

**RIGIDITÉ DES  $SL_2(\mathbb{R})$ -ORBITES DANS LES ESPACES DE  
MODULES DE SURFACES PLATES**  
[d'après Eskin, Mirzakhani et Mohammadi]

par **Jean-François QUINT**

**INTRODUCTION**

Soient  $g$  un entier  $\geq 1$  et  $\alpha = (\alpha_1, \dots, \alpha_n)$  une partition de l'entier  $2g - 2$ . À cette donnée combinatoire, on associe une strate de différentielles abéliennes  $\mathcal{H}(\alpha)$ . Les éléments de  $\mathcal{H}(\alpha)$  sont les classes d'isomorphismes de paires  $(M, \omega)$  où  $M$  est une surface de Riemann de genre  $g$  et  $\omega$  est une 1-forme différentielle holomorphe sur  $M$  admettant exactement  $n$  zéros de multiplicités  $\alpha_1, \dots, \alpha_n$ . L'espace  $\mathcal{H}(\alpha)$  est naturellement muni d'une structure analytique complexe affine plate et d'une action du groupe  $GL_2^+(\mathbb{R})$  des matrices carrées réelles de taille 2 et de déterminant  $> 0$ . Dans une série de travaux récents, Eskin et Mirzakhani [11] et Eskin, Mirzakhani et Mohammadi [12] ont montré que les adhérences des  $GL_2^+(\mathbb{R})$ -orbites dans  $\mathcal{H}(\alpha)$  étaient des sous-variétés affines complexes, répondant ainsi affirmativement à une conjecture de McMullen.

Dans cet exposé, après avoir décrit plus précisément la structure de l'espace  $\mathcal{H}(\alpha)$ , je donnerai des indications sur la démonstration de ce théorème. Je me concentrerai essentiellement sur le résultat principal de [11], qui fait appel à des notions de théorie ergodique.

J'ai sollicité l'aide de nombreux collègues au cours de la préparation de cet exposé. Je tiens à remercier Christophe Bavard, Sylvain Crovisier, Vincent Koziarz et Carlos Matheus. Je remercie tout particulièrement Duc-Manh N'Guyen, qui a répondu avec patience à mes questions de débutant sur les surfaces de translation, Yves Benoist, avec qui nous avons beaucoup discuté de ces sujets, et Alex Eskin qui a bien voulu m'expliquer des points particulièrement délicats des démonstrations. Je remercie enfin Viviane Le Dret qui a considérablement amélioré la présentation de ce texte en y supprimant de nombreuses coquilles.

**1. STRATES DE DIFFÉRENTIELLES ABÉLIENNES**

Dans tout cet exposé, on fixe un entier  $g \geq 1$  et une surface topologique  $S$  compacte, orientable et de genre  $g$ . Nous allons définir les structures de translation à singularités coniques sur  $S$  et des espaces de modules de telles structures. Ces espaces seront munis d'une action du groupe  $GL_2^+(\mathbb{R})$ .

### 1.1. Structures plates et formes holomorphes

Une carte plate à singularité conique de  $S$  est un triplet  $(U, \varphi, \psi)$  où  $U$  est un ouvert de  $S$ ,  $\varphi$  est un homéomorphisme de  $U$  vers un ouvert de  $\mathbb{C}$  contenant 0 et  $\psi : U \rightarrow \mathbb{C}$  est l'application  $x \mapsto \varphi(x)^{\alpha+1}$ , où  $\alpha$  est un entier  $\geq 0$ , uniquement déterminé par  $\varphi$  et  $\psi$ .

Deux cartes plates à singularité conique  $(U_1, \varphi_1, \psi_1)$  et  $(U_2, \varphi_2, \psi_2)$  sont dites compatibles par translation s'il existe  $c$  dans  $\mathbb{C}$  tel que, pour tout  $x$  dans  $U_1 \cap U_2$ , on ait  $\psi_2(x) = \psi_1(x) + c$ . En particulier, dans ce cas, le changement de carte

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \rightarrow \varphi_2(U_1 \cap U_2)$$

est une application holomorphe.

Un atlas de translation à singularités coniques sur  $S$  est un ensemble  $\mathcal{A}$  de cartes plates à singularité conique, compatibles par translation entre elles, qui est maximal pour l'inclusion et tel que  $\bigcup_{(U, \varphi, \psi) \in \mathcal{A}} U = S$ . Un tel atlas étant donné, on dit que la paire  $(S, \mathcal{A})$  est une surface de translation à singularités coniques. Une singularité de  $(S, \mathcal{A})$  est un élément  $x$  de  $S$  tel qu'il existe une carte  $(U, \varphi, \psi)$  de  $\mathcal{A}$  avec  $x \in U$ ,  $\varphi(x) = 0$  et  $\alpha \geq 1$  où  $\alpha$  est l'entier tel que  $\psi = \varphi^{\alpha+1}$ . On dit que  $\alpha$  est l'ordre de la singularité  $x$ . L'ensemble des singularités de  $(S, \mathcal{A})$  est fini.

Comme les changements de carte de  $\mathcal{A}$  sont holomorphes, la structure de translation induit sur  $S$  une structure de surface de Riemann. De plus,  $S$  est munie de la 1-forme holomorphe  $\omega$  telle que, sur une carte  $(U, \varphi, \psi)$  de  $\mathcal{A}$ , on ait  $\omega = d\psi$ . D'après le théorème de Riemann-Roch, si  $\omega$  a  $n$  zéros, de multiplicité  $\alpha_1, \dots, \alpha_n$ , on a  $\alpha_1 + \dots + \alpha_n = 2g - 2$ . En d'autres termes, si  $x_1, \dots, x_n$  sont les singularités de  $(S, \mathcal{A})$  et si  $\alpha_1, \dots, \alpha_n$  sont les ordres de  $x_1, \dots, x_n$ , on a  $\alpha_1 + \dots + \alpha_n = 2g - 2$ .

Réciproquement, la donnée d'une structure de surface de Riemann sur  $S$  et d'une 1-forme holomorphe non nulle relativement à celle-ci détermine uniquement une structure de surface de translation à singularités coniques sur  $S$ .

*Remarque 1.1.* — Il est sans doute plus simple de définir une structure de surface de translation comme la donnée d'une structure de surface de Riemann et d'une 1-forme holomorphe non nulle. Le lourd formalisme des atlas de translation que nous venons d'introduire prendra son sens plus loin lorsque nous définirons une action de  $\mathrm{GL}_2(\mathbb{R})$  sur l'ensemble de toutes les surfaces plates.

### 1.2. L'espace des structures de translation

Soit  $\alpha = (\alpha_1, \dots, \alpha_n)$  une suite d'entiers  $\geq 1$  telle que  $\alpha_1 + \dots + \alpha_n = 2g - 2$ . Nous notons  $\mathcal{P}(\alpha)$  l'ensemble des atlas de translation à singularités coniques d'ordre  $\alpha_1, \dots, \alpha_n$  sur  $S$ . Notons que la donnée d'un tel atlas détermine une orientation de  $S$ . Nous fixons une fois pour toutes une orientation de  $S$  et nous notons  $\mathcal{P}^+(\alpha) \subset \mathcal{P}(\alpha)$  l'ensemble des structures de translation compatibles avec cette orientation fixée.

Notons  $\mathcal{G}$  le groupe des homéomorphismes de  $S$  qui préservent l'orientation et  $\mathcal{G}^\circ \subset \mathcal{G}$  le groupe des homéomorphismes isotopes à l'identité, de sorte que  $\Gamma = \mathcal{G}/\mathcal{G}^\circ$  est le

groupe modulaire de  $S$ . Le groupe  $\mathcal{G}$  agit naturellement sur  $\mathcal{P}(\alpha)$  (en préservant  $\mathcal{P}^+(\alpha)$ ) : pour  $g$  dans  $\mathcal{G}$  et  $\mathcal{A}$  dans  $\mathcal{P}(\alpha)$ , on note  $g\mathcal{A}$  l'atlas dont les cartes sont les  $(gU, \varphi \circ g^{-1}, \psi \circ g^{-1})$  où  $(U, \varphi, \psi)$  est une carte de  $\mathcal{A}$ .

Nous définissons l'espace  $\mathcal{H}(\alpha)$  comme le quotient de  $\mathcal{P}^+(\alpha)$  par  $\mathcal{G}$ . On note aussi  $\tilde{\mathcal{H}}(\alpha)$  le quotient de  $\mathcal{P}^+(\alpha)$  par  $\mathcal{G}^\circ$ , de sorte que  $\mathcal{H}(\alpha) = \Gamma \backslash \tilde{\mathcal{H}}(\alpha)$ . On dit que  $\mathcal{H}(\alpha)$  est une strate de différentielles abéliennes.

### 1.3. Le cas de genre 1

Supposons que  $g = 1$ , de sorte que  $S$  est un tore et que  $\alpha$  est nécessairement la suite vide. Choisissons un revêtement universel  $\tilde{S} \rightarrow S$  de  $S$  de groupe de revêtement  $\Lambda \simeq \mathbb{Z}^2$ . La donnée d'une structure plate sur  $S$  revient à la donnée d'une application développante qui est ici un difféomorphisme  $D : \tilde{S} \rightarrow \mathbb{C}$  et d'un morphisme d'holonomie (qui est ici d'image discrète et cocompacte)  $h : \Lambda \rightarrow \mathbb{C}$  tels que, pour tous  $g$  dans  $\Lambda$  et  $x$  dans  $\tilde{S}$ , on ait  $D(gx) = h(g) + D(x)$  (plus exactement, l'application  $D$  est déterminée à translation près par un élément constant de  $\mathbb{C}$  par la structure plate).

On dispose d'un isomorphisme naturel entre les groupes de cohomologie  $H^2(S, \mathbb{Z})$  et  $H^2(\Lambda, \mathbb{Z})$  et l'orientation de  $S$  détermine un générateur  $c$  de ce groupe cyclique. Fixons deux générateurs  $g_1$  et  $g_2$  de  $\Lambda$ , de sorte que  $c(g_1, g_2) > 0$ .

La donnée de l'homomorphisme d'holonomie caractérise complètement la structure plate à isotopie près. Plus précisément, l'espace  $\tilde{\mathcal{H}}(\emptyset)$  s'identifie à l'espace des homomorphismes  $h : \Lambda \rightarrow \mathbb{C}$  d'image discrète et cocompacte et tels que la paire  $(h(g_1), h(g_2))$  soit orientée positivement dans  $\mathbb{C}$ . En d'autres termes, à travers ce choix de générateurs, l'espace  $\tilde{\mathcal{H}}(\emptyset)$  s'identifie à l'ensemble des couples  $(z_1, z_2)$  dans  $\mathbb{C} \times \mathbb{C}$  avec  $\text{Im}(\bar{z}_1 z_2) > 0$ .

Le groupe  $\Gamma$  s'identifie à  $\text{SL}_2(\mathbb{Z})$ . On vérifie qu'il agit sur  $\tilde{\mathcal{H}}(\emptyset)$  de sorte que, si  $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  est dans  $\text{SL}_2(\mathbb{Z})$  et  $(z_1, z_2)$  est dans  $\tilde{\mathcal{H}}(\emptyset)$ , on a

$$\gamma(z_1, z_2) = (dz_1 - cz_2, -bz_1 + az_2).$$

Identifions  $\mathbb{C}$  à  $\mathbb{R}^2$  en munissant  $\mathbb{C}$  de la base  $(1, i)$  de  $\mathbb{C}$ . Alors on peut voir  $\tilde{\mathcal{H}}(\emptyset)$  comme l'espace des bases orientées de  $\mathbb{R}^2$ . Soit  $\text{GL}_2^+(\mathbb{R})$  le groupe des matrices carrées de taille 2 et de déterminant positif. Le groupe  $\text{GL}_2^+(\mathbb{R})$  agit sur  $\tilde{\mathcal{H}}(\emptyset)$  de sorte que

$$g(z_1, z_2) = (gz_1, gz_2),$$

pour  $g$  dans  $\text{GL}_2^+(\mathbb{R})$  et  $(z_1, z_2)$  dans  $\tilde{\mathcal{H}}(\emptyset)$ . Cette action est simplement transitive, si bien que le choix du point  $(1, i)$  dans  $\tilde{\mathcal{H}}(\emptyset)$  permet de voir cet ensemble comme une copie de  $\text{GL}_2^+(\mathbb{R})$ . L'action de  $\text{SL}_2(\mathbb{Z})$  sur cet espace se lit maintenant comme son action à droite sur  $\text{GL}_2^+(\mathbb{R})$ , si bien qu'on peut assimiler  $\mathcal{H}(\emptyset)$  au quotient  $\text{GL}_2^+(\mathbb{R})/\text{SL}_2(\mathbb{Z})$ .

### 1.4. Action du groupe $\text{GL}_2^+(\mathbb{R})$

Supposons à nouveau que  $g$  est un entier  $\geq 1$  quelconque. Nous allons maintenant décrire une action du groupe  $\text{GL}_2^+(\mathbb{R})$  sur  $\mathcal{H}(\alpha)$  qui généralise celle qui a été introduite ci-dessus quand  $g = 1$ .

Pour cela, commençons par essayer de construire une action de  $\mathrm{GL}_2^+(\mathbb{R})$  sur l'ensemble  $\mathcal{C}$  des cartes plates à singularité conique de  $S$ . Pour  $(U, \varphi, \psi)$  dans  $\mathcal{C}$  et  $g$  dans  $\mathrm{GL}_2^+(\mathbb{R})$ , on souhaite que  $g(U, \varphi, \psi)$  soit de la forme  $(U, \varphi', g\psi)$  (où  $g\psi$  est la fonction  $x \mapsto g\psi(x)$  sur  $U$ ). Si  $\psi = \varphi$ , nous pouvons poser  $\varphi' = \varphi$ . En revanche, si  $\psi = \varphi^{\alpha+1}$  avec  $\alpha \geq 1$  (c'est-à-dire si la carte contient une singularité), la fonction  $\varphi'$  doit satisfaire l'équation  $(\varphi')^{\alpha+1} = g\psi$  et cette équation possède au moins  $\alpha + 1$  solutions. Pour lever cette ambiguïté, nous allons remplacer le groupe  $\mathrm{GL}_2^+(\mathbb{R})$  par son revêtement universel.

Rappelons que  $\mathrm{GL}_2^+(\mathbb{R})$  est connexe et se rétracte continûment sur  $\mathrm{SO}(2)$ . En particulier, son groupe fondamental est isomorphe à  $\mathbb{Z}$ . Fixons un revêtement universel  $\pi : \widetilde{\mathrm{GL}}_2^+(\mathbb{R}) \rightarrow \mathrm{GL}_2^+(\mathbb{R})$  de noyau  $Z$ . L'homomorphisme continu de groupes

$$\theta : \mathbb{R} \rightarrow \mathrm{SO}(2), t \mapsto \begin{pmatrix} \cos(2\pi t) & -\sin(2\pi t) \\ \sin(2\pi t) & \cos(2\pi t) \end{pmatrix}$$

se relève en un homomorphisme continu  $\tilde{\theta} : \mathbb{R} \rightarrow \widetilde{\mathrm{GL}}_2^+(\mathbb{R})$  et on a  $\tilde{\theta}(\mathbb{Z}) = Z$ . On pose  $c = \tilde{\theta}(1)$  : c'est le générateur de  $Z$  associé à l'orientation naturelle de  $\mathbb{C}$ .

Des raisonnements élémentaires de théorie des revêtements permettent d'établir le

LEMME 1.2. — *Soit  $\alpha$  un entier naturel. Il existe une unique action continue par homéomorphismes de  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$  sur  $\mathbb{C}$ , notée  $(g, z) \mapsto g \cdot_\alpha z$ , telle que, pour tous  $g$  dans  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$  et  $z$  dans  $\mathbb{C}$ , on ait*

$$(g \cdot_\alpha z)^{\alpha+1} = \pi(g)(z^{\alpha+1})$$

(où, dans le second membre,  $\mathrm{GL}_2^+(\mathbb{R})$  agit sur  $\mathbb{C} \simeq \mathbb{R}^2$  par l'action linéaire classique). Pour  $g$  dans  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$ , l'application  $z \mapsto g \cdot_\alpha z$  est homogène de degré 1. Pour tout  $z$  dans  $\mathbb{C}$ , on a

$$(1) \quad c \cdot_\alpha z = \exp\left(\frac{2i\pi}{\alpha+1}\right) z.$$

Remarque 1.3. — Soit  $\alpha \geq 1$ . Alors, pour  $g$  dans  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$ , l'application  $z \mapsto g \cdot_\alpha z$  est  $\mathcal{C}^\infty$  sur  $\mathbb{C} \setminus \{0\}$ , mais elle n'est pas différentiable en 0 dès que  $\pi(g) \notin \mathbb{R}_+^* \mathrm{SO}(2)$ . En effet, comme cette application est homogène de degré 1, si elle était différentiable en 0, elle serait linéaire.

La définition de l'action naturelle de  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$  sur l'ensemble  $\mathcal{C}$  des cartes plates à singularité conique devient maintenant claire : pour  $g$  dans  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$  et  $(U, \varphi, \psi)$  dans  $\mathcal{C}$ , on pose  $g(U, \varphi, \psi) = (U, g \cdot_\alpha \varphi, g\psi)$  où  $\alpha$  est tel que  $\psi = \varphi^{\alpha+1}$ . On vérifie aisément que cette action préserve la relation de compatibilité des cartes. En particulier, si nous fixons une suite  $\alpha = (\alpha_1, \dots, \alpha_n)$  d'entiers  $\geq 1$  avec  $\alpha_1 + \dots + \alpha_n = 2g - 2$ , l'action de  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$  sur  $\mathcal{C}$  induit une action sur l'espace  $\mathcal{P}^+(\alpha)$  des atlas compatibles avec l'orientation et de singularités d'ordre  $(\alpha_1, \dots, \alpha_n)$ . Pour conclure, il nous reste à démontrer le

LEMME 1.4. — Le groupe  $Z$  agit trivialement sur  $\mathcal{P}^+(\alpha)$ , si bien que l'action de  $\widetilde{\mathrm{GL}}_2^+(\mathbb{R})$  sur cet ensemble factorise en une action de  $\mathrm{GL}_2^+(\mathbb{R})$ .

*Démonstration.* — Soit  $\mathcal{A}$  dans  $\mathcal{P}^+(\alpha)$  et montrons que  $c\mathcal{A} = \mathcal{A}$ . Pour  $(U, \varphi, \psi)$  dans  $\mathcal{A}$ , soit  $\beta$  l'entier tel que  $\psi = \varphi^{\beta+1}$ . D'après (1), on a  $c(U, \varphi, \psi) = (U, \exp\left(\frac{2i\pi}{\beta+1}\right)\varphi, \psi)$  et donc  $c(U, \varphi, \psi)$  est compatible avec toutes les cartes de  $\mathcal{A}$ . Par maximalité de l'atlas  $\mathcal{A}$ , on a bien  $c\mathcal{A} = \mathcal{A}$ .  $\square$

*Remarque 1.5.* — Soit  $\mathcal{A}$  dans  $\mathcal{P}^+(\alpha)$ ; alors, d'après la remarque 1.3, pour  $g$  dans  $\mathrm{GL}_2^+(\mathbb{R}) \setminus \mathbb{R}_+^* \mathrm{SO}(2)$ , les structures plates  $\mathcal{A}$  et  $g\mathcal{A}$  n'induisent pas la même structure différentielle sur la surface  $S$ . C'est cette difficulté qui nous a conduit à introduire la notion de surface de translation dans un cadre topologique.

L'action de  $\mathrm{GL}_2^+(\mathbb{R})$  que nous venons de définir est naturelle. Plus précisément, elle commute à l'action du groupe  $\mathcal{G}$  des homéomorphismes directs de  $S$ . Elle induit donc une action de  $\mathrm{GL}_2^+(\mathbb{R})$  sur l'espace  $\tilde{\mathcal{H}}(\alpha)$  qui commute à l'action du groupe modulaire  $\Gamma$  et une action de  $\mathrm{GL}_2^+(\mathbb{R})$  sur  $\mathcal{H}(\alpha) = \Gamma \backslash \tilde{\mathcal{H}}(\alpha)$ .

### 1.5. Application développante et topologie des strates

Terminons cette section en introduisant une topologie naturelle sur l'espace  $\mathcal{H}(\alpha)$ .

Fixons une fois pour toutes un revêtement universel  $\pi : \tilde{S} \rightarrow S$ . Si  $\mathcal{A}$  est un élément de  $\mathcal{P}^+(\alpha)$ , il existe une application continue  $D_{\mathcal{A}} : \tilde{S} \rightarrow \mathbb{C}$  telle que, pour toute carte  $(U, \varphi, \psi)$  dans  $\mathcal{A}$ , si  $V$  est un ouvert connexe de  $\tilde{S}$  tel que  $\pi$  induise un homéomorphisme de  $V$  sur  $U$ , la fonction  $D_{\mathcal{A}} - \psi \circ \pi$  est constante sur  $V$ . La fonction  $D_{\mathcal{A}}$  est unique à addition d'une constante près. Par abus de langage, on dit que  $D_{\mathcal{A}}$  est l'application développante de  $\mathcal{A}$ . Pour  $g$  dans  $\mathrm{GL}_2^+(\mathbb{R})$ , la développante de  $g\mathcal{A}$  est  $gD_{\mathcal{A}}$ .

Nous munissons l'espace  $\mathcal{P}^+(\alpha)$  de la topologie qui y est induite par la topologie de la convergence uniforme sur les compacts pour les applications développantes. Plus précisément, c'est la topologie pour laquelle une base de voisinages d'un élément  $\mathcal{A}_0$  de  $\mathcal{P}^+(\alpha)$  est constituée de l'ensemble des  $\mathcal{A}$  dans  $\mathcal{P}^+(\alpha)$  pour lesquels il existe  $c$  dans  $\mathbb{C}$  avec

$$\sup_{x \in K} |D_{\mathcal{A}}(x) - D_{\mathcal{A}_0}(x) - c| \leq \varepsilon,$$

où  $\varepsilon > 0$  et le compact  $K$  de  $\tilde{S}$  sont fixés. Les groupes  $\mathcal{G}$  et  $\mathrm{GL}_2^+(\mathbb{R})$  agissent alors par transformations continues sur  $\mathcal{P}^+(\alpha)$ . On munit  $\mathcal{H}(\alpha)$  et  $\tilde{\mathcal{H}}(\alpha)$  des topologies quotient.

On peut montrer que l'action de  $\Gamma$  sur  $\tilde{\mathcal{H}}(\alpha)$  est propre : en effet, l'application qui, à une structure de translation, associe la structure complexe sous-jacente induit une application continue  $\Gamma$ -équivariante de  $\tilde{\mathcal{H}}(\alpha)$  vers l'espace de Teichmüller de la surface  $S$  et l'action de  $\Gamma$  sur l'espace de Teichmüller est propre [31, Th. 4.10.5].

## 2. LA STRUCTURE DE VARIÉTÉ AFFINE DES STRATES

Dans toute la suite de cet exposé, on fixe une suite  $\alpha = (\alpha_1, \dots, \alpha_n)$  d'entiers  $\geq 1$  avec  $\alpha_1 + \dots + \alpha_n = 2g - 2$  et on étudie la strate  $\mathcal{H}(\alpha)$ . Nous allons munir cet ensemble d'une structure analytique complexe affine et plate, invariante par l'action du groupe  $\mathrm{GL}_2^+(\mathbb{R})$ .

Commençons par préciser ce que nous entendons ici par structure affine plate.

### 2.1. $(H, V)$ -structures plates

Soient  $M$  un espace topologique,  $V$  un espace affine (au sens de la géométrie affine élémentaire) et  $H$  un sous-groupe fermé du groupe des automorphismes affines de  $V$ .

Une carte de  $M$  modélée sur  $V$  est un couple  $(U, \varphi)$  où  $U$  est un ouvert de  $M$  et  $\varphi$  est un homéomorphisme de  $U$  sur un ouvert de  $V$ . Deux cartes  $(U_1, \varphi_1)$  et  $(U_2, \varphi_2)$  sont dites  $H$ -compatibles s'il existe  $h$  dans  $H$  tel que, pour tout  $x$  dans  $U_1 \cap U_2$ , on ait  $\varphi_2(x) = h\varphi_1(x)$ . Un  $(H, V)$ -atlas sur  $M$  est un ensemble  $\mathcal{A}$  de cartes modélées sur  $V$  et deux à deux compatibles, qui est maximal pour l'inclusion et tel que

$$\bigcup_{(U, \varphi) \in \mathcal{A}} U = M.$$

Une  $(H, V)$ -structure plate sur  $M$  est la donnée d'un tel atlas. Elle induit sur  $M$  une structure de variété analytique réelle.

Dans les structures plates que nous allons construire sur les strates, l'espace  $V$  aura une forme particulière que nous allons à présent décrire.

Donnons-nous un espace vectoriel réel de dimension finie  $W$ , muni d'une forme bilinéaire alternée  $\varpi$  (cette forme sera définie grâce à un cup-produit sur un espace de cohomologie). On pose  $V = W \oplus W$ . On considère l'ouvert  $V_\varpi \subset V$  constitué de l'ensemble des éléments  $(w_1, w_2)$  de  $V$  tels que  $\varpi(w_1, w_2) > 0$ . On note  $H$  le groupe des automorphismes linéaires de  $V$  qui sont de la forme  $(w_1, w_2) \mapsto (gw_1, gw_2)$  où  $g$  est un automorphisme linéaire de  $W$  qui préserve  $\varpi$ , de sorte que  $V_\varpi$  est stable par  $H$ . Nous dirons qu'une  $(H, V)$ -structure plate est  $\varpi$ -adaptée si toutes ses cartes sont à valeurs dans  $V_\varpi$ .

La donnée d'une telle structure sur  $M$  définit un germe d'action analytique de  $\mathrm{GL}_2^+(\mathbb{R})$  sur  $M$ , à stabilisateurs discrets. En effet, le groupe  $\mathrm{GL}_2(\mathbb{R})$  agit naturellement sur  $V$  qu'on peut identifier à  $\mathbb{R}^2 \otimes W$ . Plus précisément, pour tous  $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  dans  $\mathrm{GL}_2(\mathbb{R})$  et  $(w_1, w_2)$  dans  $V$ , on a

$$g(w_1, w_2) = (aw_1 + bw_2, cw_1 + dw_2).$$

Cette action commute à l'action de  $H$ . Elle préserve l'ouvert  $V_\varpi$  et le stabilisateur d'un élément de  $V_\varpi$  est trivial. Pour  $v = (w_1, w_2)$  dans  $V_\varpi$ , l'espace tangent en  $v$  à sa  $\mathrm{GL}_2^+(\mathbb{R})$ -orbite est le sous-espace vectoriel  $\mathcal{F}_v$  de  $v$  engendré par les vecteurs  $(w_1, 0)$ ,  $(w_2, 0)$ ,  $(0, w_1)$  et  $(0, w_2)$ .

Étant donnée une  $(H, V)$ -structure  $\varpi$ -adaptée sur  $M$ , nous pouvons alors lui associer le feuilletage de  $M$  en  $\mathrm{GL}_2^+(\mathbb{R})$ -orbites locales, c'est-à-dire le feuilletage dont la distribution tangente  $\mathcal{F}$  est telle que, pour toute carte  $(U, \varphi)$  de la structure, pour tout  $x$  dans  $U$ , on a  $\mathcal{F}_x = d\varphi(x)^{-1}\mathcal{F}_{\varphi(x)}$ . À chaque élément  $X$  de l'algèbre de Lie de  $\mathrm{GL}_2^+(\mathbb{R})$ , nous pouvons associer le champ de vecteurs  $\mathcal{V}^X$  sur  $M$  tel que, pour toute carte  $(U, \varphi)$ , pour tout  $x$  dans  $U$ , on ait  $\mathcal{V}_x^X = d\varphi(x)^{-1}(X\varphi(x))$  (c'est-à-dire que  $\mathcal{V}^X$  est le champ tangent à l'intersection de la carte avec la courbe  $t \mapsto \exp(tX)v$ ).

*Remarque 2.1.* — Si la variété  $M$  n'est pas compacte il n'y a pas de raison que les champs  $\mathcal{V}^X$  définissent une action globale du groupe  $\mathrm{GL}_2^+(\mathbb{R})$  ou d'un de ses revêtements. Ce sera néanmoins le cas dans les structures plates que nous allons construire sur les strates, puisque cette action globale de  $\mathrm{GL}_2^+(\mathbb{R})$  aura été construite a priori.

Notons que,  $V$  étant muni de la structure complexe  $J$  telle que  $J(w_1, w_2) = (-w_2, w_1)$ , une  $(H, V)$ -structure plate induit aussi une structure de variété analytique complexe.

Supposons que  $\Gamma$  est un groupe discret agissant proprement par difféomorphismes sur  $M$  en préservant une  $(H, V)$ -structure plate. Alors, si l'action de  $\Gamma$  sur  $M$  est sans point fixe, la variété  $\Gamma \backslash M$  est naturellement munie d'une  $(H, V)$ -structure plate. Dans le cas général, on appellera encore  $(H, V)$ -structure plate sur l'espace  $\Gamma \backslash M$  la donnée d'une  $(H, V)$ -structure plate  $\Gamma$ -invariante sur  $M$ .

## 2.2. L'espace modèle des strates

Rappelons que nous avons fixé un multi-indice  $\alpha = (\alpha_1, \dots, \alpha_n)$  avec  $\alpha_1 + \dots + \alpha_n = 2g - 2$ . Nous allons maintenant chercher à construire une structure plate du type décrit dans le paragraphe précédent sur la strate  $\mathcal{H}(\alpha)$ . Pour cela, nous commençons par construire ici l'espace vectoriel qui nous servira de modèle.

Choisissons une fois pour toutes un  $n$ -uplet  $\Sigma = (x_1, \dots, x_n)$  de points distincts de  $S$ . Nous disposons d'une suite exacte longue de cohomologie :

$$0 \rightarrow H^0(S, \mathbb{R}) \rightarrow H^0(\Sigma, \mathbb{R}) \rightarrow H^1(S, \Sigma, \mathbb{R}) \xrightarrow{p} H^1(S, \mathbb{R}) \rightarrow 0$$

(où par abus de notation, on identifie  $\Sigma$  et l'ensemble  $\{x_1, \dots, x_n\}$ ). En particulier, l'espace de cohomologie relative  $H^1(S, \Sigma, \mathbb{R})$  est de dimension  $2g + n - 1$ . Nous allons lui faire jouer le rôle de l'espace  $W$  du paragraphe précédent. Pour cela, nous allons le munir d'une forme bilinéaire alternée.

Rappelons que l'espace  $H^1(S, \mathbb{R})$  est naturellement muni du cup-produit, qui y définit une application bilinéaire alternée non dégénérée

$$H^1(S, \mathbb{R}) \times H^1(S, \mathbb{R}) \rightarrow H^2(M, \mathbb{R}).$$

En fixant une orientation de la surface  $S$ , nous avons fixé un générateur du groupe cyclique  $H^2(M, \mathbb{Z}) \subset H^2(M, \mathbb{R})$ . Ceci détermine donc une identification de  $H^2(M, \mathbb{R})$  avec  $\mathbb{R}$  à travers laquelle nous pouvons considérer le cup-produit comme une forme bilinéaire alternée non dégénérée

$$\varpi : H^1(S, \mathbb{R}) \times H^1(S, \mathbb{R}) \rightarrow \mathbb{R}.$$

Par abus de notation, pour  $x, y$  dans  $H^1(S, \Sigma, \mathbb{R})$ , nous noterons  $\varpi(x, y)$  pour  $\varpi(p(x), p(y))$ .

Posons  $W = H^1(S, \Sigma, \mathbb{R})$ ,  $V = W \oplus W = H^1(S, \Sigma, \mathbb{R}^2)$  et notons toujours  $H$  le groupe des automorphismes linéaires de  $V$  qui sont de la forme  $(w_1, w_2) \mapsto (gw_1, gw_2)$  où  $g \in \text{GL}(W)$  préserve la forme bilinéaire  $\varpi$ . Nous cherchons à munir  $\tilde{\mathcal{H}}(\alpha)$  et  $\mathcal{H}(\alpha)$  d'une  $(H, V)$ -structure  $\varpi$ -adaptée, au sens du paragraphe 2.1. Nous allons construire cette structure sur un revêtement fini de  $\tilde{\mathcal{H}}(\alpha)$  dans lequel nous n'allons considérer que des structures dont les singularités se situent en  $(x_1, \dots, x_n)$ .

Notons  $\mathcal{P}_\Sigma^+(\alpha) \subset \mathcal{P}^+(\alpha)$  l'ensemble des structures de translation à singularités coniques sur  $S$ , compatibles avec l'orientation, et dont les singularités sont  $x_1, \dots, x_n$  et ont pour ordre  $\alpha_1, \dots, \alpha_n$ . Il est stable par l'action de  $\text{GL}_2^+(\mathbb{R})$ . Soit  $\mathcal{G}_{\alpha, \Sigma}$  l'ensemble des homéomorphismes directs de  $g$  de  $S$  tels que  $g\Sigma = \Sigma$  et que, pour tous  $1 \leq i, j \leq n$ , si  $gx_i = x_j$ , on a  $\alpha_i = \alpha_j$ . Soient encore  $\mathcal{G}_{\alpha, \Sigma}^\circ \subset \mathcal{G}_{\alpha, \Sigma}$  le groupe des homéomorphismes de  $S$  qui sont isotopes à l'identité dans  $\mathcal{G}_{\alpha, \Sigma}$  et  $\Gamma_{\alpha, \Sigma} = \mathcal{G}_{\alpha, \Sigma} / \mathcal{G}_{\alpha, \Sigma}^\circ$ . Notons que le groupe  $\Gamma_{\alpha, \Sigma}$  agit naturellement sur  $W = H^1(S, \Sigma, \mathbb{R})$  en préservant la forme  $\varpi$ . Il agit donc naturellement comme un sous-groupe de  $H$  sur  $V = H^1(S, \Sigma, \mathbb{R}^2)$ .

On pose  $\mathcal{H}_\Sigma(\alpha) = \mathcal{G}_{\alpha, \Sigma}^\circ \backslash \mathcal{P}_\Sigma^+(\alpha)$ . Comme  $\mathcal{G} = \mathcal{G}^\circ \mathcal{G}_{\alpha, \Sigma}$  (tout homéomorphisme est isotope à un homéomorphisme qui préserve  $\Sigma$ ), l'ensemble  $\Gamma_{\alpha, \Sigma} \backslash \mathcal{H}_\Sigma(\alpha)$  s'identifie à  $\mathcal{H}(\alpha)$ . Comme  $\mathcal{G}_{\alpha, \Sigma}^\circ$  est d'indice fini dans  $\mathcal{G}^\circ \cap \mathcal{G}_{\alpha, \Sigma}$ , l'ensemble  $\tilde{\mathcal{H}}(\alpha)$  s'identifie au quotient de  $\mathcal{H}_\Sigma(\alpha)$  par un groupe fini.

### 2.3. Application de période

Maintenant que nous avons construit l'espace  $\mathcal{H}_\Sigma(\alpha)$ , nous allons montrer qu'il porte une structure plate naturelle modelée sur  $V = H^1(S, \Sigma, \mathbb{R}^2)$ . Pour cela, nous allons construire une application dite application de période  $\Pi : \mathcal{H}_\Sigma(\alpha) \rightarrow H^1(S, \Sigma, \mathbb{R}^2)$ .

Rappelons que nous avons fixé un revêtement universel  $\pi : \tilde{S} \rightarrow S$ . À un élément  $\mathcal{A}$ , nous avons associé une application développante  $D_{\mathcal{A}} : \tilde{S} \rightarrow \mathbb{C}$  définie à addition d'une constante près. Si  $\gamma : [0, 1] \rightarrow S$  est un chemin continu, choisissons un relevé continu  $\tilde{\gamma} : [0, 1] \rightarrow \tilde{S}$  de  $\gamma$ . Alors, le nombre

$$c_{\mathcal{A}}(\gamma) = D_{\mathcal{A}}(\tilde{\gamma}(1)) - D_{\mathcal{A}}(\tilde{\gamma}(0))$$

ne dépend ni du choix de la développante  $D_{\mathcal{A}}$ , ni du choix du relevé  $\tilde{\gamma}$ . Comme, en outre,  $c_{\mathcal{A}}(\gamma)$  s'annule lorsque le chemin  $\gamma$  est constant, l'application  $c_{\mathcal{A}}$  définit par linéarité un élément, encore noté  $c_{\mathcal{A}}$ , de l'espace  $C^1(S, \Sigma, \mathbb{R}^2)$  des cochaînes de  $S$  relativement à  $\Sigma$  à coefficients dans  $\mathbb{C} \simeq \mathbb{R}^2$ , encore notée  $c_{\mathcal{A}}$ . On vérifie aisément que cette cochaîne est un cocycle.

Par construction, l'application  $\mathcal{A} \mapsto c_{\mathcal{A}}$  est équivariante sous l'action de  $\text{GL}_2^+(\mathbb{R})$  (qui agit sur les coefficients des cochaînes par son action naturelle sur  $\mathbb{R}^2$ ) et sous l'action du groupe  $\mathcal{G}_{\alpha, \Sigma}$  (qui agit naturellement dans  $C^1(S, \Sigma, \mathbb{R}^2)$ ). En particulier, comme  $\mathcal{G}_{\alpha, \Sigma}^\circ$  agit trivialement dans  $H^1(S, \Sigma, \mathbb{R}^2)$ , la restriction de cette application à  $\mathcal{P}_\Sigma^+(\alpha)$  définit une application

$$\Pi : \mathcal{H}_\Sigma(\alpha) \rightarrow H^1(S, \Sigma, \mathbb{R}^2)$$

qu'on appelle application de période de la strate  $\mathcal{H}_\Sigma(\alpha)$ . Elle est  $\Gamma_{\alpha,\Sigma}$ -équivariante et  $\mathrm{GL}_2^+(\mathbb{R})$ -équivariante (quand  $\mathrm{GL}_2^+(\mathbb{R})$  agit sur  $\mathrm{H}^1(S, \Sigma, \mathbb{R}^2) = \mathbb{R}^2 \otimes \mathrm{H}^1(S, \Sigma, \mathbb{R})$  à travers l'action naturelle sur  $\mathbb{R}^2$  et l'action triviale sur  $\mathrm{H}^1(S, \Sigma, \mathbb{R})$ ).

L'application de période va nous permettre de définir une structure affine sur la strate grâce au remarquable

**THÉORÈME 2.2** (Veech [33]). — *L'application  $\Pi : \mathcal{H}_\Sigma(\alpha) \rightarrow \mathrm{H}^1(S, \Sigma, \mathbb{R}^2)$  est un homéomorphisme local.*

Ainsi l'application période définit-elle sur l'espace topologique  $\mathcal{H}_\Sigma(\alpha)$  une  $(H, V)$ -structure affine plate qui est l'image inverse par  $\Pi$  de la structure tautologique de  $\mathrm{H}^1(S, \Sigma, \mathbb{R}^2)$ . Comme  $\Pi$  est  $\Gamma_{\alpha,\Sigma}$ -équivariante et que  $\Gamma_{\alpha,\Sigma}$  agit sur  $V$  à travers des éléments du groupe  $H$ , cette structure descend à  $\mathcal{H}(\alpha)$  en une  $(H, V)$ -structure affine plate.

Il nous reste à vérifier que cette structure est  $\varpi$ -adaptée, c'est-à-dire que l'application  $p \circ \Pi : \mathcal{H}_\Sigma(\alpha) \rightarrow \mathrm{H}^1(S, \mathbb{R}^2)$  prend ses valeurs dans l'ensemble des paires  $(w_1, w_2)$  dans  $\mathrm{H}^1(S, \mathbb{R}) \oplus \mathrm{H}^1(S, \mathbb{R})$  avec  $\varpi(w_1, w_2) > 0$ . Fixons donc un atlas  $\mathcal{A}$  dans  $\mathcal{P}_\Sigma^+(\alpha)$ . Munissons  $S$  de la structure différentielle réelle sous-jacente à la structure plate à singularités coniques  $\mathcal{A}$ . L'application  $D_{\mathcal{A}} : \tilde{S} \rightarrow \mathbb{C} \simeq \mathbb{R}^2$  est lisse. Notons, comme il est d'usage,  $dx$  et  $dy$  les 1-formes différentielles associées aux formes linéaires coordonnées de  $\mathbb{R}^2$ . Les 1-formes  $D_{\mathcal{A}}^* dx$  et  $D_{\mathcal{A}}^* dy$  peuvent être considérées comme des 1-formes sur  $S$ . Par construction,  $p(\Pi(\mathcal{A}))$  est l'image  $(w_1, w_2)$  dans  $\mathrm{H}^1(S, \mathbb{R}) \oplus \mathrm{H}^1(S, \mathbb{R})$  de la paire de 1-formes  $(D_{\mathcal{A}}^* dx, D_{\mathcal{A}}^* dy)$  et donc on a

$$(2) \quad \varpi(w_1, w_2) = \int_S D_{\mathcal{A}}^*(dx \wedge dy).$$

Comme  $D_{\mathcal{A}}$  est un difféomorphisme local de  $\tilde{S} \setminus \pi^{-1}\Sigma$  vers  $\mathbb{R}^2$  et que l'orientation de  $\mathcal{A}$  est compatible avec celle de  $S$ , cette intégrale est  $> 0$ , ce qu'il fallait démontrer.

L'équivariance de l'application de période sous le groupe  $\mathrm{GL}_2^+(\mathbb{R})$  implique que celui-ci préserve la structure affine plate que nous venons de construire sur  $\mathcal{H}(\alpha)$ . Plus précisément, les orbites de cette action sont les feuilles maximales du feuilletage naturel d'une variété munie d'une  $(H, V)$ -structure  $\varpi$ -adaptée, tel qu'il a été introduit au paragraphe 2.1.

## 2.4. Fibré de Hodge

Nous terminons cette section en définissant un fibré  $\mathrm{GL}_2^+(\mathbb{R})$ -équivariant au-dessus de  $\mathcal{H}(\alpha)$  qui jouera un rôle fondamental dans l'étude dynamique de l'action de  $\mathrm{GL}_2^+(\mathbb{R})$ .

Commençons par revenir temporairement au formalisme du paragraphe 2.1. Étant donnée une variété  $M$ , munie d'une  $(H, V)$ -structure plate  $\varpi$ -adaptée, nous pouvons lui associer un fibré vectoriel plat, de fibre  $W$  au-dessus de  $M$ . Plus précisément, après avoir fixé un revêtement universel  $\tilde{M} \rightarrow M$ , de groupe de Galois  $\Lambda$ , la donnée d'une  $(H, V)$  structure plate permet de définir une application développante  $D : \tilde{M} \rightarrow V$  et un morphisme d'holonomie  $h : \Lambda \rightarrow H$  avec  $D(\lambda x) = h(\lambda)D(x)$ , pour  $x$  dans  $M$  et

$\lambda$  dans  $\Lambda$ . Alors, le fibré tangent à  $M$  s'identifie au quotient  $M \times_{\Lambda} V$  de  $\tilde{M} \times V$  par l'action de  $\Lambda$  donnée par  $\lambda(x, v) = (\lambda x, h(\lambda)v)$ .

Supposons que la  $(H, V)$  structure est  $\varpi$ -adaptée, où  $\varpi$  est comme au paragraphe 2.1 ; dans ce cas, la développante  $D$  prend ses valeurs dans l'ouvert  $V_{\varpi}$ . Supposons en outre, pour éviter les lourdes notations superflues dans le cas des strates, que le germe d'action de  $\mathrm{GL}_2^+(\mathbb{R})$  associé à la  $(H, V)$ -structure s'intègre en une action globale de  $\mathrm{GL}_2^+(\mathbb{R})$  sur  $\tilde{M}$ . Dans ce cas, les structures algébriques dont nous avons muni  $V$  fournissent à la fois des actions linéaires de  $\Lambda$  sur  $W$  et sur l'espace quotient  $W^{\varpi} = W/\ker \varpi$ , ce qui nous permet de définir les fibrés plats  $M \times_{\Lambda} W$  et  $W^{\varpi}$ . Ces fibrés plats sont équipés d'une action naturelle de  $\mathrm{GL}_2^+(\mathbb{R})$ , qui est induite de son action sur  $\tilde{M} \times W$  (resp.  $\tilde{M} \times W^{\varpi}$ ) définie par  $g(x, w) = (gx, w)$ . En particulier, le fibré tangent de  $M$  peut s'écrire comme le produit tensoriel  $\mathbb{R}^2 \otimes (M \times_{\Lambda} W)$  du fibré trivial  $M \times \mathbb{R}^2$  avec  $M \times_{\Lambda} W$  et on vérifie aisément que l'action tangente à l'action de  $\mathrm{GL}_2^+(\mathbb{R})$  est le produit tensoriel de son action sur  $\mathbb{R}^2$  et de l'action décrite ci-dessus sur  $M \times_{\Lambda} W$ .

Toujours dans ce cadre abstrait, on peut encore définir des sous-fibrés  $\mathrm{GL}_2^+(\mathbb{R})$ -équivariants de  $M \times_{\Lambda} W$ . Comme ci-dessus, nous allons définir ces sous-fibrés en construisant des sous-fibrés  $(\mathrm{GL}_2^+(\mathbb{R}), \Lambda)$ -équivariants de  $\tilde{M} \times W$ . Pour  $x$  dans  $\tilde{M}$ , écrivons  $D(x) = (w_1, w_2)$ , où  $w_1$  et  $w_2$  sont dans  $W$  et notons  $U(x)$  le plan de  $W$  engendré par  $w_1$  et  $w_2$  dans  $W$ . La restriction de  $\varpi$  à  $U(x)$  est non dégénérée. Soient aussi  $U_{\perp}(x)$  l'orthogonal de  $U$  pour  $\varpi$  et  $U_{\perp}^{\varpi}(x)$  l'image de  $U_{\perp}(x)$  dans  $W^{\varpi}$ . Nous notons encore (abusivement, car ces fibrés ne sont pas tous plats),  $M \times_{\Lambda} U$ ,  $M \times_{\Lambda} U_{\perp}$  et  $M \times_{\Lambda} U_{\perp}^{\varpi}$  les fibrés obtenus à partir de ces distributions de sous-espaces vectoriels par passage au quotient sous l'action de  $\Lambda$ .

Alors le fibré en plans  $\mathrm{GL}_2(\mathbb{R})$ -équivariant  $M \times_{\Lambda} U \subset M \times_{\Lambda} W$  est trivial et l'action de  $\mathrm{GL}_2(\mathbb{R})$  s'y identifie à son action produit sur  $M \times \mathbb{R}^2$ . Le fibré  $M \times_{\Lambda} U_{\perp}$  est un supplémentaire  $\mathrm{GL}_2(\mathbb{R})$ -équivariant de  $M \times_{\Lambda} U$ .

Nous conserverons cette terminologie pour décrire les quotients  $\Gamma \backslash M$ , où  $\Gamma$  est un groupe discret de difféomorphismes agissant sans point fixe sur  $M$ . En particulier, nous appellerons alors abusivement fibré tangent de  $M$  l'espace  $\Gamma \backslash (M \times_{\Lambda} V)$ .

Revenons au cas des strates de différentielles abéliennes. Par définition, l'espace  $W^{\varpi}$  s'identifie à l'espace de cohomologie  $H^1(S, \mathbb{R})$ . On appellera le fibré associé sur  $\mathcal{H}(\alpha)$  fibré de Hodge. On notera désormais  $\mathcal{H}(\alpha) \times_{\Gamma} H_{\perp}^1$  (resp.  $\mathcal{H}(\alpha) \times_{\Gamma} p(H_{\perp}^1)$ ) le fibré noté  $M \times_{\Lambda} U_{\perp}$  (resp.  $M \times_{\Lambda} U_{\perp}^{\varpi}$ ) ci-dessus.

*Remarque 2.3.* — Dans le cas où  $g = 1$ , on a  $H_{\perp}^1 = \{0\}$ , si bien que l'action tangente de  $\mathrm{GL}_2^+(\mathbb{R})$  s'identifie à son action produit sur  $\mathcal{H}(\emptyset) \times (\mathbb{R}^2 \otimes \mathbb{R}^2)$ . Cela peut se montrer directement grâce à des arguments élémentaires de théorie de Lie et au fait que  $\mathcal{H}(\emptyset)$  peut se voir comme  $\mathrm{GL}_2^+(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$ .

### 3. DYNAMIQUE DE L'ACTION DE $\mathrm{GL}_2^+(\mathbb{R})$

Maintenant que nous avons introduit les strates de différentielles abéliennes et décrit les structures qui existent naturellement sur ces espaces, nous pouvons énoncer les théorèmes de rigidité pour la dynamique de l'action de  $\mathrm{GL}_2^+(\mathbb{R})$  sur ces strates qui ont été récemment prouvés par Eskin, Mirzakhani et Mohammadi.

#### 3.1. Le théorème topologique d'Eskin-Mirzakhani-Mohammadi

Nous conservons les notations introduites précédemment. Nous disposons donc de la strate  $\mathcal{H}(\alpha)$  et nous l'avons munie d'une  $(H, V)$ -structure affine plate qui est  $\varpi$ -adaptée.

Nous appellerons sous-variété linéaire de  $\mathcal{H}_\Sigma(\alpha)$  une sous-variété connexe  $\mathcal{M}$  de  $\mathcal{H}_\Sigma(\alpha)$  telle que, pour tout  $x$  dans  $\mathcal{M}$ , il existe un ouvert  $U$  de  $x$  tel que l'application de période  $\Pi$  réalise un difféomorphisme de  $U$  vers un ouvert de  $V = \mathrm{H}^1(S, \Sigma, \mathbb{R}^2)$  et que  $\Pi(U \cap \mathcal{M})$  soit l'intersection de  $\Pi(U)$  et d'un sous-espace vectoriel  $V'$  de  $V$ . Par connexité, le sous-espace  $V'$  ne dépend que de  $\mathcal{M}$ .

Nous appellerons sous-variété linéaire de  $\mathcal{H}(\alpha)$  l'image dans  $\mathcal{H}(\alpha)$  d'une sous-variété linéaire de  $\mathcal{H}_\Sigma(\alpha)$ . Soit  $B \subset \mathrm{GL}_2^+(\mathbb{R})$  le sous-groupe des matrices triangulaires supérieures.

**THÉORÈME 3.1** (Eskin-Mirzakhani-Mohammadi [12]). — *Soit  $x$  dans  $\mathcal{H}(\alpha)$ . Alors l'adhérence  $\overline{Bx}$  de l'orbite de  $x$  sous  $B$  est égale à l'adhérence  $\overline{\mathrm{GL}_2^+(\mathbb{R})x}$  de son orbite sous  $\mathrm{GL}_2^+(\mathbb{R})$  et cet ensemble est une sous-variété linéaire de  $\mathcal{H}(\alpha)$ .*

Quand  $g = 2$  (et donc  $\alpha = (1, 1)$  ou  $\alpha = (2)$ ), la propriété de rigidité des adhérences de  $\mathrm{GL}_2^+(\mathbb{R})$ -orbites est due à McMullen [26], qui en donne en outre une classification précise.

#### 3.2. Le théorème métrique d'Eskin-Mirzakhani

Le théorème 3.1 est une conséquence d'un résultat de classification de mesures invariantes. Cela découle d'une procédure devenue classique depuis les travaux de Ratner [29, 30] en dynamique homogène, procédure que nous expliquerons plus loin.

Énonçons à présent ce résultat. Pour cela, nous devons introduire des mesures associées aux sous-variétés linéaires. Soient  $\mathcal{M} \subset \mathcal{H}_\Sigma(\alpha)$  une sous-variété linéaire et  $V' \subset V$  le sous-espace vectoriel associé. Choisissons une mesure de Lebesgue  $\nu'$  sur  $V'$ . Il existe alors une unique mesure de Radon  $\nu_{\mathcal{M}}$  sur  $\mathcal{M}$  telle que, pour tout  $x$  dans  $\mathcal{M}$ , il existe un ouvert  $U$  de  $\mathcal{H}_\Sigma(\alpha)$  où la restriction de  $\Pi$  induit un difféomorphisme vers un ouvert de  $V$  de sorte que la mesure  $\Pi_*(\nu_{\mathcal{M}}|_U)$  soit égale à la restriction de  $\nu'$  à  $\Pi(U)$ . La mesure  $\nu_{\mathcal{M}}$  est déterminée par  $\mathcal{M}$  à multiplication par un nombre  $> 0$  près.

Soient  $k \geq 0$  un entier et  $\nu$  une mesure de Radon sur  $\mathcal{H}(\alpha)$ . Nous pouvons associer à  $\nu$  de manière naturelle une mesure de Radon  $\Gamma_{\alpha, \Sigma}$ -invariante  $\tilde{\nu}$  sur  $\mathcal{H}_\Sigma(\alpha)$ . Nous dirons que  $\nu$  est linéaire (de dimension  $k$ ) si, pour tout  $x$  dans le support de  $\tilde{\nu}$ , il existe une sous-variété linéaire  $\mathcal{M}$  de dimension  $k$  de  $\mathcal{H}_\Sigma(\alpha)$  et un ouvert  $U$  de  $\mathcal{H}_\Sigma(\alpha)$  contenant  $x$  tels que  $\tilde{\nu}|_U = \nu_{\mathcal{M}}|_U$ .

*Remarque 3.2.* — Quand  $g = 1$ , l'action de  $\mathrm{GL}_2^+(\mathbb{R})$  sur  $\mathcal{H}(\emptyset)$  s'identifie à son action sur  $\mathrm{GL}_2^+(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$ . Celle-ci ne préserve pas de mesure borélienne finie. Si l'on cherche à étudier les propriétés ergodiques de l'action, il convient donc de se restreindre à l'action du groupe  $\mathrm{SL}_2(\mathbb{R})$  sur  $\mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$ .

Dans le cas général, nous allons maintenant décrire une fonction  $\mathcal{H}(\alpha) \rightarrow \mathbb{R}_+^*$ , dont les lignes de niveau seront  $\mathrm{SL}_2(\mathbb{R})$ -invariantes. Pour  $\mathcal{A}$  dans  $\mathcal{P}^+(\alpha)$ , définissons l'aire de la surface de translation  $(S, \mathcal{A})$  comme le nombre apparaissant dans (2), c'est-à-dire l'aire d'un domaine fondamental de  $\tilde{S}$  pour l'image inverse par la développante  $D_{\mathcal{A}}$  de la forme volume standard de  $\mathbb{R}^2$ . Cette fonction d'aire est  $\mathcal{G}$ -invariante et factorise donc en une fonction  $a : \mathcal{H}(\alpha) \rightarrow \mathbb{R}_+^*$ . On peut aussi dire que, si  $x$  est un élément de  $\mathcal{H}_{\Sigma}(\alpha)$  et si  $\Pi(x)$  est de la forme  $(w_1, w_2)$  avec  $w_1$  et  $w_2$  dans  $W = \mathrm{H}^1(S, \Sigma, \mathbb{R})$ , alors  $a(x) = \varpi(w_1, w_2)$ . Pour  $g$  dans  $\mathrm{GL}_2^+(\mathbb{R})$  et  $x$  dans  $\mathcal{H}(\alpha)$ , on a  $a(gx) = (\det g)a(x)$ .

Nous définissons  $\mathcal{H}_1(\alpha)$  comme l'ensemble des  $x$  dans  $\mathcal{H}(\alpha)$  avec  $a(x) = 1$ . L'application qui a un point  $x$  de  $\mathcal{H}(\alpha)$  associe le couple  $(\sqrt{a(x)}, \frac{1}{\sqrt{a(x)}}x)$  identifie  $\mathcal{H}(\alpha)$  et  $\mathbb{R}_+^* \times \mathcal{H}_1(\alpha)$ .

Soient  $k \geq 0$  un entier et  $\nu$  une mesure de Radon sur  $\mathcal{H}_1(\alpha)$ . Nous dirons que  $\nu$  est affine (de dimension  $k$ ) si la mesure  $t^k dt \otimes \nu$  sur  $\mathcal{H}(\alpha) \simeq \mathbb{R}_+^* \times \mathcal{H}_1(\alpha)$  est linéaire de dimension  $k + 1$ . Nous dirons aussi qu'un sous-ensemble  $\mathcal{M}$  de  $\mathcal{H}_1(\alpha)$  est une sous-variété affine si l'ensemble  $\mathbb{R}_+^* \times \mathcal{M} \subset \mathcal{H}(\alpha)$  est une sous-variété linéaire. Le support d'une mesure affine est une sous-variété affine.

*Remarque 3.3.* — Les termes de mesure affine et de sous-variété affine ne paraissent pas très appropriés. Leur emploi semble néanmoins faire consensus chez les auteurs.

Nous notons  $P = B \cap \mathrm{SL}_2(\mathbb{R})$  le groupe triangulaire supérieur dans  $\mathrm{SL}_2(\mathbb{R})$ .

**THÉORÈME 3.4** (Eskin-Mirzakhani [11]). — *Soit  $\nu$  une mesure de probabilité  $P$ -invariante et ergodique sur  $\mathcal{H}_1(\alpha)$ . Alors  $\nu$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante et affine.*

En genre 2, la classification des mesures  $\mathrm{SL}_2(\mathbb{R})$ -invariantes est aussi due à McMullen [26].

### 3.3. Un autre cas de rigidité : la théorie de Ratner

Les théorèmes 3.1 et 3.4 répondent à des conjectures formulées depuis plusieurs années par des spécialistes du sujet par analogie avec les célèbres théorèmes de Ratner. Nous rappelons ici brièvement les énoncés de ces théorèmes.

Soit  $G$  un groupe de Lie d'algèbre de Lie  $\mathfrak{g}$ . Étant donné un élément  $X$  de  $\mathfrak{g}$ , on dit que  $X$  est ad-nilpotent si l'endomorphisme  $\mathrm{ad}_X$  de  $\mathfrak{g}$  est nilpotent. On dit alors que le groupe à un paramètre  $t \mapsto \exp(tX)$  est Ad-unipotent.

**THÉORÈME 3.5** (Ratner [29]). — *Soient  $\Gamma$  un sous-groupe discret de  $G$ ,  $H$  un sous-groupe fermé de  $G$  engendré par des groupes à un paramètre Ad-unipotents et  $\nu$  une mesure de probabilité borélienne  $H$ -invariante et ergodique sur  $G/\Gamma$ . Alors il existe un sous-groupe  $L$  de  $G$  et un point  $x$  de  $G$  tels que  $\nu(Lx) = 1$  et que  $\nu$  soit  $L$ -invariante.*

Notons que, si  $H$  est l'image d'un homomorphisme  $\mathrm{SL}_2(\mathbb{R}) \rightarrow G$ , il est bien engendré par des éléments Ad-unipotents.

Par la procédure déjà mentionnée plus haut, on déduit de ce théorème métrique un énoncé topologique :

**THÉORÈME 3.6** (Ratner [30]). — *Supposons que  $\Gamma$  est un réseau de  $G$  et que  $H$  est comme ci-dessus. Si  $x$  est un point de  $G/\Gamma$ , il existe un sous-groupe fermé  $L$  de  $G$  tel que  $\overline{Hx} = Lx$ .*

Au vu de ces énoncés, il serait légitime de se demander si dans les théorèmes 3.1 et 3.4, on peut remplacer le groupe  $P$  par le groupe  $N$  des matrices de la forme  $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ ,  $t \in \mathbb{R}$ . Dans ce cas, des contre-exemples à cette extension ont été construits par Smillie et Weiss (communication privée).

### 3.4. Équidistribution

Nous allons à présent esquisser la démonstration du théorème 3.1 à partir du théorème 3.4. Comme nous l'avons dit plus haut, nous allons suivre une démarche relativement classique qui consiste à établir à partir du théorème de classification 3.4 un résultat d'équidistribution.

Dans la théorie de Ratner, le résultat d'équidistribution concerne les orbites d'un groupe à un paramètre Ad-unipotent et il repose sur une technique d'analyse du comportement de ces flots près des sous-variétés invariantes et près de l'infini dans  $G/\Gamma$ , technique développée par Dani et Margulis [6] en ce qui concerne le comportement en l'infini.

Dans le cas qui nous occupe, on va établir une propriété d'équidistribution de toute une  $\mathrm{SL}_2(\mathbb{R})$ -orbite. Pour cela, les auteurs utilisent une méthode introduite par Eskin, Margulis et Mozes [9], d'inspiration probabiliste, que nous allons à présent décrire.

Considérons un espace topologique localement compact (séparable, union dénombrable de compacts)  $X$ , muni d'une action continue du groupe  $\mathrm{SL}_2(\mathbb{R})$ . Pour  $t, \theta$  dans  $\mathbb{R}$ , nous notons

$$a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \text{ et } r_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Si  $f : X \rightarrow \mathbb{R}_+$  est une fonction continue, on pose, pour  $x$  dans  $X$ ,

$$A_t f(x) = \int_0^{2\pi} \int_0^{2\pi} f(r_{\theta_1} a_t r_{\theta_2} x) d\theta_1 d\theta_2 = \int_{\mathrm{SO}(2)} f(gx) d\mu_t(g),$$

où  $\mu_t$  est la mesure de probabilité  $\mathrm{SO}(2)$ -invariante à gauche et à droite (on dit parfois bi- $\mathrm{SO}(2)$ -invariante) sur l'ensemble  $\mathrm{SO}(2)a_t\mathrm{SO}(2)$ .

*Remarque 3.7.* — L'espace quotient  $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}(2)$  s'identifie au plan hyperbolique  $\mathbb{H} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$  à travers l'application  $g \mapsto gi$ . L'ensemble  $\mathrm{SO}(2)a_t\mathrm{SO}(2)$  est l'image inverse du cercle hyperbolique de centre  $i$  et de rayon  $2t$  par cette application.

Nous allons donner un critère de récurrence de l'action de  $\mathrm{SL}_2(\mathbb{R})$  dans  $X$  dans lequel on considère les opérateurs  $A_t$  comme des opérateurs de Markov et on demande qu'il satisfassent une condition de récurrence classique en théorie des probabilités (voir [27, 28]). Nous dirons donc que l'action de  $\mathrm{SL}_2(\mathbb{R})$  dans  $X$  est exponentiellement récurrente s'il existe une fonction continue propre  $\mathrm{SO}(2)$ -invariante  $f : X \rightarrow \mathbb{R}_+$  et  $C > 0$  tels que, pour tout  $0 < c \leq 1$ , il existe  $t_0 > 0$  avec, pour  $t \geq t_0$  et  $x$  dans  $X$ ,

$$(3) \quad A_t f(x) \leq c f(x) + C.$$

**PROPOSITION 3.8** (Eskin-Mirzakhani-Mohammadi [12]). — *Soit  $\mathcal{M}$  une sous-variété affine  $\mathrm{SL}_2(\mathbb{R})$ -invariante de  $\mathcal{H}(\alpha)$ . Alors l'espace  $\mathcal{H}(\alpha) \setminus \mathcal{M}$  est exponentiellement récurrent pour l'action de  $\mathrm{SL}_2(\mathbb{R})$ .*

Le cas où  $\mathcal{M}$  est vide, c'est-à-dire le fait que l'action de  $\mathrm{SL}_2(\mathbb{R})$  sur  $\mathcal{H}(\alpha)$  est exponentiellement récurrente, est dû à Athreya [1].

Cette propriété de récurrence établie, un ingrédient nous manque encore pour déduire le théorème 3.1 du théorème 3.4, à savoir le résultat suivant de dénombrabilité, qui nous dira qu'une orbite n'a pas beaucoup de choix de sous-variétés affines où s'accumuler. Ce résultat est aussi un analogue d'un phénomène qui apparaît dans la théorie de Ratner [30].

**PROPOSITION 3.9** (Eskin-Mirzakhani-Mohammadi [12]). — *L'ensemble des mesures de probabilité affines  $\mathrm{SL}_2(\mathbb{R})$ -invariantes sur  $\mathcal{H}(\alpha)$  est dénombrable.*

Nous pouvons alors établir la

*Démonstration du théorème 3.1.* — Soit  $\mathfrak{M}$  l'ensemble des mesures affines  $\mathrm{SL}_2(\mathbb{R})$ -invariantes et ergodiques sur  $\mathcal{H}(\alpha)$ . Fixons  $x$  dans  $\mathcal{H}(\alpha)$ .

Commençons par remarquer qu'il existe un élément  $\nu_x$  de  $\mathfrak{M}$  dont le support  $\mathcal{M}_x$  contient  $x$  et qui est minimal pour cette propriété. En effet, choisissons simplement  $\nu_x$  de façon à ce que le support de  $\mathcal{M}_x$  soit de dimension minimale. Si  $\nu$  est un élément de  $\mathfrak{M}$ , dont le support  $\mathcal{M}$  contient  $x$ , a même dimension que  $\mathcal{M}_x$  et est inclus dans  $\mathcal{M}_x$ , comme  $\mathcal{M}$  est ouvert dans  $\mathcal{M}_x$ , on a  $\nu_x|_{\mathcal{M}} = \nu$  et donc, par ergodicité,  $\nu = \nu_x$ , ce qu'il fallait démontrer.

Montrons alors que  $\overline{\mathrm{SL}(\mathbb{R})x} = \mathcal{M}_x$ . Soit  $\nu_\infty$  une valeur d'adhérence quand  $p \rightarrow \infty$ , pour la convergence vague des mesures, de la suite de mesures de probabilité

$$\nu_p = \frac{1}{p} \sum_{k=0}^{p-1} \mu_1^{*k} * \delta_x$$

sur  $\mathcal{H}(\alpha)$ . Nous allons montrer que  $\nu_\infty = \nu_x$ , ce qui achèvera la démonstration.

Commençons par montrer que  $\nu$  est une mesure de probabilité, c'est-à-dire que la masse de  $\nu_p$  reste concentrée sur les parties compactes de  $\mathcal{H}(\alpha)$ . Appliquons la proposition 3.8 avec  $\mathcal{M} = \emptyset$ . Nous disposons alors d'une fonction continue propre  $f : \mathcal{H}(\alpha) \rightarrow \mathbb{R}_+$  et de  $C > 0$  qui satisfont (3). Nous allons estimer

$$A_1^k f(x) = \int_{\mathrm{SL}_2(\mathbb{R})} f(gx) d\mu_1^{*k}(g)$$

pour  $k$  grand. Pour cela, remarquons que, comme la mesure  $\mu_1^{*k}$  est bi-SO(2)-invariante, elle s'écrit

$$\mu_1^{*k} = \int_{\mathbb{R}_+} \mu_t d\rho_k(t)$$

où  $\rho_k$  est une mesure de probabilité sur  $\mathbb{R}_+$ . Soit  $t_0 > 0$  tel que (3) soit valide, avec  $c = 1$ . Par un théorème de Furstenberg sur les produits de matrices aléatoires [15, 5], pour tout  $t_0 > 0$ , on a

$$(4) \quad \rho_k([0, t_0]) \xrightarrow[k \rightarrow \infty]{} 0$$

(la norme d'un produit aléatoire de  $k$ -matrices tend vers l'infini avec  $k$ ). Soient  $\varepsilon > 0$  et  $K = f^{-1}([0, \frac{1}{\varepsilon}]) \subset \mathcal{H}(\alpha)$ , si bien que  $K$  est compact par propriété de  $f$ . Pour tout  $t \geq t_0$ , on a, d'après l'inégalité de Tchebychev,

$$\mu_t * \delta_x(K^c) \leq \varepsilon A_t f(x) \leq \varepsilon(f(x) + C),$$

d'où

$$\mu_1^{*k}(K^c) \leq \rho_k([0, t_0]) + \varepsilon(f(x) + C).$$

D'après (4), on a donc

$$\limsup_{p \rightarrow \infty} \nu_p(K^c) \leq \varepsilon(f(x) + C).$$

Comme ceci est vrai pour tout  $\varepsilon > 0$ ,  $\nu_\infty$  est bien une mesure de probabilité.

*Remarque 3.10.* — L'utilisation de  $c < 1$  dans la définition de la récurrence exponentielle permettrait de montrer qu'on a en réalité

$$\limsup_{p \rightarrow \infty} \nu_p(K^c) \leq \varepsilon C.$$

Donnons-nous maintenant  $\nu$  dans  $\mathfrak{M}$  dont le support  $\mathcal{M}$  est inclus proprement dans  $\mathcal{M}_x$ . Par minimalité, on a  $x \notin \mathcal{M}$ . En raisonnant comme ci-dessus, en appliquant maintenant la proposition 3.8 à  $\mathcal{M}$ , on montre que  $\nu_\infty(\mathcal{M}) = 0$ .

Nous allons à présent montrer que  $\nu_\infty$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante. Commençons par remarquer que, par construction,  $\nu_\infty$  est  $A_1^*$ -invariante (où  $A_1^*$  est l'opérateur adjoint de  $A_1$ , qui agit sur les mesures boréliennes). Par un autre théorème de Furstenberg [16, 3], une telle mesure s'écrit comme une moyenne de mesures de probabilité

$$\nu_\infty = \int_{\mathrm{SL}_2(\mathbb{R})/P} \nu_{\infty, \xi} d\sigma(\xi),$$

où  $\sigma$  est la mesure de probabilité  $\mathrm{SO}(2)$ -invariante de  $\mathrm{SL}_2(\mathbb{R})/P \simeq \mathbb{P}_{\mathbb{R}}^1$  et où, pour tout  $\xi = gP$  dans  $\mathrm{SL}_2(\mathbb{R})/P$ , la mesure  $\nu_{\infty, \xi}$  est  $gPg^{-1}$ -invariante. D'après le théorème 3.4, la mesure  $\nu_{\infty, \xi}$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante, et donc  $\nu_{\infty}$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante.

La mesure  $\nu_{\infty}$  se décompose en une moyenne de mesures  $\mathrm{SL}_2(\mathbb{R})$ -invariantes et ergodiques. Toujours d'après le théorème 3.4, ces mesures ergodiques sont des éléments de  $\mathfrak{M}$ . Comme, d'après la proposition 3.9, l'ensemble  $\mathfrak{M}$  est dénombrable, on peut écrire

$$\nu_{\infty} = \sum_{\substack{\nu \in \mathfrak{M} \\ \text{supp } \nu \subset \mathcal{M}_x}} a_{\nu} \nu.$$

Or, pour tout  $\nu$  dans  $\mathfrak{M}$ , si le support  $\mathcal{M}$  de  $\nu$  est proprement inclus dans  $\mathcal{M}_x$ , on a

$$a_{\nu} \leq \nu_{\infty}(\mathcal{M}) = 0.$$

Il vient bien  $\nu_{\infty} = \nu_x$ , ce qu'il fallait démontrer.  $\square$

### 3.5. Application aux billards

Le théorème 3.1 implique des conséquences pour le comptage des trajectoires périodiques dans les billards plans polygonaux à angles rationnels. En effet, suivant une procédure classique (voir par exemple [34]), à un tel billard, on peut associer une surface de translation. Les propriétés dynamiques du billard se réinterprètent comme des propriétés des flots de translation dans cette surface. Si le billard est rectangulaire, cette surface de translation est un tore et des raisonnements de comptage élémentaire permettent de montrer que le nombre de (cylindres de) trajectoires périodiques de longueur  $\leq T$  est équivalent à un multiple de  $T^2$ . Dans le cas général, on a le

**THÉORÈME 3.11** (Masur [24, 25]). — *Étant donné un billard plan polygonal à angles rationnels, il existe des constantes  $0 < c_1 \leq c_2$  telles que le nombre  $N(T)$  de cylindres de trajectoires périodiques de longueur  $\leq T$  vérifie*

$$c_1 T^2 \leq N(T) \leq c_2 T^2.$$

Les théorèmes 3.1 et 3.4 entraînent, suivant une procédure qui était connue [10], le

**THÉORÈME 3.12** (Eskin-Mirzakhani-Mohammadi [12]). — *Étant donné un billard plan polygonal à angles rationnels, il existe une constante  $c > 0$  telle que le nombre  $N(T)$  de cylindres de trajectoires périodiques de longueur  $\leq T$  vérifie*

$$e^{-T} \int_1^T \frac{N(S) \, dS}{S^2} \xrightarrow{T \rightarrow \infty} c.$$

La question de savoir si on a  $N(T) \sim_{T \rightarrow \infty} cT^2$  reste ouverte.

#### 4. COCYCLES LINÉAIRES

Maintenant que nous avons vu que le théorème 3.1 découlait du théorème 3.4 par une procédure relativement classique – quoique pas forcément facile à mettre en œuvre – nous allons tâcher de donner des éléments de la démonstration du remarquable théorème 3.4.

Comme nous l’avons dit plus haut, le théorème 3.4 se veut un analogue, pour l’action de  $\mathrm{SL}_2(\mathbb{R})$  sur les strates de différentielles abéliennes, du théorème de Ratner 3.5. Au demeurant, sa démonstration emprunte certaines des méthodes développées pour l’étude de la dynamique dans les espaces homogènes [4, 23].

La stratégie générale de ce type de théorème est la suivante. On dispose d’un groupe  $H$  agissant par difféomorphismes sur une variété  $X$ , munie d’une mesure de probabilité  $\nu$ . On cherche à montrer que  $\nu$  doit satisfaire certaines conditions géométriques (être affine, être homogène sous l’action d’un sous-groupe fermé). Pour cela, on applique des théorèmes abstraits de théorie ergodique qui garantissent que, pour  $\nu$ -presque tout point de  $x$ , certaines suites de sous-ensembles  $H_{n,x}$  de l’orbite  $Hx$  de  $x$  ont tendance à s’équidistribuer dans  $X$ . On applique alors ces théorèmes à deux points  $x$  et  $y$  très proches de  $X$ , dans l’optique de trouver  $h$  et  $n$  dans  $H$  tel que, d’une part  $hx \in H_{n,x}$  et  $hy \in H_{n,y}$  et, d’autre part, la distance entre  $hx$  et  $hy$  soit de taille prescrite (et non plus arbitrairement petite). La propriété d’équidistribution de  $H_{n,x}$  et  $H_{n,y}$  doit quant à elle permettre de forcer  $hx$  et  $hy$  à être dans des ensembles de grosse mesure où certaines fonctions mesurables  $H$ -équivariantes deviennent continues en restriction (de tels ensembles existent d’après le théorème de Lusin). Le fait que  $x$  et  $y$  soient proches et que  $hx$  et  $hy$  soient à une distance contrôlée permet de montrer que certains objets associés à  $\nu$  et à la géométrie de l’espace possèdent des propriétés d’invariance par les transformations qui permettent de passer de  $hx$  à  $hy$ .

Je vais tâcher de rendre ce dernier paragraphe plus clair dans la suite du texte. Néanmoins, le lecteur aura peut-être compris qu’un problème important du sujet consiste à contrôler à quelle vitesse, partant de deux points proches  $x$  et  $y$ , les orbites  $Hx$  et  $Hy$  s’éloignent l’une de l’autre. Pour cela on analyse l’action tangente du groupe sur le fibré tangent  $TX$ .

C’est ici qu’apparaît une différence très importante entre le cas des espaces homogènes et celui des strates de différentielles abéliennes.

En effet, si  $G$  est un groupe de Lie et  $\Gamma$  un sous-groupe discret de  $G$ , l’action tangente à l’action de  $G$  dans  $G/\Gamma$  est constante. Plus précisément, si  $\mathfrak{g}$  est l’algèbre de Lie de  $G$ , en identifiant  $\mathfrak{g}$  et l’algèbre de Lie des champs de vecteurs invariants à droite sur  $G$ , on obtient un isomorphisme de  $TG$  sur  $G \times \mathfrak{g}$ , à travers lequel l’action à droite d’un élément  $g$  de  $G$  se lit comme la transformation

$$(h, X) \mapsto (hg, X)$$

et son action à gauche comme la transformation

$$(h, X) \mapsto (gh, \text{Ad}(g)X),$$

où  $\text{Ad}$  désigne l'action adjointe de  $G$  dans  $\mathfrak{g}$ . Cette structure étant invariante à droite, elle descend à  $G/\Gamma$  et donne donc un isomorphisme entre  $TG/\Gamma$  et  $G/\Gamma \times \mathfrak{g}$ , à travers lequel l'action tangente à l'action de  $g$  se lit toujours comme

$$(x, X) \mapsto (gx, \text{Ad}(g)X).$$

En particulier, le comportement asymptotique de cette action tangente le long d'une orbite  $g^n x$ ,  $n \geq 0$ , ne dépend pas de  $x$  et est déterminé par la réduction de Jordan de l'endomorphisme linéaire  $\text{Ad}(g)$ .

Dans les strates de différentielles abéliennes, le fibré tangent ne possède pas de trivialisations  $\text{SL}_2(\mathbb{R})$ -équivariante. En revanche, la structure affine plate est  $\text{SL}_2(\mathbb{R})$ -invariante, si bien que l'action tangente est en quelque sorte localement constante. Plus précisément, si  $(g_t)_{0 \leq t \leq 1}$  est une courbe continue tracée sur  $\text{SL}_2(\mathbb{R})$  avec  $g_0 = e$  et si  $x$  et  $y$  sont deux points proches de  $\mathcal{H}(\alpha)$  tels que, pour tout  $0 \leq t \leq 1$ ,  $g_t x$  et  $g_t y$  restent proches, on peut naturellement identifier les espaces tangents à  $\mathcal{H}(\alpha)$  en  $x$  et  $y$  (resp. en  $g_1 x$  et en  $g_1 y$ ) de façon à ce que la différentielle de l'action de  $g_1$  soit donnée par la même application linéaire en  $x$  et en  $y$ .

L'utilisation de cette idée joue un rôle fondamental dans le travail d'Eskin et Mirzakhani. Elle s'y articule avec la théorie d'Osseledets, c'est-à-dire l'étude générale du comportement asymptotique de l'action tangente au-dessus d'une mesure invariante pour l'action d'un groupe sur une variété différentielle. Nous allons commencer par rappeler les éléments de cette théorie.

#### 4.1. Cocycles au-dessus d'une action ergodique

Nous rappelons ici le langage usuel des cocycles en théorie ergodique.

Soit  $G$  un groupe topologique localement compact (séparable et union dénombrable de compacts) agissant de façon mesurable sur un espace de probabilité de Lebesgue  $(X, \nu)$  en préservant la mesure  $\nu$  (rappelons qu'un espace de Lebesgue est un espace probabilisé isomorphe à l'union disjointe d'un intervalle, muni de la mesure de Lebesgue et d'un nombre au plus dénombrable d'atomes). On suppose que  $\nu$  est  $G$ -ergodique. Soit  $H$  un autre groupe topologique localement compact (séparable et union dénombrable de compacts).

Un cocycle à valeurs dans  $H$  au-dessus de l'action de  $G$  sur  $X$  est une application mesurable

$$\sigma : G \times X \rightarrow H$$

telle que, pour tous  $g_1, g_2$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait

$$(5) \quad \sigma(g_1 g_2, x) = \sigma(g_1, g_2 x) \sigma(g_2, x).$$

*Remarque 4.1.* — Dans la suite, nous nous autoriserons éventuellement à inverser les quantificateurs dans la formule 5 et dans les formules analogues. Cela est possible grâce à des théorèmes généraux sur les actions de groupes localement compacts séparables et union dénombrables de compacts sur des espaces de Lebesgue (voir par exemple [35]).

*Exemple 4.2.* — Supposons que  $X$  est une variété différentielle de dimension  $d$  sur laquelle  $G$  agit par difféomorphismes. Alors le fibré tangent de  $X$  est mesurablement trivial : il existe un isomorphisme mesurable de fibrés de  $TX$  sur  $X \times \mathbb{R}^d$ . L'action tangente de  $G$  sur  $TX$  est alors définie par une application mesurable  $\sigma : G \times X \rightarrow \mathrm{GL}_d(\mathbb{R})$ . La formule de dérivation des applications composées se traduit par le fait que l'application  $\sigma$  est un cocycle. Cet argument s'étend à toute action fibrée de  $G$  sur un fibré vectoriel au-dessus de  $X$ .

Soit  $Y$  un ensemble muni d'une action de  $H$ . Alors, à un cocycle  $\sigma : G \times X \rightarrow H$ , on peut associer l'action croisée sur l'ensemble  $X \times Y$  définie par  $g(x, y) = (gx, \sigma(g, x)y)$ ,  $g \in G$ ,  $x \in X$ ,  $y \in Y$ . La relation 5 garantit précisément que cette formule définit une action.

Deux cocycles  $\sigma, \sigma' : G \times X \rightarrow H$  sont cohomologues s'il existe une application  $\varphi : X \rightarrow H$  avec, pour tout  $g$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ ,

$$\sigma'(g, x) = \varphi(gx)\sigma(g, x)\varphi(x)^{-1}.$$

Dans ce cas, les actions croisées associées à  $\sigma$  et  $\sigma'$  sont conjuguées.

## 4.2. Adhérence de Zariski

Nous allons associer à un cocycle à valeurs dans un groupe algébrique  $H$  une classe de conjugaison de sous-groupes algébriques de  $H$ . Par groupe algébrique, nous entendrons toujours ici le groupe des points réels d'un groupe algébrique affine défini sur  $\mathbb{R}$ , c'est-à-dire, en langage élémentaire, un sous-groupe d'un groupe de matrices, défini par des équations polynomiales en les coordonnées.

Soient  $G$  et  $(X, \nu)$  comme avant et  $H$  un groupe localement compact, séparable et union dénombrable de compacts. Soient  $\sigma : G \times X \rightarrow H$  un cocycle et  $L$  un sous-groupe fermé de  $H$ . Si  $\sigma$  est cohomologue à un cocycle à valeurs dans  $L$ , il existe une application  $\varphi : X \rightarrow H/L$  telle que, pour tout  $g$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ ,  $\varphi(gx) = \sigma(g, x)\varphi(x)$ , c'est-à-dire que la fibration  $X \times H/L \rightarrow X$  possède une section  $G$ -équivariante pour l'action croisée. Réciproquement, si une telle application  $\varphi$  existe, l'existence d'une section borélienne pour l'application quotient  $H \rightarrow H/L$  garantit que  $\sigma$  est cohomologue à un cocycle à coefficients dans  $L$ .

**PROPOSITION 4.3** (Zimmer [35]). — *Soient  $H$  un groupe algébrique réel et  $\sigma : G \times X \rightarrow H$  un cocycle. Il existe un sous-groupe algébrique  $L$  de  $H$  tel que  $\sigma$  soit cohomologue à un cocycle à coefficients dans  $L$  et que, si  $L'$  est un autre sous-groupe algébrique avec cette propriété,  $L'$  contient un conjugué de  $L$  (c'est-à-dire qu'on a  $hLh^{-1} \subset L'$  pour un  $h$  dans  $H$ ).*

On dit que le groupe  $L$  est l'adhérence de Zariski-Zimmer du cocycle  $\sigma$  (il y a là un abus de langage, dans la mesure où le groupe  $L$  n'est défini qu'à conjugaison près).

*Démonstration.* — Cela découle de la remarque suivante : si  $L_1$  et  $L_2$  sont des sous-groupes algébriques de  $L$ , tels que  $\sigma$  soit cohomologue d'une part à un cocycle à coefficients dans  $L_1$  et d'autre part à un cocycle à coefficients dans  $L_2$ , alors  $\sigma$  est cohomologue à un cocycle à coefficients dans un sous-groupe de la forme  $L_1 \cap hL_2h^{-1}$  pour un élément  $h$  de  $H$ .

En effet, pour  $i = 1, 2$ , il existe une application  $\varphi_i$  de  $X$  dans  $H/L_i$  telle que, pour tout  $g$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait  $\varphi_i(gx) = \sigma(g, x)\varphi_i(x)$ . On pose  $\varphi = (\varphi_1, \varphi_2)$  et on munit  $Y = H/L_1 \times H/L_2$  de l'action produit de  $H$ . On a encore  $\varphi(gx) = \sigma(g, x)\varphi(x)$ . Or, comme  $Y$  est une variété quasi-projective, munie d'une action algébrique de  $H$ , le quotient  $H \backslash Y$  est à base dénombrable. Par ergodicité de l'action de  $G$  sur  $X$ , l'application  $\varphi$  prend donc ses valeurs dans une seule  $H$ -orbite dans  $Y$ . Celle-ci est de la forme  $H/L$ , où  $L$  est le stabilisateur d'un point de  $Y$ . Ces stabilisateurs sont bien de la forme  $h_1L_1h_1^{-1} \cap h_2L_2h_2^{-1}$ .  $\square$

### 4.3. Actions de groupes moyennables

Nous allons à présent utiliser la notion d'adhérence de Zariski-Zimmer et des raisonnements de théorie des groupes liénaires pour montrer que, si le groupe  $G$  est moyennable, un cocycle à valeurs dans un groupe algébrique  $H \subset \mathrm{GL}_d(\mathbb{R})$  est essentiellement cohomologue à un cocycle à valeurs dans un groupe de matrices qui se triangularise par blocs, avec des blocs qui sont des similitudes. C'est ce résultat, appliqué à l'action du sous-groupe  $P$  de  $\mathrm{SL}_2(\mathbb{R})$  dans  $\mathcal{H}(\alpha)$ , qui nous permettra de remplacer les arguments reposant sur la réduction de Jordan des endomorphismes linéaires utilisés d'habitude dans l'étude des systèmes dynamiques sur les espaces homogènes.

L'hypothèse que  $G$  est moyennable sera utilisée sous la forme du lemme suivant, qui est une version fibrée du fait que toute action continue de  $G$  sur un espace compact y préserve une mesure de probabilité.

LEMME 4.4. — *Soient  $\sigma : G \times X \rightarrow H$  un cocycle et  $Y$  un espace métrique compact, muni d'une action continue de  $H$ . Supposons que  $G$  est moyennable. Alors l'action croisée de  $G$  sur  $X \times Y$  possède une mesure invariante dont la projection sur  $X$  est  $\nu$ .*

*Démonstration.* — Nous devons construire une application mesurable  $x \mapsto \rho_x$  de  $X$  vers l'espace des mesures de probabilité boréliennes sur  $Y$  telle que, pour tout  $g$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait  $\rho_{gx} = \sigma(g, x)_*\rho_x$ . Nous allons montrer qu'une telle application existe en la réalisant comme un point fixe pour une certaine action continue de  $G$  sur une partie convexe compacte d'un espace vectoriel topologique localement convexe. Pour cela, considérons l'espace de Banach

$$E = L^1(X, \nu, \mathcal{C}^0(Y))$$

des classes essentielles d'applications  $x \mapsto f_x$  de  $X$  dans  $\mathcal{C}^0(Y)$  qui sont mesurables (pour la structure borélienne de l'espace de Banach séparable  $\mathcal{C}^0(Y)$ ) et telles que

$\int_X \|f_x\|_\infty d\nu(x) < \infty$ . Identifions l'espace dual de  $\mathcal{C}^0(Y)$  à l'espace  $\mathcal{M}(Y)$  des mesures boréliennes complexes sur  $Y$  et notons  $\mathcal{P}(Y) \subset \mathcal{M}(Y)$  l'ensemble des mesures de probabilité. Comme  $(X, \nu)$  est un espace de Lebesgue et comme  $\mathcal{C}^0(Y)$  est séparable, l'espace dual  $E'$  de  $E$  s'identifie naturellement à l'espace des classes essentielles d'applications  $x \mapsto \rho_x$  de  $X$  dans  $\mathcal{M}(Y)$  qui sont mesurables (pour la structure borélienne induite par la topologie faible-\* de  $\mathcal{M}(Y)$ ) et essentiellement bornées. Notons  $E'_1 \subset E'$  le sous-ensemble de  $E'$  constitué des applications  $x \mapsto \rho_x$  qui prennent essentiellement leurs valeurs dans  $\mathcal{P}(Y)$ . Alors, d'après le théorème de Banach-Alaoglu,  $E'_1$  est une partie convexe compacte de  $E'$  pour la topologie faible-\*. Comme l'action croisée de  $G$  sur  $X \times Y$  induit une action faible-\* continue sur  $E'$  qui préserve  $E'_1$ ,  $G$  fixe un point dans  $E'_1$ , ce qu'il fallait démontrer.  $\square$

Dans nos constructions, l'espace  $Y$  sera un certain espace projectif  $\mathbb{P}_{\mathbb{R}}^{d-1}$ . Nous appliquerons alors aux mesures construites un lemme classique sur la structure du stabilisateur d'une mesure sur  $\mathbb{P}_{\mathbb{R}}^{d-1}$  dans  $\mathrm{GL}_d(\mathbb{R})$ . Rappelons qu'un sous-groupe  $L$  de  $\mathrm{GL}_d(\mathbb{R})$  agit irréductiblement sur  $\mathbb{R}^d$  si et seulement si  $\{0\}$  et  $\mathbb{R}^d$  sont les seuls sous-espaces  $L$ -invariants de  $\mathbb{R}$ .

LEMME 4.5 (Furstenberg). — *Soit  $L$  un sous-groupe de  $\mathrm{GL}_d(\mathbb{R})$  qui agit irréductiblement sur  $\mathbb{R}^d$ . Supposons que  $L$  préserve une mesure de probabilité sur  $\mathbb{P}_{\mathbb{R}}^{d-1}$ . Alors, il existe un sous-groupe d'indice fini  $M$  de  $L$  et une décomposition de  $\mathbb{R}^d$  en somme directe  $\mathbb{R}^d = V_1 \oplus \cdots \oplus V_k$  telle que, pour  $1 \leq i \leq k$ ,  $M$  préserve  $V_i$  et que l'image de  $M$  dans  $\mathrm{PGL}(V_i)$  soit compacte.*

*Démonstration.* — Commençons par montrer que, si  $L$  est un sous-groupe de  $\mathrm{GL}_d(\mathbb{R})$  qui préserve une mesure de probabilité  $\rho$  sur  $\mathbb{P}_{\mathbb{R}}^{d-1}$ , et si l'image de  $L$  dans  $\mathrm{PGL}_d(\mathbb{R})$  n'est pas compacte, la mesure  $\rho$  est concentrée sur la réunion de deux sous-espaces projectifs propres de  $\mathbb{P}_{\mathbb{R}}^{d-1}$ . En effet, par hypothèse,  $L$  contient une suite d'éléments  $g_n$  dont la décomposition de Cartan est de la forme

$$k_n \begin{pmatrix} a_{1,n} & & 0 \\ & \ddots & \\ 0 & & a_{d,n} \end{pmatrix} l_n$$

(avec  $k_n, l_n$  dans  $\mathrm{O}(d)$  et  $a_{1,n} \geq \cdots \geq a_{d,n} > 0$ ) et qui tend vers l'infini dans  $\mathrm{PGL}_d(\mathbb{R})$ . Quitte à extraire une sous-suite, on peut supposer que, pour certains  $k, l$  dans  $\mathrm{O}(d)$  et  $1 \leq i \leq d-1$ , on a  $k_n \xrightarrow[n \rightarrow \infty]{} k$ ,  $l_n \xrightarrow[n \rightarrow \infty]{} l$  et

$$\frac{a_{i,n}}{a_{i+1,n}} \xrightarrow[n \rightarrow \infty]{} \infty.$$

Alors, nécessairement,  $\rho$  est concentrée sur l'ensemble

$$k\mathbb{P}(\mathbb{R}^i \times \{0\}) \cup l^{-1}\mathbb{P}(\{0\} \times \mathbb{R}^{d-i}).$$

Supposons à présent que  $L$  agit irréductiblement sur  $\mathbb{R}^d$  et soit  $\mathcal{V}$  l'ensemble des sous-espaces vectoriels non nuls  $V$  de  $\mathbb{R}^d$  tels que  $\rho(\mathbb{P}(V)) > 0$  et que la dimension de  $V$  soit minimale pour cette propriété. Pour tous  $V_1 \neq V_2$  dans  $\mathcal{V}$ , on a

$$\rho(\mathbb{P}(V_1) \cup \mathbb{P}(V_2)) = \rho(\mathbb{P}(V_1)) + \rho(\mathbb{P}(V_2))$$

et donc l'ensemble  $\mathcal{W}$  des éléments de  $\mathcal{V}$  dont l'image dans  $\mathbb{P}_{\mathbb{R}}^{d-1}$  est de mesure maximale est fini. Comme  $L$  préserve  $\mathcal{W}$ , les éléments de  $\mathcal{W}$  engendrent  $\mathbb{R}^d$ . Pour  $W$  dans  $\mathcal{W}$ , le stabilisateur de  $W$  dans  $L$  préserve la restriction de  $\rho$  à  $\mathbb{P}(W)$  et cette mesure ne charge pas de sous-espace projectif de  $\mathbb{P}(W)$ . Ce groupe a donc une image compacte dans  $\mathrm{PGL}(W)$ . Le lemme en découle.  $\square$

Pour achever la réduction générale des cocycles au-dessus des actions de groupes moyennables, nous aurons encore besoin d'un outil pour contourner le problème qui se pose quand le sous-groupe d'indice fini  $M$  du lemme 4.5 est un sous-groupe propre de  $L$ . Nous appellerons extension finie de l'action de  $G$  sur  $X$  une action de  $G$  sur un espace de la forme  $\tilde{X} = X \times F$  qui est l'action croisée définie par un cocycle à valeurs dans le groupe des permutations de l'ensemble  $F$ . En d'autres termes, il s'agit d'une action sur  $X \times F$  telle que la projection sur  $X$  soit  $G$ -équivariante. Nous munissons  $\tilde{X}$  de la mesure de produit de  $\nu$  et de la mesure de comptage normalisée de  $F$ . Tout cocycle au-dessus de  $X$  peut alors se voir comme un cocycle au-dessus de  $\tilde{X}$ . Par contre deux cocycles peuvent être cohomologues au-dessus de  $\tilde{X}$  sans l'être au-dessus de  $X$ .

Soient  $L$  un sous-groupe de  $\mathrm{GL}_d(\mathbb{R})$  et  $\{0\} = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_k = \mathbb{R}^d$  un drapeau de  $\mathbb{R}^d$ . Nous dirons que le groupe  $L$  est triangulaire à blocs conformes par rapport au drapeau  $(V_i)$  si pour tout  $1 \leq i \leq k$ , on a  $LV_i = V_i$  et l'image de  $L$  dans  $\mathrm{PGL}(V_i)$  est compacte. En d'autres termes, il existe une base adaptée au drapeau  $(V_i)$  dans laquelle les matrices des éléments de  $L$  sont de la forme

$$(6) \quad \begin{pmatrix} e^{\lambda_1} U_1 & * & * & * \\ 0 & e^{\lambda_2} U_2 & * & * \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & e^{\lambda_k} U_k \end{pmatrix},$$

avec  $\lambda_i \in \mathbb{R}$ ,  $U_i \in \mathrm{O}(d_i)$ ,  $d_i = \dim V_i - \dim V_{i-1}$ ,  $1 \leq i \leq k$ .

Des deux lemmes précédents, nous déduisons la

**PROPOSITION 4.6.** — *Supposons que le groupe  $G$  est moyennable. Soit  $\sigma : G \times X \rightarrow \mathrm{GL}_d(\mathbb{R})$  un cocycle. Il existe une action de  $G$  sur une extension finie  $\tilde{X}$  telle que l'adhérence de Zariski-Zimmer de  $\sigma$ , vu comme un cocycle au-dessus de  $\tilde{X}$  soit triangulaire à blocs conformes. En d'autres termes, au-dessus de  $\tilde{X}$ , le cocycle  $\sigma$  est cohomologue à un cocycle à valeurs dans le groupe des matrices de la forme (6).*

*Démonstration.* — Soit  $L$  l'adhérence de Zariski-Zimmer de  $\sigma$ . Nous pouvons supposer que  $\sigma$  prend ses valeurs dans  $L$ . Soit  $W$  un sous-quotient irréductible de l'action de  $L$  dans  $\mathbb{R}^d$ , c'est-à-dire que  $W = U/V$  où  $U \supset V$  sont des sous-espaces  $L$ -invariants de  $\mathbb{R}^d$  et que l'action de  $L$  dans  $\mathbb{R}^d$  est irréductible. D'après le lemme 4.4, il existe une

application mesurable  $x \mapsto \rho_x$  de  $X$  dans l'espace  $\mathcal{P}(\mathbb{P}(W))$  des mesures de probabilité sur  $\mathbb{P}(W)$  telle que, pour tout  $g$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait  $\rho_{gx} = \sigma(g, x)_* \rho_x$ . D'après [35], l'espace quotient  $L \backslash \mathcal{P}(\mathbb{P}(W))$  est à base dénombrable. Par conséquent, par ergodicité, l'application  $x \mapsto \rho_x$  prend essentiellement ses valeurs dans une seule  $L$ -orbite qui est de la forme  $L/L_W$ , où  $L_W$  est le stabilisateur d'une mesure de probabilité de  $\mathbb{P}(W)$  dans  $L$ . En d'autres termes, le cocycle  $\sigma$  est cohomologue à un cocycle  $\sigma_W$  à coefficients dans  $L_W$  et en particulier, par définition de  $L$ ,  $L_W$  est Zariski dense dans  $L$  (en fait, on peut même montrer que  $L_W = L$ , mais cela ne nous servira pas). En particulier,  $L_W$  agit irréductiblement sur  $W$ .

D'après le lemme 4.5, il existe une décomposition  $W = W_1 \oplus \cdots \oplus W_k$  de  $W$  en somme directe et des normes euclidiennes  $\|\cdot\|_1, \dots, \|\cdot\|_k$  sur  $W_1, \dots, W_k$  telles que, si  $M_W$  est le sous-groupe de  $L$  constitué des éléments de  $L$  qui, pour  $1 \leq i \leq k$ , préservent  $W_i$  en y agissant de façon  $\|\cdot\|_i$ -conforme, alors  $L_W \cap M_W$  est d'indice fini dans  $L_W$ . Si  $M_W = L_W$ , nous avons terminé l'analyse de ce sous-quotient  $W$ . Dans le cas général, c'est ici qu'il nous faut introduire une extension finie de  $X$ . Nous posons  $\tilde{X}_W = X \times L_W/M_W$  et nous le munissons de l'action croisée associée au cocycle  $\sigma_W$ . Alors, par construction, le cocycle  $\sigma_W$ , et donc le cocycle  $\sigma$ , vu comme un cocycle au-dessus de  $\tilde{X}_W$ , est cohomologue à un cocycle à coefficients dans  $M_W$ .

Le résultat en découle en définissant  $\tilde{X}$  comme le produit fibré au-dessus de  $X$  de tous les systèmes  $\tilde{X}_W$ , où  $W$  parcourt l'ensemble des sous-quotients irréductibles de  $L$  dans  $\mathbb{R}^d$ .  $\square$

#### 4.4. Théorème d'Osseledets et exposants de Lyapunov

Dans le paragraphe précédent, nous avons développé un théorème de structure pour les cocycles au-dessus d'une action d'un groupe moyennable  $G$ . Nous allons à présent étudier des phénomènes qui apparaissent quand  $G$  est  $\mathbb{Z}$  ou  $\mathbb{R}$ , du fait de l'existence de la relation d'ordre sur ces groupes. Pour simplifier l'exposé et éviter des difficultés techniques, nous supposerons qu'on a  $G = \mathbb{Z}$ , le cas réel étant analogue.

Nous nous donnons donc un automorphisme ergodique  $T : X \rightarrow X$  de l'espace de Lebesgue  $(X, \nu)$ . Un cocycle  $\sigma : \mathbb{Z} \times X \rightarrow H$  est alors complètement déterminé par l'application  $\sigma(1, \cdot)$ .

Dans ce paragraphe, nous parlerons donc de cocycles au lieu d'applications  $\sigma : X \rightarrow H$  quand nous utiliserons ces applications pour construire des automorphismes qui sont des produits croisés  $(x, y) \mapsto (Tx, \sigma(x)y)$ . Pour une telle application, nous noterons, pour  $x$  dans  $X$  et  $n$  dans  $\mathbb{N}$ ,

$$\sigma_n(x) = \sigma(T^{n-1}x) \cdots \sigma(x)$$

et, pour  $n$  entier  $\leq -1$ ,

$$\sigma_n(x) = \sigma(T^{-n}x)^{-1} \cdots \sigma(T^{-1}x)^{-1}.$$

Nous munirons systématiquement  $\mathbb{R}^d$  de la norme euclidienne usuelle (les constructions auxquelles nous allons procéder ne dépendent pas de ce choix). Nous avons immédiatement le

LEMME 4.7. — *Soit  $\sigma : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  un cocycle. Pour tout  $\lambda$  dans  $\mathbb{R}$ , pour tout  $x$  dans  $X$ , l'ensemble*

$$V_\lambda(x) = \left\{ v \in \mathbb{R}^d \left| \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|\sigma_n(x)v\| \leq \lambda \right. \right\}$$

*est un sous-espace vectoriel. On a  $V_\lambda(Tx) = \sigma(x)V_\lambda(x)$ .*

En particulier, la dimension de  $V_\lambda$  est une fonction invariante. Par ergodicité, elle est constante. Nous notons  $d_\lambda$  sa valeur. La fonction  $\lambda \mapsto d_\lambda$  est croissante.

Supposons dorénavant  $\int_X \log \|\sigma(x)\| d\nu(x) < \infty$ , de sorte que, d'après le théorème de Birkhoff,  $d_\lambda = d$  pour  $\lambda$  suffisamment grand et aussi (mais ce ne serait pas indispensable)  $\int_X \log \|\sigma(x)^{-1}\| d\nu(x) < \infty$ , de sorte que, pour la même raison,  $d_\lambda = 0$  pour  $\lambda$  suffisamment petit. Nous dirons d'un cocycle  $\sigma : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  tel que

$$(7) \quad \int_X \log \max(\|\sigma(x)\|, \|\sigma(x)^{-1}\|) d\nu(x) < \infty$$

qu'il est intégrable.

La fonction croissante à valeurs entières  $\lambda \mapsto d_\lambda$  a un nombre fini de discontinuités en des réels  $\lambda_1 < \dots < \lambda_k$ . On note  $0 < d_1 < \dots < d_k = d$  les valeurs de cette fonction en ces discontinuités. Les nombres  $\lambda_1 < \dots < \lambda_k$  sont appelés exposants de Lyapunov du cocycle  $\sigma$ . On pose, pour  $x$  dans  $X$  et  $1 \leq i \leq k$ ,  $V_i(x) = V_{\lambda_i}(x)$ . On dit que le drapeau  $V_1(x) \subsetneq \dots \subsetneq V_k(x)$  est le drapeau de Lyapunov du cocycle.

Soit  $H \subset \mathrm{GL}_d(\mathbb{R})$  le stabilisateur d'un drapeau  $\{0\} = W_0 \subsetneq W_1 \subsetneq W_2 \subsetneq \dots \subsetneq W_k = \mathbb{R}^d$  avec  $\dim W_i = d_i$ ,  $1 \leq i \leq k$ . Le lemme 4.7 implique que  $\sigma$  est cohomologue à un cocycle à coefficients dans  $H$ . Nous allons appliquer la proposition 4.6 à l'action de ce cocycle dans  $W_i/W_{i-1}$ .

Notons que, dans ces réductions, on peut faire en sorte que la propriété d'intégrabilité (7) soit conservée. En effet, si  $Q$  est le stabilisateur d'un drapeau dans  $\mathrm{GL}_d(\mathbb{R})$ , on a, par le procédé d'orthogonalisation de Gramm-Schmidt,  $\mathrm{GL}_d(\mathbb{R}) = O(d)Q$ , si bien que si  $\sigma$  est cohomologue à un cocycle à coefficients dans  $Q$ , on peut toujours écrire, pour  $x$  dans  $X$ ,

$$\sigma(x) \in \varphi(Tx)Q\varphi(x)^{-1}$$

où  $\varphi$  prend ses valeurs dans  $O(d)$ , si bien que les propriétés d'intégrabilité de  $\sigma$  sont conservées.

De même si maintenant  $\sigma$  est cohomologue à un cocycle conforme, il existe une application  $\varphi : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  et une fonction  $\gamma : X \rightarrow \mathbb{R}$  telles que, pour  $\nu$ -presque tout  $x$  dans  $X$ ,

$$\sigma(x) \in e^{\gamma(x)}\varphi(Tx)O(d)\varphi(x)^{-1}.$$

Écrivons  $\varphi(x) = e^{\theta(x)}\varphi_1(x)$  avec  $|\det \varphi_1(x)| = 1$  et posons  $\gamma_1(x) = \gamma(x) + \theta(Tx) - \theta(x)$ . On a encore

$$\sigma(x) \in e^{\gamma_1(x)}\varphi_1(Tx)\mathrm{O}(d)\varphi_1(x)^{-1}$$

et en particulier

$$\gamma_1(x) = \frac{1}{d} \log |\det \sigma(x)|,$$

si bien que  $\gamma_1$  est une fonction intégrable.

Notons que cette réduction ne perturbe pas l'étude du comportement exponentiel de la norme des vecteurs dans les actions linéaires croisées en raison du

LEMME 4.8. — *Soit  $\sigma : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  un cocycle intégrable. Supposons que  $\sigma$  est cohomologue à un cocycle à coefficients dans  $\mathrm{O}(d)$ . Alors, pour  $\nu$ -presque tout  $x$  dans  $X$ , pour tout  $v$  dans  $\mathbb{R}^d$ , on a*

$$\frac{1}{n} \log \|\sigma_n(x)v\| \xrightarrow[n \rightarrow \infty]{} 0.$$

*Démonstration.* — On pose, pour  $n$  dans  $\mathbb{N}$ ,  $f_n = \log \|\sigma_n\|$ . Alors, on a  $f_{m+n} \leq f_n \circ T^m + f_m$ . D'après le théorème ergodique sous-additif de Kingman [32], pour  $\nu$ -presque tout  $x$  dans  $X$ , la suite  $\frac{1}{n}f_n(x)$  converge vers une limite  $\ell$  indépendante de  $x$ . Soient  $\varphi : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  et  $\theta : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  tels que, pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait  $\sigma(x) = \varphi(Tx)\theta(x)\varphi(x)^{-1}$ . Il vient  $\sigma_n(x) = \varphi(T^n x)\theta_n(x)\varphi(x)^{-1}$  et donc, il existe une suite infinie d'entiers  $n$  le long de laquelle  $f_n(x)$  reste bornée. Il vient  $\ell = 0$ .

En raisonnant de même, on montre que, pour  $\nu$ -presque tout  $x$ ,  $\frac{1}{n} \log \|\sigma_n(x)^{-1}\|$  tend vers 0 quand  $n$  tend vers l'infini. Le résultat en découle.  $\square$

Nous avons alors le

THÉORÈME 4.9 (Osseledets). — *Soit  $\sigma : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  un cocycle intégrable et conservons les notations introduites ci-dessus. Il existe des familles mesurables  $x \mapsto W_i(x) \subset V_i(x)$  définies sur  $X$  de sous-espaces vectoriels de dimension  $d_i - d_{i-1}$ ,  $1 \leq i \leq k$ , telles que, pour  $\nu$ -presque tout  $x$ ,  $W_i(Tx) = \sigma(x)W_i(x)$  et  $V_i(x) = W_i(x) \oplus V_{i-1}(x)$ . Pour tout  $v$  dans  $W_i(x)$ , on a*

$$\frac{1}{n} \log \|\sigma_n(x)v\| \xrightarrow[n \rightarrow \infty]{} \lambda_i \text{ et } \frac{1}{n} \log \|\sigma_n(x)v\| \xrightarrow[n \rightarrow -\infty]{} -\lambda_i.$$

Remarque 4.10. — En particulier, pour tout  $v$  dans  $V_i(x) \setminus V_{i-1}(x)$ , on a

$$\frac{1}{n} \log \|\sigma_n(x)v\| \xrightarrow[n \rightarrow \infty]{} \lambda_i.$$

Remarque 4.11. — Les exposants de Lyapunov du cocycle  $(\sigma \circ T^{-1})^{-1}$  au-dessus de l'automorphisme  $T^{-1}$  sont  $-\lambda_k < -\lambda_{k-1} < \dots < -\lambda_1$  et son drapeau de Lyapunov en  $x \in X$  est

$$W_k(x) \subsetneq W_{k-1}(x) \oplus W_k(x) \subsetneq \dots \subsetneq W_1(x) \oplus \dots \oplus W_k(x) = \mathbb{R}^d.$$

En particulier, les familles  $W_i$  sont uniques.

*Remarque 4.12.* — Le drapeau de Lyapunov ne dépend que du futur de la trajectoire. Plus précisément, soit  $\pi : (X, \nu, T) \rightarrow (Y, \xi, S)$  un facteur de  $(X, \nu, T)$ , c'est-à-dire que  $(Y, \xi)$  est un espace de Lebesgue, que  $\pi$  est une application mesurable qui envoie  $\nu$  sur  $\xi$ , que  $S : Y \rightarrow Y$  est une transformation mesurable qui préserve  $\xi$  et que  $\pi T = S\pi$ . Nous ne supposons pas nécessairement que  $S$  est inversible. Alors, si  $\sigma$  est défini sur  $Y$ , c'est-à-dire si  $\sigma$  est de la forme  $\tau \circ \pi$  avec  $\pi : Y \rightarrow \text{GL}_d(\mathbb{R})$ , les espaces  $V_i$ ,  $1 \leq i \leq k$ , sont aussi définis sur  $Y$ . En revanche, la construction des supplémentaires  $W_i$  fait appel au passé de la dynamique et ceux-ci ne sont pas en général définis sur  $Y$ . Ceci représente souvent une difficulté dans les applications pratiques de ce théorème, comme nous le verrons plus loin.

La démonstration du théorème 4.9 repose essentiellement sur la proposition 4.6 et sur le

LEMME 4.13. — *Soit  $\sigma : X \rightarrow \text{GL}_d(\mathbb{R})$  un cocycle intégrable. On suppose que  $\sigma$  est de la forme*

$$\sigma = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

où  $A$  est de taille  $(r, r)$ ,  $B$  de taille  $(r, d-r)$  et  $C$  de taille  $(d-r, d-r)$ , pour un entier  $1 \leq r \leq d-1$ . On suppose en outre qu'il existe des réels  $\lambda, \mu$  avec  $\lambda + \mu < 0$  et, pour  $\nu$ -presque tout  $x$  dans  $X$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|A(x)^{-1} \cdots A(T^{n-1}x)^{-1}\| \leq \lambda$$

et

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|C(T^{n-1}x) \cdots C(x)\| \leq \mu.$$

Notons  $U = \mathbb{R}^r \times \{0\}$ . Alors, il existe une famille mesurable de sous-espaces vectoriels  $x \mapsto V(x)$  de dimension  $d-r$  de  $\mathbb{R}^d$  telle que, pour  $\nu$ -presque tout  $x$ , on ait  $\mathbb{R}^d = U \oplus V(x)$  et  $\sigma(x)V(x) \subset V(Tx)$ .

*Démonstration.* — Pour  $\nu$ -presque tout  $x$ , notons  $S(x)$  la matrice définie par

$$S(x) = - \sum_{n=0}^{\infty} A(x)^{-1} \cdots A(T^n x)^{-1} B(T^n x) C(T^{n-1} x) \cdots C(x)$$

(où la série converge par hypothèse), de sorte que

$$A(x)S(x) + B(x) = S(Tx)C(x)$$

et donc

$$\begin{pmatrix} A(x) & B(x) \\ 0 & C(x) \end{pmatrix} \begin{pmatrix} S(x) \\ 1 \end{pmatrix} = \begin{pmatrix} S(Tx) \\ 1 \end{pmatrix} C(x).$$

En particulier, si  $p(x)$  désigne l'opérateur de matrice

$$\begin{pmatrix} 0 & S(x) \\ 0 & 1 \end{pmatrix},$$

$\sigma(x)$  envoie l'image de  $p(x)$  dans l'image de  $p(Tx)$ , ce qui achève la démonstration.  $\square$

*Démonstration.* — Quitte à remplacer  $\sigma$  par un cocycle cohomologue, on peut supposer que le drapeau  $V_1 \subsetneq \cdots \subsetneq V_k = \mathbb{R}^d$  est constant. Comme on l'a expliqué plus haut, on peut conserver la propriété d'intégrabilité (7) dans cette transformation.

Pour  $1 \leq i \leq k$ , appliquons maintenant la proposition 4.6 au cocycle induit par  $\sigma$  dans  $V_i/V_{i-1}$ . Quitte à remplacer  $(X, \nu, T)$  par une extension finie  $(\tilde{X}, \tilde{\nu}, \tilde{T})$  et  $\sigma$  par un cocycle cohomologue, on peut supposer que, dans une certaine base de  $V_i/V_{i-1}$ , adaptée à un drapeau

$$V_{i-1} = V_{i0} \subsetneq V_{i1} \subsetneq \cdots \subsetneq V_{i\ell_i} = V_i,$$

pour  $\nu$ -presque tout  $x$ , la matrice de  $\sigma(x)$  est de la forme :

$$A_i(x) = \begin{pmatrix} A_{i1}(x) & * & * & * \\ 0 & A_{i2}(x) & * & * \\ 0 & & 0 \cdots & \vdots \\ 0 & 0 & 0 & A_{i\ell_i}(x) \end{pmatrix},$$

avec  $A_{ij}(x) = e^{\lambda_{ij}(x)} U_{ij}(x)$  où  $\lambda_{ij}(x) \in \mathbb{R}$  et  $U_{ij}(x) \in O(r_{ij})$ ,  $1 \leq j \leq \ell_i$ . On peut en outre supposer que les fonctions  $\lambda_{ij}$  sont intégrables. Alors, d'après le théorème de Birkhoff, on a

$$\frac{1}{n} \log \|A_{ij}(x)^{-1} \cdots A_{ij}(T^{n-1}x)^{-1}\| \xrightarrow{n \rightarrow \infty} - \int_{\tilde{X}} \lambda_{ij} d\tilde{\nu}$$

et

$$\frac{1}{n} \log \|A_{ij}(T^{n-1}x) \cdots A_{ij}(x)\| \xrightarrow{n \rightarrow \infty} \int_{\tilde{X}} \lambda_{ij} d\tilde{\nu},$$

si bien que, quitte à appliquer plusieurs fois le lemme 4.13, on peut supposer qu'on a

$$\int_{\tilde{X}} \lambda_{i1} d\tilde{\nu} \leq \cdots \leq \int_{\tilde{X}} \lambda_{i\ell_i} d\tilde{\nu}.$$

Alors, pour  $1 \leq j \leq \ell_i$ , pour tout vecteur  $v$  de  $V_{ij} \setminus V_{i(j-1)}$ , on a

$$\frac{1}{n} \log \|\sigma_n(x)v\| \xrightarrow{n \rightarrow \infty} \int_{\tilde{X}} \lambda_{ij} d\tilde{\nu}$$

et donc, par définition du drapeau  $(V_i)$ ,

$$\int_{\tilde{X}} \lambda_{i1} d\tilde{\nu} = \cdots = \int_{\tilde{X}} \lambda_{i\ell_i} d\tilde{\nu} = \lambda_i.$$

Il vient, pour tout vecteur  $v$  de  $V_i \setminus V_{i-1}$ ,

$$\frac{1}{n} \log \|\sigma_n(x)v\| \xrightarrow{n \rightarrow \infty} \lambda_i.$$

Enfin, on a maintenant, pour  $\nu$ -presque tout  $x$  dans  $X$ ,

$$\frac{1}{n} \log \|A_i(T^{-1}x) \cdots A_i(T^{-n}x)^{-1}\| \xrightarrow{n \rightarrow \infty} \lambda_i$$

et

$$\frac{1}{n} \log \|A_i(T^{-n}x)^{-1} \cdots A_i(T^{-1}x)^{-1}\| \xrightarrow[n \rightarrow \infty]{} -\lambda_i.$$

On peut alors appliquer plusieurs fois le lemme 4.13 à la transformation  $T^{-1}$  et au cocycle  $(\sigma \circ T^{-1})^{-1}$  pour conclure à l'existence des supplémentaires  $W_i$ .  $\square$

## 5. ACTIONS TRIANGULAIRES

Notons dorénavant

$$A = \left\{ a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} \middle| t \in \mathbb{R} \right\}$$

et

$$N = \left\{ n_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \middle| s \in \mathbb{R} \right\}.$$

Revenons maintenant à la démonstration du théorème 3.4. On dispose d'une action du groupe  $P = AN$  sur un espace probabilisé  $(X, \nu)$ , à savoir la strate  $\mathcal{H}(\alpha)$ , munie d'une mesure de probabilité invariante ergodique. On cherche à montrer que cette mesure est en réalité  $\mathrm{SL}_2(\mathbb{R})$ -invariante et affine. Pour cela, la stratégie consistera, comme dans la théorie de Ratner [29], en utilisant la géométrie de l'espace (soit, essentiellement, la structure plate construite à la section 2), à montrer que la mesure  $\nu$  possède des propriétés d'invariance supplémentaires. Plus précisément, nous allons montrer que les mesures conditionnelles de  $\nu$  le long de feuilletages de  $X$  par des sous-variétés affines ont des propriétés d'invariance par des groupes unipotents de transformations affines des feuilles de plus en plus grands. Ces feuilles sont invariantes par l'action de  $N$  qui y agit par transformations affines unipotentes et le premier des groupes unipotents qui apparaissent est  $N$  lui-même. Par un procédé de dérive, on augmente la taille du groupe unipotent qui préserve les mesures conditionnelles, jusqu'à ce qu'il possède une propriété de maximalité. Alors, on exploite cette propriété de maximalité pour affiner les informations sur les mesures conditionnelles de  $\nu$  (comme dans [23, Sect. 10]). Ces informations impliqueront notamment que la mesure  $\nu$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante.

Dans une seconde étape, on utilise le fait que  $\nu$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante et les propriétés des mesures conditionnelles de  $\nu$  auxquelles il a été fait allusion ci-dessus pour montrer que  $\nu$  est affine.

Revenons pour l'instant à la première étape. On dispose d'une action de  $P = AN$  sur  $X$  qui préserve une mesure  $\nu$ . Comme on l'a déjà évoqué, nous allons tâcher de suivre les orbites de points très proches  $x$  et  $y$  de  $X$  jusqu'à obtenir des points  $x'$  et  $y'$  qui soient à distance macroscopique l'un de l'autre. On aura alors besoin de bien comprendre le vecteur qui, dans une carte affine, permet de passer de  $x'$  à  $y'$ . Pour cela, il nous faut donc avoir une bonne connaissance de l'évolution de certains cocycles linéaires au-dessus de l'action de  $P$  sur  $X$ .

À cet effet, Eskin et Mirzakhani développent dans [11, Sect. 8-10] un certain nombre de concepts nouveaux pour les cocycles au-dessus d'une action de  $P$ . Nous allons tâcher d'en donner ici une présentation succincte.

### 5.1. Partie inerte de la décomposition de Lyapunov

Donnons-nous donc un espace de probabilité de Lebesgue  $(X, \nu)$ , muni d'une action du groupe  $P = AN$  (qui est résoluble, donc moyennable) à stabilisateurs discrets.

Commençons par construire des partitions adaptées à l'étude de l'action. Rappelons qu'une partition  $\xi$  de l'espace de Lebesgue  $X$  en ensembles mesurables est dite être une partition mesurable si l'espace quotient  $\xi \backslash X$ , muni de la mesure image de  $\nu$  est encore un espace de Lebesgue. On identifie les partitions qui diffèrent sur un ensemble de mesure nulle. Étant données deux (classes de) partitions mesurables  $\xi$  et  $\eta$ , on dit que  $\xi$  est plus fine que  $\eta$  et on note  $\eta \prec \xi$  si les atomes de  $\xi$  sont contenus dans ceux de  $\eta$ , c'est-à-dire si  $\xi(x) \subset \eta(x)$  pour  $\nu$ -presque tout  $x$  dans  $X$ .

Si  $\xi$  est une partition mesurable de  $X$ , pour  $t$  dans  $\mathbb{R}$ , on note  $a_t \xi$  la partition  $x \mapsto a_t \xi(a_{-t}x)$ . La partition  $\xi$  est dite invariante par  $A$  si  $\xi \prec a_t \xi$  pour tout  $t \geq 0$ . On dit que  $\xi$  est subordonnée à l'action de  $N$  si, pour  $\nu$ -presque tout  $x$  dans  $X$ , on a  $\xi(x) \subset Nx$  et si l'image inverse de  $\xi(x)$  dans  $N$  par l'application orbitale est un ensemble borné qui contient le point base en son intérieur.

*Remarque 5.1.* — D'ordinaire, on dit qu'une partition mesurable  $\xi$  de l'espace de Lebesgue  $(X, \nu)$  est invariante par l'automorphisme  $T$  si, pour  $\nu$ -presque tout  $x$ , on a  $\xi(x) \subset T^{-1}\xi(Tx)$ . J'adopte ici la convention inverse pour respecter les notations des auteurs.

LEMME 5.2. — *Il existe une partition mesurable  $A$ -invariante subordonnée à l'action de  $N$ .*

*Démonstration.* — Cet énoncé est un avatar de l'existence de partition de Markov pour les dynamiques hyperboliques. Sous cette forme, c'est une généralisation directe de [23, Prop. 9.2] ou [11, Lem. B.1].  $\square$

Nous munirons les atomes d'une telle partition de la mesure de probabilité induite par la mesure de Haar  $N$  que nous noterons  $\nu_x^\xi$ .

La théorie rappelée dans la section 4 et en particulier au paragraphe 4.3 s'applique à l'action du groupe  $P$ . Suivant [11, Sect. 8-10], nous allons appliquer cette théorie aux cocycles tangents apparaissant dans les actions sur les strates de différentielles abéliennes.

Supposons toujours  $(X, \nu)$  muni d'une action de  $P$  et donnons-nous donc un cocycle  $\sigma : P \times X \rightarrow \mathrm{GL}_d(\mathbb{R})$ . La restriction de  $\sigma$  à  $A$  définit un cocycle linéaire  $\sigma_A$  au-dessus du système dynamique  $(X, T, \nu)$ , auquel nous pouvons appliquer la théorie d'Osseledets du paragraphe 4.4. Nous allons étudier comment  $N$  agit sur la décomposition de  $\mathbb{R}^d$  associée à ce cocycle dans le théorème 4.9 (nous supposons dorénavant que le cocycle

$(t, x) \mapsto \sigma(a_t, x)$  est intégrable). Pour  $x$  dans  $X$  et  $y = px$  dans  $Px$ , on notera parfois  $\sigma(x, y)$  pour  $\sigma(p, x)$ .

Pour contrôler localement l'action de  $N$ , nous allons introduire l'hypothèse (qui sera vérifiée pour les systèmes qui nous intéressent) que le cocycle  $\sigma$  ne varie pas sur les feuilles stables locales. En termes plus précis, nous choisissons une partition mesurable  $\xi$  qui soit  $A$ -invariante et subordonnée à  $N$  et nous faisons l'hypothèse que, pour  $\nu$ -presque tout  $x$  dans  $X$ , pour  $\nu_x^\xi$ -presque tout  $y$  de  $\xi(x)$ , on a  $\sigma(x, y) = e$ .

Notons alors  $\lambda_1 > \dots > \lambda_k$  les exposants de Lyapunov de  $\sigma_A$  et, pour  $x$  dans  $X$ ,  $V_1(x) \subset \dots \subset V_k(x)$  le drapeau de Lyapunov de  $\sigma_A$  en  $x$ . Notons que, comme on l'a remarqué au paragraphe 4.4, vu que  $\sigma_A$  est  $\xi$ -mesurable et que  $\xi$  est  $A$ -invariante, les familles  $x \mapsto V_i(x)$  sont  $\xi$ -mesurables.

*Remarque 5.3.* — Toujours pour respecter les conventions des auteurs, j'appelle ici drapeau de Lyapunov le drapeau qu'au paragraphe 4.4 j'ai désigné comme le drapeau de Lyapunov du cocycle  $x \mapsto \sigma(a_{-1}, x)$  au-dessus de la dynamique  $T = a_{-1}$ .

En revanche, soient  $x \mapsto W_i(x)$ ,  $1 \leq i \leq k$ , les familles de sous-espaces vectoriels définies par le théorème 4.9 appliqué à  $\sigma_A$ , de sorte que  $W_i$  est un supplémentaire de  $V_{i-1}$  dans  $V_i$ . Il n'y a aucune raison a priori pour que les applications  $x \mapsto W_i(x)$  soient  $\xi$ -mesurables. Pour  $x$  dans  $X$ , nous définissons donc  $E_i(x)$  comme le plus grand sous-espace vectoriel de  $\mathbb{R}^d$  tel qu'on ait  $\sigma(x, y)E_i(x) \subset W_i(y)$  pour presque tout  $y$  de  $\xi(x)$  (cet espace peut très bien être réduit à  $\{0\}$ , sauf pour  $E_1$ , qui est égal à  $V_1$ ). Nous dirons, comme Eskin et Mirzakhani, que  $E_1(x), \dots, E_k(x)$  sont les espaces de Lyapunov inertes du cocycle  $\sigma$  en  $x$  et que  $E(x) = E_1(x) \oplus \dots \oplus E_k(x)$  est le sous-espace inerte du cocycle.

On vérifie le

LEMME 5.4 ([11, Lem. 8.3]). — *La suite de sous-espaces inertes  $E_1, \dots, E_k$  ne dépend pas du choix de la partition  $\xi$ .*

Énonçons à présent le résultat intermédiaire de [11] qui fait jouer un rôle au sous-espace inerte, dans le cadre de la première étape de la démonstration du théorème 3.4. On dispose d'une action de  $P$  sur  $(X, \nu)$  et on cherche à montrer que les mesures conditionnelles de  $\nu$  le long de certains feuilletages possèdent une propriété d'invariance. Pour ce faire, comme dans la démonstration du théorème de Ratner [29], il nous faudra pouvoir construire, à partir de deux points proches  $x$  et  $y$ , des points  $x'$  et  $y'$  de leurs orbites sous  $P$  qui sont à distance macroscopique l'un de l'autre. On aura en outre besoin de bien connaître, dans une carte affine, le vecteur qui permet de passer de  $x'$  à  $y'$ . La proposition suivante permettra de montrer que ce vecteur a tendance à être proche d'un sous-espace inerte.

Nous conservons les notations introduites dans ce paragraphe.

PROPOSITION 5.5 ([11, Prop. 8.5]). — *Supposons que tous les exposants de Lyapunov du cocycle  $\sigma_A$  soient  $> 0$ . Il existe alors  $\alpha > 0$  tel que, pour tout  $\delta > 0$ , il existe un ensemble mesurable  $K \subset X$ , de mesure  $\geq 1 - \varepsilon$  et  $L_0 > 0$  tels que, pour tous  $x$  dans  $K$ ,  $v$  dans  $\mathbb{R}^d$  et  $L > L_0$ , il existe  $L < t < 2L$  tel que, pour tout  $y$  dans un sous-ensemble de  $\nu_{a_{-t}x}^\xi$ -mesure  $\geq 1 - \delta$  de  $\xi(a_{-t}x)$ , on a*

$$d\left(\frac{\sigma(a_{s(t,y)}, y)v}{\|\sigma(a_{s(t,y)}, y)v\|}, E(a_{s(t,y)}y)\right) \leq e^{-\alpha t},$$

où  $s(t, y)$  est le plus grand réel  $s$  tel que  $\|\sigma(a_s, y)v\| \leq \|v\|$ .

En d'autres termes, on lance la dynamique de  $A$  pendant un temps  $-t$  à partir de  $x$ ; le vecteur  $v$  se déplace en  $\sigma(a_{-t}, x)v$  qui a tendance à être petit puisque les exposants de Lyapunov de  $\sigma_A$  sont  $> 0$ . Puis on perturbe un peu  $a_{-t}x$ , en le déplaçant par un élément  $n$  de  $N$  tel que  $y = nx$  appartienne à  $\xi(a_{-t}x)$  et enfin on relance la dynamique dans l'autre sens pendant un temps  $s$ , jusqu'à ce que le vecteur  $\sigma(a_s n a_{-t}, x)v$  reprenne une taille macroscopique. Alors, pour la plupart des  $n$ , ce vecteur macroscopique est très proche de  $E(x)$ .

## 5.2. Exposants synchronisés

D'après la proposition 5.5, lors des opérations que nous allons effectuer dans les orbites, les vecteurs que nous souhaitons suivre ont tendance à se coucher dans les directions inertes, c'est-à-dire que nous contrôlons bien leur position par rapport à la décomposition d'Osseledets. Nous souhaitons maintenant préciser cette information en étudiant la position de ces vecteurs par rapport aux décompositions issues de la proposition 4.6, appliquée aux espaces  $E_i(x)$ .

Quitte à remplacer le système  $(X, \nu)$  par une extension finie, cette proposition nous fournit, pour  $1 \leq i \leq k$ , un drapeau équivariant

$$\{0\} = E_{i0}(x) \subsetneq E_{i1}(x) \subsetneq \cdots \subsetneq E_{i\ell_i}(x) = E_i(x)$$

et, pour  $1 \leq j \leq \ell_i$ , un cocycle scalaire  $\lambda_{ij} : P \times X \rightarrow \mathbb{R}$  tel que, dans  $E_{ij}(x)/E_{i(j-1)}(x)$ , l'action de  $\sigma$  soit conforme d'exposant  $e^{\lambda_{ij}(x)}$ . Notons que, comme  $\sigma$  est  $\xi$ -mesurable, on peut supposer qu'il en est de même pour  $\lambda_{ij}$ , c'est-à-dire que, pour  $\nu$ -presque tout  $x$  dans  $X$ , pour  $\nu_x^\xi$ -presque tout  $y$  dans  $\xi(x)$ , on a  $\lambda_{ij}(x, y) = 0$ .

Nous supposons dorénavant que, comme dans la proposition 5.5, les exposants de Lyapunov  $\lambda_1 > \cdots > \lambda_k$  du cocycle  $\sigma_A$  sont  $> 0$ , de sorte que, d'après le théorème 4.9, pour tous  $i, j$ , on a

$$\int_X \lambda_{ij}(a_t, x) d\nu(x) = \lambda_i > 0 \quad \text{pour tout } t > 0.$$

Nous allons dans ce paragraphe décrire une relation d'équivalence appelée synchronisation sur les cocycles  $\xi$ -mesurables et intégrables  $\lambda : P \times X \rightarrow \mathbb{R}$  d'intégrale  $> 0$ ,

c'est-à-dire tels que

$$\int_X \lambda(a_t, x) d\nu(x) > 0 \quad \text{pour tout } t > 0.$$

Ces constructions sont effectuées dans [11, Sect. 10]. La motivation pour la définition de cette relation d'équivalence peut être éclairée par l'exemple suivant.

*Exemple 5.6.* — Donnons-nous une mesure de probabilité  $\mu$  sur  $\mathbb{R}$  à support compact et centrée (différente de la masse de Dirac en  $\{0\}$ ). Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires indépendantes et de loi  $\mu$ . On pose  $S_n = X_1 + \dots + X_n$  et on note  $M_n$  la suite de matrices aléatoires

$$M_n = \begin{pmatrix} e^{1+S_n} & 0 \\ 0 & e^{1-S_n} \end{pmatrix}.$$

Pour  $\varepsilon > 0$ , on note  $n(\varepsilon)$  l'entier aléatoire qui est le plus grand entier  $n$  tel que le vecteur  $v_n(\varepsilon) = M_n \begin{pmatrix} \varepsilon \\ \varepsilon \end{pmatrix}$  soit de norme  $\leq 1$ .

On se convainc alors aisément que les deux composantes de  $v_n(\varepsilon)$  ne deviennent pas grandes en même temps, c'est-à-dire qu'avec une probabilité qui tend vers 1 quand  $\varepsilon \rightarrow 0$ , le vecteur  $v_{n(\varepsilon)}(\varepsilon)$  est proche de l'un des axes de coordonnées.

Eskin et Mirzakhani établissent un résultat analogue pour des suites de la forme  $\sigma(a_n, ux)v$  où  $v$  est un petit vecteur et où  $u$  joue le rôle de l'alea dans l'exemple précédent. Ils montrent que, pour une certaine quantité de  $u$ , ces vecteurs se rapprochent en direction d'une réunion de sous-espaces vectoriels de  $E(a_n ux)$ . Ces sous-espaces sont définis en regroupant des sous-espaces des  $E_{ij}(x)$  précisément quand les cocycles  $\lambda_{ij}$  sont synchronisés.

Définissons à présent cette relation. Soit donc  $\lambda : P \times X \rightarrow \mathbb{R}$  un cocycle  $\xi$ -mesurable tel que  $\lambda_A$  soit intégrable et d'intégrale  $> 0$ . Pour  $x$  dans  $X$ , nous posons

$$\mathcal{F}_\lambda[x] = \{y \in Px \mid \lambda(x, y) = 0\}.$$

Nous dirons que deux cocycles  $\lambda$  et  $\mu$  tels que  $\lambda_A$  et  $\mu_A$  soient intégrables et d'intégrale 0 sont synchronisés s'il existe des cocycles  $\lambda'$  et  $\mu'$  cohomologues à  $\lambda$  et  $\mu$  tels que, pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait

$$\mathcal{F}_{\lambda'}[x] = \mathcal{F}_{\mu'}[x].$$

Il ne s'agit pas là de la définition originale de [11, Sect. 10]. Rappelons brièvement celle-ci qui est équivalente.

Soit toujours  $\lambda$  comme ci-dessus. Quitte à remplacer  $\lambda$  par un cocycle cohomologue, on peut supposer qu'il existe  $\varepsilon > 0$  tel que, pour tout  $t > 0$ , pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait  $\lambda(a_t, x) \geq \varepsilon t$ . Un tel cocycle définit un changement de temps du flot  $A$ . Plus précisément, pour  $u$  dans  $\mathbb{R}$  et  $x$  dans  $X$ , notons  $\tau_\lambda(u, x)$  l'unique réel tel que

$$\lambda(a_{\tau_\lambda(u, x)}, x) = u$$

et  $a_u^\lambda x = a_{\tau_\lambda(u,x)}x$ . Alors, la relation de cocycle implique que la famille  $(a_u^\lambda)$  est un flot sur  $X$ . Notons qu'il ne préserve pas nécessairement  $\nu$ , mais qu'il agit de façon absolument continue. Ce flot est la reparamétrisation du flot  $(a_t)$  le long de laquelle le cocycle  $\lambda$  devient simplement la coordonnée de temps.

Comme  $\lambda$  est  $\xi$ -mesurable, pour tout  $u \leq 0$ ,  $\tau_\lambda(u, \cdot)$  est constant sur  $\xi(x)$  et donc la partition  $\xi$  est invariante par le flot  $(a_t^\lambda)$ . Pour  $u \geq 0$  et  $x$  dans  $X$ , définissons  $\mathcal{F}_\lambda[x, u]$  comme l'ensemble des  $y$  dans  $Px$  tels que  $\xi(a_{-u}^\lambda x) = \xi(a_{-u}^\lambda y)$  : on peut le voir comme un morceau de feuille fortement instable pour le flot reparamétré  $(a_t^\lambda)$ . On vérifie que  $\mathcal{F}_\lambda[x, u]$  est l'ensemble des  $y$  dans  $Px$  tels que  $\lambda(x, y) = 0$  et que  $\xi(a_{-t}x) = \xi(a_{-t'}y)$  pour certains  $t$  et  $t'$  tels que  $\lambda(a_{-t}, x) \geq -u$  et  $\lambda(a_{-t'}, y) \geq -u$ . On pose

$$\mathcal{F}_\lambda[x] = \bigcup_{u \geq 0} \mathcal{F}_\lambda[x, u] = \{y \in Px \mid \lambda(x, y) = 0\},$$

qu'on voit comme l'intersection de la feuille fortement instable de  $x$  pour le flot  $(a_u^\lambda)$  et de l'orbite  $Px$  de  $x$ . On munit  $\mathcal{F}_\lambda[x, u]$  de la mesure  $\nu_{x,u}^\lambda$  qui est l'image par  $a_u^\lambda$  de la mesure  $\nu_{a_{-u}^\lambda x}^\xi$ .

Soit toujours  $\mu : P \times X \rightarrow \mathbb{R}$  un autre cocycle  $\xi$ -mesurable tel que  $\mu_A$  soit intégrable et d'intégrale  $> 0$ . Alors  $\lambda$  et  $\mu$  sont synchronisés si, pour  $\nu$ -presque tout  $x$  dans  $X$ , il existe  $C > 0$  et  $\theta > 0$  tel que, pour  $u$  suffisamment grand, on ait  $|\mu(x, y)| \leq c$  pour  $y$  dans un ensemble de  $\nu_{x,u}^\lambda$ -mesure  $\geq \theta$  de  $\mathcal{F}_\lambda[x, u]$ . On peut montrer qu'on peut alors choisir  $\theta$  arbitrairement proche de 1.

*Remarque 5.7.* — L'équivalence entre ces deux définitions peut se voir comme un analogue du fait général suivant : si  $(X, T, \nu)$  est un système dynamique ergodique et si  $f$  est une fonction réelle intégrable sur  $X$ , alors  $f$  est cohomologue à 0 (c'est-à-dire  $f$  s'écrit  $g \circ T - g$  pour une fonction  $g$ ) si et seulement si il existe  $C > 0$  et  $\theta > 0$  tel que, pour  $\nu$ -presque tout  $x$  dans  $X$ , on ait

$$|f(x) + \cdots + f(T^n x)| \leq C$$

pour  $n$  dans un ensemble d'entiers de densité  $\geq \theta$ .

### 5.3. Sous-espace borné

Continuons à suivre les constructions de [11, Sect. 10]. Nous introduisons maintenant une dernière notion abstraite qui nous sera utile dans l'étude des cocycles linéaires au-dessus d'une action ergodique de  $P$ , et plus précisément pour l'étude de ces cocycles dans les composantes inertes  $E_i$ , en lien avec la décomposition de Zimmer.

En effet, cette décomposition produit des cocycles triangulaires à blocs conformes de la forme (6). Supposons pour simplifier que nous avons un bloc de taille  $(p+q) \times (p+q)$

$$\begin{pmatrix} e^\lambda U & B \\ 0 & e^\mu V \end{pmatrix}$$

où  $U$  et  $V$  sont des cocycles à valeurs dans  $O(p)$  et  $O(q)$ ,  $\lambda$  et  $\mu$  sont des cocycles scalaires intégrables d'intégrale  $> 0$  et  $B$  est une fonction matricielle sur  $P \times X$ . Alors,

le comportement de la norme du vecteur  $\sigma(g, x)v$ , où  $g$  est un grand élément de  $P$  et  $v$  est un vecteur de  $\{0\} \times \mathbb{R}^q$ , n'est pas nécessairement contrôlé par le cocycle  $\mu(g, x)$ . Nous allons introduire un sous-espace de  $\{0\} \times \mathbb{R}^q$  où c'est le cas.

Reprenons les notations du paragraphe précédent pour  $\xi, \sigma, \lambda_{ij}$ , etc.

Soient  $1 \leq i \leq k$  et  $1 \leq j \leq \ell_i$ . Pour  $x$  dans  $X$ , on note  $\mathcal{F}_{ij}[x]$  pour  $\mathcal{F}_{\lambda_{ij}}[x]$ . On définit la sous-distribution bornée  $E_{ij,bdd} \subset E_{ij}$  comme la plus grande famille mesurable  $x \mapsto E_{ij,bdd}(x)$  de sous-espaces de  $E_{ij}(x)$  qui soit  $\sigma$ -équivariante (c'est-à-dire telle que  $E_{ij,bdd}(gx) = \sigma(g, x)E_{ij,bdd}(x)$  pour tout  $g$  dans  $P$  et  $\nu$ -presque tout  $x$ ) et dans laquelle  $\sigma$  est cohomologue à un cocycle dont la restriction à  $\mathcal{F}_{ij}$  est à valeurs dans un groupe compact. En d'autres termes, il existe une famille mesurable  $x \mapsto \varphi_{x,ij}$  de produits scalaires sur  $E_{ij,bdd}(x)$  telle que, pour  $\nu$ -presque tout  $x$  dans  $X$ , pour tout  $y$  de  $\mathcal{F}_{ij,bdd}[x]$ , pour tous  $v, w$  dans  $E_{ij,bdd}(x)$ , on a

$$\varphi_{y,ij}(\sigma(x, y)v, \sigma(x, y)w) = \varphi_{x,ij}(v, w)$$

et  $E_{ij,bdd}$  est la plus grande distribution possédant cette propriété.

*Remarque 5.8.* — L'existence de  $E_{ij,bdd}$  peut se voir comme une généralisation du fait suivant : si  $(X, T, \nu)$  est un système dynamique mesuré ergodique et  $\sigma : X \rightarrow \mathrm{GL}_d(\mathbb{R})$  est un cocycle, il existe une plus grande distribution  $\sigma$ -équivariante  $x \mapsto V(x)$  de sous-espaces vectoriels de  $\mathbb{R}^d$  dans laquelle  $\sigma$  préserve un produit scalaire.

*Remarque 5.9.* — À nouveau, il ne s'agit pas exactement de la définition de  $E_{ij,bdd}$  donnée dans [11, Sect. 10]. L'équivalence entre les deux définitions résulte encore d'un raisonnement analogue à celui qui a été décrit dans la remarque 5.7.

Enfin, dans le fibré  $E$ , nous pouvons rassembler les vecteurs pour lesquels le comportement asymptotique de la norme sous l'action du cocycle est essentiellement du même type : pour  $1 \leq i \leq k$  et  $1 \leq j \leq \ell_i$ , nous définissons  $E_{[ij],bdd}$  comme la somme directe

$$\bigoplus_{pq \sim ij} E_{pq,bdd}$$

où  $pq$  parcourt l'ensemble des paires telles que  $\lambda_{pq}$  et  $\lambda_{ij}$  soit synchronisées. Nous pouvons alors énoncer le résultat de [11] qui caractérise dans notre situation l'analogie du phénomène décrit dans l'exemple 5.6.

**PROPOSITION 5.10** ([11, Prop. 10.1]). — *Supposons que tous les exposants de Lyapunov du cocycle  $\sigma_A$  soient  $> 0$ . Alors, il existe  $\theta > 0$  et, pour tout  $\delta > 0$ , il existe un ensemble mesurable  $K \subset X$ , de mesure  $\geq 1 - \varepsilon$  et  $L_0 > 0$  tels que, pour tous  $x$  dans  $K$ ,  $v$  dans  $E(x)$  et  $t > L_0$ , pour tout  $y$  dans un sous-ensemble de  $\nu_{a_{-t}x}^\xi$ -mesure  $\geq \theta$  de  $\xi(a_{-t}x)$ , on a*

$$d \left( \frac{\sigma(a_{s(t,y)}, y)v}{\|\sigma(a_{s(t,y)}, y)v\|}, \bigcup_{ij} E_{[ij],bdd}(a_{s(t,y)}y) \right) \leq \delta,$$

où  $s(t, y)$  est le plus grand réel  $s$  tel que  $\|\sigma(a_s, y)v\| \leq \|v\|$ .

Ainsi, la proposition 5.5 a permis de montrer que les vecteurs qui apparaissent dans les constructions appartiennent au sous-espace inerte  $E$ . La proposition 5.10 permet de montrer qu'une partie d'entre eux appartient à l'un des  $E_{[ij],bdd}$ .

#### 5.4. Actions triangulaires

Les objets que nous venons de présenter sont ceux qui apparaissent au début du procédé inductif qui permet d'établir la première étape de la démonstration du théorème 3.4. Lors de ce procédé inductif, comme nous l'avons déjà évoqué, on améliore les propriétés d'invariance de la mesure  $\nu$  le long de certains feuilletages. Au cours de la récurrence, on est alors amené à remplacer les orbites du groupe  $N$  par des sous-variétés de dimension supérieure de la strate  $\mathcal{H}(\alpha)$ . Ces sous-variétés sont naturellement paramétrées par des espaces homogènes de groupes de Lie nilpotents simplement connexes. Le fait que la mesure  $\nu$  était  $N$ -invariante est alors remplacé par le fait que la mesure conditionnelle de  $\nu$  le long d'une de ces sous-variétés est la mesure de Haar de l'espace homogène qui les paramètre.

Décrivons plus précisément cette structure. Étant donné un espace de Lebesgue  $(X, \nu)$ , nous appellerons action triangulaire sur  $X$  (j'invente la terminologie) la donnée

- (i) d'une action mesurable qui préserve  $\nu$  du groupe  $A = \{a_t \mid t \in \mathbb{R}\}$  sur  $X$ ,
- (ii) d'un groupe de Lie nilpotent simplement connexe  $U$ , d'algèbre de Lie  $\mathfrak{u}$ ,
- (iii) d'un cocycle  $\theta : A \times X \rightarrow \text{Aut}(U)$ ,
- (iv) d'une famille mesurable  $\theta$ -équivariante  $x \mapsto U_x$  de sous-groupes fermés connexes de  $U$ ,
- (v) d'une application mesurable  $\pi : U \times X \rightarrow X$  (qui n'est pas nécessairement une action de  $U$  sur  $X$ ),

tels que, pour  $\nu$ -presque tout  $x$  dans  $X$ , pour tous  $t$  dans  $\mathbb{R}$ ,  $u$  dans  $U$  et  $v$  dans  $U_x$ ,

$$\begin{aligned} a_t \pi(u, x) &= \pi(\theta_{t,x}(u), x) \\ \pi(e, x) &= x \\ \pi(uv, x) &= \pi(u, x) \\ \pi(U, \pi(u, x)) &= \pi(U, x) \end{aligned}$$

et que

- (i) pour  $\nu$ -presque tout  $x$  dans  $X$ ,  $\pi$  induit une injection  $U/U_x \rightarrow X$  et la mesure conditionnelle de  $\nu$  le long de  $\pi(U, x) \simeq U/U_x$  s'identifie à la mesure  $U$ -invariante de  $U/U_x$ ;
- (ii) les exposants de Lyapunov du cocycle  $\theta$ , vu comme un cocycle à valeurs dans le groupe  $\text{Aut}(\mathfrak{u}) \subset \text{GL}(\mathfrak{u})$ , soient tous  $> 0$ .

La propriété concernant les mesures conditionnelles fait sens car les axiomes impliquent que les ensembles  $\pi(U, x)$  sont les classes d'une relation d'équivalence mesurée sur  $X$ .

Cette structure peut se comprendre comme suit : on dispose d’une partition de  $X$  en ensembles qui sont paramétrés par des espaces homogènes du groupe de Lie nilpotent  $U$ . Ces ensembles ne sont a priori pas les orbites d’une action de  $U$ , mais ils sont dilatés par l’action de  $(a_t)$ . Les mesures conditionnelles de  $\nu$  dans cette partition se lisent comme des mesures de Haar. En particulier, plus le groupe  $U$  est grand, plus la mesure  $\nu$  possède d’invariance (dans certains feuilletages).

Une action du groupe  $P$  définit naturellement une telle structure.

*Remarque 5.11.* — Les ensembles que j’ai notés  $\pi(U, x)$  sont notés  $U^+[x]$  dans [11, Sect. 3-12] et appelés espaces affines généralisés (generalized affine subspaces).

*Exemple 5.12.* — Donnons un exemple classique où une telle structure apparaît naturellement. Soient  $G = \mathrm{SL}_2(\mathbb{C})$ ,  $M$  le groupe des matrices diagonales à coefficients de module 1 dans  $G$ ,  $U$  le groupe des matrices unipotentes triangulaires supérieures dans  $G$  et  $\Gamma$  un réseau de  $G$ . Nous considérons  $A$  comme un sous-groupe de  $G$ .

Posons  $X = M \backslash G / \Gamma$  et munissons-le de la mesure de probabilité  $\nu$  qui est l’image de la mesure de probabilité  $G$ -invariante de  $Y = G / \Gamma$ . L’espace  $Y$  est muni d’une action de  $AU$  dans laquelle  $A$  dilate donc les orbites de  $U$ . Comme  $M$  normalise  $U$ , la partition en  $U$ -orbites de  $Y$  induit une partition de  $X$ . Le choix d’une section mesurable de la projection naturelle  $Y \rightarrow X$  donne un paramétrage par  $U$  des atomes de cette partition. En revanche, ceux-ci ne peuvent pas s’écrire comme les orbites d’une action de  $U$  sur  $X$ .

L’ensemble des concepts introduits dans cette partie pour les actions de  $P = AN$  (partitions invariantes, cocycles, décomposition d’Osseledets, réduction de Zimmer, cocycles synchronisés, sous-espaces bornés) peuvent être développés dans le cadre des actions triangulaires : c’est l’objet de [11, Sect. 3-4, 8-10]. Au cours du procédé de récurrence qui permet de démontrer [11, Th. 2.1], on remplace le groupe  $U$  par un sur-groupe. Nous allons à présent tâcher de donner plus d’indications sur cette récurrence.

## 6. LA DÉRIVE D’ESKIN-MIRZAKHANI

Nous quittons à présent le cadre abstrait de la section 5 et nous revenons à l’étude de l’action de  $P$  et de  $\mathrm{SL}_2(\mathbb{R})$  sur la strate  $X = \mathcal{H}_1(\alpha)$ , munie de sa structure plate modelée sur  $H^1(S, \Sigma, \mathbb{R}^2)$ , introduite à la section 2.

### 6.1. Théorème de Forni

Un point essentiel des propriétés dynamiques de ces actions réside dans la possibilité de définir un feuilletage fortement instable pour l’action de  $A$  uniquement en termes de cette structure plate. Cela découle d’un résultat remarquable de Forni, que nous allons énoncer.

Rappelons que la structure plate de  $\mathcal{H}(\alpha)$  permet de construire un certain nombre de fibrés vectoriels  $\mathrm{GL}_2^+(\mathbb{R})$ -équivariants sur  $\mathcal{H}_1(\alpha)$ , comme cela a été expliqué au paragraphe 2.4, et notamment le fibré de Hodge de fibre  $H^1(S, \mathbb{R})$  et son sous-fibré de codimension 2,  $p(H_\perp^1)$ .

Étant donné un sous-groupe  $H$  de  $\mathrm{SL}_2(\mathbb{R})$  et une mesure de probabilité  $H$ -invariante  $\nu$  sur  $\mathcal{H}_1(\alpha)$ , le choix d'une trivialisatation mesurable du fibré de Hodge définit un cocycle  $H \times X \rightarrow \mathrm{GL}(H^1(S, \mathbb{R}))$ . Par construction, ce cocycle prend ses valeurs dans le groupe symplectique de la forme symplectique  $\varpi$  induite par le cup-produit (voir le paragraphe 2.2). Le choix d'une autre trivialisatation produit un cocycle cohomologue. Ce cocycle (ou plutôt cette classe de cohomologie) est appelé cocycle de Kontsevich-Zorich dans la littérature et a été introduit dans [18].

Le théorème de Forni décrit ses exposants de Lyapunov :

**THÉORÈME 6.1** (Forni, [13]). — *Soit  $\nu$  une mesure de probabilité borélienne  $A$ -invariante et ergodique sur  $\mathcal{H}_1(\alpha)$ . Les exposants de Lyapunov du cocycle de Kontsevich-Zorich par rapport à  $\nu$  sont de la forme*

$$\lambda_1 = 1 > \lambda_2 \geq \dots \geq \lambda_g \geq 0 \geq -\lambda_g \geq \dots \geq -\lambda_2 > -\lambda_1 = -1.$$

*Remarque 6.2.* — Dans cette formule, on compte les exposants de Lyapunov avec leur multiplicité.

Les propriétés de symétrie de la suite des exposants de Lyapunov découlent directement du fait que le cocycle préserve une forme symplectique. L'information importante donnée par le théorème est que le premier exposant est simple, c'est-à-dire que  $\lambda_2 < 1$ . Elle peut se reformuler ainsi : tous les exposants de Lyapunov de l'action de  $A$  dans le fibré  $p(H_\perp^1)$  sont  $< 1$ .

## 6.2. Géométrie du feuilletage instable

Nous allons décrire ici un feuilletage de  $\mathcal{H}_1(\alpha)$  en sous-variétés affines de  $\mathcal{H}(\alpha)$ . Ce feuilletage est  $A$ -équivariant et le théorème de Forni permet de le voir comme le feuilletage fortement instable pour l'action de  $A$ .

Rappelons qu'au paragraphe 2.4, nous avons décomposé le fibré vectoriel  $\mathcal{H}(\alpha) \times_\Gamma H^1(S, \Sigma, \mathbb{R})$  en la somme d'un fibré qui s'identifie de façon  $\mathrm{GL}_2^+(\mathbb{R})$ -équivariante à  $\mathcal{H}(\alpha) \times \mathbb{R}^2$  et du fibré  $H_\perp^1$ . Comme le fibré tangent de  $\mathcal{H}(\alpha)$  est le produit tensoriel de  $\mathbb{R}^2$  avec  $\mathcal{H}(\alpha) \times_\Gamma H^1(S, \Sigma, \mathbb{R})$ , ceci donne une décomposition  $\mathrm{GL}_2^+(\mathbb{R})$ -équivariante

$$(8) \quad T\mathcal{H}(\alpha) \simeq (\mathcal{H}(\alpha) \times (\mathbb{R}^2 \otimes \mathbb{R}^2)) \oplus (\mathbb{R}^2 \otimes H_\perp^1).$$

Notons  $e_1 = (1, 0)$  et  $e_2 = (0, 1)$  la base canonique de  $\mathbb{R}^2$ . On vérifie que, dans la décomposition (8), le fibré tangent de  $\mathcal{H}_1(\alpha)$  est

$$(9) \quad T\mathcal{H}_1(\alpha) \simeq (\mathcal{H}_1(\alpha) \times (\mathbb{R}e_1 \otimes e_1 \oplus \mathbb{R}(e_1 \otimes e_2 + e_2 \otimes e_1) \oplus \mathbb{R}e_2 \otimes e_2)) \oplus (\mathbb{R}^2 \otimes H_\perp^1).$$

En particulier, introduisons les sous-fibrés

$$(10) \quad W^+ = (\mathcal{H}_1(\alpha) \times (\mathbb{R}e_1 \otimes e_1)) \oplus (\mathbb{R}e_1 \otimes \mathbb{H}_\perp^1)$$

$$(11) \quad \text{et } W^- = (\mathcal{H}_1(\alpha) \times (\mathbb{R}e_2 \otimes e_2)) \oplus (\mathbb{R}e_2 \otimes \mathbb{H}_\perp^1).$$

Alors ces sous-fibrés sont  $A$ -équivariants. La droite tangente à l'action de  $N$  est la droite  $\mathcal{H}_1(\alpha) \times (\mathbb{R}e_1 \otimes e_1) \subset W^+$ . Enfin, d'après (9), on a la décomposition

$$(12) \quad T\mathcal{H}_1(\alpha) = (\mathcal{H}_1(\alpha) \times \mathbb{R}(e_1 \otimes e_2 + e_2 \otimes e_1)) \oplus W^+ \oplus W^-,$$

dans laquelle la composante de dimension 1 est la droite tangente à l'action de  $A$ .

Si l'action de  $A$  dans  $W^+$  (resp.  $W^-$ ) était uniformément dilatante, (12) impliquerait que le flot  $A$  est un flot d'Anosov (à ceci près que  $\mathcal{H}_1(\alpha)$  n'est pas compact). Le théorème de Forni se traduit directement en une version faible de ces propriétés de dilatation :

**COROLLAIRE 6.3.** — *Soit  $\nu$  une mesure de probabilité borélienne  $A$ -invariante et ergodique sur  $\mathcal{H}_1(\alpha)$ . Les exposants de Lyapunov de  $A$  dans  $W^+$  (resp.  $W^-$ ) par rapport à  $\nu$  sont tous  $> 0$  (resp.  $< 0$ ).*

*Démonstration.* — Rappelons que nous disposons d'une suite exacte

$$0 \rightarrow \mathrm{H}^0(\Sigma, \mathbb{R})/\mathrm{H}^0(M, \mathbb{R}) \rightarrow \mathrm{H}^1(S, \Sigma, \mathbb{R}) \xrightarrow{p} \mathrm{H}^1(S, \mathbb{R}) \rightarrow 0$$

grâce à laquelle nous pouvons construire une suite exacte de fibrés vectoriels sur  $\mathcal{H}(\alpha)$ . Le théorème de Forni implique que tous les exposants de Lyapunov du fibré

$$p(\mathrm{H}_\perp^1) \subset \mathcal{H}(\alpha) \times_\Gamma \mathrm{H}^1(S, \mathbb{R})$$

sont de module  $< 1$ . Il en va de même pour les exposants de Lyapunov du fibré vectoriel  $\mathcal{H}(\alpha) \times_\Gamma \mathrm{H}^0(\Sigma, \mathbb{R})$ . En effet, l'espace  $\mathrm{H}^0(\Sigma, \mathbb{R})$  s'identifie à l'espace vectoriel libre engendré par  $\Sigma$ . Dans cet espace, le groupe modulaire  $\Gamma_{\alpha, \Sigma}$  agit par permutation des éléments de la base, et donc préserve une métrique. Par construction, le fibré plat  $\mathcal{H}(\alpha) \times_{\Gamma_{\alpha, \Sigma}} \mathrm{H}^0(\Sigma, \mathbb{R})$  possède donc une métrique  $\mathrm{SL}_2(\mathbb{R})$ -invariante et en particulier, les exposants de Lyapunov de  $A$  y sont tous nuls. Par définition de  $W^+$ , la suite des exposants de Lyapunov de  $A$ , comptés avec multiplicité, y est donc de la forme

$$2 > 1 + \lambda_2 \geq \dots \geq 1 + \lambda_g \geq 1 \dots \geq 1 \geq 1 - \lambda_g \geq \dots \geq 1 - \lambda_2$$

(où 1 apparaît  $|\Sigma| - 1 = n - 1$  fois) et ces nombres sont tous  $> 0$ . □

### 6.3. $\mathrm{SL}_2(\mathbb{R})$ -invariance et mesures conditionnelles

Nous allons à présent énoncer le théorème 2.1 de [11] que nous avons jusqu'à présent désigné comme la première étape de la démonstration du théorème principal 3.4.

Dans ce but, procédons encore à quelques constructions géométriques en termes du fibré vectoriel  $W^+$  (resp.  $W^-$ ). On vérifie aisément en utilisant les cartes de la structure affine plate de  $\mathcal{H}(\alpha)$  que la distribution  $W^+$  est intégrable. Pour  $x$  dans  $\mathcal{H}(\alpha)$ , on note  $W^+[x]$  la feuille associée; c'est une sous-variété affine de direction  $W^+(x)$  qui contient la droite affine  $Nx$ .

Étant donnée une mesure de probabilité  $\nu$  sur  $X$ , on peut alors construire une famille mesurable  $x \mapsto \nu_{W^+}(x)$  où, pour  $\nu$ -presque tout  $x$ ,  $\nu_{W^+}(x)$  est une mesure de Radon sur la feuille  $W^+[x]$  (munie de sa topologie de feuille), telle que pour toute partition mesurable  $\xi$  de  $(X, \nu)$  subordonnée au feuilletage  $W^+$ , pour  $\nu$ -presque tout  $x$  dans  $X$ , la mesure conditionnelle de  $\nu$  sur l'atome  $\xi(x)$  est proportionnelle à la restriction de  $\nu_{W^+}(x)$  à  $\xi(x)$ . Cette famille est unique, à multiplication par une fonction  $> 0$  près. Par abus de langage, on appelle encore  $\nu_{W^+}(x)$  la mesure conditionnelle de  $\nu$  le long de  $W^+[x]$ .

**THÉORÈME 6.4** ([11, Th. 2.1]). — *Soit  $\nu$  une mesure de probabilité  $P$ -invariante et ergodique sur  $X = \mathcal{H}_1(\alpha)$ . Alors  $\nu$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante et il existe une distribution mesurable définie  $\nu$ -presque partout et  $\mathrm{SL}_2(\mathbb{R})$ -équivariante  $x \mapsto \mathcal{L}(x)$  de sous-espaces du fibré  $\mathcal{H}(\alpha) \times_{\Gamma_{\alpha, \Sigma}} \mathrm{H}^1(S, \Sigma, \mathbb{R})$  telle que, pour  $\nu$ -presque tout  $x$  dans  $X$ , la mesure conditionnelle  $\nu_{W^+}(x)$  est une mesure de Lebesgue de l'espace affine*

$$x + (\mathbb{R}^2 \otimes \mathcal{L}(x) \cap W^+(x)) \subset W^+[x].$$

Notons que, vu que la mesure  $\mathrm{SL}_2(\mathbb{R})$  et la distribution  $\mathcal{L}$  sont  $\mathrm{SL}_2(\mathbb{R})$ -équivariantes, le théorème implique la propriété symétrique pour  $\nu$ -presque tout  $x$  dans  $X$ , et la mesure conditionnelle  $\nu_{W^-}(x)$  est une mesure de Lebesgue de l'espace affine

$$x + (\mathbb{R}^2 \otimes \mathcal{L}(x) \cap W^-(x)) \subset W^-[x].$$

En réalité, au cours de la démonstration, on commence par établir cette propriété de symétrie puis on en déduit le théorème.

Plus précisément, comme nous l'avons déjà mentionné plus haut, la démonstration du théorème 6.4 repose sur un argument de récurrence dans lequel on améliore peu à peu les propriétés d'invariance de la mesure  $\nu$  le long du feuilletage  $W^+$ . Cet argument est contenu dans [11, Prop. 12.1]. Une fois que ces propriétés d'invariance sont devenues maximales en un certain sens, on utilise cette propriété de maximalité pour conclure que les mesures conditionnelles de  $\nu$  le long de  $W^-$  ont les mêmes propriétés d'invariance, par un raisonnement entropique directement inspiré de [23, Sect. 10].

#### 6.4. Invariance par des groupes unipotents

Dans ce paragraphe, nous allons énoncer précisément [11, Prop. 12.1].

Commençons par détailler les propriétés de symétrie qui existent entre les feuilletages  $W^+$  et  $W^-$ . Fixons une mesure  $A$ -invariante et ergodique  $\nu$  sur  $X$ .

Soit  $\mathcal{L}^- \subset W^-$  une distribution  $A$ -équivariante de sous-espaces vectoriels, définie  $\nu$ -presque partout. Alors, comme

$$W^- = (\mathcal{H}_1(\alpha) \times (\mathbb{R}e_2 \otimes e_2)) \oplus (\mathbb{R}e_2 \otimes \mathrm{H}_\perp^1)$$

et que, d'après le théorème 6.1, les exposants de Lyapunov de  $A$  dans  $\mathbb{R}e_2 \otimes \mathrm{H}_\perp^1$  sont tous  $> -2$ , on a

$$\mathcal{L}^- \subset \mathbb{R}e_2 \otimes \mathrm{H}_\perp^1 \text{ ou } \mathcal{L}^- = (\mathcal{H}_1(\alpha) \times (\mathbb{R}e_2 \otimes e_2)) \oplus (\mathcal{L}^- \cap \mathbb{R}e_2 \otimes \mathrm{H}_\perp^1).$$

Dans le premier cas, on écrit  $\mathcal{L}^- = \mathbb{R}e_2 \otimes \mathcal{L}_\perp$  et on pose  $\mathcal{L}^+ = \mathbb{R}e_1 \otimes \mathcal{L}_\perp$ . Dans le second cas, on écrit

$$\mathcal{L}^- = (\mathcal{H}_1(\alpha) \times (\mathbb{R}e_2 \otimes e_2)) \oplus (\mathbb{R}e_2 \otimes \mathcal{L}_\perp)$$

et on pose

$$\mathcal{L}^+ = (\mathcal{H}_1(\alpha) \times (\mathbb{R}e_1 \otimes e_1)) \oplus (\mathbb{R}e_1 \otimes \mathcal{L}_\perp).$$

Cette correspondance établit une bijection entre les distributions  $A$ -équivariantes de  $W^+$  et celles de  $W^-$ , qui respecte la symétrie apparaissant dans (10).

**PROPOSITION 6.5** ([11, Prop. 12.1]). — *Soient  $\nu$  une mesure  $P$ -invariante et ergodique sur  $X$ . Pour  $\nu$ -presque tout  $x$  dans  $X$ , soit  $\mathcal{L}^-(x)$  la direction du sous-espace affine de  $W^-[x]$  engendré par le support de  $\nu_{W^-}(x)$ , c'est-à-dire le plus petit sous-espace vectoriel de  $W^-(x)$  tel que  $\nu_{W^-}$  soit supportée par  $x + \mathcal{L}^-(x)$ . Soit  $\mathcal{L}^+ \subset W^+$  comme ci-dessus. Alors, pour  $\nu$ -presque tout  $x$  dans  $X$ , la mesure  $\nu_{W^+}(x)$  est invariante par les translations par les éléments de  $\mathcal{L}^+(x)$ .*

Cet énoncé n'est pas exactement celui de [11, Prop. 12.1], mais il en est déduit au début de [11, Sect. 13]. Il est l'analogie d'un énoncé apparaissant dans la démonstration du théorème de Ratner par Margulis et Tomanov [23, Cor. 8.4]. On en déduit dans [11, Sect. 13] que  $\nu_{W^+}(x)$  et  $\nu_{W^-}(x)$  sont les mesures de Lebesgue des espaces affines  $x + \mathcal{L}^+(x)$  et  $x + \mathcal{L}^-(x)$  par un raisonnement entropique analogue à celui de [23, Sect. 10].

*Remarque 6.6.* — Le fait que  $\nu$  est  $N$ -invariante implique que l'entropie de  $a_1$  par rapport à  $\nu$  est  $> 0$ . On en déduit que, nécessairement, les conditionnelles  $\nu_{W^-}$  ne sont pas des masses de Dirac, et donc que  $\mathcal{L}^-$  est de dimension  $> 0$ . Cet argument entropique est une version simplifiée de celui de [11, Sect. 13].

Décrivons maintenant la structure de la démonstration de la proposition 6.5. Il s'agit de montrer que, si le support de la mesure  $\nu_{W^-}(x)$  est trop gros, la mesure  $\nu_{W^+}(x)$  possède des propriétés d'invariance. L'idéal serait de montrer que  $\nu_{W^+}(x)$  est invariante par un groupe de translations de l'espace affine  $W^+[x]$ . Malheureusement, on n'y arrive pas directement, mais on commence par montrer que cette mesure possède des propriétés d'invariance sous l'action de groupes unipotents de transformations affines de  $W^+[x]$ .

Expliquons précisément quelles transformations affines peuvent apparaître au cours de cette construction. Notons  $\lambda_1 = 2 > \lambda_2 > \dots > \lambda_k > 0$  les exposants de Lyapunov de l'action de  $A$  dans  $W^+$  par rapport à  $\nu$  (comptés sans multiplicité) et

$$\{0\} = W_0 \subsetneq W_1 \subsetneq W_2 \subsetneq \dots \subsetneq W_k = W^+$$

le drapeau de Lyapunov associé. Notons que  $V_1 = \mathbb{R}e_1 \otimes e_1$  et que, plus généralement, pour  $i \geq 1$ ,  $W_i = \mathbb{R}e_1 \otimes e_1 \oplus W_{i,\perp}$  pour un  $W_{i,\perp} \subset H_\perp^1$ . Comme Eskin et Mirzakhani dans [11, Sect. 6], nous notons, pour  $\nu$ -presque tout  $x$  de  $X$ ,  $Q_+(x)$  le groupe des éléments  $g$  de  $\mathrm{GL}(W^+(x))$  qui préservent le drapeau de Lyapunov de  $x$  et tels que, pour  $1 \leq i \leq k$ , pour tout  $v$  dans  $W_i(x)$ ,  $gv \in v + W_{i-1}(x)$  (dans le langage de la théorie des groupes algébriques,  $Q_+(x)$  est le radical unipotent du groupe parabolique

qui stabilise le drapeau de Lyapunov). Nous notons  $\mathcal{G}_+(x)$  le groupe des automorphismes affines de  $W^+[x]$  dont la partie linéaire appartient à  $Q_+(x)$ . Notons que, comme  $\nu$  est  $N$ -invariante, on a  $N \subset \mathcal{G}_+(x)$  pour  $\nu$ -presque tout  $x$  (ou abusivement, on a identifié  $N$  au groupe d'automorphismes qu'il induit sur  $W^+[x]$ ). On considérera toujours qu'on a  $W^+(x) \subset \mathcal{G}_+(x)$ , en identifiant les vecteurs aux translations associées.

Avant de montrer que la mesure  $\nu_{W^+}(x)$  est une mesure de Lebesgue, on montre qu'elle possède des propriétés d'invariance par des sous-groupes de  $\mathcal{G}_+(x)$ . Plus précisément, on construit une partition mesurable  $\eta_{W^+[x]}$  de la feuille  $W^+[x]$  et un sous-groupe fermé connexe  $U^+(x)$  de  $\mathcal{G}_+(x)$  (je ne sais pas pourquoi le  $+$  passe de bas en haut, mais je tâche de respecter les conventions des auteurs) tels que, pour  $\nu$ -presque tout  $x$ , la mesure conditionnelle  $\nu_{W^+, \eta}(x)$  de  $\nu_{W^+}(x)$  le long de  $\eta_{W^+(x)}$  est  $U^+(x)$ -invariante. La famille  $\eta$  de partitions mesurables sur les feuilles  $W^+[x]$  est équivariante sous  $A$  (au sens où  $a_t \eta_{W^+[x]}(x) = \eta_{W^+[a_t x]}(a_t x)$ , pour  $t$  dans  $\mathbb{R}$  et  $\nu$ -presque tout  $x$  dans  $X$ ), si bien qu'on peut supposer que la famille de groupes  $x \mapsto U^+(x)$  est  $A$ -équivariante.

La proposition 6.5 se déduit de la

**PROPOSITION 6.7** ([11, Prop. 12.1]). — *Soient  $\nu$  une mesure  $P$ -invariante et ergodique sur  $X$ . Supposons données une famille mesurable  $A$ -équivariante  $\eta_{W^+[x]}$  de partitions mesurables sur les feuilles  $W^+[x]$  et une famille mesurable  $A$ -équivariante  $x \mapsto U^+(x) \subset \mathcal{G}_+(x)$ , telles que, pour  $\nu$ -presque tout  $x$  de  $X$ , on ait  $N \subset U^+(x)$  et que la mesure conditionnelle  $\nu_{W^+, \eta}(x)$  de  $\nu_{W^+}(x)$  le long de  $\eta_{W^+(x)}$  est  $U^+(x)$ -invariante. Soient  $\mathcal{L}^-(x) \subset W^-(x)$  l'espace vectoriel direction du sous-espace affine de  $W^-[x]$  engendré par le support de  $\nu_{W^-}(x)$  et  $\mathcal{L}^+(x) \subset W^+(x)$  le sous-espace vectoriel associé par la symétrie de (10). Alors, si  $\mathcal{L}^+(x)$  n'est pas presque sûrement contenu dans  $U^+(x)$ , il existe des familles  $\eta_{new}$  et  $U_{new}^+$  ayant les mêmes propriétés et telles que, pour  $\nu$ -presque tout  $x$ ,  $U^+(x)$  soit un sous-groupe propre de  $U_{new}^+(x)$ .*

C'est la présence de ces groupes nilpotents  $U^+$  qui rend nécessaire l'introduction des actions triangulaires dont nous avons fait mention au paragraphe 5.4. En effet, l'équivariance de la famille  $U^+$  sous l'action de  $A$  implique, par des arguments analogues à ceux employés au paragraphe 4.2 (le quotient d'une variété algébrique par l'action d'un groupe algébrique est à base dénombrable), que les groupes  $U^+(x)$  sont tous conjugués à un même groupe  $U$ . La propriété que le cocycle  $\theta$  du paragraphe 5.4 a tous ses exposants de Lyapunov strictement positifs se déduit de ce que  $U^+$  est une sous-distribution de  $\mathcal{G}_+$  dans laquelle on sait par construction que tous les exposants de Lyapunov de  $A$  sont strictement positifs.

## 6.5. La dérive

Nous allons dans ce paragraphe tenter de présenter l'argument de dérive qui permet d'augmenter la taille du groupe  $U^+$  dans la proposition 6.7. Il repose sur un principe général dont l'introduction me semble due à Katok et Spatzier [17], qui a été beaucoup utilisé dans des travaux récents sur les systèmes dynamiques dans les espaces homogènes [7, 22, 8, 4] et que nous allons expliquer.

Supposons donnée une action d'un groupe  $G$  sur un espace de probabilité de Lebesgue  $(X, \nu)$ . On donne une partition  $x \mapsto W[x]$  de  $X$  en ensembles mesurables, chacun de ces ensembles étant muni d'une structure géométrique (dans le cas des espaces homogènes, ces ensembles sont les orbites d'une action d'un groupe de Lie connexe  $H$ ; dans le cas des strates, ce sont des espaces affines). On suppose que, pour tout  $g$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ , on a  $W[gx] = gW[x]$ , c'est-à-dire que  $G$  permute les atomes de la partition, et on suppose aussi que  $g$  préserve la structure géométrique de ces ensembles (dans le cas des espaces homogènes,  $G$  normalise le groupe  $H$ ; dans le cas des strates,  $G = AN$  agit par un cocycle à coefficients dans le groupe affine).

À cette donnée, on peut associer une famille  $x \mapsto \nu_W(x)$  de mesures de Radon sur  $W[x]$  (mesures de Radon au sens de la topologie intrinsèque de  $W[x]$ , topologie sous-jacente à la structure géométrique). Les  $\nu_W(x)$  sont définies par la propriété que, si  $\xi$  est une partition mesurable subordonnée à la partition  $W$  (c'est-à-dire que, pour  $\nu$ -presque tout  $x$ ,  $\xi(x)$  est un ensemble relativement compact de  $W^+[x]$  qui contient  $x$  en son intérieur), pour  $\nu$ -presque tout  $x$  dans  $X$ , la mesure conditionnelle de  $\nu$  sur  $\xi(x)$  est proportionnelle à la restriction de  $\nu_W(x)$ . Notons que les propriétés d'équivariance de l'action de  $G$  impliquent que, pour  $g$  dans  $G$ , pour  $\nu$ -presque tout  $x$  dans  $X$ ,  $\nu_W(gx)$  est proportionnelle à  $g_*\nu_W(x)$ .

On cherche à montrer que  $\nu_W$  a des propriétés d'invariance reliées à la structure géométrique de  $W$  (dans le cas des espaces homogènes,  $\nu_W(x)$  est invariante par un sous-groupe de  $H$ ; dans le cas des strates, par un sous-groupe du groupe affine). Pour cela, la démarche est la suivante. On commence par construire, en utilisant le théorème de Lusin, un sous-ensemble compact  $Y$  de  $X$  (s'il n'y a pas de topologie, on en rajoute une qui soit compatible avec la structure borélienne) où l'application  $x \mapsto \nu_W(x)$  est continue. On cherche à construire des suites  $(x_n)$  et  $(y_n)$  dans  $Y$  et  $(g_n)$  dans  $G$  telles que  $d(x_n, y_n) \rightarrow 0$ ,  $g_n x_n$  et  $g_n y_n$  appartiennent encore à  $Y$  et tendent vers des éléments  $x_\infty$  et  $y_\infty \in W[x_\infty]$  avec  $x_\infty \neq y_\infty$ . Alors, les propriétés d'équivariance et de continuité devraient pouvoir permettre de montrer que la mesure conditionnelle  $\nu_W(x_\infty)$  possède une invariance par une transformation de  $W[x_\infty]$  qui envoie  $x_\infty$  sur  $y_\infty$  (une translation du groupe, une transformation affine).

Pour garantir que  $g_n x_n$  et  $g_n y_n$  appartiennent à  $Y$ , on emploie des théorèmes ergodiques sur l'action de  $G$ . Le plus difficile est de garantir à la fois que  $g_n x_n$  et  $g_n y_n$  appartiennent à  $Y$  et que la distance  $d(g_n x_n, g_n y_n)$  reste minorée et majorée.

La méthode employée par Eskin et Mirzakhani s'inspire de celle introduite dans [4], dans laquelle ils introduisent de nouveaux degrés de liberté qui en améliorent considérablement la souplesse. Détaillons maintenant leur construction. Nous revenons donc au cas où  $X$  est la strate  $\mathcal{H}_1(\alpha)$ , munie d'une mesure  $P$ -invariante ergodique  $\nu$ . Nous supposons que nous disposons d'une famille mesurable  $A$ -équivariante  $x \mapsto U^+(x)$  de groupes de transformations affines de  $W^+[x]$ , comme dans la proposition 6.7.

On se donne alors des points  $q$  et  $q' \in W^-(q)$  avec  $q' \neq q$ , qui vérifient des propriétés d'équidistribution, découlant du théorème de Birkhoff pour le flot  $A$ , et qui garantissent

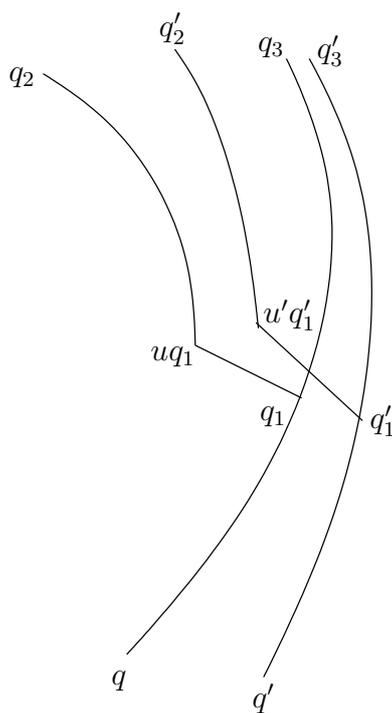


FIGURE 1. La construction de la dérive

que les orbites  $\{a_t q | t \geq 0\}$  et  $\{a_t q' | t \geq 0\}$  passent beaucoup de temps dans des ensembles compacts de grosse mesure où des applications a priori mesurables sont continues (les mesures conditionnelles, la décomposition d'Osseledets du cocycle de Kontsevich-Zorich, etc.). Comme on l'a fait observer dans la remarque 6.6, le fait qu'on puisse trouver deux tels points sur une même feuille de  $W^-$  découle de ce que, comme la mesure est  $N$ -invariante, l'entropie de  $A$  est  $> 0$ .

On laisse s'écouler le flot pendant un temps  $\ell$  (à choisir ultérieurement) et on obtient deux points  $q_1 = a_\ell q$  et  $q'_1 = a_\ell q'$ . On perturbe alors  $q_1$  (resp.  $q'_1$ ) en lui appliquant un petit élément  $u$  (resp.  $u'$ ) de  $U^+(q_1)$  (resp.  $U^+(q'_1)$ ). On a maintenant  $u'q'_1 \notin W^+(uq_1)$ , si bien que le flot a tendance à éloigner ces points. On lance alors le flot pendant un temps  $\tau$  jusqu'à ce que les points  $q_2 = a_\tau uq_1$  et  $q'_2 = a_\tau u'q'_1$  se trouvent à une distance macroscopique l'un de l'autre. Alors des arguments reposant sur les propositions 5.5 et 5.10 permettent de montrer que, en choisissant correctement  $u$  et  $u'$ , on peut garantir que, dans une carte affine contenant  $q_2$  et  $q'_2$ , le vecteur  $q'_2 - q_2$  qui permet de passer de  $q_2$  à  $q'_2$  est en direction très proche de  $E_{[ij],bdd}(q_2)$ , pour un certain  $(i, j)$  (nous utilisons les notations de la section 5). Notons que, pour appliquer ces raisonnements, il faut pouvoir garantir que  $q_1$  et  $q'_1$  appartiennent à des ensembles prescrits de mesure proche de 1 : ce sera possible, à condition de choisir  $\ell$  dans un ensemble de réels qui possède une densité proche de 1.

De même, comme on l’a fait observer plus haut dans la description générale de la méthode, il nous faut pouvoir garantir que les points  $q_2$  et  $q'_2$  appartiennent à des ensembles prescrits de mesure proche de 1. Ici apparaît une idée nouvelle remarquable. Eskin et Mirzakhani observent que le temps  $\tau$  est essentiellement une fonction bilipschitzienne de  $\ell$  (les énoncés précis sur ce point essentiel sont donnés dans [11, Sect. 7]). Par conséquent, demander que  $\tau$  évite un ensemble de temps qui a une densité proche de 0 revient à demander que  $\ell$  évite un ensemble de temps qui a une densité proche de 0.

Dans [11, Sect. 11], Eskin et Mirzakhani introduisent, pour chaque  $(i, j)$ , une relation d’équivalence  $A$ -équivariante  $\mathcal{C}_{ij}$  sur les feuilles de  $W^+$ , dont les atomes sont, en un certain sens, tangents à la distribution  $E_{ij}$ . Le fait que  $q'_2 - q_2$  se rapproche d’un certain  $E_{[ij],bdd}(q_2)$  garantit que, en passant à la limite, on arrivera à construire des points  $\tilde{q}_2$  et  $\tilde{q}'_2$  distincts, qui appartiennent au même atome de  $\mathcal{C}_{ij}$ .

Nous voulons utiliser cette construction pour en déduire que la mesure conditionnelle de  $\nu_{W^+}(\tilde{q}_2)$  le long de  $\mathcal{C}_{ij}$  possède des propriétés d’invariance par des éléments de  $\mathcal{G}_+(\tilde{q}_2)$ . Nous allons pour ce faire construire des points  $q_3$  et  $q'_3$  reliés dynamiquement à  $q_2$  et  $q'_2$  et très proches l’un de l’autre. Pour cela, nous lançons encore la dynamique de  $A$  à partir de  $q_1$  et  $q'_1$  pendant un temps  $t$  : nous arrivons alors en des points  $q_3 = a_t q_1$  et  $q'_3 = a_t q'_1$  qui sont très proches, puisque  $q'_1 \in W^-[q_1]$ .

Dans [11, Prop. 11.4], Eskin et Mirzakhani montrent qu’en choisissant correctement les paramètres dans cette construction, on peut en déduire une propriété d’invariance de la mesure conditionnelle le long de  $\mathcal{C}_{ij}$ . Tâchons de présenter cette idée.

Nous voulons comparer les mesures conditionnelles de  $\mathcal{C}_{ij}$  en  $q_2, q'_2, q_3$  et  $q'_3$ . Pour cela, transportons ces mesures sur les espaces tangents. Pour  $(i, j)$  donné, pour  $\nu$ -presque tout  $x$  dans  $X$ , notons  $\nu_{ij}(x)$  la mesure sur  $W^+(x)$  qui est l’image inverse de la mesure conditionnelle de  $\nu_{W^+}(x)$  le long de  $\mathcal{C}_{ij}(x)$  par l’application  $W^+(x) \rightarrow W^+[x], v \mapsto x + v$ .

Alors, le lien entre la partition  $\mathcal{C}_{ij}$  et la distribution  $E_{[ij],bdd}$  (que nous n’avons pas explicité) implique que la mesure  $\nu_{ij}(q_2)$  est l’image de  $\nu_{ij}(uq_1)$  par une application linéaire qui est essentiellement une similitude de rapport  $\exp(\lambda_{ij}(uq_1, q_2))$ . De même,  $\nu_{ij}(q_3)$  est l’image de  $\nu_{ij}(q_1)$  par une application linéaire qui est essentiellement une similitude de rapport  $\exp(\lambda_{ij}(q_1, q_3))$ . Si  $u$  n’est pas trop grand, on peut garantir que  $\nu_{ij}(uq_1)$  et  $\nu_{ij}(q_1)$  sont proches (il suffit de demander que  $q_1$  et  $uq_1$  appartiennent à un ensemble de mesure proche de 1 où l’application  $\nu_{ij}$  est continue). Comme on peut aussi supposer que  $\lambda_{ij}(q_1, uq_1)$  est uniformément majoré, on en déduit que, si  $t$  est tel que  $\lambda_{ij}(q_2, q_3)$  est borné, la mesure  $\nu_{ij}(q_3)$  est l’image de la mesure  $\nu_{ij}(q_3)$  par une application linéaire dont la norme et la norme de l’inverse sont uniformément bornées. De même, si  $t$  est tel que  $\lambda_{ij}(q'_2, q'_3)$  est borné, la mesure  $\nu_{ij}(q_3)$  est l’image de la mesure  $\nu_{ij}(q_3)$  par une application linéaire dont la norme et la norme de l’inverse sont uniformément bornées. Dans ce cas, comme  $q_3$  et  $q'_3$  sont proches, si on arrive à garantir qu’ils appartiennent à un ensemble de continuité de  $\nu_{ij}$ ,  $\nu_{ij}(q_3)$  est proche de  $\nu_{ij}(q'_3)$  et donc, en passant à la limite  $\nu_{ij}(\tilde{q}'_2)$  est l’image de  $\nu_{ij}(\tilde{q}_2)$  par une application linéaire. Or, par construction,

$\nu_{ij}(\tilde{q}_2)$  est l'image de  $\nu_{ij}(\tilde{q}_2)$  par une application affine qui ne fixe pas 0, puisque ces deux mesures sont les images inverses dans  $W^+(\tilde{q}_2)$  et  $W^+(\tilde{q}'_2)$  de la même mesure sur  $W^+[\tilde{q}_2] = W^+[\tilde{q}'_2]$ . Il en découle que  $\nu_{ij}(\tilde{q}_2)$  est invariante par une transformation affine non triviale, et donc que la mesure conditionnelle le long de  $\mathcal{C}_{ij}(\tilde{q}_2)$  est invariante par une transformation affine non triviale (qui envoie  $\tilde{q}_2$  sur  $\tilde{q}'_2$ ). Notons que, dans cette démonstration, toutes les mesures conditionnelles sont définies à multiplication par une constante près, mais une fois que l'invariance (à une constante près) est établie, des arguments standards permettent d'évacuer cette difficulté.

Revenons sur les hypothèses que nous avons faites : nous avons demandé que  $\lambda_{ij}(q_2, q_3)$  et  $\lambda_{ij}(q_2, q_3)$  restent uniformément bornés et que  $q_3$  et  $q'_3$  appartiennent à un ensemble prescrit de mesure proche de 1. Les auteurs obtiennent la première propriété en faisant en sorte que  $\lambda_{ij}(uq_1, q_2) \approx \lambda_{ij}(u'q'_1, q'_2)$  et que  $\lambda_{ij}(q_1, q_3) \approx \lambda_{ij}(q'_1, q'_3)$ . Si c'est le cas, fixons alors la valeur de  $t$  en demandant que

$$\lambda_{ij}(q_2, q_3) = 0.$$

Alors, d'une part,  $\lambda_{ij}(q_2, q_3)$  est borné et, d'autre part,  $t$  devient une fonction essentiellement bilipschitzienne de  $\tau$ . Comme  $\tau$  dépend de la même manière de  $\ell$ ,  $t$  est une fonction essentiellement bilipschitzienne de  $\ell$ . Demander que  $q_3$  et  $q'_3$  évitent un ensemble de petite mesure revient alors à demander que  $\ell$  évite un ensemble de nombres réels de faible densité, ce qu'on fait.

Il nous reste à donner des indications sur la façon dont les auteurs s'assurent que  $\lambda_{ij}(uq_1, q_2) \approx \lambda_{ij}(u'q'_1, q'_2)$  et que  $\lambda_{ij}(q_1, q_3) \approx \lambda_{ij}(q'_1, q'_3)$ . C'est là une des grandes difficultés de la méthode. Dans l'étude des systèmes dynamiques dans les espaces homogènes, dont Eskin et Mirzakhani s'inspirent, cette difficulté est complètement absente, puisque les cocycles ne dépendent pas du point base. Dans le cas des strates, les auteurs arrivent à résoudre ce problème en utilisant que les cocycles linéaires qui interviennent proviennent de l'action de groupes sur des fibrés plats et donc, que dans des cartes affines, ces cocycles ne dépendent pas du point base. Ces méthodes sont employées dans la démonstration de [11, Prop. 4.4] (qui reprend des idées de Ledrappier [19]) qui permet de contrôler le comportement de la réduction de Zimmer à l'intérieur d'un sous-espace de Lyapunov donné quand on change de point base. Dans [11, Sect. 11], on établit par des méthodes proches des résultats qui permettront de contrôler le comportement de la décomposition de Lyapunov. À titre d'exemple, citons le

LEMME 6.8 ([11, Lem. 11.12]). — *Soit  $M$  une variété différentielle,  $(g_t)$  un flot sur  $M$  et  $E$  un fibré vectoriel plat  $(g_t)$ -équivariant. Supposons  $M$  munie d'une mesure de probabilité borélienne  $(g_t)$ -invariante ergodique et soit  $x \mapsto \mathcal{V}_1(x) \oplus \cdots \oplus \mathcal{V}_r(x)$  la décomposition d'Osseledets de l'action de  $(g_t)$  dans  $E$ , associée aux exposants de Lyapunov  $\lambda_1 > \cdots > \lambda_r$ . Il existe alors  $\alpha, \varepsilon > 0$  ayant la propriété suivante : pour tout  $\delta > 0$ , il existe  $C > 0$  et un ensemble mesurable  $Y \subset X$  de mesure  $\geq 1 - \delta$  tel que, pour tous  $x$  et  $y$  dans  $Y$ , pour tout  $s > 0$ , si  $d(g_t x, g_t y) \leq \varepsilon$  pour  $|t| \leq s$ , alors, pour tout*

$1 \leq i \leq r$ , on a

$$d(\mathbb{P}(\mathcal{V}_i(x)), \mathbb{P}(\mathcal{V}_i(y))) \leq Ce^{-\alpha s}.$$

*Démonstration.* — Nous choisissons  $Y$  de sorte qu'il existe  $\theta > 0$  et  $0 < \rho < \frac{1}{2} \min_{1 \leq i \leq r-1} \lambda_i - \lambda_{i+1}$  tels que, pour tout  $x$  dans  $Y$ , pour tout  $t$  réel, pour tout  $1 \leq i \leq r$ , pour tout  $v$  dans  $\mathcal{V}_i(x)$ , on a

$$\theta \exp((\lambda_i - \rho)t) \|v\| \leq \|g_t v\| \leq \theta^{-1} \exp((\lambda_i + \rho)t)$$

et que, en outre, pour tous  $1 \leq i < j \leq r$  les espaces projectifs  $\mathbb{P}(\mathcal{V}_i(x))$  et  $\mathbb{P}(\mathcal{V}_j(x))$  restent à distance  $\geq \theta$  l'un de l'autre.

Si  $x$  et  $y$  sont comme dans l'énoncé, la structure plate du fibré et le fait que les orbites de  $x$  et de  $y$  restent proches permet d'identifier la fibre  $E_y$  à  $E_x$  (resp.  $E_{g_s y}$  à  $E_{g_s x}$ , resp.  $E_{g_{-s} y}$  à  $E_{g_{-s} x}$ ) de façon à ce que l'action dans ces espaces de  $g_s$  (resp.  $g_{-s}$ ) se lise comme la même application linéaire  $A_s : E_x \rightarrow E_{g_s x}$  (resp.  $A_s : E_x \rightarrow E_{g_s x}$ ). Alors, dans  $\mathcal{V}_1(x)$ ,  $A_s$  multiplie la norme par un facteur au moins de l'ordre de  $\exp((\lambda_1 - \rho)s)$  et, dans  $\mathcal{V}_2(x) \oplus \cdots \oplus \mathcal{V}_r(x)$ , comme dans  $\mathcal{V}_2(y) \oplus \cdots \oplus \mathcal{V}_r(y)$  la norme de  $A_s$  est dominée par  $\exp((\lambda_2 + \rho)s) \ll \exp((\lambda_1 - \rho)s)$ . Par un raisonnement élémentaire d'algèbre linéaire, on en déduit que les espaces projectifs

$$\mathbb{P}(\mathcal{V}_2(x) \oplus \cdots \oplus \mathcal{V}_r(x)) \text{ et } \mathbb{P}(\mathcal{V}_2(y) \oplus \cdots \oplus \mathcal{V}_r(y))$$

sont à distance  $\ll \exp(-\alpha s)$  pour un certain  $\alpha > 0$ . Quitte à diminuer  $\alpha$ , en raisonnant de la même façon avec  $A_{-s}$ , on obtient

$$d(\mathbb{P}(\mathcal{V}_1(x)), \mathbb{P}(\mathcal{V}_1(y))) \ll e^{-\alpha s}.$$

Le cas général s'obtient de manière analogue.  $\square$

Pour terminer cette discussion sur l'argument de dérive, signalons que nous n'avons donné qu'une version incomplète : en effet, rien ne garantit dans nos arguments que la transformation affine qui préserve la mesure  $\mathcal{C}_{ij}(\tilde{q}_2)$  n'appartient pas au groupe  $U^+(\tilde{q}_2)$ . Pour obtenir cette propriété supplémentaire il nous faut utiliser une hypothèse de la proposition 6.7, à savoir le fait que  $\mathcal{L}^+(x)$  n'est pas contenu dans  $U^+(x)$ .

Si  $U^+(x)$  est un sous-espace vectoriel de  $W^+(x)$ , il nous faut reprendre les constructions en faisant en sorte que, dans la preuve, ce ne soit pas la distance entre  $q'_2$  et  $q_2$  qui devienne macroscopique, mais la distance entre  $q'_2$  et l'espace affine  $q_2 + U^+(q_2) \subset W^+[q_2]$ . Un peu de travail est alors nécessaire, dans la mesure où la distribution  $U^+ \subset W^+$  ne possède pas nécessairement de supplémentaire  $A$ -équivariant. Il faut donc lui choisir une transversale et tenir compte de possibles déviations par rapport à cette transversale

Dans le cas général, on dispose simplement d'une distribution de sous-algèbres de Lie  $\mathfrak{u}^+$  de la distribution  $\mathfrak{g}_+$  des algèbres de Lie de  $\mathcal{G}_+$ , et nous devons garantir que  $q'_2$  se trouve à distance minorée de  $U^+(q_2)q_2 \subset W^+[q_2]$ . Dans toute la preuve, il faut alors remplacer les estimations sur les exposants de Lyapunov, la réduction de Zimmer, etc., du fibré  $W^+$  par les objets analogues dans un supplémentaire de  $\mathfrak{u}^+$  dans  $\mathfrak{g}^+$ . Un grand

nombre des difficultés posées par cette extension sont traitées par les résultats de [11, Sect. 6].

### 6.6. L'argument entropique

Nous indiquons brièvement ici comment la proposition 6.5 permet de conclure la démonstration du théorème 6.4. D'après cette proposition, nous disposons de distributions  $AN$ -équivariantes de sous-espaces vectoriels  $\mathcal{L}^- \subset W^-$  et  $\mathcal{L}^+ \subset W^+$  telles que, pour  $\nu$ -presque tout  $x$ ,  $\nu_{W^-}(x)$  est concentrée sur  $\mathcal{L}^-(x) + x \subset W^-[x]$  et  $\nu_{W^+}(x)$  est  $\mathcal{L}^+(x)$ -invariante. Comme au début du paragraphe 6.4, nous notons  $\mathcal{L}_\perp \subset H_\perp^1$  la distribution  $A$ -équivariante telle que  $\mathcal{L}^- \cap \mathbb{R}e_2 \otimes H_\perp^1 = \mathbb{R}e_2 \otimes \mathcal{L}_\perp$ . Enfin, notons, pour  $\nu$ -presque tout  $x$ ,  $U(x) \supset \mathcal{L}_\perp(x)$  le sous-espace vectoriel de  $H_\perp^1(x)$  tel que  $\mathbb{R}e_1 \otimes e_1 \oplus \mathbb{R}e_1 \otimes U(x)$  soit la composante connexe du stabilisateur de  $\nu_{W^+}(x)$  dans  $W^+(x)$  ( $\mathbb{R}e_1 \otimes e_1$  stabilise  $\nu_{W^+}(x)$  car c'est la direction tangente à l'action de  $N$  et  $N$  préserve  $\nu$ ).

Notons  $(\lambda_i)_{i \in I}$  les exposants de Lyapunov du cocycle de Kontsevich-Zorich dans  $U$  comptés avec multiplicité, et soit  $J \subset I$  tel que les exposants de Lyapunov dans  $\mathcal{L}^+$  soient  $(\lambda_i)_{i \in J}$ . D'après [11, Th. A.3] (qui s'inspire d'idées de Forni [13], reprises par Forni, Matheus et Zorich [14]), comme la distribution  $U$  est  $AN$ -équivariante, on a

$$\sum_{i \in I} \lambda_i \geq 0.$$

Nous pouvons estimer l'entropie de  $A$  pour la mesure  $\nu$  en termes des mesures conditionnelles de  $\nu$  le long des feuilletages stables et instables. Ce calcul provient de la formule d'entropie de Ledrappier-Young [20, 21], utilisée pour des problèmes de rigidité par Margulis et Tomanov [23] et adaptée ici par Eskin et Mirzakhani [11, Th. B.7]. Le principe de cette formule est que l'entropie est la somme des exposants de Lyapunov positifs (ou négatifs) multipliés par une dimension de la mesure le long de feuilletages associés aux exposants. Dans ce cadre, cette formule nous donne, vu que, pour  $\nu$ -presque tout  $x$  dans  $X$ ,  $\nu_{W^-}(x)$  est concentrée sur  $\mathcal{L}^-(x) + x$ ,

$$h(a_1, \nu) \leq 2 + \sum_{i \in J} (1 - \lambda_i) \leq 2 + \sum_{i \in I} (1 - \lambda_i) \leq 2 + |I|$$

(où nous avons utilisé que les  $\lambda_i$  sont  $\leq 1$  et que  $\sum_{i \in I} \lambda_i \geq 0$ ). Mais, comme, pour  $\nu$ -presque tout  $x$ ,  $\nu_{W^+}(x)$  est  $U(x)$ -invariante et  $N$ -invariante, la même formule implique que

$$h(a_1, \nu) \geq 2 + \sum_{i \in I} (1 + \lambda_i) \geq 2 + |I|.$$

Toutes les inégalités sont donc des égalités. En particulier, on a  $\mathcal{L}_\perp = U$ ,  $\mathbb{R}e_2 \otimes e_2 \subset \mathcal{L}^-$  et le cas d'égalité dans la formule de Ledrappier-Young implique que, pour  $\nu$ -presque tout  $x$  dans  $X$ ,  $\nu_{W^-}(x)$  est invariante par les translations de  $\mathcal{L}^-(x)$ . La mesure  $\nu$  est donc  $N_-$ -invariante (où  $N_-$  est le groupe des matrices triangulaires inférieures unipotentes dans  $SL_2(\mathbb{R})$ ) et, comme  $N$  et  $N_-$  engendrent  $SL_2(\mathbb{R})$ ,  $\nu$  est  $SL_2(\mathbb{R})$ -invariante. De même, la distribution  $\mathcal{L}^\perp$  est à la fois  $N_-$  et  $N$ -équivariante : elle est donc  $SL_2(\mathbb{R})$ -équivariante.

### 6.7. Marches aléatoires

Évoquons brièvement la fin de la démonstration du théorème 3.4, une fois que le théorème 6.4 est établi. Ces arguments sont développés dans [11, Sect. 14-16].

On dispose d'une mesure de probabilité  $P$ -invariante et ergodique  $\nu$  sur  $X = \mathcal{H}_1(\alpha)$ . On cherche à montrer que  $X$  est affine. D'après le théorème 6.4, on sait que  $\nu$  est  $\mathrm{SL}_2(\mathbb{R})$ -invariante et on dispose d'une distribution mesurable  $\mathcal{L}$  de sous-espaces vectoriels de  $T\mathcal{H}(\alpha)$ , définie  $\nu$ -presque partout, qui est  $\mathrm{SL}_2(\mathbb{R})$ -équivariante, et telle que, pour  $\nu$ -presque tout  $x$ , la mesure conditionnelle  $\nu_{W^+}(x)$  est la mesure de Lebesgue de  $x + (\mathcal{L}(x) \cap W^+(x))$ . La distribution  $\mathcal{L}$  est le candidat à être la distribution tangente du sous-espace affine support de  $\nu$ .

Pour montrer que  $\nu$  est affine, il suffit de montrer qu'il existe un ensemble de  $x$  de mesure  $> 0$  pour lesquels il existe un ouvert  $\Omega$  de  $\mathcal{L}(x)$  avec  $\nu(\Omega + x) > 0$ . En supposant le contraire, Eskin et Mirzakhani arrivent à construire, dans tout ensemble de mesure  $> 0$ , des éléments  $x$  et  $y$  avec  $y \in W^+[x]$  mais  $y \notin x + (W^+(x) \cap \mathcal{L}(x))$ , ce qui est contradictoire.

Cette construction repose sur un argument de dérive analogue au précédent, mais pour un système dynamique différent.

Construisons ce système dynamique. On fixe une mesure de probabilité  $\mathrm{SO}(2)$ -invariante à droite et à gauche  $\mu$  sur  $\mathrm{SL}_2(\mathbb{R})$  qui est à support compact et de la classe de Lebesgue. On note  $B_+$  (resp.  $B_-$ ) l'espace des suites  $(b_0, b_1, \dots)$  (resp.  $\dots, b_{-2}, b_{-1}$ ) d'éléments de  $\mathrm{SL}_2(\mathbb{R})$ , qu'on munit de la mesure produit  $\beta_+ = \mu^{\otimes \mathbb{N}}$  (resp.  $\beta_- = \mu^{\otimes \mathbb{Z}^-}$ ). On note  $B = B_- \times B_+$  l'espace des suites indexées par  $\mathbb{Z}$ , qu'on munit aussi de la mesure produit  $\beta = \mu^{\otimes \mathbb{Z}}$ . Enfin, on note  $T : B \rightarrow B$  le décalage des suites.

Notons que si  $b = (b_-, b_+)$  est un élément de  $B$ , on peut considérer l'ensemble des suites de la forme  $(b_-, a)$  avec  $a$  dans  $B_+$  comme une feuille instable locale de  $b$  pour l'action de  $T$  sur  $B$ . L'idée de la seconde étape de la démonstration consiste à faire jouer à cette feuille instable locale le rôle joué par  $U^+$  dans la première étape.

Plus précisément, introduisons le système dynamique

$$T^X : B \times X \rightarrow B \times X, (b, x) \mapsto (Tb, b_0x).$$

La transformation mesurable  $T^X$  préserve la mesure produit  $\beta \otimes \nu$ , qui est ergodique par des raisonnements généraux.

La deuxième étape de la démonstration s'établit alors par un procédé analogue à celui qui permet d'obtenir la première, en remplaçant les transformations  $q_1 \rightarrow uq_1$  et  $q'_1 \rightarrow u'q'_1$  par des transformations du type  $q_1 = (b_1, x_1) \rightarrow ((b_{1,-}, a), x)$  et  $q_1 = (b_1, x'_1) \rightarrow ((b_{1,-}, a), x')$  avec  $a$  dans  $B_+$ .

Certaines des difficultés qui sont apparues pendant la première étape (nécessité d'introduire le sous-espace inerte, les sous-espaces bornés, etc.) ne se présentent pas, car le cocycle de Kontsevich-Zorich possède de meilleures propriétés ergodiques au-dessus de mesures  $\mathrm{SL}_2(\mathbb{R})$ -invariantes qu'au-dessus des mesures  $A$ -invariantes. La démonstration utilise en particulier le

THÉORÈME 6.9 ([11, Th. A.6]). — *Soit  $\nu$  une mesure de probabilité  $\mathrm{SL}_2(\mathbb{R})$ -invariante et ergodique sur  $X = \mathcal{H}_1(\alpha)$ . Alors l'adhérence de Zariski du cocycle de Kontsevich-Zorich dans  $H^1(S, \mathbb{R})$  au-dessus de  $\nu$  est un groupe semi-simple.*

Il n'est pas clair que l'adhérence de Zariski du cocycle dans  $H^1(S, \Sigma, \mathbb{R}^2)$  soit semi-simple. Pour pallier les problèmes posés par cette difficulté, les auteurs utilisent un résultat sur la géométrie des mesures invariantes. On note  $\mathcal{F} \subset H^1(S, \mathbb{R})$  la somme des distributions  $\mathrm{SL}_2(\mathbb{R})$ -équivariantes définies  $\nu$ -presque partout où tous les exposants de Lyapunov de  $A$  sont nuls.

THÉORÈME 6.10 (Avila, Eskin, Möller [2]). — *Soit  $\nu$  une mesure de probabilité  $\mathrm{SL}_2(\mathbb{R})$ -invariante et ergodique sur  $X = \mathcal{H}_1(\alpha)$ . Alors, il existe un ensemble  $Y \subset X$  de mesure 1 pour  $\nu$  tel que, pour tout  $x$  dans  $Y$ , pour tout  $y$  dans  $Y$  suffisamment proche de  $x$ , on a  $p(y - x) \in \mathbb{R}^2 \otimes \mathcal{F}(x)^\perp$ .*

Dans cette formule,  $\mathcal{F}(x)^\perp$  désigne l'orthogonal de  $\mathcal{F}(x)$  pour la forme symplectique  $\varpi$ .

## RÉFÉRENCES

- [1] J. S. ATHREYA – *Quantitative recurrence and large deviations for Teichmüller geodesic flow*, *Geom. Dedicata* **119** (2006), 121–140.
- [2] A. AVILA, A. ESKIN, M. MÖLLER – *Symplectic and isometric  $\mathrm{SL}(2, \mathbb{R})$ -invariant subbundles of the Hodge bundle*, arXiv :1209.2854.
- [3] R. AZENCOTT – *Espaces de Poisson des groupes localement compacts*, *Lecture Notes in Math.*, Vol. 148, Springer-Verlag, Berlin-New York, 1970, ix+141 pp.
- [4] Y. BENOIST, J.-F. QUINT – *Mesures stationnaires et fermés invariants des espaces homogènes*, *Ann. of Math. (2)* **174** (2011), 1111–1162.
- [5] Ph. BOUGEROL, J. LACROIX – *Products of random matrices with applications to Schrödinger operators*, *Progress in Probability and Statistics* 8, Birkhäuser Boston, Inc., Boston, MA, 1985, xii+283 pp.
- [6] S. G. DANI, G. A. MARGULIS – *Asymptotic behaviour of trajectories of unipotent flows on homogeneous spaces*, *Proc. Indian Acad. Sci. Math. Sci.* **101** (1991), 1–17.
- [7] M. EINSIEDLER, A. KATOK – *Invariant measures on  $G/\Gamma$  for split simple Lie groups  $G$* , *Comm. Pure Appl. Math.* **56** (2003), 1184–1221.
- [8] M. EINSIEDLER, A. KATOK, E. LINDENSTRAUSS – *Invariant measures and the set of exceptions to Littlewood's conjecture*, *Ann. of Math. (2)* **164** (2006), 513–560.
- [9] A. ESKIN, G.A. MARGULIS, S. MOZES – *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, *Ann. of Math. (2)* **147** (1998), 93–141.

- [10] A. ESKIN, H. MASUR – *Asymptotic formulas on flat surfaces*, Ergodic Theory Dynam. Systems **21** (2001), 443-478.
- [11] A. ESKIN, M. MIRZAKHANI – *Invariant and stationary measures for the  $SL(2, \mathbb{R})$  action on moduli space*, preprint, 2013.
- [12] A. ESKIN, M. MIRZAKHANI, A. MOHAMMADI – *Isolation, equidistribution and orbit closures for the  $SL(2, \mathbb{R})$  action on moduli space*, preprint, 2013.
- [13] G. FORNI – *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Ann. of Math. (2) **155** (2002), 1-103.
- [14] G. FORNI, C. MATHEUS, A. ZORICH – *Lyapunov spectrum of invariant subbundles of the Hodge bundle*, Ergodic Theory Dynam. Systems **34** (2014), 353-408.
- [15] H. FURSTENBERG – *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428.
- [16] H. FURSTENBERG – *A Poisson formula for semi-simple Lie groups*, Ann. of Math. (2) **77** (1963), 335–386.
- [17] A. KATOK, R. SPATZIER – *Invariant measures for higher-rank hyperbolic abelian actions*, Ergodic Theory Dynam. Systems **16** (1996), 751-778.
- [18] M. KONTSEVICH, A. ZORICH – *Lyapunov exponents and Hodge theory*, arXiv :hep-th/9701164.
- [19] F. LEDRAPPIER – *Positivity of the exponent for stationary sequences of matrices*, Lyapunov exponents (Bremen, 1984), 56-73, Lecture Notes in Math. 1186, Springer, Berlin, 1986.
- [20] F. LEDRAPPIER, L.-S. YOUNG – *The metric entropy of diffeomorphisms I, Characterization of measures satisfying Pesin’s entropy formula*, Ann. of Math. (2) **122** (1985), 509-539.
- [21] F. LEDRAPPIER, L.-S. YOUNG – *The metric entropy of diffeomorphisms II, Relations between entropy, exponents and dimension*, Ann. of Math. (2) **122** (1985), 540-574.
- [22] E. LINDENSTRAUSS – *Invariant measures and arithmetic quantum unique ergodicity*, Ann. of Math. (2) **163** (2006), 165-219.
- [23] G. A. MARGULIS, G. M. TOMANOV *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. **116** (1994), 347–392.
- [24] H. MASUR – *Lower bounds for the number of saddle connections and closed trajectories of a quadratic differential*, Holomorphic functions and moduli, Vol. I (Berkeley, CA, 1986), 215-228, Math. Sci. Res. Inst. Publ., 10, Springer, New York, 1988.
- [25] H. MASUR – *The growth rate of trajectories of a quadratic differential*, Ergodic Theory Dynam. Systems **10** (1990), 151-176.
- [26] C. McMULLEN – *Dynamics of  $SL_2(\mathbb{R})$  over moduli space in genus two*, Ann. of Math. (2) **165** (2007), 397–456.

- [27] S. MEYN, R. TWEEDIE – *Markov chains and stochastic stability*, Second edition, Cambridge Univ. Press, Cambridge, 2009, xxviii+594 pp.
- [28] E. NUMMELIN – *General irreducible Markov chains and nonnegative operators*, Cambridge Tracts in Mathematics 83, Cambridge Univ. Press, Cambridge, 1984, xi+156 pp.
- [29] M. RATNER – *On Raghunathan’s measure conjecture*, Ann. of Math. (2) **134** (1991), 545–607.
- [30] M. RATNER – *Raghunathan’s topological conjecture and distributions of unipotent flows*, Duke Math. J. **63** (1991), 235–280.
- [31] M. SEPPALA, T. SORVALI – *Geometry of Riemann surfaces and Teichmüller spaces*, North-Holland Mathematics Studies 169, North-Holland Publishing Co., Amsterdam, 1992, ii+263 pp.
- [32] J. M. STEELE – *Kingman’s subadditive ergodic theorem*, Ann. Inst. H. Poincaré Probab. Statist. **25** (1989), 93–98.
- [33] W. VEECH – *Moduli spaces of quadratic differentials*, J. Analyse Math. **55** (1990), 117–171.
- [34] J.-C. YOCCOZ – *Interval exchange maps and translation surfaces*, Homogeneous flows, moduli spaces and arithmetic, 1-69, Clay Math. Proc., **10**, Amer. Math. Soc., Providence, RI, 2010.
- [35] R. ZIMMER – *Ergodic theory and semisimple groups*, Monographs in Mathematics 81, Birkhäuser Verlag, Basel, 1984, x+209 pp.

Jean–François QUINT

Université de Bordeaux

Institut de Mathématiques de Bordeaux

351 cours de la Libération

F–33405 Talence Cedex

*E-mail* : [jquint@math.u-bordeaux1.fr](mailto:jquint@math.u-bordeaux1.fr)