* Amer. Math. Soc. (2000), 161-174.

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

[29] R. Muchnik, Semigroup actions on $\mathbb{T}^n$. *Geom. Dedicata* **110** (2005), 1-47.

[30] M. Ratner, On Raghunathan's measure conjecture, *Ann. of Math.* **134** (1991), 545-607.

[31] M. Ratner, Raghunathan's topological conjecture and distributions of unipotent flows, *Duke Math. J.* **63** (1991), 235-280.

[32] V.A. Rohlin, On the fundamental ideas of measure theory, *Math. Sbornik* **25**(1949), 107-150.

[33] V.A. Rohlin, Lectures on the entropy theory of measure-preserving transformations, *Russian Math. Surveys* **22** (1967), 1-54.

[34] 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.

[35] R. Zimmer, *Ergodic theory and semisimple groups*, Birkhäuser, Boston, 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